Correlations between MO Eigenvectors and the Thermochemistry of Simple Organic Molecules, Related to Empirical Bond Additivity Schemes

A thesis
submitted in partial fulfilment
of the requirements for the degree
of
Doctor of Philosophy in Chemistry
at
The University of Waikato
by
Matthew C.J. Lee

The University of Waikato
2008

Abstract

A bondingness term is further developed to aid in heat of formation ($\Delta_f H^o$) calculations for C, N, O and S containing molecules. Bondingness originated from qualitative investigations into the antibonding effect in the occupied MOs of ethane. Previous work used a single parameter for bondingness to calculate $\Delta_f H^o$ in an alkane homologous series using an additivity scheme. This work modifies the bondingness algorithm and uses the term to parameterise a test group of 345 molecules consisting of 17 subgroups that include alkanes, alkenes, alkynes, alcohols, ethers, aldehydes, ketones, carboxylic acids, esters, amines, amides, diazenes, nitriles, nitroalkanes, nitrates, thiols and benzenoids. Comparing experimental with calculated $\Delta_f H^o$ values, a standard deviation for the residuals of 6.3 kJ mol^{-1} can be achieved using bondingness with a simple steric repulsion term (SSR) in a bond additivity scheme, and a standard deviation of 5.2 kJ mol^{-1} can be achieved using a Lennard-Jones potential. The method is compared with the group method of Pedley, which for a slightly smaller set of 338 molecules, a subset of the test set of 345 molecules, gives a standard deviation of 7.0 kJ mol^{-1}.

Bondingness, along with SSR or a Lennard-Jones potential, is parameterised in the lowest level of *ab initio* (HF-SCF) or semiempirical quantum chemical calculations. It therefore may be useful in determining the $\Delta_f H^o$ values for the largest molecules that are amenable to quantum chemical calculation.

As part of our analysis we calculated the difference between the lowest energy conformer and the average energy of a mixture populated with higher energy conformers. This is the difference between the experimental $\Delta_f H^o$ value and the $\Delta_f H^o$ calculated for a single conformer. Example calculations which we have followed are given by Dale and Eliel *et al.*. Dale calculates the energy difference for molecules as large as hexane using relative energies based on the number of 1,4 *gauche* interactions. We have updated these values with constant increments ascertained by Klauda *et al.* as well as *ab initio* MP2 cc-pVDZ relative energies and have included calculations for heptane and octane.

Acknowledgements

The author wishes to thank his supervisor
Derek W. Smith
for his intrinsic role in this work.

My parents Wallace and Nancy have been the greatest support and best cheerleaders. In fact without them I would not have even bothered because it would have seemed unfeasible. If I may quote Bette Midler and say you're the wind beneath my wings. I can't forget the rest of my family: my brothers and sisters who recognising me as a poor student have been generous on many occasions. Thanks for the 2 gig storage device Richard.

Some friends have been very helpful as well, and if I may continue to speak in the language of one receiving a prestigious award then I must say it like this.

Some of my friends whose help was immediately associated with the production of this thesis are: Andrew Bryan for his computer support. Thanks for teaching me PERL buddy, and writing the first version of the state manager, and sorting out that crazy OLE storage system that Spartan used, as well as those books; Michael Hallet, for lending me his PERL books and suggestions, and Wendy Bryce, thanks for proof reading this whole thing you did a great job.

Thank you God for my parents and friends and everything really. Looking forward to Your return Jesus, as Lord and King, and in power and majesty this time!

Table of Contents

List of Tables, Graphs and Figures

List of Abbreviations

5Z	quintuple zeta (basis set)
AM1	Austin model 1
AO	atomic orbital
AOMO	the sum of $B*(\pi)$ over all the OMOs
a.u.	atomic units
B	bondingness
B*	antibondingness
$B(\pi)$	π bondingness
$B*(\pi)$	π antibondingness
$B(\sigma)$	σ bondingness
$B*(\sigma)$	σ antibondingness
BO	bond order
CASSCF	complete active-space self-consistent field
CC	coupled cluster
CI	configuration interaction
CRSE	conventional ring strain energy
CSE	conventional strain energy
CSF	configuration state function
$\Delta_f H^o$	heat of formation (formation enthalpy)
DFT	density functional theory
ΔH_a	heat of atomisation (atomisation enthalpy)
DZ	double zeta (basis set)
E(A-B)	thermochemical bond energy term between bonded atoms A and B
EHT	extended Hückel theory
G2	Gaussian-2
G3	Gaussian-3
GGA	generalised gradient approximation
GTO	Gaussian type orbital
HAOMO	highest antibondingness occupied molecular orbital
HF	Hartree-Fock
HOMO	highest occupied molecular orbital
$H_{vib}(0)$	zero-point vibrational energy
LCAO	linear combination of atomic orbitals

LDA	local density approximation
LSDA	local spin density approximation
LSE	least squares estimate
LUMO	lowest unoccupied molecular orbital
MCSCF	multiconfiguration self-consistent field
MM	molecular mechanics
MMFF94	Merck Molecular Force Field 1994
MO	molecular orbital
MINDO/3	modified intermediate neglect of diatomic overlap
MNDO	modified neglect of diatomic overlap
MP2	Møller-Plesset second-order perturbational theory
MP3	Møller-Plesset third-order perturbational theory
MP4	Møller-Plesset fourth-order perturbational theory
MRCI	multireference configuration interaction
nb-B(π)	nonbonded B(π)
OMO	occupied molecular orbital
PDDG	pairwise distance directed Gaussian
PES	potential energy surface
PM3	MNDO, Parametric Method Number 3 (MNDO-PM3)
QZ	quadruple zeta (basis set)
r^2	coefficient of determination
RASSCF	restricted active-space self-consistent field
[S]	conventional steric correction term (strain energy)
SCF	self-consistent field
SE	steric energy
SI	International System of Units
sp^2sp^3	number of sp^2 to sp^3 bonded carbons
SSR	simple steric repulsion
STO	Slater type orbital
SV	split valence (basis set)
TZ	triple zeta (basis set)
$\lvert\bar{x}\rvert$	average error
ZPE	zero-point energy

1 Introduction

The energy of a molecule may be defined by the energy difference for the molecule relative to its constituents in defined states. Three commonly used terms are the enthalpy of formation ($\Delta_f H^o$), the atomisation enthalpy (ΔH_a) and the total energy (ΔE_{Tot}). The $\Delta_f H^o$ is the difference between all the molecule's atoms in the elemental standard state as reactants, and those same atoms bound as the molecule in the state specified at 298 K. For the common states: $\Delta_f H^o$(g) for the gaseous state, $\Delta_f H^o$(s) for the solid state and $\Delta_f H^o$(l) for the liquid state. Also morphologies of a species must be specified *e.g.* diamond or graphite for the solid state of carbon. The ΔH_a is the energy difference between the molecule in the specified state as the reactant, and all of its constituent atoms in the gaseous state at 298 K. The ΔE_{Tot} is the difference between the molecule's atoms completely ionised at infinite separation, from each other and their electrons, as the reactants and the nonvibrating molecule at 0 K. The $\Delta_f H^o$ and ΔH_a are easily interconvertible. However the ΔE_{Tot} is calculated within the Born-Oppenheimer approximation. This assumes that the nuclear coordinates can be fixed, because the speed of the electrons is much faster than the motions of the vibrating nuclei and therefore the electronic energy adjusts to nuclear displacements very quickly. The molecule is however vibrating at 0 K. The difference between the vibrating and theoretical nonvibrating molecule is the zero-point vibrational energy [$H_{vib}(0)$]. Also the thermal energy [$\Delta H(T)$] at 298 K must be corrected for. Both $H_{vib}(0)$ and $\Delta H(T)$ are difficult to calculate precisely but can be approximated by the parameterisation of an empirical bond or group additivity model[1]. Enthalpy terms are defined at constant pressure as a diathermal process, and thus the difference between them and ΔE_{Tot} values must be adjusted with a pressure and volume term (PV) where $\Delta H = \Delta E + P\Delta V$.

1.1 Molecular Mechanics

Experimental $\Delta_f H^o$ values are not available for the majority of chemical compounds, and a value must be estimated from experimental $\Delta_f H^o$ values of closely related compounds. The $\Delta_f H^o$ can then be calculated empirically, either by an additivity scheme using group methods[2-32] or with a molecular mechanics force field[33]. Force fields have their origin in vibrational analysis where the force fields are generally more rigorous than those of molecular mechanics. Many terms and principles are common to both fields of study. The principal problem in vibrational analysis is the determination of the force field from vibrational frequency data. Molecular mechanics (MM) uses force fields based on mainly other sources, optimised to reproduce molecular geometry and other chemical properties. To understand MM it is useful to consider some basic principles in the construction of the F matrix (matrix of force constants) used in vibrational analysis. The vibrational potential energy about the position of minimum energy for a diatomic molecule can be represented by a Maclaurin series shown in *eq.* (1).

$$V = V_o + \left(\frac{\mathrm{d}V}{\mathrm{d}q}\right)_0 q + \frac{1}{2}\left(\frac{\mathrm{d}^2 V}{\mathrm{d}q^2}\right)_0 q^2 + \frac{1}{6}\left(\frac{\mathrm{d}^3 V}{\mathrm{d}q^3}\right)_0 q^3 + \cdots \qquad \ldots(1)$$

Here q is the internal coordinate of a diatomic molecule. In this instance the coordinate is a bond-stretching coordinate. The internal coordinate in terms of Cartesian coordinates has the origin placed at the centre of mass between atom A and atom B, where atoms are coincident with the z axis. The equilibrium positions of atoms A and B are represented by z_{A} and z_{B} respectively. If the change in atomic positions is given by z'_{A} and z'_{B} then $q = (z'_{\mathrm{A}} - z'_{\mathrm{B}}) - (z_{\mathrm{A}} - z_{\mathrm{B}})$. The system of coordinates used is optional and not all force fields use internal coordinates. Indeed there is an advantage in using Cartesian coordinates over internal coordinates for cyclic molecules. The constant V_o is the minimum potential which can be set at zero to ascertain vibrational energy only. The subscript zero indicates a value at the position of minimum potential energy where from the definition of a potential minimum the first derivative is zero $[(\mathrm{d}V/\mathrm{d}q)_o = 0]$. If we presume molecular vibrations are only small deviations from equilibrium bond lengths, we can assume the potential energy is closely approximated by a harmonic potential, and ignore third order derivatives and higher. This is the harmonic approximation. In this approximation the force constants are set equal to the second order derivative, and are said to be quadratic *i.e.* $f = \mathrm{d}^2V/\mathrm{d}q^2$. The harmonic approximation applied to the general molecule then gives *eq.* (2).

$$2V = \left(\frac{\partial^2 V}{\partial q_1^2}\right)_0 q_1^2 + \left(\frac{\partial^2 V}{\partial q_2^2}\right)_0 q_2^2 + \left(\frac{\partial^2 V}{\partial q_3^2}\right)_0 q_3^2 \cdots$$

$$+\left(\frac{\partial^2 V}{\partial q_1 \partial q_2}\right)_0 q_1 q_2 + \left(\frac{\partial^2 V}{\partial q_2 \partial q_1}\right)_0 q_2 q_1 + \left(\frac{\partial^2 V}{\partial q_1 \partial q_3}\right)_0 q_1 q_3 \cdots$$

$$2V = \sum_i \left(\frac{\partial^2 V}{\partial q_i^2}\right) q_i^2 + 2\sum_{i<j}\left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right) q_i q_j \qquad \ldots(2)$$

In *eq.* (2) internal coordinates represent the change in relative position of atoms with respect to bond lengths or inter-bond angles, for which if there are *n* atoms in a molecule there are 3*n* – 6 internal coordinates for nonlinear, and 3*n* – 5 internal coordinates for linear molecules. The internal coordinates (vector **q**) can be ascertained from the transformation from Cartesian coordinates by left multiplication of the appropriate *n* by 3*n* – (6 or 5) matrix (**B**) with the 3*n* column vector of Cartesian displacement coordinates (**x**) shown in *eq.* (3). If the derivatives are replaced with the force constants related by *eq.* (4) then for the nonlinear case the F matrix is given by *eq.* (5).

$$\mathbf{q} = \mathbf{Bx} \qquad \ldots(3)$$

$$f_{\mathrm{ij}} = \left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right) \qquad \ldots(4)$$

$$\mathbf{F} = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1,3n-6} \\ f_{21} & f_{22} & & \\ \vdots & & \ddots & \\ f_{3n-6,1} & & & f_{3n-6,.3n-6} \end{pmatrix} \qquad \ldots(5)$$

By matrix notation the F matrix is related to potential energy by *eq.* (6).

$$2V = \mathbf{q'Fq} \qquad \ldots(6)$$

Vibrational frequencies are ascertained by solving the secular equation, which is derived from the classical equation of motion which also involves kinetic energy

(*T*). This requires that kinetic and potential energy be expressed with a common coordinate system. The Wilson GF formalism expresses the secular determinant in terms of the eigenvalues of the vector $\boldsymbol{\lambda}$ and the matrices **F** and **G** shown in *eq.* (8). Where **G** is given by *eq.* (7) expressed in terms of the matrix **B** of *eq.* (3) and **M,** which is a diagonal matrix of the mass of the atom associated with the corresponding coordinate. Each eigenvalue λ_i of the vector $\boldsymbol{\lambda}$ is related to frequency by $\lambda_i = 4\pi^2\nu_i^2$.

$$\mathbf{G} = \mathbf{B}\mathbf{M}^{-1}\mathbf{B}' \qquad \ldots(7)$$

$$|\mathbf{F}\mathbf{G} - \boldsymbol{\lambda}\mathbf{I}| = 0 \qquad \ldots(8)$$

Determination of the force constants from vibrational frequencies is difficult, as there is a greater proportion of force constants to vibrational frequencies *i.e.* for a nonlinear molecule there are $\frac{1}{2}(3n - 6)(3n - 5)$ force constants and $(3n - 6)$ frequencies per molecule. Therefore, complete determination of an F matrix requires analysis of many isotopically substituted spectra.

Systematic simplifications have been developed to minimise the number of off-diagonal terms used in the F matrix. In the central force field, which is specified only in terms of interatomic distances, off-diagonal terms (cross terms) are usually ignored. The diagonal force field created results in a model that does not distinguish between bonded and nonbonded interactions. This model is in accord with the ionic model of bonding, but inadequate for molecules with covalent bonds. If the force field is specified in terms of torsional angles and interatomic angles as well as interatomic distances, a valence force field is then specified. If cross terms are ignored, the diagonal terms of the F matrix can be determined from a set of closely related molecules. However transferability of these force constants is limited to closely related molecules due to the neglect of cross terms. It has been shown by Schachtschneider and Snyder that some off diagonal terms are more important than others *viz.* cross terms are bigger when two internal coordinates end on a common atom or nearest neighbour atom[34]. Force fields extended in this way still lack nonbonded interactions. These include London or van der Waals forces, hyperconjugation, dipole interactions, electronegativity, and electrostatic or coulombic interactions.

It has been suggested by Hagler[35] that class 1 force fields contain only harmonic terms and only explicit diagonal elements in the force constant matrix, class 2 adds cubic and higher terms and explicit off-diagonal elements in the force constant matrix, and Allinger *et al.*[36] propose that class 3 should have the mechanical terms of a class 2 as well as chemical phenomena like electronegativity and hyperconjugation.

A force field specifies among other things, potential energy. When the potential energy is related to nuclear positions of the molecule, a potential energy surface (PES) can be defined. The PES is a multidimensional nonlinear function of the molecular nuclear positions. Molecular geometry is ascertained by repeated energy calculations and subsequent geometry adjustment by steepest decent techniques to arrive at a minimum on the potential energy surface. This process is called geometry optimisation. If a geometry is specified, then only one energy calculation at the specified nuclear positions is undergone, and this is referred to as a single point energy calculation. Geometry optimisation may be viewed as a number of successive single point energy calculations to minimise the steric energy by adjusting the molecular geometry. In contrast, spectroscopic analysis uses force fields optimised to reproduce experimental vibrational frequencies. The sum of the potential energy functions at the resultant molecular geometry gives a steric energy (SE).

In this work we distinguish steric energy calculated by MM from strain energy, where strain energy is ascertained by conventional methods in accord with a particular definition, for which there are many. For instance, a definition may nominate *n*-alkanes, isobutane and neopentane as strainless reference standards. From the reference standards, strainless increments for C-H and C-C bonds and primary, tertiary and quaternary carbons are ascertained by least squares fit (a secondary term is redundant). From the sum of strainless increments in a molecule, a theoretical unstrained energy is ascertained, from which the difference with the actual value may give the strain energy. Alternatively, the chair conformation of cyclohexane may be the reference standard for the CH_2 increment in cyclic compounds, from which a conventional ring strain energy may be determined in other cyclic alkanes. Other strain energies may be ascertained by the sum of gauche 1,4 interactions in a molecule and appropriate conversion factor[6] (we use the conversion factor -2.5 kJ mol^{-1}). A more complex scheme based on a similar concept is proposed by Skinner[5]. All these methods may be termed conventional strain energy (CSE).

The experimental $\Delta_f H^o$ values represent the average value of a mixture of the conformers present. The mole fractions of each conformer by statistical mechanical analysis is in accordance with a Boltzmann distribution calculated by *eq.* (9).

$$N_i = \frac{g_i e^{-\Delta G_i / \mathrm{RT}}}{\sum_i g_i e^{-\Delta G_i / \mathrm{RT}}} \qquad \ldots(9)$$

Here N_i is the mole fraction of the i^{th} conformer, g_i is a statistical weighting for stereo isomers with identical energy, ΔG_i is the Gibbs free energy excluding entropy effects, R is the universal gas constant and T is temperature in Kelvins. However the energy of each conformer at 298 K is represented by the average value of a mixture of that conformer in different excited vibrational states. The vibrational energy U_{vib} including $H_{vib}(0)$ can be calculated from the fundamental frequencies by *eq.* (10),

$$U_{vib} = \sum_{i=1}^{3n-6} h\nu_i \left(\frac{1}{2} + \frac{1}{e^{h\nu_i / k\mathrm{T}} - 1} \right) \qquad \ldots(10)$$

where h is Planck's constant, ν_i is the i^{th} fundamental frequency, k is Boltzmann's constant and T is temperature in Kelvins. The translational and rotational degrees of freedom are fully excited corresponding to the values 3RT for nonlinear, and 2.5RT for linear molecules. The $\Delta_f H^o$ is commonly approximated by calculating only the lowest energy conformer, which corresponds to the largest mole fraction, and deriving group increments or bond increments with structural parameters (primary, tertiary and quaternary C) that are unique to a particular force field to be used in conjunction with that force field's SE. These terms are shown in *eq.* (11),

$$\Delta_f H^o = \mathrm{BE} + \mathrm{SE} + 4\mathrm{RT} \text{ (3.5RT for linear molecules)} \qquad \ldots(11)$$

where BE represents the sum of bond or group increments. The SE is the difference between the geometry optimised molecule and the impossible theoretical molecule with bond lengths and angles all ideal, and with each atom at infinite separation with regard to intramolecular forces like van der Waals forces and Coulombic interactions. MM may also be used to calculate vibrational data and this is made use of by the MM programmes that calculate the most accurate $\Delta_f H^o$. These calculations that explicitly calculate U_{vib} from vibrational frequencies use BE values fitted to the energy at the bottom of the well on a PES. Only some MM force fields calculate $\Delta_f H^o$ *e.g.* all MM2[33, 37], MM3[38-40] and MM4[36, 41-63] by Allinger *et al.*. MM4 is a class 3 molecular force field. In a recent paper MM4 has been fit to fluorinated hydrocarbon data[63]. This article is written to accommodate readers of multiple disciplines, and an effort is made to explain some esoteric terms and notation, and probably represents the last update to the force field before it is made available through Wavefunction (a commercial software development company). Wavefunction provides the chemical software package Spartan which includes MM, semiempirical quantum chemical and *ab initio* quantum chemical models. The MM force fields in Spartan are a very simple SYBYL force field[64] and the Merck molecular force field 94 (MMFF94)[65-69]. MMFF94 does not calculate $\Delta_f H^o$ values but can calculate fundamental frequencies.

MMFF94 uses a buffered 14 7[70] function to treat nonbonded van der waals forces for both intramolecular and intermolecular interactions. The buffered 14 7 potential was developed to treat the rare gas data, calculated from the most accurate quantum chemical potentials the author was aware of at the time. The quantum chemical potentials were in turn derived from a simultaneous fit to several kinds of experimental data. This differs from other empirical potentials because these are typically fit to intermolecular interactions in a way that errors and omissions made in the description of other physical terms might be counterbalanced among the terms. The buffered 14 7 function has the general form given in *eq.* (12) and alternate form *eq.* (13).

$$E_{vdW} = \varepsilon_{ij}\left(\frac{1+\delta}{\rho_{ij}+\delta}\right)^{(n-m)}\left(\frac{1+\gamma}{\rho_{ij}{}^{m}+\gamma}-2\right) \qquad \text{...(12)}$$

$$E_{vdW} = \varepsilon_{ij}\left(\frac{1.07R_{ij}^{*}}{R_{ij}+0.07R_{ij}^{*}}\right)^{7}\left(\frac{1.12R_{ij}^{*7}}{R_{ij}^{7}+0.12R_{ij}^{*7}}-2\right) \qquad \ldots(13)$$

Here R^{*}_{ij} is the equilibrium distance, R_{ij} is the distance, $\rho_{ij} = R_{ij}/R^{*}_{ij}$ and ε_{ij} is the well depth between atoms i and j. The function is buffered by the constants δ and γ. If $\delta = 0.07$, $\gamma = 0.12$, $m = 7$ and $n = 14$ in *eq.* (12) then the function is equivalent to *eq.* (13). Buffering the function keeps the potential finite as $R_{ij} \to 0$ and avoids the too-strong divergence found in the unbuffered Lennard-Jones potential.

1.2 Additivity Schemes

In additivity schemes to a first approximation the atomisation energy of a molecule is approximated as the sum of constant transferable bond energy terms between any two atoms A and B [E°(A-B)]. The value of the bond energy is determined by the two atoms of the bond, and the order of the bond as specified in a valence bond structure. This approximation works for the higher members of a homologous series like the *n*-alkanes (C6 and above) but fails for structural isomers and the lower members of the series. This indicates the C-C bond energy [E(C-C)] varies according to the groups attached. Put another way, the concept of a constant transferable bond energy holds so long as the nearest neighbours remain the same.

To account for the thermochemical differences between structural isomers, one needs to take into account the chemical environment of the bond, and this is what the group methods do. On the basis that the energy of a bond is constant as long as the nearest neighbours are the same, a molecular fragment about a bond or atom that includes all nearest neighbour atoms may be assigned a group parameter. Whether parameters are designated to a molecular fragment or bonds distinguished by the chemical environment of the neighbour atoms, the results are equivalent when the same number of parameters are used. The three common group methods by Laidler[2], Benson[3] and Allen[4] have been shown to be equivalent by Cox and Pilcher[6]. If only alkanes are considered, four parameters are required, and may be identified with primary, secondary, tertiary and quaternary carbons. Also some account must be made for steric strain. This is usually done by multiplying the number of *gauche* 1,4 interactions in the molecule by an appropriate coefficient. The performance of a group method is as good as if not superior to other methods, if parameters are derived from accurate experimental data. A common difficulty encountered however, is a molecule may contain a molecular fragment for which a group parameter cannot be derived from the available experimental values.

Data of $\Delta_f H^o$ has best been organised into related groups by Pedley[15] and in the same work suggests his own group method. Selected values taken from Pedley and the performance of his scheme are the benchmark for this work.

1.3 Ab Initio

If no data exist for any species similar to a molecule under examination, a high level correlated *ab initio* calculation can be performed at a high cost in computer time. However this is only possible for very small molecules, with no more than ten first row atoms. To reduce computer time for larger molecules, approximations of severity commensurate with molecular size must be made. A quantum chemical model satisfies the Schrödinger equation shown in *eq.* (14).

$$\hat{\mathrm{H}}\Psi = E\Psi \qquad \ldots(14)$$

By the Schrödinger equation the function Ψ is operated on, in this case by $\hat{\mathrm{H}}$ (the Hamiltonian operator) to yield itself multiplied by the scalar value E which is the eigenvalue of the function Ψ, which is necessarily an eigenfunction. The term Hamiltonian refers to the operator itself, but sometimes refers to the left hand side of *eq.* (14). The eigenvalue of the Hamiltonian operator is the energy of the system. The Schrödinger equation can be solved exactly for the one electron case as in the H and He^+ atoms. However approximations must be made in the many electron case. The approximations made or the terms used or not used constructing an operator define a model. Common *ab initio* methods[71-75] range from low level Hartree-Fock to high level correlated methods.

1.3.1 The Linear Combination of Atomic Orbitals Approximation

Molecular orbital (MO) theory is based on atomic orbital (AO) theory. The functions of molecules in MO theory are constructed from a linear combination of atomic orbitals (LCAO). In AO theory the AOs are wave functions with states defined by the quantum numbers *n*, *l* and *m*. In AO theory the AOs can be solved numerically given the spherical symmetry of the atom; in the LCAO approach of MO theory the AOs are predetermined spatial wave functions (basis functions) that mimic the form of the AOs of AO theory. The basis functions in the LCAO are the basis set of the system.

1.3.2 Extended Hückel Theory

Drastic approximations are made in methods such as extended Hückel theory (EHT), where the defining approximation is: the repulsive interaction between electrons in different MOs is ignored. The consequence of this is that the same orbital energies are ascertained for a molecule and its corresponding radical ions. However it is apparent that an electron in the anion is repelled by one more electron than in the neutral molecule, and by two more than in the corresponding cation. Methods with this approximation are called one electron methods and are mainly used for transition metal complexes and metallic substances.

1.3.3 The Self-Consistent Field

One electron methods would be correct if electrons in a molecule were attracted to the nuclei, while at the same time did not repel each other. This model gives predicted energies for molecules that are much too low. To account for electron-electron repulsion the term $1/r_{ij}$ (where r_{ij} is the distance between the i^{th} and j^{th} electrons) must be incorporated into the Hamiltonian. This then gives energies too high, because the functions in the LCAO are arrived at from AOs where electron repulsion was ignored. The LCAO functions therefore distribute too much electron density toward the nuclei. The true function can be more closely approximated by using more diffuse functions, and these are arrived at by self-consistent field (SCF) methods. In the SCF approach, an electron in its molecular orbital is optimised within the time averaged “smeared out” electric field of all the other electrons, by adjusting nuclear charge according to the degree of shielding due to the other electrons. This results in a function less contracted about the nuclei. The same is done with every electron and the whole process is repeated until a subsequent optimisation with every electron results in very small or no change from the previous optimisations and the method is said to have gone self-consistent.

1.3.4 The Hartree-Fock Approximation

If the spin of the electron is to be considered, functions in the LCAOs are the product of a spin function (α or β) and an orbital function and are called spin orbitals. To obey the Pauli principle, the spin orbitals are written in the form of a Slater determinant so that the overall wavefunction is antisymmetric with respect to interchange of electron coordinates. Then if the Fock operator is used which includes a core Hamiltonian, Coulomb operator and exchange operator, we have arrived at the Hartree-Fock (HF) approximation. To arrive at a model energy using the Fock operator and a set of basis functions employing the LCAO approximation the Roothaan-Hall equations[76, 77] are required. Methods that solve the Roothaan-Hall equations are called HF models.

The HF method is variational *i.e.* the HF energy is a bound for the exact Schrödinger energy, also the method is size consistent *i.e.* the magnitude of the error is commensurate with the size of the molecule.

1.3.5 Correlation

In the HF method the motions of the electrons are independent of each other. However between two electrons if one electron is on one side of an orbital, the other electron prefers the other side, so that the motions of the electrons are said to be correlated. Compared with the uncorrelated HF model, inter-electron repulsion is reduced in correlated models. Unfortunately an equation that describes the exact correlation explicitly is unknown. Methods that incorporate correlation energy include density functional theory (DFT), configuration interaction (CI) and Møller-Plesset (MP). CI and MP models add flexibility to a HF model by mixing ground state and excited state wavefunctions. DFT incorporates approximate correlation energy by a more explicit approach.

1.3.6 Basis Set

Slater type orbitals (STOs) are a set of approximate atomic orbitals with states given by the quantum numbers n, l and m. The exact spatial wavefunction can in theory be represented by all the STOs of states given by the complete set of quantum numbers, corresponding to the complete basis set. This infinite set of functions is unattainable and the number of basis functions must be reduced to a finite set, which consequently incurs a basis set truncation error. To a first approximation the exact spatial wave equation may be represented by a minimal basis set. This would include the minimum number of basis functions to accommodate all the electrons for each atom. *i.e.* a single 1s function per hydrogen or helium; five functions for each first row atom: 1s, 2s, $2p_x$, $2p_y$, $2p_z$ and nine functions for every second row atom: 1s, 2s, $2p_x$, $2p_y$, $2p_z$, 3s, $3p_x$, $3p_y$, $3p_z$ the basis set of the remaining main group rows are built up in the same way *mutatis mutandis*. The minimal basis set is significantly improved by the double zeta basis set (DZ basis set) where each basis function in the minimal basis set is replaced with two basis functions. A triple zeta basis set (TZ basis set) replaces each basis function of a minimal basis set with three basis functions. Likewise quadruple, quintuple zeta and higher (QZ, 5Z, 6Z…) basis sets correspond to the respective increased basis set size and give improved flexibility in the wavefunction.

With three or more atoms, the evaluation of the two electron integrals of the Fock matrix is impractical using STOs. However Gaussian type orbitals (GTOs) have the property that the product of two Gaussians centred on different atoms are equivalent to one Gaussian centred between the atoms. Thus two electron integrals for three and four atoms can be reduced to integrals over two points positioned on or between the appropriate atoms. This significantly reduces the number of two electron integrals. However GTOs have a poor representation of orbitals at or near the nucleus where STOs have a cusp that is not present in the GTOs. The accuracy of the orbital is improved by grouping GTOs together termed contracted Gaussian functions. A contracted Gaussian is a linear combination of the original Gaussians (primitive Gaussians). The STO-*n*G basis set uses *n* primitive Gaussians to construct a contracted Gaussian to mimic each STO in a minimal basis set. The use of contracted Gaussians almost offsets the advantages from using GTOs, so while maximising *n* gives improved accuracy, minimising *n* reduces cost. However improvement by increasing *n* reduces as *n* increases. The STO-3G basis set is commonly used and is considered to be a good compromise between cost and accuracy. The STO-*n*G basis set consists of functions that are spherical or are in sets of functions which taken together are spherical.

Consequently spherically shaped molecular environments are handled better than aspherical ones. This may be alleviated by using DZ or TZ basis sets. A compromise for DZ and TZ is a split valence (SV) basis set for which two sets of valence basis functions and one set of core basis functions are used. For instance the p orbitals of the σ component in a bond needs to be more contracted than the p orbitals of the π component in the bond. The split valence 3-21G basis set represents inner-shell atomic orbitals with one contracted Gaussian of three primitives and valence shell orbitals with two contracted Gaussians of two and one primitives. Likewise the 6-31G basis set uses a contracted Gaussian of six primitives for the inner-shell orbitals and valence shell orbitals with two contracted Gaussians constructed with three and one primitives, while the 6-311G basis set has valence shell orbitals with valence functions split into three *mutatis mutandis*.

Thus far SV, DZ, TZ and larger basis sets have their orbital functions centred on the atoms of a molecule. Providing polarisation functions allows electrons to be distributed away from the nuclear positions. The 6-31G* and 6-311G* basis sets include d-type functions on non-hydrogen main group elements. For 6-31G** and 6-311G**, p-type functions are added for hydrogen as well. The 3-21G(*) basis set only has polarisation functions for second row main group elements and higher. The Gaussian exponents of polarisation functions are chosen to minimise the energy for a group of representative molecules.

Most basis sets have Gaussian exponents and linear expansion coefficients based on HF calculations. The correlation consistent polarised valence double, triple and quadruple zeta (cc-pVDZ, cc-pVTZ and cc-pVQZ) basis sets are formulated to minimise CISD ground state atom energies and should be better suited than a basis set like 6-31G* to capture correlation energy at least for the atoms.

1.3.7 Configuration Interaction

In a configuration interaction (CI) model for a particular basis set the ground state spin orbitals are determined as the HF spin orbitals in the form of a Slater determinant. However many excited state determinantal wavefunctions may be formed by promotion of any number of electrons to the virtual orbitals of the ground state wavefunction. An excited state wavefunction, or linear combination of a small number of functions conforming to the appropriate symmetry form a configuration state function (CSF). The CI energy is then a linear combination of every CSF and a mixing coefficient. In the full CI calculation the entire combination of electron promotions is evaluated excluding redundant CSFs on symmetry grounds. For a basis set this gives the exact basis set correlation energy. The full CI calculation is impractical for most systems and a truncated CI model must be used. If the number of excited state wavefunctions are limited (truncated) to only those that involve single electron promotions (CIS) this leads to no improvement over the HF wavefunction, but is applicable for calculations on excited states. The simplest procedure leading to an improvement over the HF wavefunction is CI involving only double electron promotions CID. This is improved on by singles and doubles CISD, which is further improved with the correction for triples CISD(T) an approximation for CISDT and so the cost increases with accuracy and number of electron promotions considered.

Truncated CI is not size consistent. By this it is meant that the model energy is not proportional to the number of electrons of similar systems. To demonstrate this consider the calculation for a helium dimer. For a CISD calculation on a single helium with only two electrons this is a full CI calculation and corresponds to half the energy of the dimer at infinite separation. However a CISD calculation on the dimer with small separation (small r) requires triple and quadruple excited CSFs. This is no longer a full CI calculation. At some point as r increases, how the calculation of the energy of the system is treated alters in a discontinuous way.

1.3.8 The Multiconfiguration SCF

In CI the coefficients in the determinant of each CSF are determined from an initial HF calculation; these remain fixed while the mixing coefficients are allowed to vary. The Multiconfiguration self-consistent field (MCSCF) method optimises the coefficients in the determinant of a CSF as well as the mixing coefficients. MCSCF is computationally more costly than CI with the same number of CSFs but allows for a more accurate calculation with less CSFs. Further development in multiconfiguration methods has led to complete active-space self-consistent field (CASSCF) and restricted active-space self-consistent field (RASSCF) methods.

In a multireference method, for instance, multireference configuration interaction (MRCI), subsequent to an MCSCF calculation, determinants with sufficiently large mixing coefficients are selected for a set of reference determinants. The reference set would include both singly and doubly excited determinants. A CI calculation with single and double excitations from the reference set then incorporates triple and quadruple excitations. The quadruple excitations are important when reducing the size consistency error. Therefore in an MRCI calculation the size consistency error is usually significantly reduced.

1.3.9 Møller-Plesset Perturbation Theory

A size consistent method based on perturbation theory introduced by Møller and Plesset[78] was developed by Pople and co-workers[79]. The Møller-Plesset method defines the unperturbed system (MP0) as a sum of HF one-electron energies. To the unperturbed system first, second, third and fourth order corrections (MP1, MP2, MP3 and MP4 respectively) extend the amount of correlation energy determined by the model beginning with MP2. MP1 introduces the coulomb and exchange operators in the series and is equivalent to the HF-SCF method. MP2 introduces interaction of doubly excited configurations with ground configurations and is the first in the series to account for some of the correlation energy. As well as interaction with ground configurations MP3 allows interactions between the doubly excited configurations. MP4 adds interactions that include single, double, triple and quadruple excitations. MP2, MP3 and MP4 capture roughly 80%, 95% and 99% of the correlation energy respectively, with MP2 being appropriate for geometry optimisation on molecules of moderate size.

1.3.10 Coupled Cluster models

Coupled cluster (CC) approaches use an exponential excitation operator. By the nature of the operator each class of excitations (singles, doubles *etc.*) is included to all orders. This means that the operator for doubles includes products of doubles, which correspond to a subset of the quadruple excitations of CI. The products of the double excitation determinants are not connected or linked, this is different from the set of CI quadruple excitations where connected excitations are included. The cluster operator that includes all classes of excitation with a complete basis set gives the exact wave function. This of course cannot be realised. The wavefunction for CC truncated to only double excitations (CCD) includes doubly, quadruply, hextuply and so on, excited determinants of which only doubly excited excitations are connected. The inclusion of the higher order terms is what makes the method size consistent. A disconnected quadruple term does not treat quadruple excitations as exactly as a connected quadruply excited determinant. However the disconnected product of determinants usually gives a good approximation for the corresponding connected determinant. CCD is more accurate and computationally expensive than CID. However it is only slightly less accurate than CISDTQ while being significantly less expensive computationally as well as size consistent. The cluster operator is commonly truncated to include single as well as double excitations (CCSD) as including singles incurs only a small extra computational cost. A method shown to be even more accurate is to subsequently include a contribution due to triple excitations by perturbation theory CCSD(T).

1.3.11 Density Functional Theory

All the correlated procedures discussed so far construct CSFs starting with the HF approximation. In 1964 a proof by Hohenberg and Kohn[80] showed that the ground state energy (E_0) and all the other ground state electronic properties are uniquely determined by the ground state electron density function (ρ_0). The following year Kohn and Sham developed one electron equations from which electron density could be ascertained[81]. By the Kohn-Sham formalism E_0 is a sum of kinetic energy, E_T; electron-nuclear interaction energy, E_V; Coulomb energy, E_J and exchange/correlation energy, E_{XC}. The electronic energy is said to be a functional of the electron density. The exact functional dependence on ρ_0 is known for E_V and E_J, but E_{XC} must be approximated and E_T is not a functional of electron density. The Hohenberg-Kohn theorem shows a unique value for E_0 depends on ρ_0, however it does not indicate the form of the functional dependence otherwise E_{XC} could be ascertained exactly.

The E_{XC} functional is usually separated into an exchange functional and a correlation functional to give exchange and correlation energies respectively. The name of a DFT method designates the pairing of these two functionals used in the calculation *e.g.* The BLYP functional uses the exchange functional developed by Becke[82] and the correlation functional developed by Lee, Yang and Parr[83].

There are three types of functional in current use: (i) local density models use the local density approximation (LDA) and the local spin density approximation (LSDA), (ii) non-local density models with the generalised gradient approximation (GGA) and (iii) non-local density models that use the exact HF exchange term. The first two types are the pure density functional methods while the third type are the hybrid methods. Pure density functional methods use special algorithms for the exchange term which for larger molecules are significantly computationally faster than HF models. The hybrid models like B3LYP can never be faster than a HF calculation.

1.3.12 Composite methods

Analogous to molecular mechanics, geometry optimisation proceeds by successive single point energy calculations with subsequent geometry adjustment to arrive at a minimum on a potential energy surface. If each single point energy calculation is computationally expensive, geometry optimisation can be impractical if the single point energy is barely feasible. However the geometry need not necessarily be arrived at at the same level of theory. A geometry may be ascertained by a lower level calculation and a single point energy calculation performed at a higher level. Composite methods automatically extrapolate correction terms by varying the basis set and choosing higher and lower levels of theory to approximate corrections for such things as basis set truncation error. Composite methods minimise computational cost by using the lowest level adequate at each step in the procedure, to reproduce experimental data within or close to the limits of the experimental error of a test set of molecules.

G4 is the latest in a series of G*n* (*n* = 1,2,3,4)[84-87] composite procedures where in G4 parameters have been extended for the first, second and third row compounds. Only the first and second rows were examined in the procedures prior to G4. Each successive method in the series adopts modifications that improve upon the earlier methods. For instance connected electron promotions in singles and doubles with triples approximated coupled cluster theory CCSD(T) is the highest level for a single point energy calculation used in G4[87], whereas quadratic configuration interaction QCISD(T) is used in G3[86], an approximation to CCSD(T). Analogous to the naming scheme of the G*n* series is the W*n* series developed at the Weizmann Institute of Science, the latest in this series being W4[88]. W4 calculates at a higher level of theory than G4, accommodating connected quadruple and quintuple promotions with a double zeta basis set (CCSDTQ5/DZ). Appropriate for the higher level of calculation this has only been examined with regard to the first and second row compounds. Karton *et al.*[88] propose the progression W1 → W2.2 → W3.2 → W4lite → W4 progressing from low to highest cost and with the ability to recycle all steps from W2.2 into W3.2 and W3.2 into W4. The W4 procedure is feasible if a single point calculation on the molecule at the CCSDTQ level with the cc-pVDZ basis set is feasible (CCSDTQ/cc-pVDZ). The information before and after a slash respectively refers to model and the basis set used for a calculation. This includes geometry optimisation at that level. If geometry optimisation is performed at a different level, the model and basis set at which the geometry optimisation was performed is appended subsequent to a double slash. This indicates that a single point energy at the level specified in front of a double slash is performed on a geometry that has been optimised at the level subsequent to the double slash. It is routine to calculate a single point energy at a high level on a geometry optimised at a lower level. A space can be used instead of the single slash as separator between model and basis set. Minimal basis sets are inappropriate for correlated models so if a STO-3G basis set is specified it is assumed the HF model is used.

Recently Tasi *et al.*[89] have attributed the systematic error in G2 $\Delta_f H^o$ values to uncertainty in the $\Delta_f H^o$ of the carbon atom; 711.194 ± 0.45 kJ mol^{-1}. They suggest the true value might be 711.65 kJ mol^{-1}.

1.3.13 Heat of Formation by Isodesmic Reactions

If there is a good value for a closely related molecule but not enough for a homologous series it may be possible to relate the heat of formation by an isodesmic reaction. Isodesmic reactions have the same number of bond types on the left and right sides of a chemical equation and in these equations correlation energy associated with bond type cancels. Using an isodesmic reaction a relative energy is calculated with a lower level correlated method like MP2, or density functional theory (DFT), or non-correlated methods CIS and HF. For instance, $\Delta_f H^o$ for propene can be calculated from the isodesmic reaction $CH_3\text{-}CH{=}CH_2 + CH_4 \rightarrow CH_2{=}CH_2 + CH_3\text{-}CH_3$. If the experimental values for methane, ethane and ethene are known this gives an equation in two unknowns $\Delta_f H^o$ C_3H_6 (*g*) and the heat of reaction (ΔH_r). A reasonable heat of reaction can be calculated with SCF methods because the correlation energy is roughly additive between like bonds and since the number and type of bonds are conserved in isodesmic reactions the correlation energies cancel. Thus the sum of the energies of the products minus the sum of the energies of the reactants as calculated by SCF methods can give a reasonable value for the heat of the isodesmic reaction and thus solve for the unknown $\Delta_f H^o$ (C_3H_6, g).

1.4 Semiempirical Quantum Mechanics

For large chemical systems as encountered in organic chemistry and biochemistry and pharmaceutical research, the semiempirical neglect of diatomic differential overlap methods (NDDO) provide an alternative to molecular mechanics. These methods have a minimum basis set that ignores core electrons with the common approximation that atomic orbitals on neighbouring atoms do not overlap. NDDO methods are parameterised to experimental data to reproduce equilibrium geometries, heats of formation, dipole moments and ionisation potentials. The common methods are Austin model 1 (AM1[90]), modified neglect of differential overlap (MNDO[91] and MNDO/d[92]) and parametric method 3 (PM3[93]). PM3 has been updated by PM5 but the method has not been published. Jorgensen *et al.* have included extra terms in the core repulsion formula to apply a pairwise distance directed Gaussian function (PDDG) between bonded atoms[94-96]. AM1 has been reparameterised to a training set of 1736 molecules and is now called RM1[97]. RM1 is easily implemented in programmes that already have AM1 as no line of code needs to be changed except for the values of the parameters. Because NDDO methods are parameterised from experimental data they calculate $\Delta_f H^o$ directly without calculating $\Delta H(T)$. However NDDO $\Delta_f H^o$ values are not accurate enough to correctly order the stability of structural isomers. Jorgensen *et al.*[98] have recently compared semiempirical MO methods.

1.5 Systematic Corrections

The semiempirical methods and *ab initio* methods can be improved with the use of atom, bond or group equivalents methods. In the simplest of these only atom equivalents are ascertained. Wiberg[99, 100] and subsequently Ibrahim and Schleyer[101] independently ascertained group equivalents for HF methods.

Allinger *et al.* included two more terms TOR and POP. TOR is a correction for low lying torsional vibrations not accounted for in the harmonic approximation of vibrational frequency calculations and is approximated by a coefficient with the number of single bonds in a molecule about which there is free rotation, excluding methyl groups. POP is a correction for excess energy in $\Delta_f H^o$ due to population of higher energy conformers.

Herndon[102] ascertained atom equivalents by least squares estimates for the ΔE_{Tot} and number of carbon and hydrogen atoms over a group of 65 saturated and unsaturated as well as strained hydrocarbons for the HF model. Liu and Chen[103] retrained the Herndon test group for DFT and MP2 single point energy calculations with large basis sets geometry-optimised and thermally corrected with a smaller basis set. This was done with similar regression analysis as Herndon, but included a regression constant.

Habibollahzadeh *et al.*[104] ascertained valency dependent atom equivalents for DFT when ΔE_{Tot} is corrected with $\Delta H(T)$ calculated and geometry optimised with the 6-31G(*d,p*) basis set.

Mole *et al.*[105] ascertained atom equivalents for six DFT models using a test group of 23 molecules and showed B3LYP to perform best.

Repasky *et al.*[106] used a training set of 329 molecules and a test set of 583 molecules including the training set to ascertain 61 group equivalents including TOR for AM1, MNDO and PM3. AM1 and PM3 performed about the same with PM3 having a slightly better mean absolute error.

Delley[107] has compared 25 electronic structure models over test groups ranging from a subset of 234 molecules for MP2 to the complete set of 592 molecules and atoms for a number of DFT models.

In previous work[24], we fitted forty alkanes from methane to octane including structural isomers to a bond additivity scheme. The C-C bond energy [E(C-C)] was varied about the intrinsic C-C bond energy [E°(C-C)] according to the degree of π-antibonding summed by an algorithm to give what we called π-antibondingness [B*(π)] where B*(π) = -B(π). A sigma equivalent σ-bondingness = B(σ) = -B*(σ) is ascertained as a preliminary step in the calculation of B(π) but was not used. The model was fitted to the smaller molecules from C1 to C5 excluding 2-methylpropane (the first of the small molecules with a 1,4-*gauche* steric interaction) and replacing it with hexane. This gave a training set of 12 unstrained molecules from which we ascertained parameters that were applied to the remaining 28 molecules. This value was subsequently adjusted by subtracting the steric energy by conventional methods *i.e.* to sum the number of 1,4-*gauche* steric interactions per molecule attributing 2.5 kJ mol^{-1} per steric interaction. Thus accounting for the variation in energy between structural isomers with a single parameter for B(π). Over the same molecules this had been done with a minimum of two parameters by Skinner[10]. Since we published[24], Wodrich and von Schleyer[31] have proposed a scheme for alkanes, alkenes, alkynes and alkyl radicals where all parameters are ascertained without data fitting and is based on conventional reasoning. Their scheme determines parameters from five bond separation energies of the simplest molecules and a uniform attenuation treatment. Attenuation effects are common in chemistry. Consider pentane, isopentane and neopentane; there is increased branching about the second carbon atom respectively in these molecules, however with each successive addition of a methyl group to the second carbon the amount by which the molecule is stabilised is reduced or attenuated. Wodrich and von Schleyer apply an attenuation term for C-C-C interaction in branching and hyperconjugation in alkyl radicals and alkenes.

1.6 Bondingness

A B(π) value can be calculated from the AO coefficients in a MO between two atoms by adding the B(σ) term to the dot product of the p AO coefficients on each atom. The B(σ) is an orthogonal transformation of the AO coefficients. The transformation being the projection of the p AOs along the abstract vector between the atoms. The projection formula can be employed to ascertain the σ component of the p orbitals on each atom where the sign of the abstract vector between the atoms (**r**) is reversed for each atom, where **c** is a vector of p AO coefficients for the i^{th} and j^{th} atoms as shown in *eq.* (15). A positive value by *eq.* (15) corresponds to a bonding σ interaction, while a negative value is an antibonding σ interaction. From a B(σ) value a B(π) value is then calculated by *eq.* (16).

$$B(\sigma) = \frac{\mathbf{r}}{|\mathbf{r}|} \cdot \mathbf{c}_i \times \frac{-\mathbf{r}}{|\mathbf{r}|} \cdot \mathbf{c}_j \qquad \ldots(15)$$

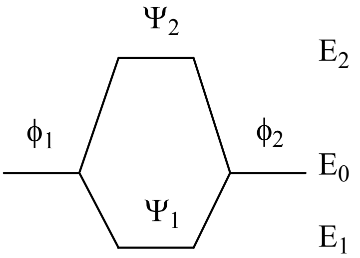
$$B(\pi) = \mathbf{c}_i \cdot \mathbf{c}_j + B(\sigma) \qquad \ldots(16)$$

The final B(π) value for the molecule was ascertained by first summing B(π) values in occupied MOs only between bonded carbon atoms with the AO coefficients of a particular MO, to get a B(π) value for each MO for carbon-carbon bonding. The B(π) in each MO is ordered according to its corresponding MO from highest to lowest. Contiguous negative values are added to give B(π) for the molecule, unless a single negative value having zero or positive B(π) for the MOs about it, is less (more negative) than a sum of contiguously negative values. In which case the more negative value is used.

The basis for using bondingness originated from a qualitative investigation of the antibonding effect and its usefulness as a qualitative explanation for the barrier to rotation in ethane[108, 109]. The antibonding effect is that the destabilisation of the antibonding MO (Ψ_2) is always greater than the stabilisation of the bonding MO (Ψ_1) relative to the energies of the AOs combined (ϕ_1 and ϕ_2) to form the MOs. The antibonding effect can be seen in Figure 1.1.

Figure 1.1

Generic energy level diagram.



If MOs for ethane are constructed from MOs from methyl radicals with geometry combined to form ethane, three MOs of a_1 symmetry and two pairs of degenerate e symmetry MOs are obtained. The barrier to rotation is attributed to overlap repulsion between C-H bonds and the C-H bonding predominantly occurs in the e MOs. The antibonding effect in the occupied 1e' and 1e'' MOs of eclipsed ethane is greater than in the $1e_u$ and $1e_g$ MOs of staggered ethane. This is a significant π antibonding effect in the highest occupied molecular orbitals (HOMOs) e_g and e'' in the staggered and eclipsed conformations respectively. The idea that overlap repulsion between vicinal C-H bonds should be accounted for in an additivity scheme was suggested by Smith[20] in 1999 and subsequent schemes were developed in 2001[21] and 2005[24].

1.7 Notation for algorithms in B(π)

Bondingness (B), in particular π-bondingness [B(π)] is calculated over occupied MOs only between atoms related to each other by the specification (usually two carbon atoms that are bonded) preceding a separator (¨) in the notation developed in this work. B(π) values of common preceding specification calculated from the same MO are added to give what may be referred to as a MO preceding B(π) value. These MO preceding B(π) values are ordered by the MO for which they correspond, from highest to lowest. Contiguous negative values are then added as are contiguous positive values. This results in a list of alternating values either negative, positive or zeroes where no two consecutive values are both positive or both negative. In the A1 algorithm the first negative value is retained. If this value was comprised from contiguous MOs then the corresponding MOs form the set of contiguous MOs for a subsequent calculation. If the value is due to a single MO then a more negative value is sought in a subsequent value that resulted from a sum of contiguous negative values. The first subsequent value arrived at which is more negative than the first negative value then gives the set of contiguous MOs, otherwise the contiguous MO set is comprised of one, corresponding to the MO of the first negative value. If the specification following the separator is the same as that preceding it then the sum of MO preceding B(π) values for the set of MOs is returned as the B(π) for the specified notation [specific B(π)]. However if the specification following the separator differs, B(π) is summed in the MOs of the set of contiguous MOs calculated between atoms conforming to the specification following the separator. This value is then returned as the specific B(π) for that corresponding notation.

The A3 algorithm is exactly the same as the A1 algorithm between atoms with bond orders less than some predefined value (normally 1.5, 2 or 3). If the bond order exceeds the predefined value then the algorithm continues in a similar manner to A1 with the single difference the set of MOs is selected to give the largest positive value for a MO preceding B(π) value or sum of contiguous MO preceding B(π) values.

Usually a type of bond is used in the specifications about the separator *e.g.* C-C, C-O, C=C or sp^3-sp^2 (sp^3-sp^2 is used to mean a single bond next to a double bond), but B(π) can be ascertained between nonbonded atoms. Parameters are ascertained by a least squares estimate (LSE) *k* for each specific B(π), the product of *k* and B(π) giving an adjusted specific B(π) [$B(\pi)_{adj}$].

Consider how different specifications in notation alter specific B(π) by the A1 algorithm in the case for alkenes. The specific B(π) with corresponding notation C-C¨C=C is calculated by selecting contiguous MOs with negative values from B(π) summed over the C-C bonds and then in those MOs the B(π) in the C=C bonds are added. To sum over all bonds capable of π bonding either the word bond, referring to any bond, or more compactly nothing, needs to be specified before the dieresis. Therefore the notation ¨C-C sums MO bond B(π) for every bond in an MO to ascertain the highest contiguous MOs with negative values and then sums B(π) over these MOs for the C-C bonds only. When what is specified is the same on either side of the dieresis, the bond type or specification information needs only to be written once; the dieresis and repeated information are assumed. Thus the following notations for specific B(π) are equivalent: $B(\pi)^{\text{C-C¨C-C}}$ and $B(\pi)^{\text{C-C}}$. In alkanes where π bonding is only between C-C bonds $B(\pi)^{\text{¨C-C}}$ and B(π) will have the same algorithm as the algorithm for the aforementioned values.

General terms are also used to describe an algorithm. For instance if there are A, B and X type environments, there would be three specific B(π) values for the molecule or group of molecules. An algorithm might calculate A¨A, B¨B and

X¨X. Referring to alkenes the algorithm would calculate C-C¨C-C, C=C¨C=C and usually sp^2sp^3¨sp^2sp^3 chemical environments (CC_1<CC_2>1¨CC_1<CC_2>1 as explained in section 1.8). This is usually referred to as a bond¨bond algorithm where the word bond is a variable referencing the same chemical environment in the first and last part of the algorithm. When it is clear by context bond¨bond can be written bond by reasoning of the previous paragraph. The other common algorithm descriptor is ¨bond which for the same group of molecules calculates ¨C-C, ¨C=C and ¨CC_1<CC_2>1.

Another way to ascertain specific B(π) for a molecule is between bonded atoms to calculate B(π) for each MO. Within the confines of a bond the B(π) values are then ordered from highest to lowest in terms of the MO each value is associated with and as was done in the A1 algorithm an alternating list of positive, negative or zero values is constructed by summing contiguous values of the same sign. The difference in the construction of the list is that values are ascertained from one bond only and not a set of common bond types. The most negative value is then ascertained in a similar manner to the A1 algorithm for each bond. Then the values attributed to each bond are added for common bond types or environments. In this notation the specified bond types or chemical environments are preceded with ¡. This notation is employed for use by the A2 algorithm. If no bond type or environment is specified the most negative B(π) ascertained for each bond is added together. For alkanes, given that only C-C bonds are capable of a π bondingness interaction in a small basis set, the $B(\pi)^{¡C\text{-}C}$ and $B(\pi)^{¡}$ notations for specific B(π) will have the same algorithm.

By the different combinations of the notation for a specific B(π) it is evident there are many ways to quantify B(π). However in alkanes there is little distinction between the different chemical environments of each C-C bond. There is perhaps more π-antibonding in terminal C-C bonds, but the different ways of ascertaining B(π) give similar results. However for molecules with hetero atoms or with bond orders greater than one, specific B(π) values can vary a great deal according to the method implied by the variations in notation. In this work a trial and error approach is adopted to ascertain the significance of an algorithm used to calculate B(π) values.

1.8 Notation in chemical environment

A further syntax was developed for computer programmes to distinguish different chemical environments. This describes the chemical environment of a bond. The algorithm employed creates a label beginning with the two atoms of the bond followed by an underscore and the bond order of the bond[*], the adjacent bonds are then similarly labelled and placed in angled brackets followed by the number of that kind of bond adjacent to the bond in the first part of the label. C-H bonds are usually ignored. Thus the sp^2sp^3 bond of 3-methylbut-1-ene, ignoring C-H bonds, would be labelled CC_1<CC_1>2<CC_2>1. Ignoring C-C bonds this becomes CC_1<CC_2>1. This has become a convenient notation for tables and the like, and makes a few appearances in section 7 as well as Tables 3.17 and 3.18. Ignoring C-C and C-H bonds prevents us from distinguishing terminal methyl, secondary, tertiary and quarternary bonding. This is intentional as we anticipate the variation in specific B(π) to accommodate these differences and reduces clutter in the chemical environment notation.

Later in section 7 there are chemical environments with small B(π) or environments unique to one molecule only where we combine the environment notation with another similar chemical environment. For instance CC_1<CN_1>2 is an environment unique in the test set to 1,1-dinitropropane. This was combined with CC_1<CN_1>1 environments to give CC_1<CN_1>1-2.

1.9 Variation in Standard Constant Bond Energy with B(π)

To ascertain parameters for calculating the $\Delta_f H^o$ in a homologous series of forty alkanes, from methane to octane including structural isomers, two test groups of eight molecules to which parameters can be fitted are defined. The first is the homologous series of eight molecules from methane to pentane including all structural isomers. This set includes 2-methylbutane which is the first molecule to

---

[*] The number is actually a bond descriptor of type integer assigned in a quantum chemical input file where for the most part it represents the approximate bond order of the bond, except in the case where the bond order is 1.5, as is the case for benzenoid or nitro compounds. In this instance the integer 5 is used and we keep this notation in this work. Therefore the carbon-carbon bond in a benzene ring has the notation CC_5<CC_5>2.

have a 1,4-*gauche* steric interaction (test group 1). The second set is the same except that 2-methylbutane is replaced with hexane to compare parameterisation without steric interactions (test group 2). Each scheme was tested by training parameters on a smaller subset of molecules and using these parameters to extrapolate values for the remaining larger set of forty molecules. By having two test groups with test group 1 slightly influenced by a steric interaction, it is expected models that cannot accommodate steric interactions will be slightly affected when trained with the different test groups. However the first agenda is to successfully account for variation in $\Delta_f H^o$ for unstrained structural isomers due to overlap repulsion with adjusted B(π) values. Unless otherwise mentioned test group 2 is used to parameterise alkanes.

It has already been ascertained[24] that no regression constant is required for alkanes and no regression estimate is required for E°(C-C) if E°(C-C) is fixed at ½ $\Delta_f H^o$ (C, g) for diamond[110] (357.4 kJ mol^{-1}). Likewise E(C-H) is fixed at ¼ $\Delta_f H^o$ (CH_4, g) the value for methane[110] (415.87 kJ mol^{-1}). This is because with these values fixed, the regression constant is close to zero. With these fixed values in the regression analysis the following results are ascertained.

The following results previously published for alkanes[24] include columns corrected conventionally for steric interactions *i.e.* conventional strain energy (CSE) is related to the number of 1,4-*gauche* interactions, where geometries were optimised at the same level as the energy calculation. For HF calculations the minimal STO-3G basis set was used.

Table 1.1

Values from previous work.

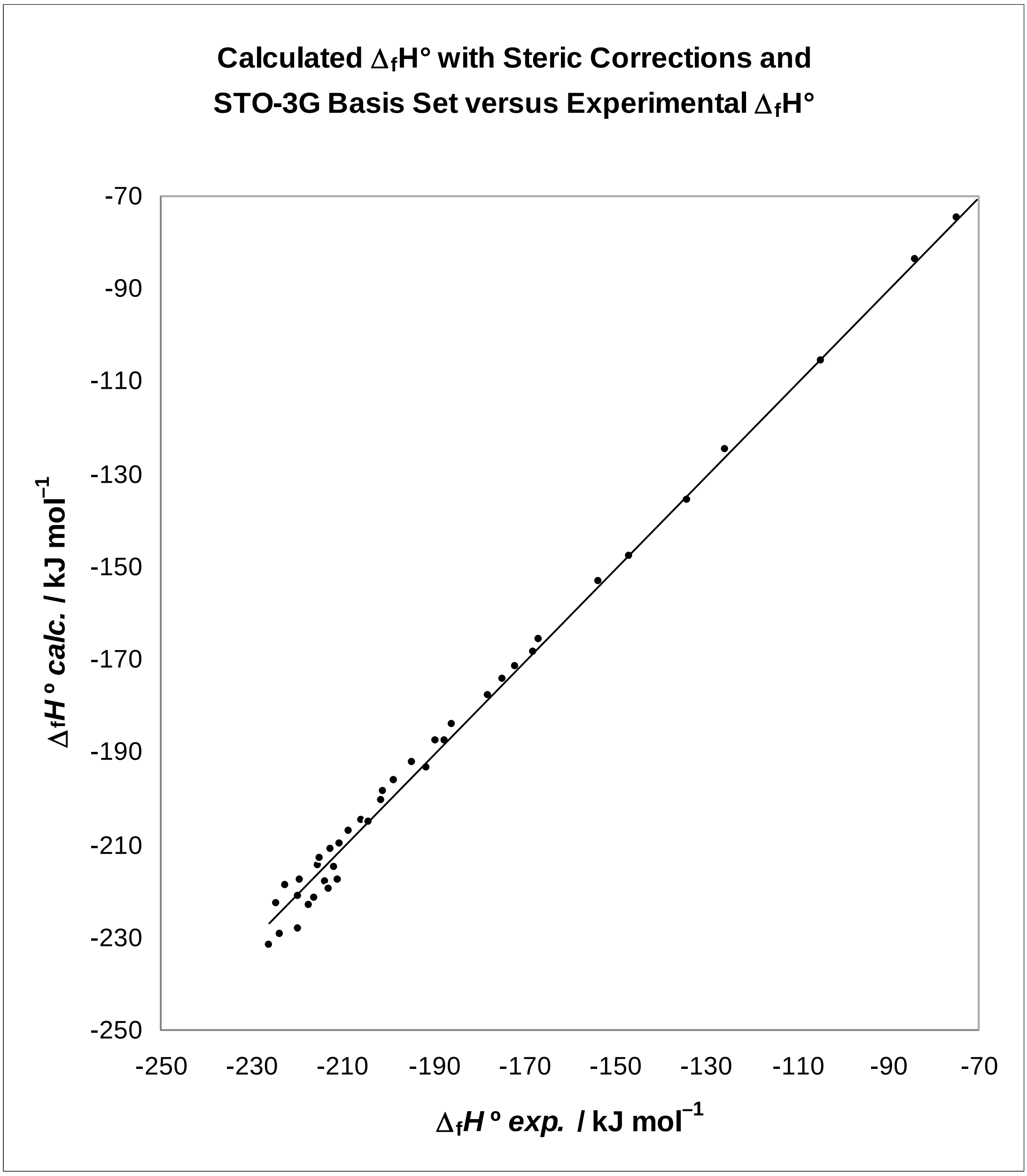
All values in kJ mol^{-1}	**$\Delta_f H^o$**	**$\Delta_f H^o$ (PM3)**	**$\Delta_f H^o$ (STO-3G)**	**$\Delta\Delta_f H^o$ PM3 B(π)**	**$\Delta\Delta_f H^o$ STO-3G B(π)**
methane	-74.8	-74.81	-74.81	0.0	0.0
ethane	-83.8	-84.02	-83.49	0.2	-0.3
propane	-104.7	-104.69	-105.38	0.0	0.7

butane	-125.7	-125.64	-124.78	-0.1	-0.9
2-methylpropane	-134.2	-133.82	-135.44	-0.4	1.2
pentane	-146.9	-149.11	-147.67	2.2	0.8
2-methylbutane	-153.6	-153.26	-153.11	-0.3	-0.5
2,2-dimethylpropane	-168.0	-164.20	-168.49	-3.8	0.5
hexane	-166.9	-167.73	-165.73	0.8	-1.2
2-methylpentane	-174.6	-174.98	-174.34	0.4	-0.3
3-methylpentane	-171.9	-180.05	-171.60	8.2	-0.3
2,2-dimethylbutane	-185.9	-184.93	-183.87	-1.0	-2.0
***gauche* 2,3-dimethylbutane**	-178.1	-180.58	-177.76	2.5	-0.3
heptane	-187.6	-191.32	-187.53	3.7	-0.1
2-methylhexane	-194.5	-194.55	-192.20	0.1	-2.3
3-methylhexane	-191.3	-195.43	-193.34	4.1	2.0
3-ethylpentane	-189.5	-191.94	-187.42	2.4	-2.1
2,2-dimethylpentane	-205.7	-214.24	-204.63	8.5	-1.1
2,3-dimethylpentane	-198.7	-198.66	-196.04	0.0	-2.7
2,4-dimethylpentane	-201.6	-199.17	-200.26	-2.4	-1.3
3,3-dimethylpentane	-201.0	-202.54	-198.37	1.5	-2.6
2,2,3-trimethylbutane	-204.4	-208.29	-205.12	3.9	0.7
octane	-208.5	-211.04	-207.08	2.5	-1.4
2-methylheptane	-215.3	-217.49	-214.58	2.2	-0.7
3-methylheptane	-212.5	-213.59	-210.92	1.1	-1.6
4-methylheptane	-211.9	-217.08	-214.84	5.2	2.9
3-ethylhexane	-210.7	-224.40	-209.83	13.7	-0.9
2,2-dimethylhexane	-224.5	-225.02	-222.85	0.5	-1.6
2,3-dimethylhexane	-213.8	-225.27	-217.96	11.5	4.2
2,4-dimethylhexane	-219.2	-219.87	-217.58	0.7	-1.6
2,5-dimethylhexane	-222.5	-220.82	-218.61	-1.7	-3.9
3,3-dimethylhexane	-219.9	-223.05	-221.12	3.1	1.2
3,4-dimethylhexane	-212.8	-220.30	-219.55	7.5	6.7
3-ethyl-2-methylpentane	-211.0	-221.91	-217.75	10.9	6.7
3-ethyl-3-methylpentane	-214.8	-226.13	-212.85	11.3	-2.0
2,2,3-trimethylpentane	-219.9	-231.72	-228.29	11.8	8.4
2,2,4-trimethylpentane	-223.9	-226.26	-229.39	2.4	5.5

2,3,3-trimethylpentane	-216.2	-226.53	-221.44	10.3	5.2
2,3,4-trimethylpentane	-217.2	-226.29	-223.21	9.1	6.0
2,2,3,3-tetramethylbutane	-226.0	-252.62	-231.63	26.6	5.6
Standard deviation				7.064	3.144

Graph 1.1

Graph published in Thermochimica Acta[24].



Graph 1 includes a solid line depicting the 1:1 line, not to be mistaken for the line of best fit. Unless otherwise specified a solid line represents a 1:1 line henceforth.

2 Systematic corrections using Bondingness

Considering the energy calculated from quantum chemical calculations, there is an apparent systematic error with the deviation in calculated energy from experimental values increasing with molecular size. This arises from such approximations as use of a finite basis set that incurs a basis set truncation error and a model's approximation of, or in the HF case, neglect of correlation energy. Some of the methods that correct for this size-consistent error are mentioned in section 1. The data in Table 2.1 are test group 2 with heptane and octane included to observe the systematic error which is approximately linear in the *n*-alkanes. It can be seen in the *n*-alkanes that the difference between experimental and calculated values δ [calculated in *eq.* (17)] is increasing as B(π) and steric energy decrease. The following method[111] uses B(π) to adjust a systematic error in PM3 $\Delta_f H^o$ values in terms of ΔH_a values. The same method is used with steric energy (from MMFF94) in section 2.2.

$$\delta = \Delta_f H^o\,(\mathrm{exp}) - \Delta_f H^o\,(\mathrm{calc})$$

and

$$\delta = \Delta H_a\,(\mathrm{exp}) - \Delta H_a\,(\mathrm{calc}) \qquad \ldots (17)$$

and by this

$$\Delta H_a\,(\mathrm{adjusted}) = \Delta H_a\,(\mathrm{calc}) + \delta \qquad \ldots (18)$$

If δ can be calculated exactly then ΔH_a (adjusted) calculated by *eq.* (18) would equal the exact experimental value.

Table 2.1

Systematic error compared with bondingness and steric energy.

molecule	$\Delta_f H^o$ Exp. kJmol^{-1}	$\Delta_f H^o$calc. PM3 kJmol^{-1}	δ kJmol^{-1}	B(π) PM3	B(π) STO-3G	Steric Energy
methane	-74.8	-54.45	-20.35	0.000	0.000	0.03
ethane	-83.8	-75.88	-7.92	-0.454	-0.330	-4.73
propane	-104.7	-98.84	-5.86	-0.717	-0.503	-4.90
butane	-125.7	-121.60	-4.10	-0.976	-0.706	-5.08
2-methylpropane	-134.2	-123.56	-10.64	-0.840	-0.580	-0.48

molecule	**$\Delta_f H^o$ Exp. kJmol^{-1}**	**$\Delta_f H^o$calc. PM3 kJmol^{-1}**	**δ kJmol^{-1}**	**B(π) PM3**	**B(π) STO-3G**	**Steric Energy**
pentane	-146.9	-144.28	-2.62	-1.193	-0.868	-5.27
2,2-dimethylpropane	-168	-149.93	-18.07	-0.941	-0.620	8.50
hexane	-166.9	-166.96	0.06	-1.490	-1.086	-5.47
heptane	-187.6	-189.64	2.04	-1.705	-1.261	-5.68
octane	-208.5	-212.32	3.82	-1.984	-1.462	-5.88

If a systematic error commensurate with molecular size for alkanes is observed then for C_nH_{2n+2} *eq.* (19) gives the difference between experimental and calculated values in terms of adjustment coefficients for the number of C-H and C-C bonds δ(C-H) and δ(C-C) respectively.

$$\delta = (2n+2)\delta(\text{C-H}) + (n-1)\delta(\text{C-C}) \qquad \ldots (19)$$

From the calculated value for methane δ(C-H) is fixed. The δ(C-C) adjustment then varies according to B(π) averaged over the number of C-C bonds by *eq.* (20).

$$\delta(\text{C-C}) = \delta°(\text{C-C}) - \delta'(\text{C-C})\,\frac{B(\pi)}{n-1} \qquad \ldots (20)$$

LSEs are ascertained by *eq.* (21), the slope giving δ'(C-C) and intercept δ°(C-C).

$$\frac{\delta}{n-1} = \delta°(\text{C-C}) - \delta'(\text{C-C})\frac{B(\pi)}{n-1} \qquad \ldots (21)$$

The above method or variations *mutatis mutandis* are used to correct energies ascertained from semiempirical and *ab initio* quantum chemical output in the remainder of this section.

2.1 Variation in PM3 Quantum Chemical Energy with B(π)

LSEs are obtained for δº(C-C) and δ'(C-C) for test group 2 in the following way. δ(C-H) has the value -5.086 kJ mol^{-1} and is ascertained by dividing the δ value [*eq.* (17)] for methane by four. Then [δ - (2n+2) δ(C-H)]/(n-1) is fitted to B(π)/(n-1). The intercept = δ°(C-C) and the slope = δ'(C-C) in *eq.* (22):

$$\Delta_f H^o_{\text{adjusted}} = \Delta_f H^o_{\text{of model}} + (2n + 2)\delta(\text{C-H}) + (n\text{-}1)\delta^o(\text{C-C}) + \delta'(\text{C-C})\ \text{B}(\pi) \quad \ldots (22)$$

The parameters of Table 2.2 are used with *eq.* (22) to give the results of Table 2.3.

Table 2.2

Parameters for systematic corrections using PM3 and STO-3G B(π).

kJ mol^{-1}	δ°(C-C)	δ'(C-C)	δ(C-H)	s
PM3	-1.624	-53.332	-5.086*	4.6
STO-3G	-0.047	-68.083	-5.086*	3.4
STO-3G	0*	-62.523	-4.772	3.0

* parameters are not LSEs.

Table 2.3

Values for $\Delta_f H^o$, calculated using parameters from Table 2.2 and $\Delta\Delta_f H^o$, the difference between calculated and experimental $\Delta_f H^o$ values. The 'a' labels refer to $\Delta_f H^o_{\text{PM3 adjusted}}$ calculated using PM3 B(π) with CSE subsequently added. The 'b' labels are the $\Delta_f H^o_{\text{PM3 adjusted}}$ values using STO-3G B(π) with CSE not required. The $\Delta_f H^o$ exp. is the experimental value.

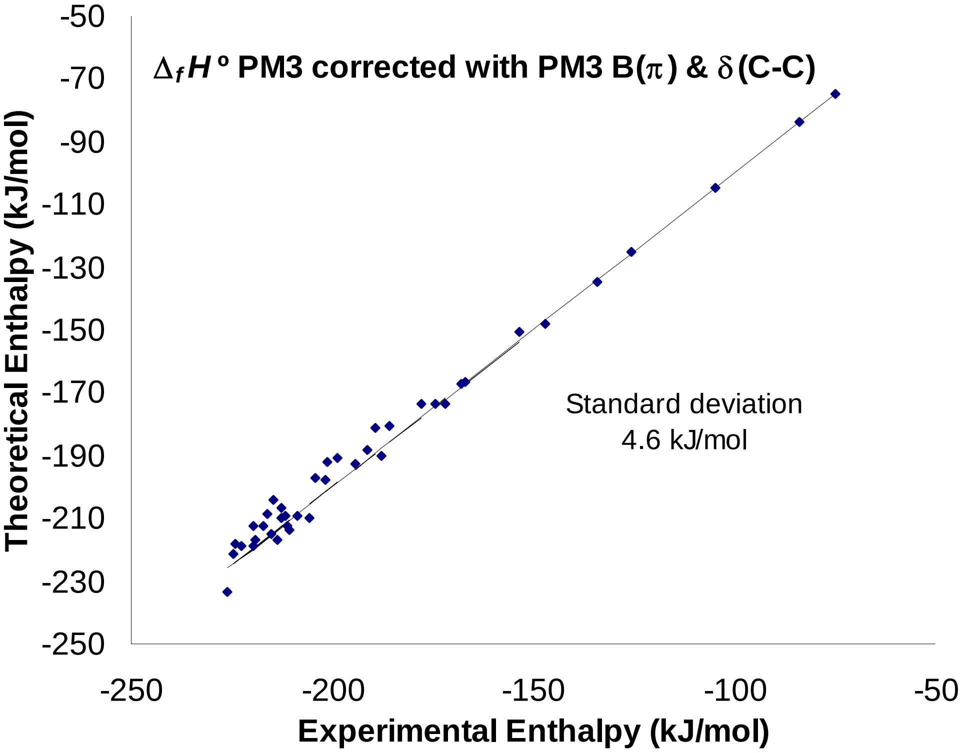
All values in kJ mol^{-1}	$\Delta_f H^o$ exp.	$\Delta_f H^o$ PM3	$\Delta_f H^{o\ a}$ PM3 B(π)	$\Delta_f H^{o\ b}$ HF B(π)	$\Delta\Delta_f H^o$ PM3 B(π)	$\Delta\Delta_f H^o$ HF B(π)
Methane	-74.8	-54.5	-74.8	-74.80	0.0	0.00
ethane	-83.8	-75.9	-83.8	-83.96	0.0	0.16
propane	-104.7	-98.8	-104.5	-105.35	-0.2	0.65
butane	-125.7	-121.6	-125.3	-124.51	-0.4	-1.19
2-methylpropane	-134.2	-123.6	-134.5	-135.10	0.3	0.90
pentane	-146.9	-144.3	-148.2	-146.42	1.3	-0.48
2-methylbutane	-153.6	-143.3	-150.6	-151.86	-3.0	-1.74
2,2-dimethylpropane	-168.0	-149.9	-167.3	-168.92	-0.7	0.92

All values in kJ mol^{-1}	$\Delta_f H^o$ exp.	$\Delta_f H^o$ PM3	$\Delta_f H^{o\ a}$ PM3 B(π)	$\Delta_f H^{o\ b}$ HF B(π)	$\Delta\Delta_f H^o$ PM3 B(π)	$\Delta\Delta_f H^o$ HF B(π)
hexane	-166.9	-167.0	-166.8	-164.43	-0.1	-2.47
2-methylpentane	-174.6	-167.3	-173.3	-173.75	-1.3	-0.85
3-methylpentane	-171.9	-163.2	-173.4	-169.44	1.5	-2.46
2,2-dimethylbutane	-185.9	-165.7	-180.3	-181.91	-5.6	-3.99
2,3-dimethylbutane	-178.1	-162.9	-173.6	-174.12	-4.5	-3.98
heptane	-187.6	-189.6	-189.9	-185.46	2.3	-2.14
2-methylhexane	-194.5	-190.0	-192.8	-191.60	-1.7	-2.90
3-methylhexane	-191.3	-184.9	-188.2	-189.44	-3.1	-1.86
3-ethylpentane	-189.5	-181.6	-181.5	-183.37	-8.0	-6.13
2,2-dimethylpentane	-205.7	-189.7	-209.8	-203.45	4.1	-2.25
2,3-dimethylpentane	-198.7	-184.7	-190.6	-193.43	-8.1	-5.27
2,4-dimethylpentane	-201.6	-191.3	-197.9	-201.46	-3.7	-0.14
3,3-dimethylpentane	-201.0	-182.7	-191.8	-195.42	-9.2	-5.58
2,2,3-trimethylbutane	-204.4	-183.1	-197.3	-201.21	-7.1	-3.19
octane	-208.5	-212.3	-209.4	-204.68	0.9	-3.82
2-methylheptane	-215.3	-212.7	-215.2	-213.11	-0.1	-2.19
3-methylheptane	-212.5	-207.5	-206.4	-207.05	-6.1	-5.45
4-methylheptane	-211.9	-207.5	-209.5	-210.22	-2.4	-1.68
3-ethylhexane	-210.7	-205.6	-213.8	-206.23	3.1	-4.47
2,2-dimethylhexane	-224.5	-212.4	-221.4	-221.57	-3.1	-2.93
2,3-dimethylhexane	-213.8	-208.0	-216.9	-215.20	3.1	1.40
2,4-dimethylhexane	-219.2	-212.8	-216.9	-219.71	-2.3	0.51
2,5-dimethylhexane	-222.5	-213.4	-218.6	-219.11	-3.9	-3.39
3,3-dimethylhexane	-219.9	-205.8	-212.5	-217.61	-7.4	-2.29
3,4-dimethylhexane	-212.8	-205.9	-210.2	-216.45	-2.6	3.65
3-ethyl-2-methylpentane	-211.0	-206.9	-212.6	-215.97	1.6	4.97
3-ethyl-3-methylpentane	-214.8	-195.5	-204.4	-204.69	-10.4	-10.11
2,2,3-trimethylpentane	-219.9	-204.5	-218.6	-224.15	-1.3	4.25
2,2,4-trimethylpentane	-223.9	-208.4	-218.2	-224.85	-5.7	0.95

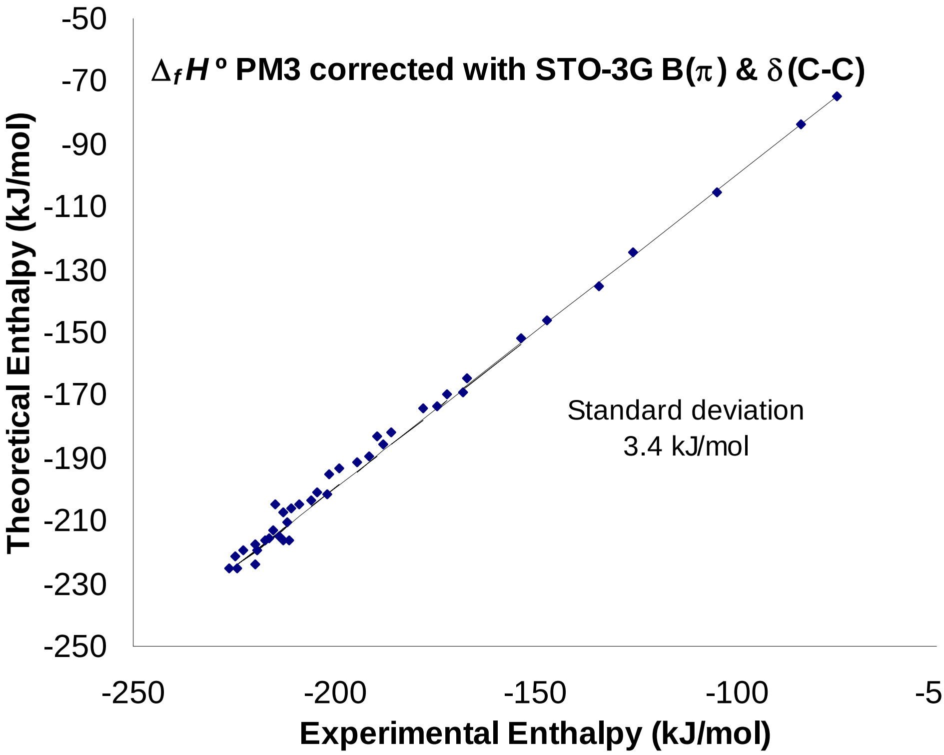
All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^o$ PM3**	**$\Delta_f H^{o\,a}$ PM3 B(π)**	**$\Delta_f H^{o\,b}$ HF B(π)**	**$\Delta\Delta_f H^o$ PM3 B(π)**	**$\Delta\Delta_f H^o$ HF B(π)**
2,3,3-trimethylpentane	-216.2	-199.3	-208.6	-215.44	-7.6	-0.76
2,3,4-trimethylpentane	-217.2	-202.6	-212.2	-216.10	-5.0	-1.10
2,2,3,3-tetramethylbutane	-226.0	-200.8	-233.2	-225.11	7.2	-0.89
Standard deviation					4.6	3.4

Graphs 2.1 and 2.2 show the results for fitting PM3 $\Delta_f H^o$ to PM3 and STO-3G B(π) respectively. The addition of CSE to corrections made with STO-3G B(π) increased the standard deviation to 9.2 kJ mol^{-1}. The results in Graph 2.2 are without CSEs added.

Graph 2.1



Graph 2.2



The LSE for δ^o(C-C) is nearly zero when fitting PM3 $\Delta_f H^o$ values corrected with STO-3G B(π) Thus δ^o(C-C) is set to zero and δ(C-H) is parameterised. We fit (2*n* + 2) and B(π) to δ and force the regression constant to zero. The following formula is employed:

$$\Delta_f H^o_{\text{PM3 adjusted}} = \Delta_f H^o_{\text{ PM3}} + (2n+2)\,\delta(\text{C-H}) + \delta'(\text{C-C})\,B(\pi) \qquad \ldots (23)$$

This gives the values δ(C-H) = -4.772 kJ mol^{-1} and δ'(C-C) = -62.523 kJ mol^{-1} giving the following calculated $\Delta_f H^o$ values in Table 2.4 with corresponding Graph 2.3. The parameters are also shown in Table 2.2.

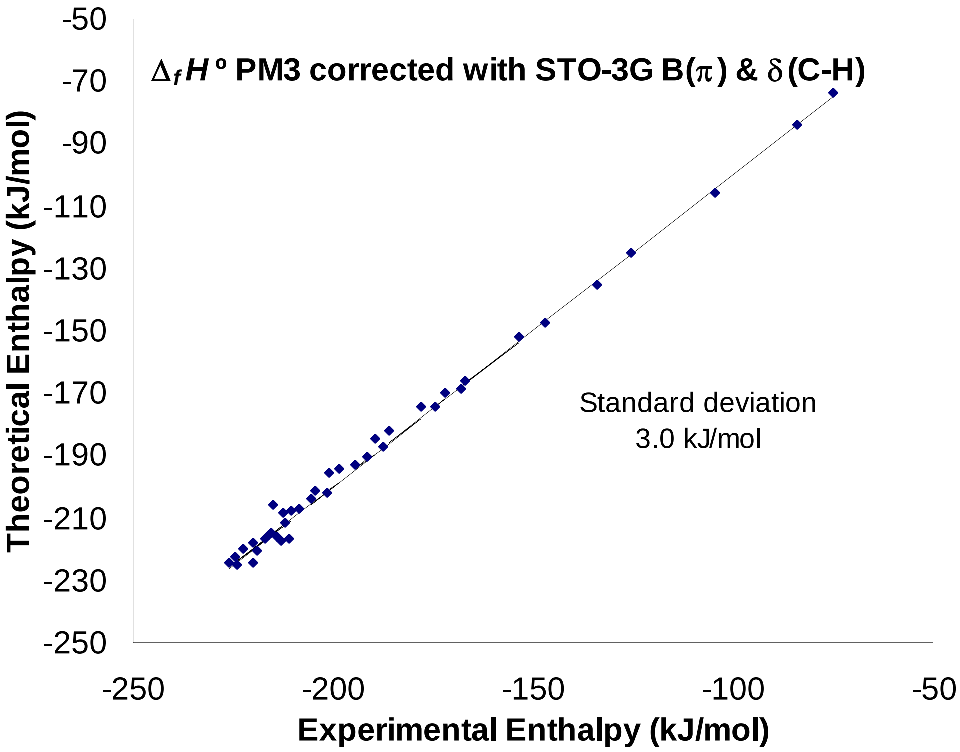
Table 2.4

Values for $\Delta_f H^o$ calculated using parameters for systematic error attributed to C-H bonds and B(π) only. The difference between calculated and experimental $\Delta_f H^o$ values is represented by $\Delta\Delta_f H^o$. The $\Delta_f H^o$ exp. is the experimental value.

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^o$ PM3 STO-3G B(π) δ(C-H)**	**$\Delta\Delta_f H^o$**
methane	-74.8	-73.5	-1.3
ethane	-83.8	-83.9	0.1
propane	-104.7	-105.5	0.8
butane	-125.7	-125.2	-0.5
2-methylpropane	-134.2	-135.0	0.8
pentane	-146.9	-147.3	0.4
2-methylbutane	-153.6	-152.2	-1.4
2,2-dimethylpropane	-168	-168.4	0.4
hexane	-166.9	-165.8	-1.1
2-methylpentane	-174.6	-174.4	-0.2
3-methylpentane	-171.9	-170.1	-1.8
2,2-dimethylbutane	-185.9	-181.8	-4.1
2,3-dimethylbutane	-178.1	-174.4	-3.7
heptane	-187.6	-187.2	-0.4
2-methylhexane	-194.5	-192.8	-1.7
3-methylhexane	-191.3	-190.4	-0.9
3-ethylpentane	-189.5	-184.6	-4.9
2,2-dimethylpentane	-205.7	-203.7	-2.0
2,3-dimethylpentane	-198.7	-194.1	-4.6
2,4-dimethylpentane	-201.6	-202.0	0.4
3,3-dimethylpentane	-201	-195.8	-5.2
2,2,3-trimethylbutane	-204.4	-201.1	-3.3
octane	-208.5	-206.8	-1.7
2-methylheptane	-215.3	-214.6	-0.7
3-methylheptane	-212.5	-208.6	-3.9
4-methylheptane	-211.9	-211.5	-0.4
3-ethylhexane	-210.7	-207.7	-3.0

All values in kJ mol^{-1}	$\Delta_f H^o$ exp.	$\Delta_f H^o$ PM3 STO-3G B(π) δ(C-H)	$\Delta\Delta_f H^o$
2,2-dimethylhexane	-224.5	-222.3	-2.2
2,3-dimethylhexane	-213.8	-216.1	2.3
2,4-dimethylhexane	-219.2	-220.7	1.5
2,5-dimethylhexane	-222.5	-220.2	-2.3
3,3-dimethylhexane	-219.9	-218.2	-1.7
3,4-dimethylhexane	-212.8	-217.1	4.3
3-ethyl-2-methylpentane	-211	-216.8	5.8
3-ethyl-3-methylpentane	-214.8	-205.5	-9.3
2,2,3-trimethylpentane	-219.9	-224.1	4.2
2,2,4-trimethylpentane	-223.9	-225.0	1.1
2,3,3-trimethylpentane	-216.2	-215.7	-0.5
2,3,4-trimethylpentane	-217.2	-216.5	-0.7
2,2,3,3-tetramethylbutane	-226	-224.6	-1.4
Standard deviation			3.0

Graph 2.3



The combination of using HF B(π) values to correct PM3 $\Delta_f H^o$ values with systematic corrections for C-H bonds (Table 2.4) gives superior results than a correction for C-C bonds using HF or PM3 adjusted B(π).

2.2 Variation in PM3 Energies with Molecular Mechanics Steric Energies

The steric energy (SE) in Molecular Mechanics is perceived as contributions from deviations from ideal bond lengths angles and torsional angles as well as nonbonded van der Waals or coulombic interactions. From the overall deviation from the ideal values of the aforementioned kinds a steric energy is ascertained. The MMFF94 SE has no units and cannot be used in thermochemical calculations as it is considered specific to a molecule (in terms of a measure of the deviation from its ideal bond lengths and angles), the exceptions being comparisons between isomers e.g. *cis* and trans and conformers[73]. However one may fit a parameterised scheme to the SE as mentioned in the introductory section on molecular mechanics (MM) and we do this here by fitting δ, the difference between experimental $\Delta_f H^o$ and PM3 $\Delta_f H^o$ values, to a function of the SE. A correlation between strain energies and δ is most evident for the *n*-alkanes where it can be seen in Table 2.5 that as δ [from *eq.* (17)] increases for the *n*-alkanes steric energy decreases.

Table 2.5

Molecule	**$\Delta_f H^o$ Exp. (kJ mol^{-1})**	**$\Delta_f H^o$calc. PM3 (kJ mol^{-1})**	**δ (kJ mol^{-1})**	**Steric Energy**
methane	-74.8	-54.45	-20.35	0.03
ethane	-83.8	-75.88	-7.92	-4.73
propane	-104.7	-98.84	-5.86	-4.90
butane	-125.7	-121.60	-4.10	-5.08
2-methylpropane	-134.2	-123.56	-10.64	-0.48
pentane	-146.9	-144.28	-2.62	-5.27

Molecule	**$\Delta_f H^o$ Exp. (kJ mol^{-1})**	**$\Delta_f H^o$calc. PM3 (kJ mol^{-1})**	**δ (kJ mol^{-1})**	**Steric Energy**
2,2-dimethylpropane	-168	-149.93	-18.07	8.50
hexane	-166.9	-166.96	0.06	-5.47
heptane	-187.6	-189.64	2.04	-5.68
octane	-208.5	-212.32	3.82	-5.88

The steric energy was first tested by a fit to δ and secondly by fitting δ to steric energy and δ(C-C) given by *eq.* (24) and (25) respectively.

$$\Delta_f H^o(C_nH_{2n+2}) = \Delta_f H^o_{PM3} + \beta_1 \text{ SE} + \beta_o \qquad \ldots (24)$$

$$\Delta_f H^o(C_nH_{2n+2}) = \Delta_f H^o_{PM3} + \beta_1 \times \text{SE} + (n\text{-}1)\ \delta(\text{C-C}) + \beta_o \qquad \ldots (25)$$

$$\Delta_f H^o(C_nH_{2n+2}) = \Delta_f H^o_{PM3} + \beta_1 \times \text{SE} + (n\text{-}1)\ \delta(\text{C-C}) + (2n + 2)\ \delta(\text{C-H}) \ldots (26)$$

Eq. (25) and (26) are equivalent with *eq.* (26) having the advantage of interpreting β_0. The parameters are given in Table 2.6 and reproduced in Section 2.4.2.2. The parameters ascertained by *eq.* (25) can be expressed with the parameters of *eq.* (26) as shown in the following equivalences where $\delta(\text{C-C})_{(25)}$ is from *eq.* (25) and $\delta(\text{C-C})_{(26)}$ is from *eq.* (26):

$$\delta(\text{C-C})_{(25)} = \delta(\text{C-C})_{(26)} + 2\delta(\text{C-H})$$

$$\beta_o = 4\delta(\text{C-H})$$

Table 2.6

LSE's	β_0 (kJ mol^{-1})	β_1 (kJ mol^{-1})	δ(C-C) (kJ mol^{-1})	δ(C-H) (kJ mol^{-1})	s (kJ mol^{-1})
***eq.* (24)**	-11.5	-1.1			9.2
***eq.* (25)**	-18.3	-1.3	2.5		4.2
***eq.* (26)**		-1.3	11.6	-4.6	4.2
***eq.* (26)**		-2.2	11.1	-5.1*	21.2
***eq.* (26)†**		-1.2	12.9	-5.1*	3.8

* fixed at the δ value for methane divided by four.

† parameterised on all 40 alkanes.

In Table 2.7 excepting the last column, the regression analysis is performed on test group 1 as it was anticipated that the steric energy might appropriately accommodate steric strain in 2-methylbutane.

Table 2.7

All values except SE are in kJ mol^{-1}	$\Delta_f H^o$ Exp.	Steric Energy	$\Delta\Delta_f H^o$ *eq.* (24)	$\Delta\Delta_f H^o$ *eqs.* (25) & (26)	$\Delta\Delta_f H^o$ *eq.* (26)†	$\Delta\Delta_f H^o$ *eq.* (26)†*
methane	-74.8	0.0	-8.8	-2.0	0.1	0.0
ethane	-83.8	-4.7	-1.5	1.8	1.0	3.9
propane	-104.7	-4.9	0.4	1.2	1.8	3.0
butane	-125.7	-5.1	2.0	0.3	2.2	1.8
2-methylpropane	-134.2	-0.5	0.4	-0.3	5.8	0.9
pentane	-146.9	-5.3	3.3	-1.0	2.3	0.3
2-methylbutane	-153.6	0.3	1.6	-1.4	6.9	-0.6
2,2-dimethylpropane	-168	8.5	2.6	1.3	17.2	1.7
hexane	-166.9	-5.5	5.7	-1.0	3.5	0.0
2-methylpentane	-174.6	0.2	4.4	-1.1	8.6	-0.5
3-methylpentane	-171.9	1.4	4.3	-0.9	9.9	-0.4
2,2-dimethylbutane	-185.9	10.2	2.4	-1.0	17.9	-1.1
2,3-dimethylbutane	-178.1	6.4	3.2	-1.0	14.4	-0.8
heptane	-187.6	-5.7	7.5	-1.8	4.1	-1.1

All values except SE are in kJ mol^{-1}	**$\Delta_f H^o$ Exp.**	**Steric Energy**	**$\Delta\Delta_f H^o$ *eq.* (24)**	**$\Delta\Delta_f H^o$ *eqs.* (25) & (26)**	**$\Delta\Delta_f H^o$ *eq.* (26)†**	**$\Delta\Delta_f H^o$ *eq.* (26)†***
2-methylhexane	-194.5	0.0	7.0	-1.0	10.0	-0.7
3-methylhexane	-191.3	1.3	6.5	-1.2	11.0	-1.0
3-ethylpentane	-189.5	2.4	6.2	-1.3	11.9	-1.2
2,2-dimethylpentane	-205.7	10.1	6.5	0.6	20.9	0.2
2,3-dimethylpentane	-198.7	7.3	5.4	-1.1	16.7	-1.2
2,4-dimethylpentane	-201.6	5.4	7.1	0.2	16.3	0.2
3,3-dimethylpentane	-201.0	11.8	5.9	0.4	22.3	0.0
2,2,3-trimethylbutane	-204.4	17.4	8.9	4.6	31.6	3.8
octane	-208.5	-5.9	9.0	-2.7	4.5	-2.3
2-methylheptane	-215.3	-0.2	8.7	-1.9	10.5	-1.8
3-methylheptane	-212.5	1.1	7.7	-2.5	11.0	-2.6
4-methylheptane	-211.9	1.1	8.4	-1.9	11.7	-1.9
3-ethylhexane	-210.7	2.1	8.7	-1.3	13.2	-1.4
2,2-dimethylhexane	-224.5	9.9	10.1	1.7	23.3	1.1
2,3-dimethylhexane	-213.8	7.2	13.4	4.5	23.6	4.0
2,4-dimethylhexane	-219.2	6.5	12.1	3.0	21.5	2.6
2,5-dimethylhexane	-222.5	5.6	8.5	-0.8	16.9	-1.2
3,3-dimethylhexane	-219.9	11.6	9.9	1.9	25.1	1.2
3,4-dimethylhexane	-212.8	8.3	13.6	4.9	25.0	4.3
3-ethyl-2-methylpentane	-211.0	8.8	16.9	8.3	29.0	7.8
3-ethyl-3-methylpentane	-214.8	13.8	7.2	-0.4	24.9	-1.2
2,2,3-trimethylpentane	-219.9	19.0	16.6	10.2	40.2	9.0
2,2,4-trimethylpentane	-223.9	18.4	15.8	9.3	38.7	8.1

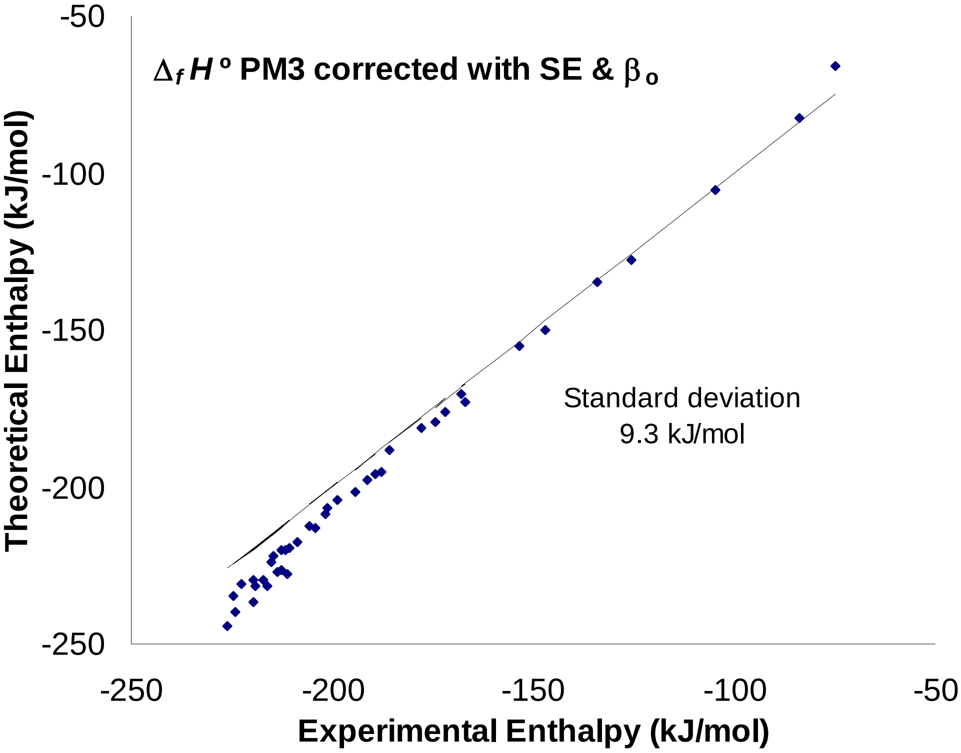
All values except SE are in kJ mol^{-1}	**$\Delta_f H^o$ Exp.**	**Steric Energy**	**$\Delta\Delta_f H^o$ *eq.* (24)**	**$\Delta\Delta_f H^o$ *eqs.* (25) & (26)**	**$\Delta\Delta_f H^o$ *eq.* (26)†**	**$\Delta\Delta_f H^o$ *eq.* (26)†***
2,3,3-trimethylpentane	-216.2	19.1	15.2	8.8	38.9	7.6
2,3,4-trimethylpentane	-217.2	14.2	12.2	4.7	30.2	3.8
2,2,3,3-tetramethylbutane	-226.0	29.8	18.3	14.2	54.0	12.2
Standard deviation			9.2	4.2	21.2	3.8

† δ(C-H) is ascertained from the δ value for methane divided by four.

* Parameters are fit to all forty alkanes. This is an instance where fitting all forty molecules has a greater than usual improvement in results at the expense of only ethane and propane of the test group 1 molecules.

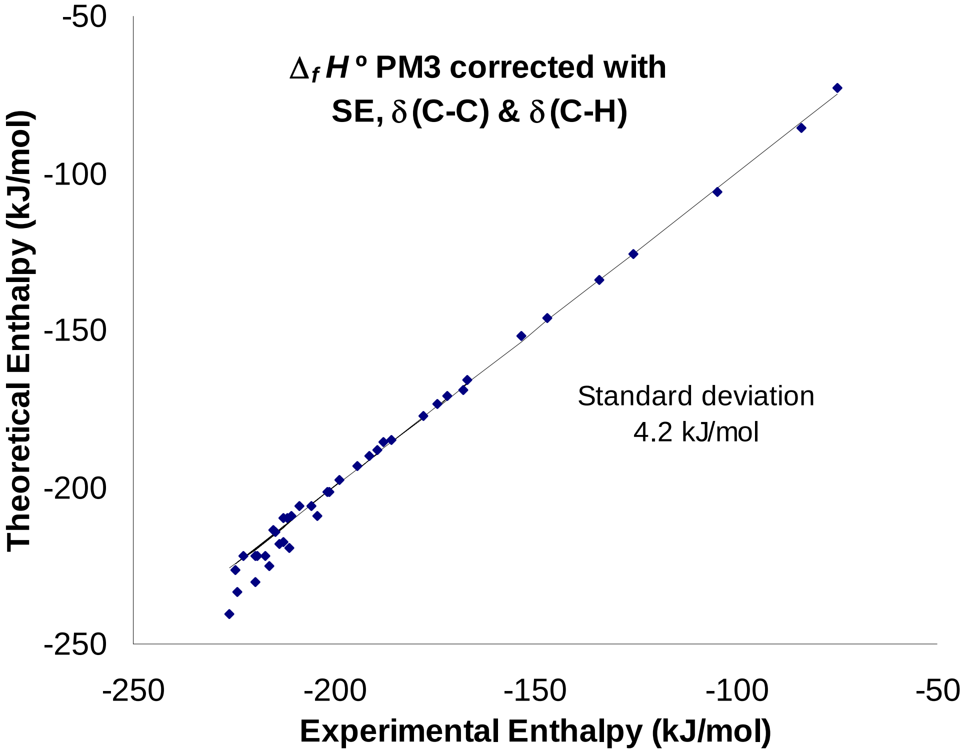
Graph 2.4

Uses *eq.* (24).



Graph 2.5

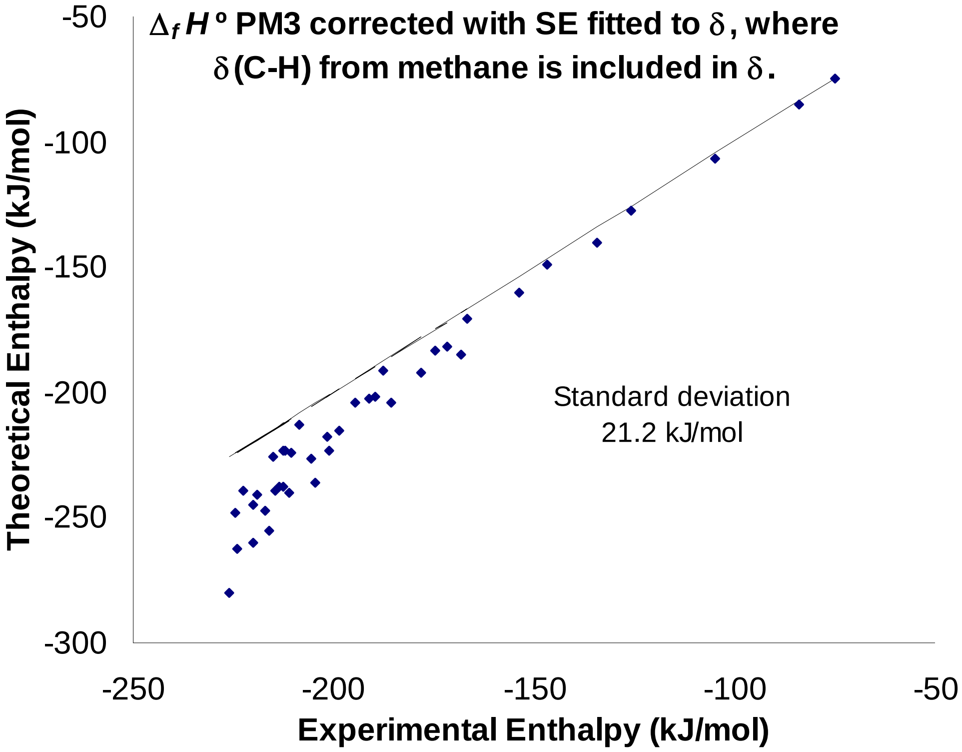
Uses either *eq.* (25) or (26).



If δ(C-H) is fixed at -5.1 kJ mol^{-1}, the δ value for methane divided by four [δ - (2*n*+2) δ(C-H)]/(*n*-1) is fitted to steric energy averaged over the C-C bond as was done for B(π) in the previous section, where δ'(C-C) and B(π) are substituted with β_1 and SE respectively in *eqs.* (20) and (21). The $\Delta_f H°$ can then be calculated by *eq.* (26) without parameterising δ(C-H). Using test group 1, a slope and intercept are ascertained corresponding to β_1 = -2.2 kJ mol^{-1} and δ(C-C) = 11.1 kJ mol^{-1} respectively. This gives a standard deviation of 21.2 kJ mol^{-1}. Parameterising over all 40 alkanes, Graph 2.7 gives a standard deviation (3.8 kJ mol^{-1}) similar to the standard deviations of the previous methods of this section.

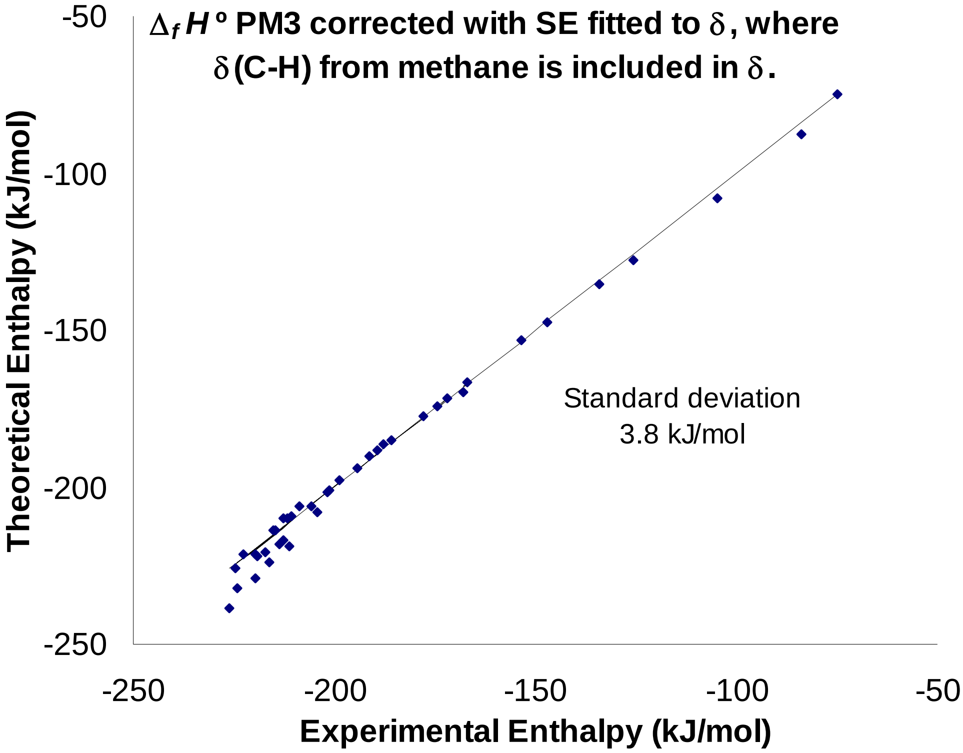
Graph 2.6

Using *eq*. (26) parameterised on test group 1 with δ(C-H) fixed.



Graph 2.7

Using *eq*. (26) parameterised on all forty molecules with δ(C-H) fixed.



All the models tested have a steric energy term with non-zero values ascertained even for conventionally unstrained molecules. If parameters are ascertained from test group 1, models in only two LSEs calculate energies too low for highly strained molecules. This is likely to be because test group 1 is too heavily weighted in unstrained molecules to correctly determine parameters for strained molecules. This can be seen in Graphs 2.4 and 2.6. The model for calculated values in Graph 2.5 has three LSEs: β_1, δ(C-C) and δ(C-H) or β_o depending on whether *eq.* (25) or (26) is used. The model for calculated values in Graph 2.7 had two LSEs for SE and C-C bonds, but was parameterised over all forty molecules. As can be seen in these two graphs, using three parameters or fitting two parameters over all forty molecules gave better estimation of $\Delta_f H^o$ in the highly strained molecules.

Henceforth, the approach where parameters are ascertained from a smaller subset of molecules of a larger set on which a model is tested, is considered less adequate for the terms and functions for which parameters are sought in the models of the remaining sections. Subsequently the entire set for which data is calculated is the set over which parameters are ascertained, unless specified otherwise.

2.3 Variation in Ab Initio Total Energy with B(π)

The energy given by HF output is the total energy (ΔE_{Tot}) and is the energy at 0 K with the Born-Oppenheimer approximation. We need to make $H_{vib}(0)$ and thermal corrections for the molecule at 298 K. ΔE_{Tot} is shown in Scheme 2.1 (a). A PV term is also required to convert energies to enthalpies by $\Delta H = \Delta E + P\Delta V$: where E is the sum of potential and kinetic energy and P and V are pressure and volume respectively. By the ideal gas approximation the $P\Delta V$ per mole is RT. The RT constant is included in the $\Delta H(T)$ energies, used along with $H_{vib}(0)$ to convert total energy to heat of formation at 298 K. The chemical equations for alkanes of Scheme 2.1 are constructed in the Born-Haber cycle for alkanes of Scheme 2.2 to show their relation to the heat of formation. In Scheme 2.2 the $\Delta_f H^o$ calculated via **A** → **B** → **C** is approximated with the calculated value for the lowest energy conformer. Between points **B** and **C** all ΔE_{Tot}, $H_{vib}(0)$ and $\Delta H(T)$ were calculated with the Spartan©[112] software package. Between points **AB** and **AC**, energies are associated with the experimental values.

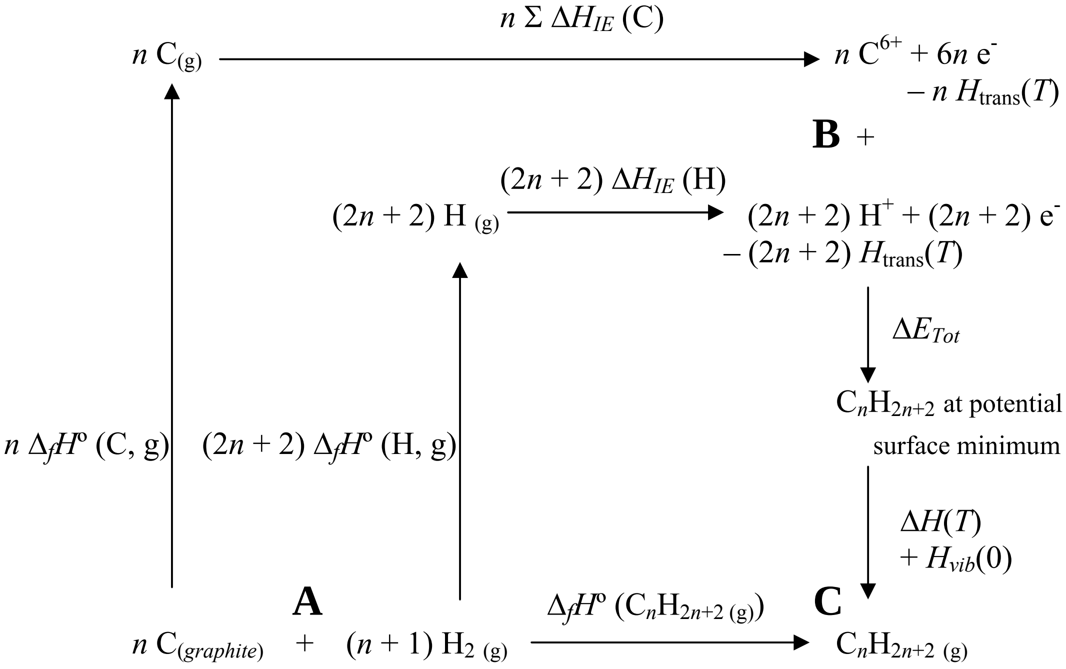
Scheme 2.1

a) ΔE_{Tot} (C_nH_{2n+2}, g): $nC^{6+} + (2n+2)H^+ + (8n+2)e^- \rightarrow C_nH_{2n+2}$

b) ΔH_a (C_nH_{2n+2}, g): $C_nH_{2n+2\ (g)} \rightarrow nC_{\ (g)} + (2n+2)H_{\ (g)}$

c) $\Delta_f H^o$ (C_nH_{2n+2}, g): $nC_{\ (graphite)} + (n+1)H_{2\ (g)} \rightarrow C_nH_{2n+2\ (g)}$

Scheme 2.2



$\Delta_f H^o$ (C, g) : 716.67 kJ mol^{-1}

$\Delta_f H^o$ (H, g) : 218.00 kJ mol^{-1}

$\Sigma\ \Delta H_{IE}$ (C) : 1086.4 + 2352.6 + 4620.5 + 6222.6 + 37830.4 + 47276.9
= 99389.4 kJ mol^{-1}

ΔH_{IE} (H) : 1312.0 kJ mol^{-1}

ΔE_{Tot} : total energy.

$H_{vib}(0)$: zero-point vibrational energy.

$\Delta H(T)$: vibrational and thermal energies.

$$\Delta H(T) = H_{\mathrm{trans}}(T) + H_{\mathrm{rot}}(T) + \Delta H_{\mathrm{vib}}(T) + RT$$

$H_{\mathrm{trans}}(T)$: Translational enthalpy : $H_{\mathrm{trans}}(T) = \frac{3}{2}RT$

$H_{\mathrm{rot}}(T)$: Rotational enthalpy : $H_{\mathrm{rot}}(T) = \frac{3}{2}\mathrm{RT}$ (RT for linear molecules)

The experimental total ionisation energy of the atom is the sum of successive ionisation energies of each electron in the atom. This is represented by $\Sigma\ \Delta H_{IE}$ (C) for the carbon atom, and for hydrogen because there is only one electron the total ionisation energy is the ionisation energy [ΔH_{IE} (H)]. The alternate route from **A** to **C** via **B** involving *ab initio* calculations as well as experimental atomic data is calculated in *eq.* (27).

$$\Delta_f H^{\circ}\ (\mathrm{C}_n\mathrm{H}_{2n+2}, \mathrm{g}) = n\ [\ \Delta_f H^{\circ}\ (\mathrm{C, g}) + \Sigma\ \Delta H_{IE}\ (\mathrm{C}) - H_{\mathrm{trans}}(T)] + (2n + 2)\ [\ \Delta_f H^{\circ}\ (\mathrm{H, g}) + \Delta H_{IE}\ (\mathrm{H}) - H_{\mathrm{trans}}(T)] + \Delta E_{Tot} + \Delta H(T) + H_{vib}(0) \qquad \ldots (27)$$

Conversely the experimental $\Delta_f H^{\circ}$ is converted to a total enthalpy at 298 K for alkanes [ΔH_{tot} (298, exp)] which corresponds to the reactions **B** to **C** via **A** in the Born-Haber cycle:

$$\Delta H^{\circ}_{Tot}\ (298, \mathrm{exp}) = \Delta_f H^{\circ} - n\ [\Sigma\ \Delta H_{IE}\ (\mathrm{C}) + \Delta_f H^{\circ}\ (\mathrm{C, g})] - (2n + 2)\ [\Delta H_{IE}\ (\mathrm{H}) + \Delta_f H^{\circ}\ (\mathrm{H, g})] \qquad \ldots (28)$$

For alkanes the calculated ΔH°_{Tot} (298) to be compared with the experimental value is therefore:

$$\Delta H^{\circ}_{Tot}\ (298, \textit{ab initio}) = \Delta E_{Tot} + \Delta H(T) + H_{vib}(0) - (3n + 2)\ H_{\mathrm{trans}}(T) \qquad \ldots (29)$$

The heat of formation is then related to this ΔH°_{Tot} by *eq.* (30).

$$\Delta_f H^{\circ} = \Delta H^{\circ}_{Tot} + n\ [\Delta_f H^{\circ}\ (\mathrm{C, g}) + \Sigma\ \Delta H_{IE}\ (\mathrm{C})] + (2n + 2)\ [\Delta_f H^{\circ}\ (\mathrm{H, g}) + \Delta H_{IE}\ (\mathrm{H})] \qquad \ldots (30)$$

With the full basis set using the HF-SCF *ab initio* model, the difference [δ of *eq.* (31)] between a calculated $\Delta_f H^{\circ}$ using *eq.* (27) (via ABC in Scheme 2.2) and experimental $\Delta_f H^{\circ}$ (via AC in Scheme 2.2) is the correlation energy (E_{corr}). Of course a finite basis set is used, so the difference is the correlation energy with a basis set truncation error (ΔE_{basis} = energy with complete basis set – energy with truncated basis set) and is exactly the same as the difference between ΔH°_{Tot} (298) calculated with the experimental values of *eq.* (28) [ΔH^{o}_{Tot} (298, exp)] and ΔH°_{Tot} (298) calculated via BC in Scheme 2.2 [ΔH^{o}_{Tot} (298, *ab initio*)]. The exact $\Delta_f H^{o}$ of the lowest energy conformer is given by *eq.* (27) if full configuration interaction (full CI) is used.

if

$$\delta = \Delta_f H^{o}\ (\text{exp}) - \Delta_f H^{o}\ (\text{HF-SCF})$$

which is the same as

$$\delta = \Delta H^{o}_{Tot}\ (298, \text{exp}) - \Delta H^{o}_{Tot}\ (298, \text{HF-SCF})$$

then

$$\delta = E_{corr} + \Delta E_{basis} \qquad \ldots(31)$$

The zero-point energy is usually accounted for, within or close to experimental error, by the parameterisation of any bond or group additivity scheme. If no parameter is sought for bond energies or anything that is a function of the number of atom types, the zero-point energy and kinetic and thermal energies must be corrected for. In the methods in this section we have for the most part avoided fitting parameters to functions of the number of atom types so that vibrational energies when explicitly calculated are not counterbalanced by the parameterisation of such functions. A precise conversion factor for converting from atomic units (a.u.) can be ascertained from the Rydberg constant as is done in appendix A.2. However small variation in the conversion factor does not affect the results of a model with parameters obtained in a least squares way, only the parameters are affected. In our regression analysis we used a conversion factor less precise than the conversion factor given in appendix A.2. Parameters in this section were ascertained by fitting with the entire set of 40 alkanes, and parameters are given for ΔE_{Tot} converted to kJ mol^{-1} with 2625.4 kJ mol^{-1} a.u.$^{-1}$.

The first method involved calculating only ΔE°_{Tot} and neglecting vibrational energies. Vibrational energies would therefore also be included in δ and this is approximated by scaling $\Delta E^{o}{}_{Tot}$ with a LSE and regression constant [*eq.* (32)] subsequent to subtracting conventional strain energy (the number of 1,4-*gauche* interactions multiplied by -2.5 kJ mol^{-1}) from the δ values. This was done for HF $\Delta E^{o}{}_{Tot}$ with results given in the ‘a’ labeled columns of Tables 2.11, 2.12 and 2.13 with the STO-3G, 3-21G(*) and 6-31G* basis sets respectively and parameters given in Table 2.8.

$$\Delta H^{o}{}_{Tot}\,(\mathrm{adj}) = \beta_1 \Delta E^{o}{}_{Tot}\,(\mathrm{HF}) + \beta_o \qquad \ldots(32)$$

Table 2.8

Parameters based on *eq.* (32) with K^{Etot} substituted for β_1.

	K^{Etot} (kJ mol^{-1})	β_o (kJ mol^{-1})	R^2	s (kJ mol^{-1})
STO-3G	1.019	-26.5	0.9458	9.0
3-21G(*)	1.012	-22.7	0.9910	3.7
6-31G*	1.007	-25.2	0.9471	8.9

Secondly as well as the scaling factor β_1 (K^{Etot}) and regression constant (β_o), $\Delta E^{o}{}_{Tot}$ (HF) was also adjusted with a term for B(π) as shown in *eq.* (33).

$$\Delta H^{o}{}_{Tot}\,(\mathrm{adj}) = \beta_1 \Delta E^{o}{}_{Tot}\,(\mathrm{HF}) + \beta_2 \mathrm{B}(\pi) + \beta_o \qquad \ldots(33)$$

This was done for HF $\Delta E^{o}{}_{Tot}$ with results given in the ‘b’ labeled columns of Tables 2.11, 2.12 and 2.13 with the STO-3G, 3-21G(*) and 6-31G* basis sets respectively and parameters given in Table 2.9.

Table 2.9

Parameters based on *eq.* (33) with K^{Etot} and $k^{\cdot\cdot}$ substituted for β_1 and β_2 respectively.

	K^{Etot} (kJ mol^{-1})	$K^{\cdot\cdot}$ (kJ mol^{-1})	β_o (kJ mol^{-1})	R^2	s (kJ mol^{-1})
STO-3G	1.019	-83.9	-24.6	0.9971	2.1
3-21G(*)	1.012	-49.2	-20.3	0.9978	1.8
6-31G*	1.007	-106.7	-22.1	0.9967	2.2

Thirdly a steric term was included to adjust ΔE^{o}_{Tot} (HF) shown in *eq.* (34).

$$\Delta H^{o}_{Tot}\,(\mathrm{adj}) = \beta_1 \Delta E^{o}_{Tot}\,(\mathrm{HF}) + \beta_2 \mathrm{B}(\pi) + \beta_3 \Sigma \mathrm{S}_{ij}\,(\mathrm{HH}) + \beta_o \qquad \ldots(34)$$

Instead of a conventional strain energy we use a more modern function for van der Waals interactions, the buffered 14 7 function[70], which we use to calculate a steric energy for only the H···H nonbonded interactions. This function has the form given in *eq.* (13) and is represented by the ΣS_{ij} (HH) term (sum of each ith with jth H atom interaction, not including interactions between geminal H atoms) in *eq.* (34).

This was done for HF ΔE^{o}_{Tot} with results given in the 'c' labeled columns of Tables 2.11, 2.12 and 2.13 with the STO-3G, 3-21G(*) and 6-31G* basis sets respectively and parameters given in Table 2.10.

Table 2.10

Parameters based on *eq.* (34) with K^{Etot} and $k^{\cdot\cdot}$ substituted for β_1 and β_2 respectively.

	K^{Etot} (kJ mol^{-1})	$K^{\cdot\cdot}$ (kJ mol^{-1})	β_3 (kJ mol^{-1})	β_o (kJ mol^{-1})	R^2	s (kJ mol^{-1})
STO-3G	1.019	-100.5	-4.9	-21.6	0.9954	2.6
3-21G(*)	1.012	-65.6	-5.6	-18.2	0.9981	1.7
6-31G*	1.007	-123.5	-9.2	-21.2	0.9963	2.3

Table 2.11

Calculations based on STO-3G energies. The a, b and c labels in column headings have the meanings: ‘a’ least squares analysis fitting slope of ΔE_{Tot} and regression constant; ‘b’ LSEs are found for ΔE_{Tot}, B(π) and regression constant; ‘c’ LSEs are found for ΔE_{Tot}, B(π),the modified buffered 14 7 steric function and regression constant. The $\Delta_f H^o$ exp. is the experimental value.

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\ a}$**	**$\Delta_f H^{o\ b}$**	**$\Delta_f H^{o\ c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
methane	-74.8	-64.2	-75.8	-76.2	-10.6	1.0	1.4
ethane	-83.8	-85.5	-82.5	-81.4	1.7	-1.3	-2.4
propane	-104.7	-109.3	-104.9	-104.4	4.6	0.2	-0.3
butane	-125.7	-133.0	-124.7	-124.2	7.3	-1.0	-1.5
2-methylpropane	-134.2	-133.5	-135.9	-136.8	-0.7	1.7	2.6
pentane	-146.9	-156.7	-148.0	-148.2	9.8	1.1	1.3
2-methylbutane	-153.6	-155.3	-154.5	-154.4	1.7	0.9	0.8
2,2-dimethylpropane	-168.0	-156.5	-168.5	-170.4	-11.5	0.5	2.4
hexane	-166.9	-180.4	-166.4	-166.3	13.5	-0.5	-0.6
2-methylpentane	-174.6	-178.6	-175.7	-176.1	4.0	1.1	1.5
3-methylpentane	-171.9	-176.3	-173.2	-172.9	4.4	1.3	1.0
2,2-dimethylbutane	-185.9	-175.8	-184.9	-184.6	-10.1	-1.0	-1.3
2,3-dimethylbutane	-178.1	-173.3	-176.3	-177.0	-4.8	-1.8	-1.1
heptane	-187.6	-204.0	-188.7	-189.0	16.4	1.1	1.4
2-methylhexane	-194.5	-202.3	-194.0	-194.0	7.8	-0.5	-0.5
3-methylhexane	-191.3	-196.6	-191.9	-192.3	5.3	0.6	1.0
3-ethylpentane	-189.5	-194.3	-186.3	-185.1	4.8	-3.2	-4.4
2,2-dimethylpentane	-205.7	-198.7	-205.3	-205.5	-7.0	-0.4	-0.2
2,3-dimethylpentane	-198.7	-193.2	-193.8	-194.5	-5.5	-4.9	-4.2
2,4-dimethylpentane	-201.6	-200.1	-202.4	-202.8	-1.5	0.8	1.2
3,3-dimethylpentane	-201.0	-195.0	-200.3	-197.5	-6.0	-0.7	-3.5
2,2,3-trimethylbutane	-204.4	-190.3	-202.4	-202.5	-14.1	-2.0	-1.9
octane	-208.5	-227.7	-208.4	-208.8	19.2	-0.1	0.3
2-methylheptane	-215.3	-226.0	-216.8	-217.4	10.7	1.5	2.1

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\ a}$**	**$\Delta_f H^{o\ b}$**	**$\Delta_f H^{o\ c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
3-methylheptane	-212.5	-220.3	-209.9	-209.9	7.8	-2.6	-2.6
4-methylheptane	-211.9	-219.9	-213.5	-214.3	8.0	1.6	2.4
3-ethylhexane	-210.7	-217.3	-208.4	-207.9	6.6	-2.3	-2.8
2,2-dimethylhexane	-224.5	-222.4	-223.9	-223.9	-2.1	-0.6	-0.6
2,3-dimethylhexane	-213.8	-216.4	-215.6	-217.1	2.6	1.8	3.3
2,4-dimethylhexane	-219.2	-220.5	-219.3	-219.3	1.3	0.1	0.1
2,5-dimethylhexane	-222.5	-224.4	-221.7	-221.5	1.9	-0.8	-1.0
3,3-dimethylhexane	-219.9	-217.9	-222.7	-220.7	-2.0	2.8	0.8
3,4-dimethylhexane	-212.8	-214.0	-217.3	-218.7	1.2	4.5	5.9
3-ethyl-2-methylpentane	-211.0	-213.0	-214.5	-215.9	2.0	3.5	4.9
3-ethyl-3-methylpentane	-214.8	-212.7	-214.3	-210.1	-2.1	-0.5	-4.7
2,2,3-trimethylpentane	-219.9	-209.6	-224.1	-225.0	-10.3	4.2	5.1
2,2,4-trimethylpentane	-223.9	-211.3	-221.8	-221.7	-12.6	-2.1	-2.2
2,3,3-trimethylpentane	-216.2	-208.6	-218.7	-216.8	-7.6	2.5	0.6
2,3,4-trimethylpentane	-217.2	-208.5	-215.4	-218.5	-8.7	-1.8	1.3
2,2,3,3-tetramethylbutane	-226.0	-200.7	-221.4	-220.5	-25.3	-4.6	-5.5
Standard deviation					9.033	2.095	2.628

Table 2.12

Calculations based on HF 3-21G(*) energies. The letters a, b and c in column headings have meanings: 'a' least squares analysis fitting slope of ΔE_{Tot} and regression constant; 'b' LSEs are found for ΔE_{Tot}, B(π) and regression constant; 'c' LSEs are found for ΔE_{Tot}, B(π), the modified buffered 14 7 steric function and regression constant. The $\Delta_f H^o$ exp. is the experimental value.

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\,a}$**	**$\Delta_f H^{o\,b}$**	**$\Delta_f H^{o\,c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
methane	-74.8	-68.9	-73.5	-74.0	-5.9	-1.3	-0.8
ethane	-83.8	-85.1	-83.7	-83.2	1.3	-0.1	-0.6
propane	-104.7	-107.4	-105.5	-105.4	2.7	0.8	0.7
butane	-125.7	-129.2	-126.0	-126.0	3.5	0.3	0.3
2-methylpropane	-134.2	-134.5	-135.5	-136.1	0.3	1.3	1.9
pentane	-146.9	-150.7	-147.5	-147.8	3.8	0.6	0.9
2-methylbutane	-153.6	-154.6	-154.2	-153.9	1.0	0.6	0.3
2,2-dimethylpropane	-168.0	-164.5	-169.2	-169.8	-3.5	1.2	1.8
hexane	-166.9	-172.3	-167.2	-167.4	5.4	0.3	0.5
2-methylpentane	-174.6	-176.0	-175.1	-175.3	1.4	0.5	0.7
3-methylpentane	-171.9	-173.9	-172.5	-172.4	2.0	0.6	0.5
2,2-dimethylbutane	-185.9	-182.2	-185.6	-184.4	-3.7	-0.3	-1.5
2,3-dimethylbutane	-178.1	-174.7	-175.5	-176.0	-3.4	-2.6	-2.1
heptane	-187.6	-193.9	-188.4	-188.9	6.3	0.8	1.3
2-methylhexane	-194.5	-197.6	-194.8	-194.7	3.1	0.3	0.2
3-methylhexane	-191.3	-192.2	-190.7	-191.3	0.9	-0.6	0.0
3-ethylpentane	-189.5	-190.1	-187.1	-186.9	0.6	-2.4	-2.6
2,2-dimethylpentane	-205.7	-203.4	-206.1	-205.3	-2.3	0.4	-0.4
2,3-dimethylpentane	-198.7	-193.4	-193.6	-194.6	-5.3	-5.1	-4.1
2,4-dimethylpentane	-201.6	-201.4	-202.6	-202.6	-0.2	1.0	1.0
3,3-dimethylpentane	-201.0	-200.3	-202.0	-198.7	-0.7	1.0	-2.3
2,2,3-trimethylbutane	-204.4	-199.5	-203.7	-203.2	-4.9	-0.7	-1.2
octane	-208.5	-215.5	-208.5	-208.9	7.0	0.0	0.4
2-methylheptane	-215.3	-219.2	-216.1	-216.3	3.9	0.8	1.0

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\,a}$**	**$\Delta_f H^{o\,b}$**	**$\Delta_f H^{o\,c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
3-methylheptane	-212.5	-213.8	-210.2	-210.5	1.3	-2.3	-2.0
4-methylheptane	-211.9	-213.5	-211.4	-212.2	1.6	-0.5	0.3
3-ethylhexane	-210.7	-211.4	-208.1	-208.4	0.7	-2.6	-2.3
2,2-dimethylhexane	-224.5	-225.0	-226.0	-225.1	0.5	1.5	0.6
2,3-dimethylhexane	-213.8	-214.7	-214.4	-216.0	0.9	0.6	2.2
2,4-dimethylhexane	-219.2	-220.3	-219.9	-219.9	1.1	0.7	0.7
2,5-dimethylhexane	-222.5	-223.1	-222.6	-221.8	0.6	0.1	-0.7
3,3-dimethylhexane	-219.9	-221.5	-223.1	-220.4	1.6	3.2	0.5
3,4-dimethylhexane	-212.8	-212.5	-213.4	-215.1	-0.3	0.6	2.3
3-ethyl-2-methylpentane	-211.0	-211.9	-212.1	-214.1	0.9	1.1	3.1
3-ethyl-3-methylpentane	-214.8	-216.7	-216.7	-212.8	1.9	1.9	-2.0
2,2,3-trimethylpentane	-219.9	-217.3	-222.6	-223.0	-2.6	2.7	3.1
2,2,4-trimethylpentane	-223.9	-218.3	-222.0	-220.4	-5.6	-1.9	-3.5
2,3,3-trimethylpentane	-216.2	-216.7	-220.0	-217.9	0.5	3.8	1.7
2,3,4-trimethylpentane	-217.2	-211.9	-214.6	-218.1	-5.3	-2.6	0.9
2,2,3,3,-tetramethylbutane	-226.0	-214.9	-222.0	-225.3	-11.1	-4.0	-0.7
Standard deviation					3.657	1.824	1.696

Table 2.13

Calculations based on HF 6-31G* energies. The a, b and c labels in column headings have meanings: ‘a’ least squares analysis fitting slope of ΔE_{Tot} and regression constant; ‘b’ LSEs are found for ΔE_{Tot}, B(π) and regression constant; ‘c’ LSEs are found for ΔE_{Tot}, B(π), the modified buffered 14 7 steric function and regression constant. The $\Delta_f H^o$ exp. is the experimental value.

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\ a}$**	**$\Delta_f H^{o\ b}$**	**$\Delta_f H^{o\ c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
methane	-74.8	-63.7	-74.3	-75.5	-11.1	-0.5	0.7
ethane	-83.8	-84.6	-81.1	-81.1	0.8	-2.7	-2.7
propane	-104.7	-108.8	-104.0	-104.2	4.1	-0.7	-0.5
butane	-125.7	-132.8	-125.1	-124.9	7.1	-0.6	-0.8
2-methylpropane	-134.2	-134.3	-136.7	-137.0	0.1	2.5	2.8
pentane	-146.9	-156.5	-148.1	-148.0	9.6	1.2	1.1
2-methylbutane	-153.6	-155.6	-154.6	-154.3	2.0	1.0	0.7
2,2-dimethylpropane	-168.0	-158.4	-171.6	-170.5	-9.6	3.6	2.5
hexane	-166.9	-180.3	-167.2	-166.6	13.4	0.3	-0.3
2-methylpentane	-174.6	-178.9	-176.4	-176.2	4.3	1.8	1.6
3-methylpentane	-171.9	-176.0	-172.4	-172.9	4.1	0.5	1.0
2,2-dimethylbutane	-185.9	-176.8	-186.6	-184.8	-9.1	0.7	-1.1
2,3-dimethylbutane	-178.1	-173.2	-175.7	-176.9	-4.9	-2.4	-1.2
heptane	-187.6	-204.0	-189.8	-189.1	16.4	2.2	1.5
2-methylhexane	-194.5	-202.7	-195.5	-194.7	8.2	1.0	0.2
3-methylhexane	-191.3	-196.2	-191.0	-191.7	4.9	-0.3	0.4
3-ethylpentane	-189.5	-193.5	-185.0	-185.6	4.0	-4.5	-3.9
2,2-dimethylpentane	-205.7	-199.7	-207.5	-205.9	-6.0	1.8	0.2
2,3-dimethylpentane	-198.7	-192.8	-193.4	-195.2	-5.9	-5.3	-3.5
2,4-dimethylpentane	-201.6	-201.4	-204.2	-203.7	-0.2	2.6	2.1
3,3-dimethylpentane	-201.0	-195.4	-200.9	-197.8	-5.6	-0.1	-3.2
2,2,3-trimethylbutane	-204.4	-190.5	-202.8	-203.8	-13.9	-1.6	-0.6
octane	-208.5	-227.8	-209.1	-207.9	19.3	0.6	-0.6
2-methylheptane	-215.3	-226.4	-217.7	-216.9	11.1	2.4	1.6

All values in kJ mol^{-1}	$\Delta_f H^o$ **exp.**	$\Delta_f H^{o\,a}$	$\Delta_f H^{o\,b}$	$\Delta_f H^{o\,c}$	$\Delta\Delta_f H^o$ **a**	$\Delta\Delta_f H^o$ **b**	$\Delta\Delta_f H^o$ **c**
3-methylheptane	-212.5	-219.9	-210.0	-210.0	7.4	-2.5	-2.5
4-methylheptane	-211.9	-219.5	-212.1	-212.8	7.6	0.2	0.9
3-ethylhexane	-210.7	-216.6	-207.4	-208.1	5.9	-3.3	-2.6
2,2-dimethylhexane	-224.5	-223.4	-227.1	-224.9	-1.1	2.6	0.4
2,3-dimethylhexane	-213.8	-216.0	-214.8	-216.9	2.2	1.0	3.1
2,4-dimethylhexane	-219.2	-221.2	-220.1	-219.9	2.0	0.9	0.7
2,5-dimethylhexane	-222.5	-225.4	-223.8	-222.1	2.9	1.3	-0.4
3,3-dimethylhexane	-219.9	-218.3	-222.6	-219.7	-1.6	2.7	-0.2
3,4-dimethylhexane	-212.8	-212.5	-212.5	-214.9	-0.3	-0.3	2.1
3-ethyl-2-methylpentane	-211.0	-211.2	-210.7	-214.3	0.2	-0.3	3.3
3-ethyl-3-methylpentane	-214.8	-212.3	-213.5	-211.3	-2.5	-1.3	-3.5
2,2,3-trimethylpentane	-219.9	-208.9	-222.3	-224.5	-11.0	2.4	4.6
2,2,4-trimethylpentane	-223.9	-212.4	-223.5	-221.1	-11.5	-0.4	-2.8
2,3,3-trimethylpentane	-216.2	-208.0	-217.8	-217.9	-8.2	1.6	1.7
2,3,4-trimethylpentane	-217.2	-207.8	-214.3	-220.7	-9.4	-2.9	3.5
2,2,3,3-tetramethylbutane	-226.0	-200.3	-220.8	-219.7	-25.7	-5.2	-6.3
Standard deviation					8.927	2.233	2.336

The form of Tables 2.8, 2.9, 2.10, 2.11, 2.12 and 2.13 can be recreated with results from calculating vibrational energies explicitly to calculate ΔH^o_{Tot} (298, *ab intio*) and substituting this for ΔE_{Tot} in *eqs.* (32), (33) and (34).

Table 2.14

Parameters based on *eq.* (32) with ΔH^{o}_{Tot} and K^{Htot} substituted for ΔE_{Tot} and β_1 respectively.

	K^{Htot} (kJ mol^{-1})	β_o (kJ mol^{-1})	R^2	s (kJ mol^{-1})
STO-3G	1.020	-85.6	0.9604	7.7
3-21G(*)	1.013	-75.1	0.9956	2.6
6-31G*	1.008	-76.6	0.9594	7.8

Table 2.15

Parameters based on *eq.* (33) with ΔH^{o}_{Tot}, K^{Htot} and $k^{\cdot\cdot}$ substituted for ΔE_{Tot}, β_1 and β_2 respectively.

	K^{Htot} (kJ mol^{-1})	$K^{\cdot\cdot}$ (kJ mol^{-1})	β_o (kJ mol^{-1})	R^2	s (kJ mol^{-1})
STO-3G	1.020	-71.0	-84.0	0.9973	2.0
3-21G(*)	1.013	-28.4	-73.7	0.9978	1.8
6-31G*	1.008	-91.7	-73.9	0.9961	2.4

Table 2.16

Parameters based on *eq.* (34) with ΔH^{o}_{Tot}, K^{Htot} and $k^{\cdot\cdot}$ substituted for ΔE_{Tot}, β_1 and β_2 respectively.

	K^{Htot} (kJ mol^{-1})	$K^{\cdot\cdot}$ (kJ mol^{-1})	β_3 (kJ mol^{-1})	β_o (kJ mol^{-1})	R^2	s (kJ mol^{-1})
STO-3G	1.020	-86.2	-5.7	-81.6	0.9958	2.5
3-21G(*)	1.013	-43.7	-5.9	-71.9	0.9980	1.7
6-31G*	1.008	-107.3	-10.0	-73.5	0.9963	2.3

Table 2.17

Calculations based on ΔH^{o}_{Tot} (298, STO-3G) energies. The a, b and c labels in column headings have the meanings: 'a' least squares analysis fitting slope of ΔH_{Tot} and regression constant; 'b' LSEs are found for ΔH_{Tot}, B(π) and regression constant; 'c' LSEs are found for ΔH_{Tot}, B(π),the modified buffered 14 7 steric function and regression constant. The $\Delta_f H^o$ exp. is the experimental value.

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\,a}$**	**$\Delta_f H^{o\,b}$**	**$\Delta_f H^{o\,c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
methane	-74.8	-65.9	-75.7	-76.3	-8.9	0.9	1.5
ethane	-83.8	-84.9	-82.4	-81.6	1.1	-1.4	-2.2
propane	-104.7	-108.4	-104.7	-104.4	3.7	0.0	-0.3
butane	-125.7	-131.9	-124.8	-124.4	6.2	-0.9	-1.3
2-methylpropane	-134.2	-133.8	-135.7	-136.5	-0.4	1.5	2.3
pentane	-146.9	-155.2	-147.9	-148.0	8.3	1.0	1.1
2-methylbutane	-153.6	-155.1	-154.4	-154.2	1.5	0.8	0.6
2,2-dimethylpropane	-168.0	-159.0	-169.1	-170.3	-9.0	1.1	2.3
hexane	-166.9	-178.6	-166.9	-166.6	11.7	0.0	-0.3
2-methylpentane	-174.6	-178.1	-175.7	-175.9	3.5	1.1	1.3
3-methylpentane	-171.9	-175.6	-173.0	-172.9	3.7	1.1	1.0
2,2-dimethylbutane	-185.9	-177.5	-185.2	-184.5	-8.4	-0.7	-1.4
2,3-dimethylbutane	-178.1	-173.8	-176.3	-177.0	-4.3	-1.8	-1.1
heptane	-187.6	-202.1	-189.1	-189.2	14.5	1.5	1.6
2-methylhexane	-194.5	-201.5	-194.5	-194.3	7.0	0.0	-0.2
3-methylhexane	-191.3	-195.3	-191.3	-191.7	4.0	0.0	0.4
3-ethylpentane	-189.5	-192.6	-185.8	-185.2	3.1	-3.7	-4.3
2,2-dimethylpentane	-205.7	-200.1	-205.7	-205.4	-5.6	0.0	-0.3
2,3-dimethylpentane	-198.7	-193.1	-193.6	-194.6	-5.6	-5.1	-4.1
2,4-dimethylpentane	-201.6	-200.5	-202.4	-202.6	-1.1	0.8	1.0
3,3-dimethylpentane	-201.0	-195.9	-200.4	-197.5	-5.1	-0.6	-3.5
2,2,3-trimethylbutane	-204.4	-192.5	-202.7	-203.0	-11.9	-1.7	-1.4
octane	-208.5	-225.3	-209.0	-209.1	16.8	0.5	0.6
2-methylheptane	-215.3	-224.9	-217.2	-217.4	9.6	1.9	2.1

All values in kJ mol^{-1}	$\Delta_f H^o$ **exp.**	$\Delta_f H^{o\ a}$	$\Delta_f H^{o\ b}$	$\Delta_f H^{o\ c}$	$\Delta\Delta_f H^o$ **a**	$\Delta\Delta_f H^o$ **b**	$\Delta\Delta_f H^o$ **c**
3-methylheptane	-212.5	-218.7	-209.9	-209.9	6.2	-2.6	-2.6
4-methylheptane	-211.9	-218.3	-212.8	-213.7	6.4	0.9	1.8
3-ethylhexane	-210.7	-215.4	-207.8	-207.7	4.7	-2.9	-3.0
2,2-dimethylhexane	-224.5	-223.5	-224.8	-224.2	-1.0	0.3	-0.3
2,3-dimethylhexane	-213.8	-216.0	-215.3	-217.1	2.2	1.5	3.3
2,4-dimethylhexane	-219.2	-220.4	-219.4	-219.3	1.2	0.2	0.1
2,5-dimethylhexane	-222.5	-224.5	-222.3	-221.7	2.0	-0.2	-0.8
3,3-dimethylhexane	-219.9	-218.5	-222.6	-220.4	-1.4	2.7	0.5
3,4-dimethylhexane	-212.8	-213.4	-216.2	-218.1	0.6	3.4	5.3
3-ethyl-2-methylpentane	-211.0	-212.6	-213.9	-215.9	1.6	2.9	4.9
3-ethyl-3-methylpentane	-214.8	-212.9	-214.2	-210.5	-1.9	-0.6	-4.3
2,2,3-trimethylpentane	-219.9	-211.3	-223.5	-224.7	-8.6	3.6	4.8
2,2,4-trimethylpentane	-223.9	-213.6	-222.5	-221.7	-10.3	-1.4	-2.2
2,3,3-trimethylpentane	-216.2	-209.9	-218.5	-216.9	-6.3	2.3	0.7
2,3,4-trimethylpentane	-217.2	-209.0	-214.8	-218.6	-8.2	-2.4	1.4
2,2,3,3-tetramethylbutane	-226.0	-204.2	-221.8	-220.9	-21.8	-4.2	-5.1
Standard deviation					7.699	1.986	2.490

Table 2.18

Calculations based on ΔH^o_{Tot} (298, 3-21G(*)) energies. The letters a, b and c in column headings have meanings: ‘a’ least squares analysis fitting slope of ΔH_{Tot} and regression constant; ‘b’ LSEs are found for ΔH_{Tot}, $B(\pi)$ and regression constant; ‘c’ LSEs are found for ΔH_{Tot}, $B(\pi)$, the modified buffered 14 7 steric function and regression constant. The $\Delta_f H^o$ exp. is the experimental value.

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\,a}$**	**$\Delta_f H^{o\,b}$**	**$\Delta_f H^{o\,c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
methane	-74.8	-70.6	-73.2	-73.8	-4.2	-1.6	-1.0
ethane	-83.8	-84.5	-83.6	-83.3	0.7	-0.2	-0.5
propane	-104.7	-106.5	-105.4	-105.5	1.8	0.7	0.8
butane	-125.7	-128.1	-126.3	-126.3	2.4	0.6	0.6
2-methylpropane	-134.2	-134.8	-135.4	-135.9	0.6	1.2	1.7
pentane	-146.9	-149.6	-147.7	-148.0	2.7	0.8	1.1
2-methylbutane	-153.6	-154.4	-154.2	-153.9	0.8	0.6	0.3
2,2-dimethylpropane	-168.0	-166.7	-169.5	-169.7	-1.3	1.5	1.7
hexane	-166.9	-171.0	-168.0	-168.1	4.1	1.1	1.2
2-methylpentane	-174.6	-175.6	-175.1	-175.1	1.0	0.5	0.5
3-methylpentane	-171.9	-173.1	-172.3	-172.3	1.2	0.4	0.4
2,2-dimethylbutane	-185.9	-183.7	-185.7	-184.2	-2.2	-0.2	-1.7
2,3-dimethylbutane	-178.1	-175.0	-175.5	-176.0	-3.1	-2.6	-2.1
heptane	-187.6	-192.2	-189.0	-189.4	4.6	1.4	1.8
2-methylhexane	-194.5	-197.0	-195.4	-195.2	2.5	0.9	0.7
3-methylhexane	-191.3	-190.9	-190.0	-190.7	-0.4	-1.3	-0.6
3-ethylpentane	-189.5	-188.5	-186.7	-186.7	-1.0	-2.8	-2.8
2,2-dimethylpentane	-205.7	-204.6	-206.2	-205.2	-1.1	0.5	-0.5
2,3-dimethylpentane	-198.7	-193.1	-193.2	-194.4	-5.6	-5.5	-4.3
2,4-dimethylpentane	-201.6	-201.8	-202.5	-202.3	0.2	0.9	0.7
3,3-dimethylpentane	-201.0	-201.0	-201.9	-198.6	0.0	0.9	-2.4
2,2,3-trimethylbutane	-204.4	-201.3	-203.7	-203.2	-3.1	-0.7	-1.2
octane	-208.5	-213.6	-209.5	-209.8	5.1	1.0	1.3
2-methylheptane	-215.3	-218.4	-216.6	-216.7	3.1	1.3	1.4

All values in kJ mol^{-1}	$\Delta_f H^o$ exp.	$\Delta_f H^{o\,a}$	$\Delta_f H^{o\,b}$	$\Delta_f H^{o\,c}$	$\Delta\Delta_f H^o$ a	$\Delta\Delta_f H^o$ b	$\Delta\Delta_f H^o$ c
3-methylheptane	-212.5	-212.3	-210.2	-210.5	-0.2	-2.3	-2.0
4-methylheptane	-211.9	-212.0	-210.7	-211.6	0.1	-1.2	-0.3
3-ethylhexane	-210.7	-209.5	-207.6	-208.0	-1.2	-3.1	-2.7
2,2-dimethylhexane	-224.5	-226.0	-226.6	-225.4	1.5	2.1	0.9
2,3-dimethylhexane	-213.8	-214.2	-214.0	-215.7	0.4	0.2	1.9
2,4-dimethylhexane	-219.2	-220.1	-219.9	-219.9	0.9	0.7	0.7
2,5-dimethylhexane	-222.5	-223.3	-223.0	-222.1	0.8	0.5	-0.4
3,3-dimethylhexane	-219.9	-221.9	-222.8	-220.0	2.0	2.9	0.1
3,4-dimethylhexane	-212.8	-211.6	-212.2	-214.0	-1.2	-0.6	1.2
3-ethyl-2-methylpentane	-211.0	-211.1	-211.2	-213.4	0.1	0.2	2.4
3-ethyl-3-methylpentane	-214.8	-216.5	-216.6	-212.8	1.7	1.8	-2.0
2,2,3-trimethylpentane	-219.9	-218.6	-221.6	-222.1	-1.3	1.7	2.2
2,2,4-trimethylpentane	-223.9	-220.2	-222.4	-220.5	-3.7	-1.5	-3.4
2,3,3-trimethylpentane	-216.2	-217.6	-219.5	-217.5	1.4	3.3	1.3
2,3,4-trimethylpentane	-217.2	-212.2	-213.7	-217.4	-5.0	-3.5	0.2
2,2,3,3,-tetramethylbutane	-226.0	-221.1	-225.2	-228.8	-4.9	-0.8	2.8
Standard deviation					2.566	1.799	1.708

Table 2.19

Calculations based on ΔH^{o}_{Tot} (298, 6-31G*) energies. The a, b and c labels in column headings have meanings: ‘a’ least squares analysis fitting slope of ΔH_{Tot} and regression constant; ‘b’ LSEs are found for ΔH_{Tot}, B(π) and regression constant; ‘c’ LSEs are found for ΔH_{Tot}, B(π), the modified buffered 14 7 steric function and regression constant. The $\Delta_f H^{o}$ exp. is the experimental value.

All values in kJ mol^{-1}	**$\Delta_f H^{o}$ exp.**	**$\Delta_f H^{o\ a}$**	**$\Delta_f H^{o\ b}$**	**$\Delta_f H^{o\ c}$**	**$\Delta\Delta_f H^{o}$ a**	**$\Delta\Delta_f H^{o}$ b**	**$\Delta\Delta_f H^{o}$ c**
methane	-74.8	-65.1	-74.2	-75.5	-9.7	-0.6	0.7
ethane	-83.8	-83.9	-80.9	-81.2	0.1	-2.9	-2.6
propane	-104.7	-107.9	-103.7	-104.1	3.2	-1.0	-0.6
butane	-125.7	-131.7	-125.0	-125.0	6.0	-0.7	-0.7
2-methylpropane	-134.2	-134.6	-136.7	-136.8	0.4	2.5	2.6
pentane	-146.9	-155.2	-148.0	-147.9	8.3	1.1	1.0
2-methylbutane	-153.6	-155.5	-154.6	-154.2	1.9	1.0	0.6
2,2-dimethylpropane	-168.0	-160.7	-172.0	-170.2	-7.3	4.0	2.2
hexane	-166.9	-178.8	-167.6	-166.8	11.9	0.7	-0.1
2-methylpentane	-174.6	-178.6	-176.4	-176.1	4.0	1.8	1.5
3-methylpentane	-171.9	-175.4	-172.2	-173.0	3.5	0.3	1.1
2,2-dimethylbutane	-185.9	-178.3	-186.7	-184.5	-7.6	0.8	-1.4
2,3-dimethylbutane	-178.1	-173.5	-175.7	-177.0	-4.6	-2.4	-1.1
heptane	-187.6	-202.4	-190.2	-189.2	14.8	2.6	1.6
2-methylhexane	-194.5	-202.2	-196.0	-195.0	7.7	1.5	0.5
3-methylhexane	-191.3	-194.9	-190.5	-191.2	3.6	-0.8	-0.1
3-ethylpentane	-189.5	-191.9	-184.6	-185.5	2.4	-4.9	-4.0
2,2-dimethylpentane	-205.7	-201.0	-207.7	-205.6	-4.7	2.0	-0.1
2,3-dimethylpentane	-198.7	-192.5	-193.0	-195.1	-6.2	-5.7	-3.6
2,4-dimethylpentane	-201.6	-201.8	-204.3	-203.5	0.2	2.7	1.9
3,3-dimethylpentane	-201.0	-196.1	-200.8	-197.6	-4.9	-0.2	-3.4
2,2,3-trimethylbutane	-204.4	-194.8	-205.3	-206.4	-9.6	0.9	2.0
octane	-208.5	-226.0	-210.0	-208.4	17.5	1.5	-0.1
2-methylheptane	-215.3	-225.7	-218.3	-217.1	10.4	3.0	1.8

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\ a}$**	**$\Delta_f H^{o\ b}$**	**$\Delta_f H^{o\ c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
3-methylheptane	-212.5	-218.5	-209.9	-210.0	6.0	-2.6	-2.5
4-methylheptane	-211.9	-218.1	-211.7	-212.4	6.2	-0.2	0.5
3-ethylhexane	-210.7	-214.8	-206.9	-207.9	4.1	-3.8	-2.8
2,2-dimethylhexane	-224.5	-224.6	-227.7	-225.0	0.1	3.2	0.5
2,3-dimethylhexane	-213.8	-215.5	-214.5	-216.7	1.7	0.7	2.9
2,4-dimethylhexane	-219.2	-221.0	-220.1	-219.9	1.8	0.9	0.7
2,5-dimethylhexane	-222.5	-225.8	-224.4	-222.4	3.3	1.9	-0.1
3,3-dimethylhexane	-219.9	-218.7	-222.4	-219.4	-1.2	2.5	-0.5
3,4-dimethylhexane	-212.8	-211.7	-211.6	-214.5	-1.1	-1.2	1.7
3-ethyl-2-methylpentane	-211.0	-210.6	-210.2	-214.3	-0.4	-0.8	3.3
3-ethyl-3-methylpentane	-214.8	-212.1	-213.1	-211.4	-2.7	-1.7	-3.4
2,2,3-trimethylpentane	-219.9	-210.1	-221.7	-224.1	-9.8	1.8	4.2
2,2,4-trimethylpentane	-223.9	-214.4	-223.9	-221.0	-9.5	0.0	-2.9
2,3,3-trimethylpentane	-216.2	-208.9	-217.2	-217.7	-7.3	1.0	1.5
2,3,4-trimethylpentane	-217.2	-208.0	-213.6	-220.7	-9.2	-3.6	3.5
2,2,3,3-tetramethylbutane	-226.0	-203.0	-220.5	-219.6	-23.0	-5.5	-6.4
Standard deviation					7.808	2.421	2.335

A simple steric repulsion (SSR) term was also tested: $A\Sigma\, r_{ij}^{-n}$ where A is a linear LSE; n is a hardness parameter ascertained by a search over a set of integers that minimises the square of the residuals for the function and r_{ij} is the distance between the i$^{\text{th}}$ and j$^{\text{th}}$ nonbonded hydrogen atoms. The hardness parameter is so named by analogy with the repulsive term in a Lennard-Jones potential where the larger the exponent the steeper (or harder) the potential at close interatomic distances. The 6, 12 form of the Lennard-Jones potential is shown in *eq.* (35).

$$E_{vdW} = 4\varepsilon\left[\left(\frac{R^*}{R}\right)^{12} - \left(\frac{R^*}{R}\right)^{6}\right] \qquad \ldots(35)$$

Where ε is the well depth, R is the distance between atoms and R^* is the distance in terms of R when the potential is zero. The potential in this form has a minimum when $R = 2^{1/6}R^*$.

Both the buffered 14 7 and SSR have been restricted to only calculate H···H nonbonded interactions excluding geminal interactions. The S_{ij} term in *eq.* (34) was then taken to represent SSR to ascertain parameters for STO-3G, 3-21G(*) and 6-31G* with both ΔE_{Tot} and ΔH^{o}_{Tot} in Tables 2.20 and 2.21 respectively, with the results for these parameters in Tables 2.22 and 2.23.

Table 2.20

Parameters based on *eq.* (34) with K^{Etot}, $k^{\cdot\cdot}$ and A substituted for β_1, β_2 and β_3 respectively and *n* is the hardness parameter.

	n	K^{Etot} (kJ mol^{-1})	$K^{\cdot\cdot}$ (kJ mol^{-1})	A (Å$^{1/n}$)	β_o (kJ mol^{-1})	R²	s (kJ mol^{-1})
STO-3G	3	1.019	-105.3	-17.1	-30.9	0.9956	2.6
3-21G(*)	10	1.012	-77.7	-6435.1	-19.5	0.9980	1.7
6-31G*	4	1.007	-145.4	-83.9	-33.8	0.9967	2.2

Table 2.21

Parameters based on *eq.* (34) with K^{Htot}, $k^{\cdot\cdot}$, A and ΔH^{o}_{Tot} substituted for β_1, β_2, β_3 and ΔE_{Tot} respectively and *n* is the hardness parameter.

	n	K^{Htot} (kJ mol^{-1})	$K^{\cdot\cdot}$ (kJ mol^{-1})	A (Å$^{1/n}$)	β_o (kJ mol^{-1})	R²	s (kJ mol^{-1})
STO-3G	4	1.020	-97.4	-52.6	-89.4	0.9961	2.4
3-21G(*)	5	1.013	-73.2	-147.6	-77.5	0.9984	1.5
6-31G*	4	1.008	-131.3	-92.8	-87.5	0.9970	2.1

Table 2.22

Calculations based on *eq.* (34) with K^{Etot}, k¨ and A substituted for β_1, β_2 and β_3 respectively. The a, b and c labels in column headings have meanings: ‘a’ ΔE_{Tot} and $B(\pi)$ were ascertained with the STO-3G basis set; ‘b’ ΔE_{Tot} and $B(\pi)$ were ascertained with the 3-21G(*) basis set; ‘c’ ΔE_{Tot} and $B(\pi)$ were ascertained with the 6-31G* basis set. The $\Delta_f H^o$ exp. is the experimental value.

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\ a}$**	**$\Delta_f H^{o\ b}$**	**$\Delta_f H^{o\ c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
methane	-74.8	-76.2	-74.5	-77.8	1.4	-0.3	3.0
ethane	-83.8	-81.4	-83.0	-81.7	-2.4	-0.8	-2.1
propane	-104.7	-104.4	-105.3	-103.9	-0.3	0.6	-0.8
butane	-125.7	-124.2	-125.8	-124.3	-1.5	0.1	-1.4
2-methylpropane	-134.2	-136.8	-136.2	-136.7	2.6	2.0	2.5
pentane	-146.9	-148.2	-147.7	-147.6	1.3	0.8	0.7
2-methylbutane	-153.6	-154.4	-153.9	-153.7	0.8	0.3	0.1
2,2-dimethylpropane	-168.0	-170.4	-169.7	-170.1	2.4	1.7	2.1
hexane	-166.9	-166.3	-167.0	-165.6	-0.6	0.1	-1.3
2-methylpentane	-174.6	-176.1	-175.4	-175.8	1.5	0.8	1.2
3-methylpentane	-171.9	-172.9	-172.2	-171.4	1.0	0.3	-0.5
2,2-dimethylbutane	-185.9	-184.6	-184.3	-184.4	-1.3	-1.6	-1.5
2,3-dimethylbutane	-178.1	-177.0	-175.9	-175.7	-1.1	-2.2	-2.4
heptane	-187.6	-189.0	-188.6	-188.3	1.4	1.0	0.7
2-methylhexane	-194.5	-194.0	-194.5	-194.0	-0.5	0.0	-0.5
3-methylhexane	-191.3	-192.3	-191.7	-192.8	1.0	0.4	1.5
3-ethylpentane	-189.5	-185.1	-187.0	-186.0	-4.4	-2.5	-3.5
2,2-dimethylpentane	-205.7	-205.5	-205.4	-205.9	-0.2	-0.3	0.2
2,3-dimethylpentane	-198.7	-194.5	-194.5	-194.8	-4.2	-4.2	-3.9
2,4-dimethylpentane	-201.6	-202.8	-202.8	-204.5	1.2	1.2	2.9
3,3-dimethylpentane	-201.0	-197.5	-198.9	-198.9	-3.5	-2.1	-2.1
2,2,3-trimethylbutane	-204.4	-202.5	-202.7	-201.1	-1.9	-1.7	-3.3
octane	-208.5	-208.8	-208.5	-206.7	0.3	0.0	-1.8
2-methylheptane	-215.3	-217.4	-216.3	-216.4	2.1	1.0	1.1

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\,a}$**	**$\Delta_f H^{o\,b}$**	**$\Delta_f H^{o\,c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
3-methylheptane	-212.5	-209.9	-210.6	-211.2	-2.6	-1.9	-1.3
4-methylheptane	-211.9	-214.3	-212.7	-214.5	2.4	0.8	2.6
3-ethylhexane	-210.7	-207.9	-208.6	-210.1	-2.8	-2.1	-0.6
2,2-dimethylhexane	-224.5	-223.9	-224.9	-225.0	-0.6	0.4	0.5
2,3-dimethylhexane	-213.8	-217.1	-216.0	-217.0	3.3	2.2	3.2
2,4-dimethylhexane	-219.2	-219.3	-219.9	-220.8	0.1	0.7	1.6
2,5-dimethylhexane	-222.5	-221.5	-221.9	-222.7	-1.0	-0.6	0.2
3,3-dimethylhexane	-219.9	-220.7	-220.8	-221.8	0.8	0.9	1.9
3,4-dimethylhexane	-212.8	-218.7	-215.4	-215.5	5.9	2.6	2.7
3-ethyl-2-methylpentane	-211.0	-215.9	-214.2	-213.0	4.9	3.2	2.0
3-ethyl-3-methylpentane	-214.8	-210.1	-212.9	-212.4	-4.7	-1.9	-2.4
2,2,3-trimethylpentane	-219.9	-225.0	-223.3	-223.0	5.1	3.4	3.1
2,2,4-trimethylpentane	-223.9	-221.7	-220.7	-223.1	-2.2	-3.2	-0.8
2,3,3-trimethylpentane	-216.2	-216.8	-217.9	-217.6	0.6	1.7	1.4
2,3,4-trimethylpentane	-217.2	-218.5	-218.2	-218.7	1.3	1.0	1.5
2,2,3,3-tetramethylbutane	-226.0	-220.5	-223.9	-219.7	-5.5	-2.1	-6.3
Standard deviation					2.628	1.735	2.233

Table 2.23

Calculations based on *eq.* (34) with K^{Htot}, $k^{\cdot\cdot}$, A and ΔE_{Tot} substituted for β_1, β_2, β_3 and ΔH^o_{Tot} respectively. The a, b and c labels in column headings have meanings: ‘a’ ΔH_{Tot} and B(π) were ascertained with the STO-3G basis set; ‘b’ ΔH_{Tot} and B(π) were ascertained with the 3-21G(*) basis set; ‘c’ ΔH_{Tot} and B(π) were ascertained with the 6-31G* basis set. The $\Delta_f H^o$ exp. is the experimental value.

All values in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**$\Delta_f H^{o\ a}$**	**$\Delta_f H^{o\ b}$**	**$\Delta_f H^{o\ c}$**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**
methane	-74.8	-77.7	-75.5	-78.0	2.9	0.7	3.2
ethane	-83.8	-82.0	-83.1	-81.9	-1.8	-0.7	-1.9
propane	-104.7	-104.3	-105.1	-103.9	-0.4	0.4	-0.8
butane	-125.7	-123.9	-125.5	-124.3	-1.8	-0.2	-1.4
2-methylpropane	-134.2	-136.3	-136.0	-136.4	2.1	1.8	2.2
pentane	-146.9	-147.7	-147.5	-147.4	0.8	0.6	0.5
2-methylbutane	-153.6	-153.9	-153.8	-153.6	0.3	0.2	0.0
2,2-dimethylpropane	-168.0	-169.9	-169.4	-169.6	1.9	1.4	1.6
hexane	-166.9	-166.0	-166.8	-165.7	-0.9	-0.1	-1.2
2-methylpentane	-174.6	-175.6	-175.2	-175.6	1.0	0.6	1.0
3-methylpentane	-171.9	-172.1	-171.7	-171.3	0.2	-0.2	-0.6
2,2-dimethylbutane	-185.9	-184.3	-184.1	-184.0	-1.6	-1.8	-1.9
2,3-dimethylbutane	-178.1	-176.6	-175.9	-175.7	-1.5	-2.2	-2.4
heptane	-187.6	-188.6	-188.3	-188.3	1.0	0.7	0.7
2-methylhexane	-194.5	-193.8	-194.5	-194.2	-0.7	0.0	-0.3
3-methylhexane	-191.3	-192.7	-192.0	-192.5	1.4	0.7	1.2
3-ethylpentane	-189.5	-185.5	-187.3	-186.1	-4.0	-2.2	-3.4
2,2-dimethylpentane	-205.7	-205.4	-205.3	-205.6	-0.3	-0.4	-0.1
2,3-dimethylpentane	-198.7	-194.0	-194.3	-194.7	-4.7	-4.4	-4.0
2,4-dimethylpentane	-201.6	-202.9	-203.1	-204.3	1.3	1.5	2.7
3,3-dimethylpentane	-201.0	-198.3	-199.6	-198.8	-2.7	-1.4	-2.2
2,2,3-trimethylbutane	-204.4	-201.4	-202.1	-203.5	-3.0	-2.3	-0.9
octane	-208.5	-208.4	-208.2	-207.0	-0.1	-0.3	-1.5
2-methylheptane	-215.3	-217.1	-216.2	-216.5	1.8	0.9	1.2

All values in kJ mol^{-1}	$\Delta_f H^o$ exp.	$\Delta_f H^{o\,a}$	$\Delta_f H^{o\,b}$	$\Delta_f H^{o\,c}$	$\Delta\Delta_f H^o$ a	$\Delta\Delta_f H^o$ b	$\Delta\Delta_f H^o$ c
3-methylheptane	-212.5	-210.7	-211.2	-211.3	-1.8	-1.3	-1.2
4-methylheptane	-211.9	-215.0	-213.2	-214.3	3.1	1.3	2.4
3-ethylhexane	-210.7	-209.0	-209.2	-210.0	-1.7	-1.5	-0.7
2,2-dimethylhexane	-224.5	-224.1	-225.1	-225.0	-0.4	0.6	0.5
2,3-dimethylhexane	-213.8	-216.7	-215.6	-217.0	2.9	1.8	3.2
2,4-dimethylhexane	-219.2	-219.7	-220.1	-220.9	0.5	0.9	1.7
2,5-dimethylhexane	-222.5	-221.9	-222.4	-223.0	-0.6	-0.1	0.5
3,3-dimethylhexane	-219.9	-222.0	-221.7	-221.7	2.1	1.8	1.8
3,4-dimethylhexane	-212.8	-218.2	-214.8	-215.2	5.4	2.0	2.4
3-ethyl-2-methylpentane	-211.0	-215.2	-213.3	-212.9	4.2	2.3	1.9
3-ethyl-3-methylpentane	-214.8	-211.3	-213.8	-212.6	-3.5	-1.0	-2.2
2,2,3-trimethylpentane	-219.9	-223.9	-222.2	-222.5	4.0	2.3	2.6
2,2,4-trimethylpentane	-223.9	-223.0	-221.6	-223.1	-0.9	-2.3	-0.8
2,3,3-trimethylpentane	-216.2	-217.0	-218.2	-217.6	0.8	2.0	1.4
2,3,4-trimethylpentane	-217.2	-217.5	-217.4	-218.5	0.3	0.2	1.3
2,2,3,3-tetramethylbutane	-226.0	-220.6	-224.0	-219.5	-5.4	-2.0	-6.5
Standard deviation					2.421	1.549	2.120

The *eqs.* (32), (33) and (34) have two, three and four LSEs respectively and in this order represent a progression of increased parameterisation. There is also a progression of increased computer cost where the basis set was increased from STO-3G to 3-21G(*) to 6-31G*. Added to this progression the explicit calculation of vibrational energies furthers computational cost. Here in *eqs.* (32), (33) and (34) ΔE_{Tot} is substituted with $\Delta H^o{}_{Tot}$, where $\Delta H^o{}_{Tot}$ includes the vibrational energies and PV term as shown in *eq.* (28). A progression in parameterisation with constant basis set can be seen in Tables 2.11, 2.12, 2.13, 2.17, 2.18 and 2.19, while the progression for basis set size with constant degree of parameterisation is shown in Tables 2.8, 2.9, 2.10, 2.14, 2.15, 2.16, 2.20, 2.21, 2.22 and 2.23.

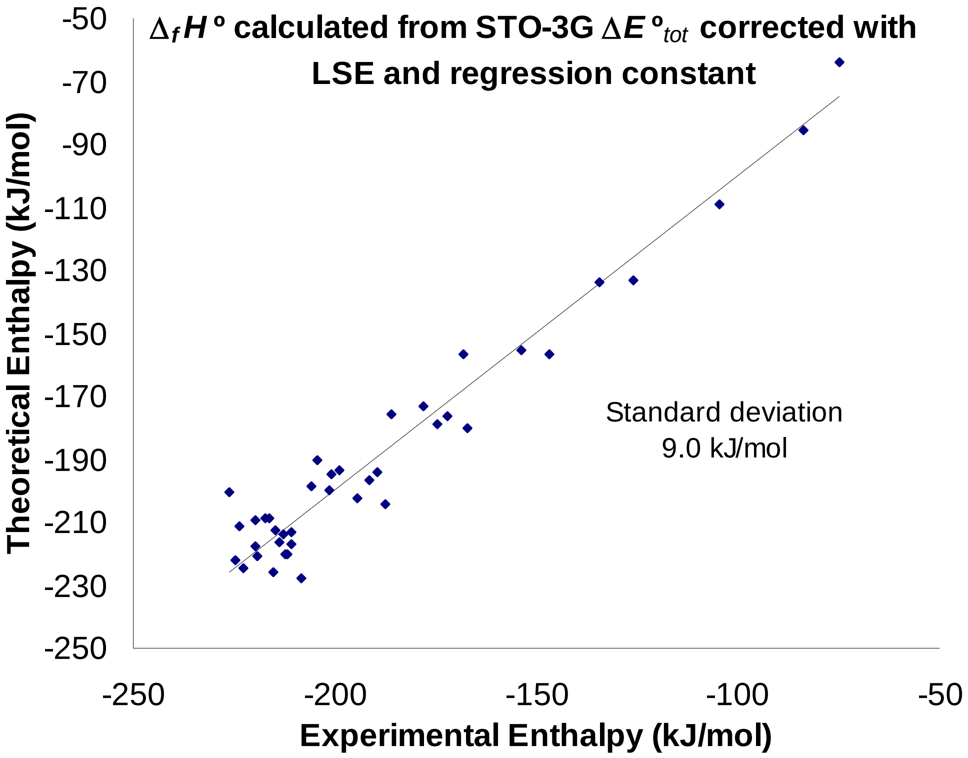
In Table 2.24 the progression with parameterisation is from left to right within the columns associated with ΔE_{Tot} and ΔH^{o}_{Tot}. Movement down the Table represents use of progressively larger basis sets. Column d represents increased parameterisation because there is an extra nonlinear hardness parameter for the S_{ij} term. It was unexpected that subsequent to parameterisation energies calculated with the 6-31G* basis set would have standard deviations roughly equivalent with the STO-3G basis set. The best results were obtained with a parameterisation of 3-21G(*) energies.

Table 2.24

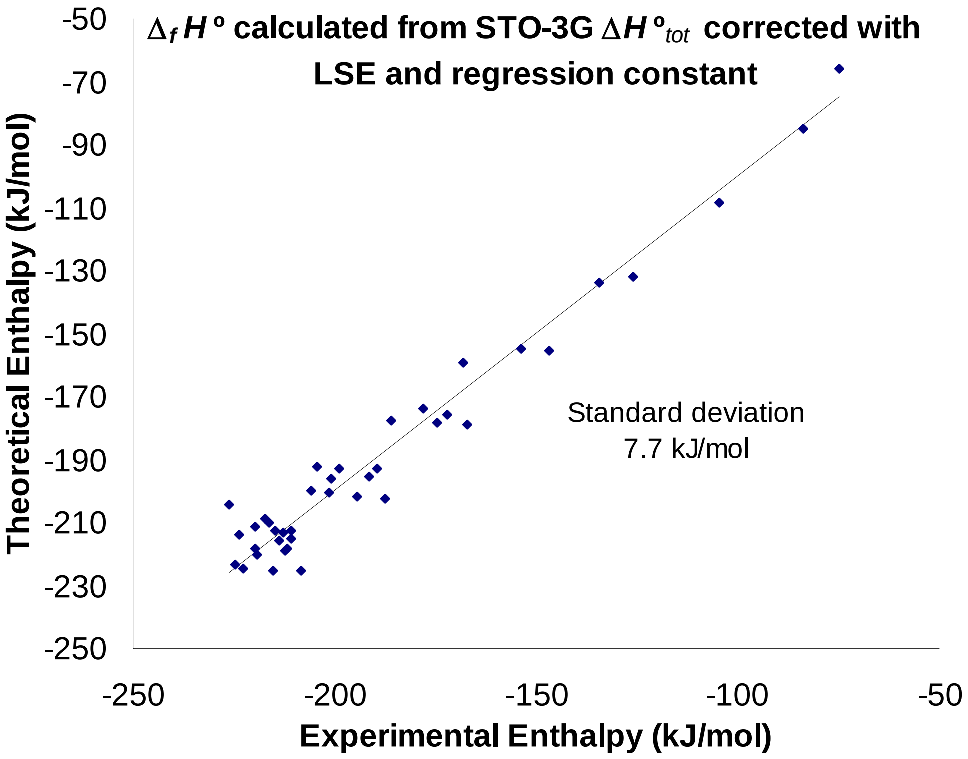
Standard deviation (s) for progression in parameterisation and basis set. Column labels a and b are the standard deviations of results ascertained by parameterisations using *eqs.* (32) and (33) respectively. Labels c and d are the standard deviations of results using *eq.* (34) for the buffered 14 7 and SSR potentials respectively.

All s have units in	**Calculated with ΔE_{Tot}**				**Calculated with ΔH^{o}_{Tot}**			
kJ mol^{-1}	a	b	c	d	a	b	c	d
STO-3G	9.0	2.1	2.6	2.6	7.7	2.0	2.5	2.4
3-21G(*)	3.7	1.8	1.7	1.7	2.6	1.8	1.7	1.5
6-31G*	8.9	2.2	2.3	2.2	7.8	2.4	2.3	2.1

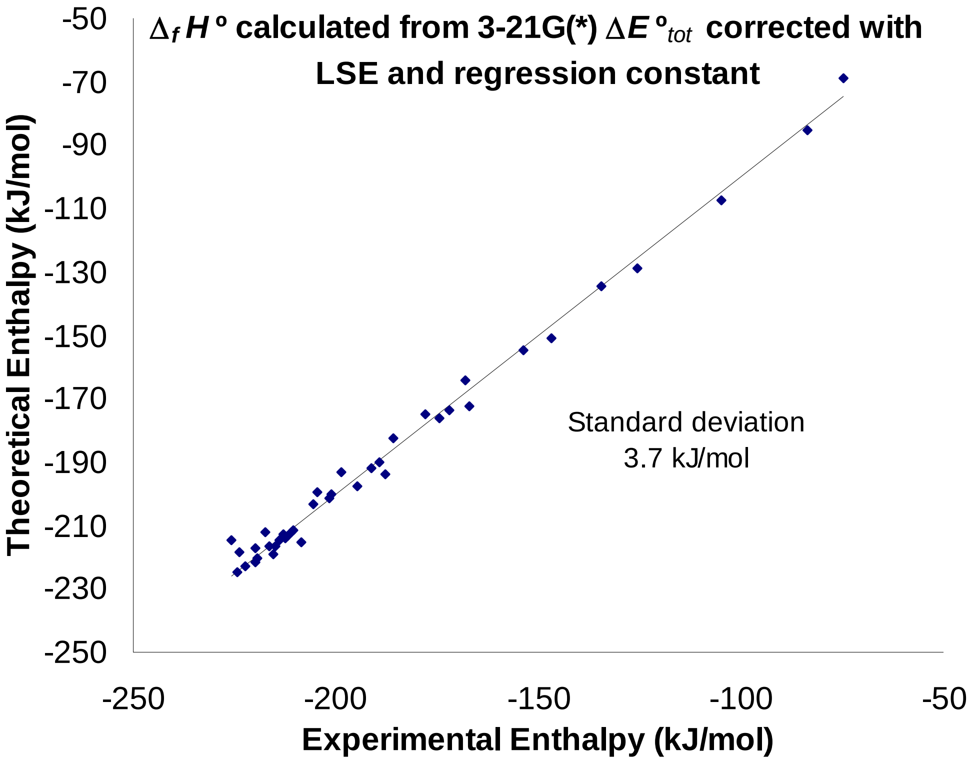
Graph 2.8



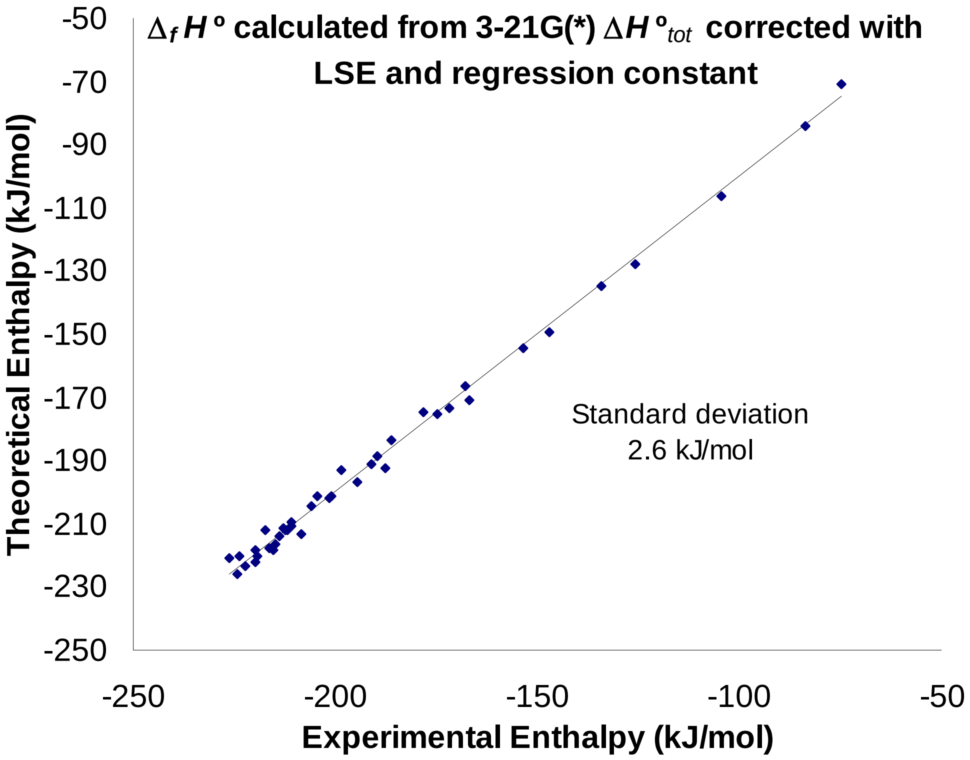
Graph 2.9



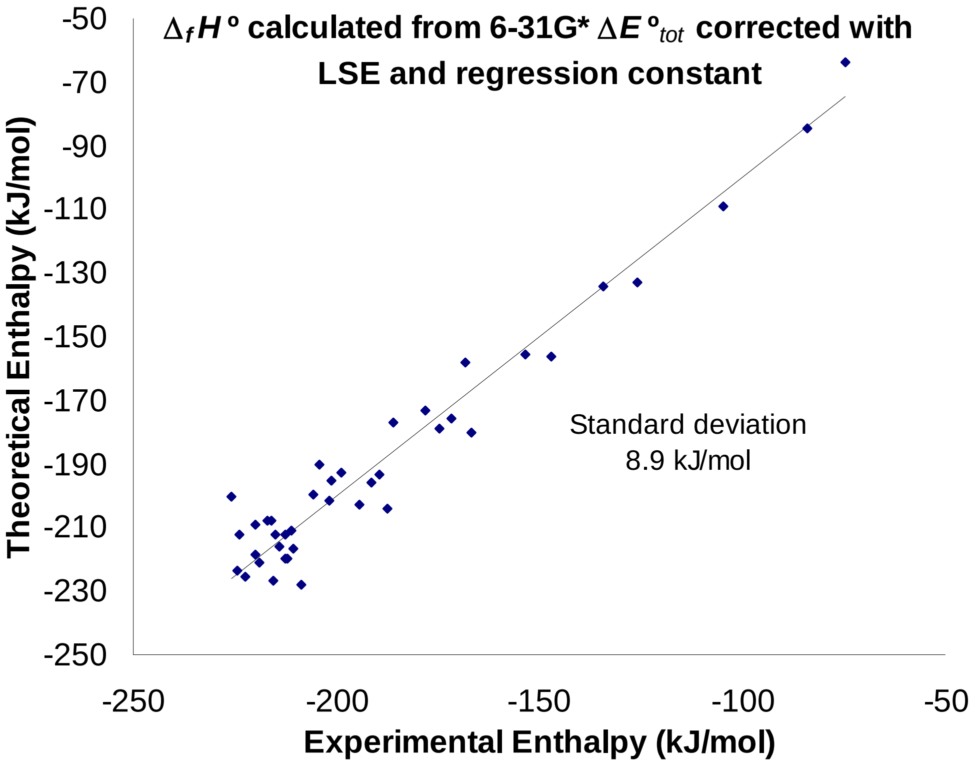
Graph 2.10



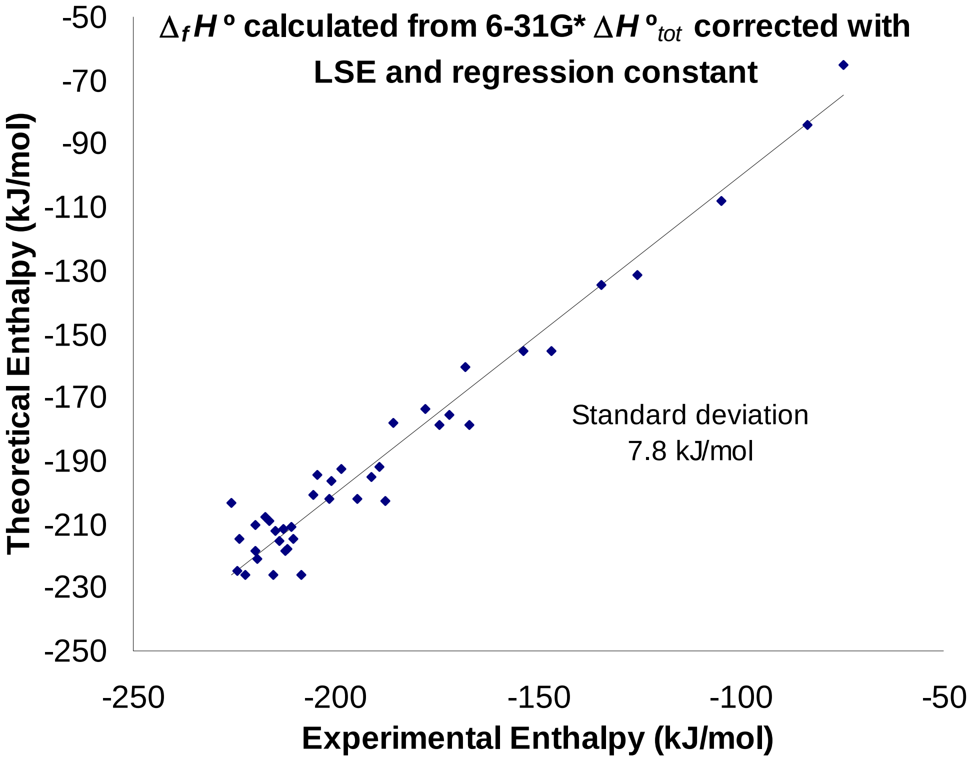
Graph 2.11



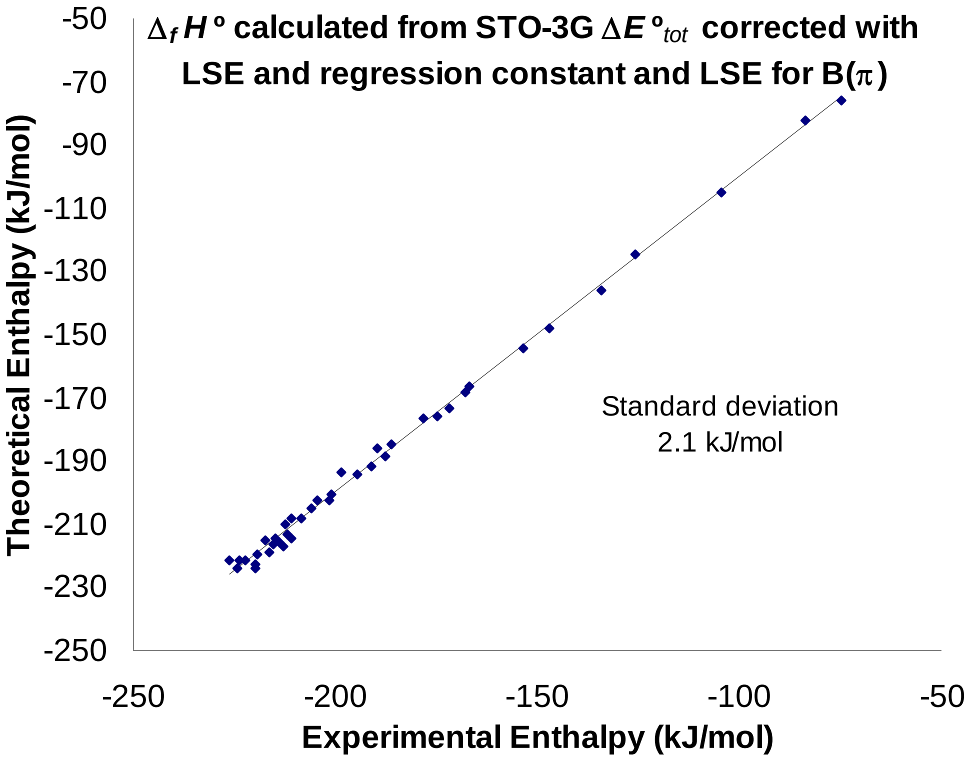
Graph 2.12



Graph 2.13



Graph 2.14



It is expected that branched structural isomers have lower energies than their straight chain or lesser branched counterparts and a stabilising factor based on branching can be included in a model calculation. In Graphs 2.8, 2.9, 2.12 and 2.13 the structural isomers are easily seen to be grouped and in particular the octane group of isomers (bottom left) has the trend where isomers that should have lower energies have higher energies (highly branched isomers) and isomers with higher energies have lower energies (less branched or straight chain isomers). HF STO-3G and 6-31G*, ΔE_{Tot} and ΔH^o_{tot} values vary incorrectly with the degree of branching within the structural isomers as seen in Graphs 2.8, 2.9, 2.12 and 2.13, whereas the energies from 3-21G(*) calculations of Graphs 2.10 and 2.11 correctly vary within each group of structural isomers. The 3-21G(*) basis set is better when parameterised than both STO-3G and 6-31G*. Including an adjusted B(π) value gives a significant improvement in the alkane homologous series. Parameterising B(π) for low level *ab initio* calculations seems to be where it has its greatest effect, particularly when what seems to be wanting is a better description of branching or molecular structure in a model, as evidenced by the much improved standard deviations of the ‘b’ columns over the ‘a’ columns of Table 2.24 and may also be observed by comparing Graph 2.8 with Graph 2.14.

The coefficient for the energies ΔE_{Tot} and ΔH^{o}_{Tot}, K^{Etot} and K^{Htot} respectively vary from 1.007 to 1.020. By these models it is estimated that there is a 0.7 - 2.0 % systematic error in HF total energies, but some of this will be due to the neglect of electron correlation. The function for correlation with basis set truncation error is not separated from a function of any other systematic errors.

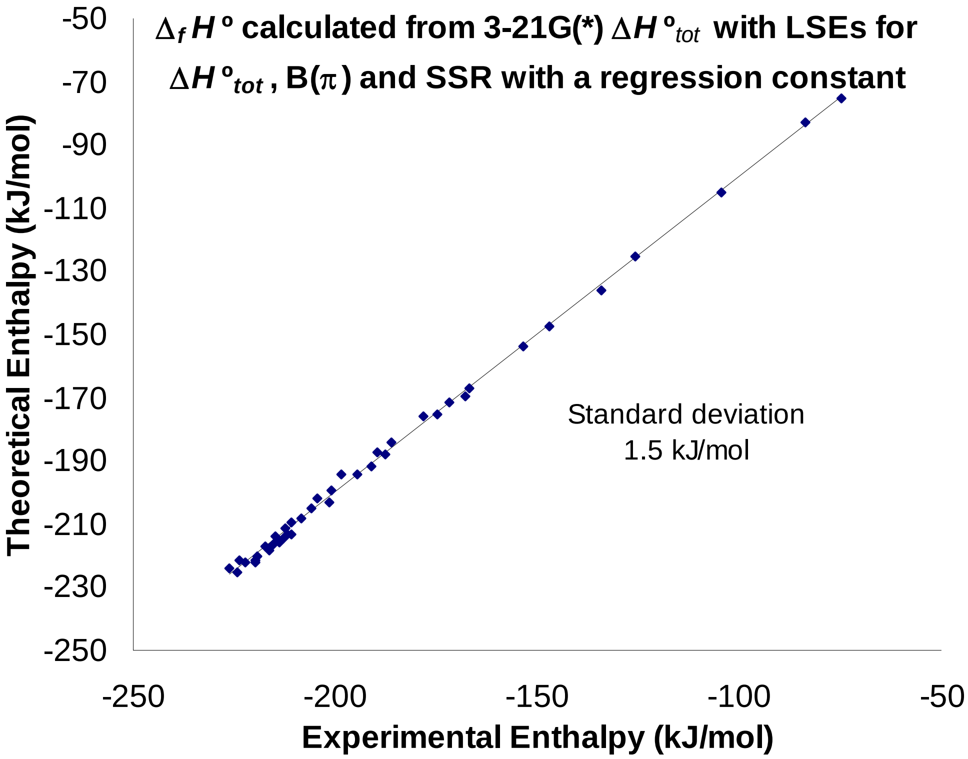
The results are somewhat improved over other methods that use bondingness. However other methods have a regression constant small enough to be eliminated (*ca.* 1-5 kJ mol^{-1} cf. 20 – 90 kJ mol^{-1} for the above method). If vibrational energies calculated with HF 3-21G(*) are scaled by 0.9207, the scaling factor for HF 3-21G[113], prior to corrections for translational and rotational energy and the *PV* term, the regression constant is reduced by almost 10 kJ mol^{-1} and the standard deviation is improved by 0.02 kJ mol^{-1}. A vibrational scaling factor does not account for the regression constant or significantly improve the standard deviation.

The best standard deviation (1.5 kJ mol^{-1}) was ascertained with LSEs for ΔH^{o}_{Tot}, B(π) and SSR with the regression constant. The results are given in the 'b' columns of Table 2.23 and the calculated $\Delta_f H^{o}$ are compared with the experimental $\Delta_f H^{o}$ in Graph 2.15. The t-statistics are evaluated for this model and are shown in Table 2.25. An observed t-value is given by the LSE divided by the standard error in the LSE (negative signs are ignored). A t-value below the t-critical value means the term associated with the LSE is ineffective in the model or is uncorrelated. The two tail t-critical value with α 0.05 and 36 degrees of freedom is 2.03 and is also shown in Table 2.25. All t-observed values are well above 2.03 and are therefore useful in estimating $\Delta_f H^{o}$.

Table 2.25

t-critical value is 2.03	**K^{Htot}**	**K¨**	**SSR**	**β_o**
LSE (kJ mol^{-1})	1.013	-73.2	-147.6	-77.5
Standard error (kJ mol^{-1})	9.2×10^{-6}	4.1	13.4	1.4
t-observed value	110485	18.0	11.0	54.5

Graph 2.15



A tractable, but *ad hoc* method was tested to compare HF ΔE_{Tot} with PM3 $\Delta_f H°$, where the total energy, B(π) and δ°(C-C) were parameterised in the same way as was done for PM3 $\Delta_f H°$ values. If test group 2 was used, then replacing $\Delta_f H°_{\text{of model}}$ with ΔE_{Tot} (STO-3G) in *eq.* (22), the following results were obtained. Using LSEs: δ°(C-C) = 20.315, δ'(C-C) = -81.34. Statistics are compared with equivalent methods using PM3 $\Delta_f H°$ values in Table 2.27.

2.4 Statistics

This section compares additivity methods with systematic corrections to quantum chemical energies.

2.4.1 Alkanes

Parameters ascertained from test group 2 yield the following statistics for the forty alkanes from C1 to C8 and all structural isomers, where k is ascertained from geometry optimised molecules at the respective levels. The statistics in columns with the label corr. of Table 2.26 are the statistics subsequent to a correction with CSE.

Table 2.26

	k (kJ mol^{-1})	R²	s (kJ mol^{-1})	R² corr.	s corr. (kJ mol^{-1})
PM3	60.019	0.9698	13.56	0.9838	7.1
STO-3G	84.154	0.9840	9.10	0.9942	3.1

2.4.2 Systematic correction methods

The methods in the following sections adjust a heat of formation or total energy as given in the Spartan© output file. This can be seen in Table 2.27. The STO-3G basis set did however provide better B(π) values for correcting PM3 $\Delta_f H^o$ values.

2.4.2.1 Variation in Quantum Chemical PM3 Energy with B(π)

This section tabulates the results of section 2.1.

Table 2.27

	d°(C-H) (kJ mol^{-1})	δ'(C-C) (kJ mol^{-1})	δ°(C-C) (kJ mol^{-1})	s	Corrected with CSE
PM3	-5.1*	-53.3	-1.6	4.6	Yes
STO-3G	-5.1*	-68.1	0.0	3.4	No
STO-3G	-4.8	-62.5	0*	3.0	No
STO-3G†	-8.8*	-81.3	20.3	2.9	Yes

* Not a least squares estimate

† Uses HF/STO-3G energy and STO-3G B(π).

2.4.2.2 Variation in Quantum Chemical PM3 Energies with Molecular Mechanics Steric Energies and Bondingness

This section tabulates results of section 2.2.

Table 2.28

LSEs	β_0 (kJ mol^{-1})	β_1 (kJ mol^{-1})	δ(C-C) (kJ mol^{-1})	δ(C-H) (kJ mol^{-1})	R²	s (kJ mol^{-1})
***eq.* (24)**	-11.5	-1.1			0.9955	9.3

LSEs	**β_0 (kJ mol^{-1})**	**β_1 (kJ mol^{-1})**	**δ(C-C) (kJ mol^{-1})**	**δ(C-H) (kJ mol^{-1})**	**R^2**	**s (kJ mol^{-1})**
***eq.* (25)**	-18.3	-1.3	2.5		0.9913	4.2
***eq.* (26)**		-1.3	11.6	-4.6	0.9913	4.2
***eq.* (26)**		-2.2	11.1	-5.1*	0.9641	21.2
***eq.* (26)†**		-1.2	12.9	-5.1*	0.9923	3.8

* calculated as discussed in section 2.2.

† parameterised on all 40 alkanes.

2.4.2.3 Quantum Chemical HF Total Energies

Table 2.29 is a copy of Table 2.24

Table 2.29

Copy of Table 2.24. Standard deviation (s) for progression in parameterisation and basis set. Column labels a and b are the standard deviations of results ascertained by parameterisations using *eqs.* (32) and (33) respectively. Labels c and d are the standard deviations of results using *eq.* (34) for the buffered 14 7 and SSR potentials respectively.

All s have units in	**Calculated with ΔE_{Tot}**				**Calculated with ΔH^{o}_{Tot}**			
kJ mol^{-1}	a	b	c	d	a	b	c	d
STO-3G	9.0	2.1	2.6	2.6	7.7	2.0	2.5	2.4
3-21G(*)	3.7	1.8	1.7	1.7	2.6	1.8	1.7	1.5
6-31G*	8.9	2.2	2.3	2.2	7.8	2.4	2.3	2.1

2.5 Section Summary

For alkanes a bond additivity scheme to estimate $\Delta_f H^o$ could be achieved with a standard deviation (s) of around 3 kJ mol^{-1} with one parameter: STO-3G adjusted B(π). Systematic corrections to quantum chemical energies required at least two parameters to achieve equivalent standard deviations in estimated $\Delta_f H^o$.

In section 2.2 some results were biased by using a training set that was a subset of the test set to ascertain parameters for the test set. This approach required that the test set be well represented by the training set. As it was sometimes uncertain as to whether this was the case it subsequently became more preferable to use the test set as the training set.

The outstanding feature of section 2.3 was the greater ability of the 3-21G(*) basis set to account for energy variation between isomers for alkanes, outperforming even 6-31G*.

3 Alkenes

The problem of how to quantify π antibonding effects in alkenes was not so obvious as it was for alkanes, as strong π bonding over the double bond conflicts with the π antibonding in the single bonds. It seemed reasonable to have separate parameters for the different kinds of bonds. There are single bonds, double bonds and single bonds adjacent to double bonds (sp^2sp^3 bonds). Again it was not clear how bond B(π) should be calculated as distinguishing bonds presented many more ways to sum the bond B(π) values. It was for the alkenes that the nomenclature was introduced in section 1.7. There were two test groups. Test group one included steric strained molecules in the group C_2 to C_5 and all the possible geometric isomers. Test group two excluded molecules with steric strain *viz. cis*-but-2-ene, *cis*-pent-2-ene, 2-methyl-but-1-ene and 2-methyl-but-2-ene.

Table 3.1

The molecules in test groups 1 and 2.

Test Group 1	Test Group 2
ethene	ethene
propene	propene
but-1-ene	but-1-ene
cis-but-2-ene	*trans*-but-2-ene
trans-but-2-ene	2-methylpropene
2-methylpropene	pent-1-ene
pent-1-ene	*trans*-pent-2-ene
cis-pent-2-ene	3-methylbut-1-ene

Test Group 1	**Test Group 2**
trans-pent-2-ene	hex-1-ene
2-methylbut-1-ene	*trans*-hex-2-ene
3-methylbut-1-ene	*trans*-hex-3-ene
2-methylbut-2-ene	3-methylpent-1-ene

Pedley's $\Delta_f H^o$ values are ascertained from enthalpies of vaporisation and combustion and an enthalpy of reaction for 1-hexene from Wiberg and Wasserman[114]. A subsequent paper by Wiberg *et al.*[115] not used by Pedley suggests improved accuracies in $\Delta_f H^o$ values particularly for *cis*-pent-2-ene and *cis*-hex-3-ene. We show the experimental error from Pedley's work which is larger for the aforementioned molecules. Wiberg writes about these methods in a chapter on experimental Thermochemistry in the book series Molecular Structure and Energetics[116]. Enthalpies of reaction ascertained by Rogers *et al.*[117-121] are also not used by Pedley.

3.1 STO-3G

The following results tested the adequacy of the B(π) of a molecule to describe variations in alkene bond energies without the assistance of steric functions and other phenomena developed later in this work. These results should serve to ascertain what algorithm best assimilates AO information to sum B(π).

In all the tables $\Delta_f H^o$ exp. is the experimental value and ± is the error in $\Delta_f H^o$ exp.

3.1.1 k

Parameters for test group one calculating molecular B(π) without distinguishing the kind of bond.

For test group 1:

Table 3.2

The column labels a, b, c and d have the following meanings: ‘a’ Values of $\Delta_f H^o$ are estimated with the single parameter k; ‘b’ Values of $\Delta_f H^o$ are estimated with the parameters k and number of sp^2sp^3 C-C bonds; ‘c’ Values of $\Delta_f H^o$ are estimated with the parameters k and number of sp^2sp^3 C-C bonds with CSE subtracted from the result; ‘d’ Values of $\Delta_f H^o$ are estimated with the parameters k and number of sp^2sp^3 C-C bonds with CSE included in the regression analysis.

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	±	**a**	**b**	**c**	**d**
ethene	52.5	0.3	4.89	2.26	2.26	1.33
propene	20.0	0.7	-2.22	-3.06	-3.06	-2.50
but-1-ene	0.1	0.9	4.14	2.83	2.83	3.02
***cis*-but-2-ene**	-7.1	1.0	1.82	3.04	-1.15	1.12
***trans*-but-2-ene**	-11.4	1.0	-3.77	-2.61	-2.61	-0.40
2-methylpropene	-16.9	0.9	-7.24	-5.99	-5.99	-3.69
pent-1-ene	-21.1	0.9	2.21	0.13	0.13	-0.32
***cis*-pent-2-ene**	-27.6	0.9	5.00	5.64	1.45	3.25
***trans*-pent-2-ene**	-31.9	1.0	-2.00	-1.49	-1.49	0.21
2-methylbut-1-ene	-35.2	0.9	-3.36	-2.76	-3.72	-1.95
3-methylbut-1-ene	-27.5	0.7	0.66	-1.20	-1.20	-1.46
2-methylbut-2-ene	-41.7	1.0	-0.13	3.21	-2.64	1.39
hex-1-ene	-43.5	1.6	-1.71	-4.61	-4.61	-5.71
***cis*-hex-2-ene**	-52.3	1.3	-2.10	-2.31	-6.50	-5.39
***trans*-hex-2-ene**	-53.9	1.5	-1.63	-1.75	-1.75	-0.57
***cis*-hex-3-ene**	-47.6	1.3	0.79	0.49	-3.69	-2.65
***trans*-hex-3-ene**	-54.4	1.3	-2.69	-2.84	-2.84	-1.68
2-methylpent-1-ene	-59.4	1.3	-5.45	-5.50	-6.46	-5.22
3-methylpent-1-ene	-49.5	1.5	2.28	-0.16	-0.16	-0.89
4-methylpent-1-ene	-51.3	1.8	0.79	-1.63	-1.63	-2.36
2-methylpent-2-ene	-66.9	1.4	-0.98	1.81	-4.04	-0.46
***cis*-3-methylpent-2-ene**	-62.3	1.4	0.07	2.70	-5.58	-2.13
***trans*-3-methylpent-2-ene**	-63.1	1.3	2.56	5.34	-2.94	0.64
***cis*-4-methylpent-2-ene**	-57.5	1.1	-2.49	-2.49	-6.67	-5.39

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	a	b	c	d
***trans*-4-methylpent-2-ene**	-61.5	1.4	-6.13	-6.11	-6.11	-4.81
2-ethylbut-1-ene	-56.0	1.4	-7.66	-7.96	-9.89	-8.85
2,3-dimethylbut-1-ene	-62.4	1.3	-5.57	-5.49	-10.09	-8.74
3,3-dimethylbut-1-ene	-60.3	1.2	5.39	3.59	-0.34	-0.56
2,3-dimethylbut-2-ene	-68.1	1.1	5.58	11.02	-7.39	-1.61
hept-1-ene	-62.3	0.9	2.78	-0.71	-0.71	-2.30
5-methylhex-1-ene	-65.7	1.0	5.72	2.52	1.13	-0.22
***cis*-3-methylhex-3-ene**	-79.4	1.1	2.37	4.23	-4.05	-1.22
***trans*-3-methylhex-3-ene**	-76.8	1.1	5.21	7.08	-1.20	1.64
2,4-dimethylpent-1-ene	-83.8	1.3	-4.06	-4.59	-5.55	-4.70
4,4-dimethylpent-1-ene	-81.6	1.9	5.21	2.71	1.58	0.80
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	19.98	20.04	-5.07	-3.74
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	0.82	0.74	-6.12	-4.90
2,4-dimethylpent-2-ene	-88.7	1.1	-0.77	1.37	-4.48	-1.43
2-ethyl-3-methylbut-1-ene	-79.5	1.4	2.86	2.45	-8.26	-7.31
2,3,3-trimethylbut-1-ene	-85.5	1.3	4.16	4.09	-11.39	-10.17
Standard deviation			4.96	5.16	4.94	3.92

If we then parameterise with test group 2 using k we get the following results:

For test group 2:

Table 3.3

The column labels e, f and g have the following meanings: 'e' Values of $\Delta_f H^o$ are estimated with the single parameter k; 'f' Values of $\Delta_f H^o$ are estimated with the parameters k and number of sp^2sp^3 C-C bonds; 'g' Values of $\Delta_f H^o$ are estimated with the parameters k and number of sp^2sp^3 C-C bonds with CSEs subtracted from the result.

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	e	f	g
ethene	52.5	0.3	5.04	0.51	0.51
propene	20.0	0.7	-1.91	-2.47	-2.47
but-1-ene	0.1	0.9	4.59	3.54	3.54
***trans*-but-2-ene**	-11.4	1	-3.37	0.28	0.28

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**e**	**f**	**g**
2-methylpropene	-16.9	0.9	-6.86	-3.12	-3.12
pent-1-ene	-21.1	0.9	2.91	1.01	1.01
***trans*-pent-2-ene**	-31.9	1	-1.39	1.56	1.56
3-methylbut-1-ene	-27.5	0.7	1.30	-0.37	-0.37
hex-1-ene	-43.5	1.6	-0.74	-3.54	-3.54
***trans*-hex-2-ene**	-53.9	1.5	-0.81	1.44	1.44
***trans*-hex-3-ene**	-54.4	1.3	-1.87	0.35	0.35
3-methylpent-1-ene	-49.5	1.5	3.10	0.80	0.80
4-methylpent-1-ene	-51.3	1.8	1.61	-0.67	-0.67
***trans*-4-methylpent-2-ene**	-61.5	1.4	-5.35	-2.95	-2.95
hept-1-ene	-62.3	0.9	3.94	0.49	0.49
***cis*-but-2-ene**	-7.1	1	2.21	5.92	1.74
***cis*-pent-2-ene**	-27.6	0.9	5.57	8.65	4.47
2-methylbut-1-ene	-35.2	0.9	-2.78	0.27	-0.70
2-methylbut-2-ene	-41.7	1	0.31	8.37	2.51
***cis*-hex-2-ene**	-52.3	1.3	-1.25	0.90	-3.29
***cis*-hex-3-ene**	-47.6	1.3	1.67	3.72	-0.46
2-methylpent-1-ene	-59.4	1.3	-4.66	-2.33	-3.29
2-methylpent-2-ene	-66.9	1.4	-0.36	7.09	1.23
***cis*-3-methylpent-2-ene**	-62.3	1.4	0.74	8.02	-0.27
***trans*-3-methylpent-2-ene**	-63.1	1.3	3.18	10.62	2.34
***cis*-4-methylpent-2-ene**	-57.5	1.1	-1.71	0.67	-3.51
2-ethylbut-1-ene	-56.0	1.4	-6.78	-4.73	-6.66
2,3-dimethylbut-1-ene	-62.4	1.3	-4.82	-2.35	-6.95
3,3-dimethylbut-1-ene	-60.3	1.2	6.01	4.41	0.48
2,3-dimethylbut-2-ene	-68.1	1.1	6.09	18.45	0.04
5-methylhex-1-ene	-65.7	1	6.79	3.65	2.27
***cis*-3-methylhex-3-ene**	-79.4	1.1	3.29	9.72	1.44
***trans*-3-methylhex-3-ene**	-76.8	1.1	6.13	12.57	4.29
2,4-dimethylpent-1-ene	-83.8	1.3	-3.11	-1.31	-2.27
4,4-dimethylpent-1-ene	-81.6	1.9	6.05	3.69	2.56
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	20.74	23.18	-1.92
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	1.62	3.92	-2.94

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	e	f	g
2,4-dimethylpent-2-ene	-88.7	1.1	0.06	6.80	0.94
2-ethyl-3-methylbut-1-ene	-79.5	1.4	3.77	5.71	-5.00
2,3,3-trimethylbut-1-ene	-85.5	1.3	4.96	7.26	-8.22
Standard deviation			5.11	6.87	3.05

3.1.2 $k^{\text{C-C}}$ and $k^{\text{C-C¨C=C}}$

If we then distinguish B(π) for double and single bonds and select MOs contiguously negative in B(π) summed over single bonds and try the same things we get the following results.

For test group 1:

Table 3.4

The column labels a, b, c and d have the following meanings: ‘a’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{C-C}}$ and $k^{\text{C-C¨C=C}}$; ‘b’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{C-C}}$, $k^{\text{C-C¨C=C}}$ and number of sp^2sp^3 C-C bonds; ‘c’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{C-C}}$, $k^{\text{C-C¨C=C}}$ and number of sp^2sp^3 C-C bonds with CSEs subtracted from the result; ‘d’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{C-C}}$, $k^{\text{C-C¨C=C}}$ and number of sp^2sp^3 C-C bonds with CSEs included in the regression analysis.

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	a	b	c	d
ethene	52.5	0.3	-1.14	-0.18	-0.18	-0.14
propene	20.0	0.7	3.67	-0.12	-0.12	-0.54
but-1-ene	0.1	0.9	9.01	3.73	3.73	3.19
***cis*-but-2-ene**	-7.1	1.0	0.28	3.90	-0.28	2.06
***trans*-but-2-ene**	-11.4	1.0	-5.57	-4.05	-4.05	-2.17
2-methylpropene	-16.9	0.9	-6.16	-4.01	-4.01	-2.09
pent-1-ene	-21.1	0.9	5.81	-0.03	-0.03	-0.31
***cis*-pent-2-ene**	-27.6	0.9	2.70	3.42	-0.76	1.16
***trans*-pent-2-ene**	-31.9	1.0	-4.34	-2.60	-2.60	-0.36
2-methylbut-1-ene	-35.2	0.9	-3.06	-1.45	-2.41	-0.31
3-methylbut-1-ene	-27.5	0.7	4.82	-0.90	-0.90	-1.27

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**a**	**b**	**c**	**d**
2-methylbut-2-ene	-41.7	1.0	-6.02	2.30	-3.56	0.79
hex-1-ene	-43.5	1.6	1.47	-5.34	-5.34	-5.47
***cis*-hex-2-ene**	-52.3	1.3	-5.30	-4.45	-8.63	-6.25
***trans*-hex-2-ene**	-53.9	1.5	-4.99	-5.07	-5.07	-2.96
***cis*-hex-3-ene**	-47.6	1.3	-2.00	-1.31	-5.49	-3.12
***trans*-hex-3-ene**	-54.4	1.3	-5.60	-4.72	-4.72	-2.37
2-methylpent-1-ene	-59.4	1.3	-5.97	-5.61	-6.58	-4.46
3-methylpent-1-ene	-49.5	1.5	5.20	-2.19	-2.19	-2.66
4-methylpent-1-ene	-51.3	1.8	3.40	-3.02	-3.02	-3.25
2-methylpent-2-ene	-66.9	1.4	-8.55	-1.32	-7.18	-2.80
***cis*-3-methylpent-2-ene**	-62.3	1.4	-6.39	2.04	-6.24	-1.53
***trans*-3-methylpent-2-ene**	-63.1	1.3	-3.90	5.79	-2.50	2.45
***cis*-4-methylpent-2-ene**	-57.5	1.1	-5.22	-5.38	-9.56	-7.54
***trans*-4-methylpent-2-ene**	-61.5	1.4	-9.00	-9.35	-9.35	-7.37
2-ethylbut-1-ene	-56.0	1.4	-8.23	-7.49	-9.42	-7.09
2,3-dimethylbut-1-ene	-62.4	1.3	-5.72	-4.79	-9.39	-7.20
3,3-dimethylbut-1-ene	-60.3	1.2	4.16	-2.02	-5.95	-6.34
2,3-dimethylbut-2-ene	-68.1	1.1	-5.02	9.20	-9.21	-2.56
hept-1-ene	-62.3	0.9	5.52	-2.22	-2.22	-2.28
5-methylhex-1-ene	-65.7	1.0	8.09	1.49	0.11	0.19
***cis*-3-methylhex-3-ene**	-79.4	1.1	-4.61	2.82	-5.46	-0.62
***trans*-3-methylhex-3-ene**	-76.8	1.1	-1.72	5.17	-3.11	1.60
2,4-dimethylpent-1-ene	-83.8	1.3	-5.44	-5.52	-6.48	-4.22
4,4-dimethylpent-1-ene	-81.6	1.9	6.27	-0.55	-1.68	-1.94
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	17.19	18.39	-6.71	-4.39
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	-0.59	-1.44	-8.30	-6.44
2,4-dimethylpent-2-ene	-88.7	1.1	-7.40	-0.96	-6.82	-2.35
2-ethyl-3-methylbut-1-ene	-79.5	1.4	1.99	2.25	-8.46	-6.19
2,3,3-trimethylbut-1-ene	-85.5	1.3	4.05	4.66	-10.82	-8.63
Standard deviation			5.99	5.05	5.82	4.06

For test group 2:

Table 3.5

The column labels e, f and g have the following meanings: ‘e’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{C\text{-}C}$ and $k^{C\text{-}C\cdot\cdot C=C}$; ‘f’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{C\text{-}C}$, $k^{C\text{-}C\cdot\cdot C=C}$ and number of sp^2sp^3 C-C bonds; ‘g’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{C\text{-}C}$, $k^{C\text{-}C\cdot\cdot C=C}$ and number of sp^2sp^3 C-C bonds with CSEs subtracted from the result.

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**e**	**f**	**g**
ethene	52.5	0.3	-0.64	-0.64	-0.05
propene	20.0	0.7	2.10	2.10	-1.27
but-1-ene	0.1	0.9	7.91	7.91	3.25
***trans*-but-2-ene**	-11.4	1	-6.60	-6.60	-1.34
2-methylpropene	-16.9	0.9	-7.41	-7.41	-1.62
pent-1-ene	-21.1	0.9	5.80	5.80	1.25
***trans*-pent-2-ene**	-31.9	1	-4.40	-4.40	1.74
3-methylbut-1-ene	-27.5	0.7	4.48	4.48	-0.15
hex-1-ene	-43.5	1.6	2.48	2.48	-2.44
***trans*-hex-2-ene**	-53.9	1.5	-4.38	-4.38	0.23
***trans*-hex-3-ene**	-54.4	1.3	-4.82	-4.82	0.93
3-methylpent-1-ene	-49.5	1.5	5.49	5.49	-0.53
4-methylpent-1-ene	-51.3	1.8	3.84	3.84	-1.02
***trans*-4-methylpent-2-ene**	-61.5	1.4	-8.66	-8.66	-4.52
hept-1-ene	-62.3	0.9	7.25	7.25	1.80
***cis*-but-2-ene**	-7.1	1	-0.51	-4.69	2.98
***cis*-pent-2-ene**	-27.6	0.9	2.31	-1.88	2.92
2-methylbut-1-ene	-35.2	0.9	-3.43	-4.39	1.39
2-methylbut-2-ene	-41.7	1	-6.76	-12.62	2.68
***cis*-hex-2-ene**	-52.3	1.3	-4.41	-8.60	-2.79
***cis*-hex-3-ene**	-47.6	1.3	-1.05	-5.23	0.44
2-methylpent-1-ene	-59.4	1.3	-5.57	-6.54	-1.61
2-methylpent-2-ene	-66.9	1.4	-8.59	-14.45	0.15
***cis*-3-methylpent-2-ene**	-62.3	1.4	-6.08	-14.36	1.76
***trans*-3-methylpent-2-ene**	-63.1	1.3	-3.60	-11.88	5.59

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	e	f	g
***cis*-4-methylpent-2-ene**	-57.5	1.1	-4.84	-9.02	-4.64
2-ethylbut-1-ene	-56.0	1.4	-7.41	-9.33	-3.71
2,3-dimethylbut-1-ene	-62.4	1.3	-5.44	-10.04	-4.56
3,3-dimethylbut-1-ene	-60.3	1.2	4.02	0.08	-4.91
2,3-dimethylbut-2-ene	-68.1	1.1	-5.46	-23.87	0.52
5-methylhex-1-ene	-65.7	1	9.62	8.24	3.89
***cis*-3-methylhex-3-ene**	-79.4	1.1	-3.34	-11.62	4.06
***trans*-3-methylhex-3-ene**	-76.8	1.1	-0.55	-8.84	6.19
2,4-dimethylpent-1-ene	-83.8	1.3	-4.38	-5.34	-0.44
4,4-dimethylpent-1-ene	-81.6	1.9	6.85	5.72	0.52
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	17.72	-7.38	-1.44
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	-0.29	-7.15	-3.58
2,4-dimethylpent-2-ene	-88.7	1.1	-6.70	-12.56	1.66
2-ethyl-3-methylbut-1-ene	-79.5	1.4	2.90	-7.81	-2.64
2,3,3-trimethylbut-1-ene	-85.5	1.3	4.50	-10.98	-5.73
Standard deviation			6.01	8.76	2.92

3.1.3 k^{C-C} and $k^{C=C}$

Only the single bonded carbons contribute an antibonding effect. Therefore summing B(π) over C-C bonds and C=C bonds ascertaining the MOs independently of each other gives the following.

For test group 1:

Table 3.6

The column labels a, b, c and d have the following meanings: ‘a’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{C\text{-}C}$ and $k^{C=C}$; ‘b’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{C\text{-}C}$, $k^{C=C}$ and number of sp^2sp^3 C-C bonds; ‘c’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{C\text{-}C}$, $k^{C=C}$ and number of sp^2sp^3 C-C bonds with CSEs subtracted from the result; ‘d’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{C\text{-}C}$, $k^{C=C}$ and number of sp^2sp^3 C-C bonds with CSEs included in the regression analysis.

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**a**	**b**	**c**	**d**
ethene	52.5	0.3	5.85	2.43	2.43	1.65
propene	20.0	0.7	-2.11	-3.18	-3.18	-2.54
but-1-ene	0.1	0.9	2.25	2.11	2.11	2.03
***cis*-but-2-ene**	-7.1	1.0	1.50	3.99	-0.20	2.12
***trans*-but-2-ene**	-11.4	1.0	-3.57	-5.66	-5.66	-3.09
2-methylpropene	-16.9	0.9	-12.07	-4.16	-4.16	-2.40
Pent-1-ene	-21.1	0.9	4.16	0.69	0.69	0.11
***cis*-pent-2-ene**	-27.6	0.9	7.72	3.00	-1.19	1.07
***trans*-pent-2-ene**	-31.9	1.0	1.83	-1.78	-1.78	0.32
2-methylbut-1-ene	-35.2	0.9	-4.57	0.17	-0.79	0.60
3-methylbut-1-ene	-27.5	0.7	2.09	-1.02	-1.02	-1.39
2-methylbut-2-ene	-41.7	1.0	-3.08	3.44	-2.42	1.51
hex-1-ene	-43.5	1.6	4.15	-2.45	-2.45	-3.59
***cis*-hex-2-ene**	-52.3	1.3	9.39	-2.77	-6.96	-4.76
***trans*-hex-2-ene**	-53.9	1.5	4.72	-3.40	-3.40	-1.63
***cis*-hex-3-ene**	-47.6	1.3	16.74	-0.49	-4.68	-2.01
***trans*-hex-3-ene**	-54.4	1.3	10.62	-3.82	-3.82	-1.30
2-methylpent-1-ene	-59.4	1.3	-2.06	-3.10	-4.06	-2.82
3-methylpent-1-ene	-49.5	1.5	6.75	-1.72	-1.72	-2.27
4-methylpent-1-ene	-51.3	1.8	0.08	-0.45	-0.45	-1.78
2-methylpent-2-ene	-66.9	1.4	1.49	0.13	-5.72	-1.65
***cis*-3-methylpent-2-ene**	-62.3	1.4	3.49	5.15	-3.13	0.61
***trans*-3-methylpent-2-ene**	-63.1	1.3	4.82	9.56	1.27	4.93

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**a**	**b**	**c**	**d**
***cis*-4-methylpent-2-ene**	-57.5	1.1	5.95	-4.90	-9.08	-6.87
***trans*-4-methylpent-2-ene**	-61.5	1.4	-1.08	-8.26	-8.26	-6.45
2-ethylbut-1-ene	-56.0	1.4	1.72	-4.90	-6.82	-5.20
2,3-dimethylbut-1-ene	-62.4	1.3	-2.90	-2.21	-6.82	-5.57
3,3-dimethylbut-1-ene	-60.3	1.2	1.97	-1.73	-5.66	-6.19
2,3-dimethylbut-2-ene	-68.1	1.1	2.30	11.97	-6.44	-0.75
Hept-1-ene	-62.3	0.9	9.15	2.64	2.64	0.81
5-methylhex-1-ene	-65.7	1.0	22.94	3.16	1.78	1.74
***cis*-3-methylhex-3-ene**	-79.4	1.1	17.65	5.84	-2.45	1.88
***trans*-3-methylhex-3-ene**	-76.8	1.1	19.12	8.00	-0.28	3.95
2,4-dimethylpent-1-ene	-83.8	1.3	1.82	-1.66	-2.62	-1.66
4,4-dimethylpent-1-ene	-81.6	1.9	0.74	2.94	1.81	0.01
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	26.44	20.30	-4.80	-2.94
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	6.03	-0.33	-7.19	-5.54
2,4-dimethylpent-2-ene	-88.7	1.1	9.02	1.18	-4.67	-0.56
2-ethyl-3-methylbut-1-ene	-79.5	1.4	10.62	5.35	-5.36	-4.04
2,3,3-trimethylbut-1-ene	-85.5	1.3	7.11	7.68	-7.80	-6.72
Standard deviation			9.14	5.49	4.47	3.36

For test group 2:

Table 3.7

The column labels e, f and g have the following meanings: ‘e’ Values of $\Delta_f H^o$ are estimated with the parameters k^{C-C} and $k^{C=C}$; ‘f’ Values of $\Delta_f H^o$ are estimated with the parameters k^{C-C}, $k^{C=C}$ and number of sp^2sp^3 C-C bonds; ‘g’ Values of $\Delta_f H^o$ are estimated with the parameters k^{C-C}, $k^{C=C}$ and number of sp^2sp^3 C-C bonds with CSEs subtracted from the result.

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**e**	**f**	**g**
ethene	52.5	0.3	8.72	8.72	0.46
propene	20.0	0.7	-1.57	-1.57	-2.29
but-1-ene	0.1	0.9	1.72	1.72	2.87
***trans*-but-2-ene**	-11.4	1	-5.33	-5.33	-1.77
2-methylpropene	-16.9	0.9	-12.61	-12.61	-0.55

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**e**	**f**	**g**
Pent-1-ene	-21.1	0.9	1.79	1.79	1.42
***trans*-pent-2-ene**	-31.9	1	-1.21	-1.21	2.06
3-methylbut-1-ene	-27.5	0.7	0.22	0.22	-0.26
hex-1-ene	-43.5	1.6	-0.10	-0.10	-1.76
***trans*-hex-2-ene**	-53.9	1.5	-0.15	-0.15	0.46
***trans*-hex-3-ene**	-54.4	1.3	5.42	5.42	0.26
3-methylpent-1-ene	-49.5	1.5	3.09	3.09	-0.90
4-methylpent-1-ene	-51.3	1.8	-2.86	-2.86	0.12
***trans*-4-methylpent-2-ene**	-61.5	1.4	-5.63	-5.63	-4.41
Hept-1-ene	-62.3	0.9	3.66	3.66	3.22
***cis*-but-2-ene**	-7.1	1	0.51	-3.67	3.59
***cis*-pent-2-ene**	-27.6	0.9	4.67	0.49	2.70
2-methylbut-1-ene	-35.2	0.9	-6.65	-7.61	2.81
2-methylbut-2-ene	-41.7	1	-5.59	-11.44	4.23
***cis*-hex-2-ene**	-52.3	1.3	4.20	0.01	-2.97
***cis*-hex-3-ene**	-47.6	1.3	11.05	6.86	-0.53
2-methylpent-1-ene	-59.4	1.3	-5.98	-6.95	-0.39
2-methylpent-2-ene	-66.9	1.4	-2.87	-8.73	1.08
***cis*-3-methylpent-2-ene**	-62.3	1.4	-0.68	-8.96	3.56
***trans*-3-methylpent-2-ene**	-63.1	1.3	1.34	-6.95	7.91
***cis*-4-methylpent-2-ene**	-57.5	1.1	1.14	-3.04	-5.11
2-ethylbut-1-ene	-56.0	1.4	-3.02	-4.95	-3.01
2,3-dimethylbut-1-ene	-62.4	1.3	-6.34	-10.95	-3.18
3,3-dimethylbut-1-ene	-60.3	1.2	-0.36	-4.30	-4.91
2,3-dimethylbut-2-ene	-68.1	1.1	-1.68	-20.09	3.12
5-methylhex-1-ene	-65.7	1	17.15	15.77	2.86
***cis*-3-methylhex-3-ene**	-79.4	1.1	10.91	2.63	4.56
***trans*-3-methylhex-3-ene**	-76.8	1.1	12.38	4.10	6.69
2,4-dimethylpent-1-ene	-83.8	1.3	-3.29	-4.25	1.04
4,4-dimethylpent-1-ene	-81.6	1.9	-2.33	-3.46	2.26
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	22.27	-2.83	-0.96
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	1.41	-5.45	-3.38
2,4-dimethylpent-2-ene	-88.7	1.1	2.96	-2.90	2.23

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	e	f	g
2-ethyl-3-methylbut-1-ene	-79.5	1.4	5.69	-5.02	-1.62
2,3,3-trimethylbut-1-ene	-85.5	1.3	3.33	-12.16	-4.19
Standard deviation			6.86	7.10	3.14

3.1.4 $k^{\text{¨C-C}}$ and $k^{\text{¨C=C}}$

Selecting MOs that had a total π antibonding effect and then distinguishing the kind of bond gave the following. This became the method of choice for adding extra parameters. This same method was used in section 3.1.5 where an extra distinction was made for sp^2sp^3 C-C bonds.

For test group 1:

Table 3.8

The column labels a, b, c and d have the following meanings: 'a' Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$ and $k^{\text{¨C=C}}$; 'b' Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$, $k^{\text{¨C=C}}$ and number of sp^2sp^3 C-C bonds; 'c' Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$, $k^{\text{¨C=C}}$ and number of sp^2sp^3 C-C bonds with CSEs subtracted from the result; 'd' Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$, $k^{\text{¨C=C}}$ and number of sp^2sp^3 C-C bonds with CSE included in the regression analysis.

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	a	b	c	d
ethene	52.5	0.3	3.65	2.21	2.21	1.25
propene	20.0	0.7	-3.05	-2.85	-2.85	-2.22
but-1-ene	0.1	0.9	4.14	2.48	2.48	2.54
***cis*-but-2-ene**	-7.1	1.0	0.83	3.89	-0.29	2.29
***trans*-but-2-ene**	-11.4	1.0	-2.33	-3.07	-3.07	-1.01
2-methylpropene	-16.9	0.9	-6.42	-6.09	-6.09	-3.83
pent-1-ene	-21.1	0.9	0.93	0.24	0.24	-0.16
***cis*-pent-2-ene**	-27.6	0.9	6.14	5.21	1.03	2.67
***trans*-pent-2-ene**	-31.9	1.0	-2.39	-1.14	-1.14	0.68
2-methylbut-1-ene	-35.2	0.9	-3.45	-2.55	-3.51	-1.66

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**a**	**b**	**c**	**d**
3-methylbut-1-ene	-27.5	0.7	-0.09	-1.30	-1.30	-1.61
2-methylbut-2-ene	-41.7	1.0	2.04	2.97	-2.88	1.06
hex-1-ene	-43.5	1.6	-4.00	-4.18	-4.18	-5.13
***cis*-hex-2-ene**	-52.3	1.3	-3.37	-1.70	-5.88	-4.55
***trans*-hex-2-ene**	-53.9	1.5	-1.44	-1.88	-1.88	-0.75
***cis*-hex-3-ene**	-47.6	1.3	-0.45	1.07	-3.12	-1.86
***trans*-hex-3-ene**	-54.4	1.3	-3.67	-2.37	-2.37	-1.03
2-methylpent-1-ene	-59.4	1.3	-6.60	-4.91	-5.87	-4.41
3-methylpent-1-ene	-49.5	1.5	2.17	-0.76	-0.76	-1.71
4-methylpent-1-ene	-51.3	1.8	-0.55	-1.58	-1.58	-2.28
2-methylpent-2-ene	-66.9	1.4	-0.95	2.55	-3.31	0.55
***cis*-3-methylpent-2-ene**	-62.3	1.4	0.26	3.31	-4.97	-1.30
***trans*-3-methylpent-2-ene**	-63.1	1.3	1.68	6.56	-1.72	2.31
***cis*-4-methylpent-2-ene**	-57.5	1.1	-1.90	-2.80	-6.98	-5.81
***trans*-4-methylpent-2-ene**	-61.5	1.4	-5.18	-6.61	-6.61	-5.49
2-ethylbut-1-ene	-56.0	1.4	-9.01	-7.33	-9.25	-7.98
2,3-dimethylbut-1-ene	-62.4	1.3	-6.13	-5.17	-9.77	-8.30
3,3-dimethylbut-1-ene	-60.3	1.2	7.23	2.13	-1.80	-2.56
2,3-dimethylbut-2-ene	-68.1	1.1	9.20	10.58	-7.83	-2.22
hept-1-ene	-62.3	0.9	0.04	-0.21	-0.21	-1.61
5-methylhex-1-ene	-65.7	1.0	2.47	3.37	1.99	0.95
***cis*-3-methylhex-3-ene**	-79.4	1.1	1.80	5.04	-3.25	-0.12
***trans*-3-methylhex-3-ene**	-76.8	1.1	5.32	7.53	-0.75	2.25
2,4-dimethylpent-1-ene	-83.8	1.3	-4.98	-4.24	-5.20	-4.22
4,4-dimethylpent-1-ene	-81.6	1.9	4.47	2.42	1.29	0.41
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	19.07	20.53	-4.57	-3.06
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	0.87	0.70	-6.16	-4.96
2,4-dimethylpent-2-ene	-88.7	1.1	0.65	1.20	-4.66	-1.66
2-ethyl-3-methylbut-1-ene	-79.5	1.4	1.85	2.88	-7.83	-6.73
2,3,3-trimethylbut-1-ene	-85.5	1.3	3.54	4.40	-11.08	-9.75
Standard deviation			5.04	5.22	4.76	3.74

For test group 2:

Table 3.9

The column labels e, f and g have the following meanings: ‘e’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$ and $k^{\text{¨C=C}}$; ‘f’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$, $k^{\text{¨C=C}}$ and number of sp^2sp^3 C-C bonds; ‘g’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$, $k^{\text{¨C=C}}$ and number of sp^2sp^3 C-C bonds with CSEs subtracted from the result.

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**e**	**f**	**g**
ethene	52.5	0.3	2.74	2.74	0.84
propene	20.0	0.7	-3.17	-3.17	-1.68
but-1-ene	0.1	0.9	5.05	5.05	2.86
***trans*-but-2-ene**	-11.4	1	-0.64	-0.64	-0.65
2-methylpropene	-16.9	0.9	-5.23	-5.23	-3.17
pent-1-ene	-21.1	0.9	1.81	1.81	1.25
***trans*-pent-2-ene**	-31.9	1	-1.24	-1.24	2.38
3-methylbut-1-ene	-27.5	0.7	0.91	0.91	-0.60
hex-1-ene	-43.5	1.6	-2.93	-2.93	-2.75
***trans*-hex-2-ene**	-53.9	1.5	0.75	0.75	0.96
***trans*-hex-3-ene**	-54.4	1.3	-2.21	-2.21	1.32
3-methylpent-1-ene	-49.5	1.5	4.18	4.18	-0.75
4-methylpent-1-ene	-51.3	1.8	0.65	0.65	-0.66
***trans*-4-methylpent-2-ene**	-61.5	1.4	-2.63	-2.63	-4.27
hept-1-ene	-62.3	0.9	1.43	1.43	1.32
***cis*-but-2-ene**	-7.1	1	0.85	-3.33	3.95
***cis*-pent-2-ene**	-27.6	0.9	8.16	3.98	3.47
2-methylbut-1-ene	-35.2	0.9	-2.19	-3.15	-0.18
2-methylbut-2-ene	-41.7	1	4.31	-1.55	2.01
***cis*-hex-2-ene**	-52.3	1.3	-2.05	-6.23	-2.00
***cis*-hex-3-ene**	-47.6	1.3	0.99	-3.20	0.71
2-methylpent-1-ene	-59.4	1.3	-5.36	-6.33	-2.01
2-methylpent-2-ene	-66.9	1.4	0.49	-5.36	2.94
***cis*-3-methylpent-2-ene**	-62.3	1.4	1.98	-6.30	1.09
***trans*-3-methylpent-2-ene**	-63.1	1.3	2.52	-5.76	5.21

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	e	f	g
***cis*-4-methylpent-2-ene**	-57.5	1.1	0.43	-3.76	-4.39
2-ethylbut-1-ene	-56.0	1.4	-7.65	-9.57	-5.34
2,3-dimethylbut-1-ene	-62.4	1.3	-4.64	-9.25	-6.30
3,3-dimethylbut-1-ene	-60.3	1.2	9.87	5.93	-2.98
2,3-dimethylbut-2-ene	-68.1	1.1	12.64	-5.77	-1.05
5-methylhex-1-ene	-65.7	1	3.22	1.84	4.00
***cis*-3-methylhex-3-ene**	-79.4	1.1	3.82	-4.46	3.08
***trans*-3-methylhex-3-ene**	-76.8	1.1	7.77	-0.52	5.08
2,4-dimethylpent-1-ene	-83.8	1.3	-3.10	-4.07	-1.70
4,4-dimethylpent-1-ene	-81.6	1.9	6.14	5.01	1.74
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	20.35	-4.76	-0.85
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	2.93	-3.94	-3.21
2,4-dimethylpent-2-ene	-88.7	1.1	3.68	-2.17	0.31
2-ethyl-3-methylbut-1-ene	-79.5	1.4	3.55	-7.16	-4.22
2,3,3-trimethylbut-1-ene	-85.5	1.3	5.14	-10.34	-7.62
Standard deviation			5.54	4.73	3.12

3.1.5 $k^{''C\text{-}C}$, $k^{''sp^2sp^3}$ and $k^{''C=C}$

For test group 1:

Table 3.10

The column labels a, b, c and d have the following meanings: ‘a’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{''C\text{-}C}$, $k^{''sp^2sp^3}$ and $k^{''C=C}$; ‘b’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{''C\text{-}C}$, $k^{''sp^2sp^3}$, $k^{''C=C}$ and number of sp^2sp^3 C-C bonds; ‘c’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{''C\text{-}C}$, $k^{''sp^2sp^3}$, $k^{''C=C}$ and number of sp^2sp^3 C-C bonds with CSEs added to the result; ‘d’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{''C\text{-}C}$, $k^{''sp^2sp^3}$, $k^{''C=C}$ and number of sp^2sp^3 C-C bonds with CSE included in the regression analysis.

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	a	b	c	d
ethene	52.5	0.3	0.95	0.95	0.59	0.47
propene	20.0	0.7	-1.93	-1.93	-1.24	-2.72
but-1-ene	0.1	0.9	2.51	2.51	3.50	3.49

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**a**	**b**	**c**	**d**
***cis*-but-2-ene**	-7.1	1.0	4.24	0.06	0.99	1.85
***trans*-but-2-ene**	-11.4	1.0	-1.66	-1.66	1.05	-0.65
2-methylpropene	-16.9	0.9	-5.73	-5.73	-3.59	-3.68
pent-1-ene	-21.1	0.9	0.90	0.90	0.99	1.85
***cis*-pent-2-ene**	-27.6	0.9	3.56	-0.63	1.25	4.62
***trans*-pent-2-ene**	-31.9	1.0	0.28	0.28	1.65	1.73
2-methylbut-1-ene	-35.2	0.9	-3.13	-4.09	-2.85	-0.29
3-methylbut-1-ene	-27.5	0.7	-2.69	-2.69	-2.56	0.54
2-methylbut-2-ene	-41.7	1.0	2.70	-3.16	0.21	1.75
hex-1-ene	-43.5	1.6	-1.92	-1.92	-2.38	-2.07
***cis*-hex-2-ene**	-52.3	1.3	0.19	-4.00	-3.28	-2.36
***trans*-hex-2-ene**	-53.9	1.5	-1.09	-1.09	0.43	1.99
***cis*-hex-3-ene**	-47.6	1.3	1.53	-2.66	-2.18	0.79
***trans*-hex-3-ene**	-54.4	1.3	-0.52	-0.52	0.41	1.12
2-methylpent-1-ene	-59.4	1.3	-4.98	-5.94	-5.43	-2.15
3-methylpent-1-ene	-49.5	1.5	-0.90	-0.90	-0.25	1.64
4-methylpent-1-ene	-51.3	1.8	-0.40	-0.40	-0.33	0.34
2-methylpent-2-ene	-66.9	1.4	1.19	-4.67	-3.00	1.84
***cis*-3-methylpent-2-ene**	-62.3	1.4	2.68	-5.60	-3.73	0.25
***trans*-3-methylpent-2-ene**	-63.1	1.3	5.11	-3.18	-2.13	3.31
***cis*-4-methylpent-2-ene**	-57.5	1.1	-2.95	-7.13	-5.49	-3.01
***trans*-4-methylpent-2-ene**	-61.5	1.4	-6.59	-6.59	-4.67	-2.62
2-ethylbut-1-ene	-56.0	1.4	-6.44	-8.36	-7.89	-5.45
2,3-dimethylbut-1-ene	-62.4	1.3	-5.34	-9.94	-9.04	-6.09
3,3-dimethylbut-1-ene	-60.3	1.2	-0.27	-4.20	-2.52	0.61
2,3-dimethylbut-2-ene	-68.1	1.1	9.60	-8.81	-4.14	-0.95
hept-1-ene	-62.3	0.9	3.50	3.50	2.87	2.26
5-methylhex-1-ene	-65.7	1.0	7.15	5.77	4.88	4.00
***cis*-3-methylhex-3-ene**	-79.4	1.1	5.42	-2.87	-1.49	2.58
***trans*-3-methylhex-3-ene**	-76.8	1.1	7.97	-0.32	1.52	5.17
2,4-dimethylpent-1-ene	-83.8	1.3	-1.82	-2.79	-1.81	-1.36
4,4-dimethylpent-1-ene	-81.6	1.9	3.19	2.06	2.45	3.48
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	18.06	-7.04	-6.77	-0.46

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**a**	**b**	**c**	**d**
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	0.25	-6.61	-5.39	-2.13
2,4-dimethylpent-2-ene	-88.7	1.1	0.57	-5.29	-2.71	1.37
2-ethyl-3-methylbut-1-ene	-79.5	1.4	4.20	-6.51	-5.76	-3.93
2,3,3-trimethylbut-1-ene	-85.5	1.3	3.88	-11.60	-10.84	-7.15
Standard deviation			4.84	4.89	4.01	2.92

For test group 2:

Table 3.11

The column labels e, f and g have the following meanings: ‘e’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$, $k^{\text{¨ }sp^2sp^3}$ and $k^{\text{¨C=C}}$; ‘f’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$, $k^{\text{¨ }sp^2sp^3}$ and $k^{\text{¨C=C}}$ with CSEs added to the estimate; ‘g’ Values of $\Delta_f H^o$ are estimated with the parameters $k^{\text{¨C-C}}$, $k^{\text{¨ }sp^2sp^3}$, $k^{\text{¨C=C}}$ and number of sp^2sp^3 C-C bonds with CSEs subtracted from the result.

All units in kJ mol^{-1}	**$\Delta_f H^o$ exp.**	**±**	**e**	**f**	**g**
ethene	52.5	0.3	0.71	0.71	0.75
propene	20.0	0.7	-1.11	-1.11	-1.62
but-1-ene	0.1	0.9	3.42	3.42	2.74
***trans*-but-2-ene**	-11.4	1	0.60	0.60	-0.67
2-methylpropene	-16.9	0.9	-3.89	-3.89	-3.06
pent-1-ene	-21.1	0.9	1.37	1.37	1.22
***trans*-pent-2-ene**	-31.9	1	1.80	1.80	2.56
3-methylbut-1-ene	-27.5	0.7	-2.39	-2.39	-0.63
hex-1-ene	-43.5	1.6	-1.59	-1.59	-2.77
***trans*-hex-2-ene**	-53.9	1.5	0.46	0.46	0.98
***trans*-hex-3-ene**	-54.4	1.3	0.81	0.81	1.50
3-methylpent-1-ene	-49.5	1.5	-0.19	-0.19	-1.00
4-methylpent-1-ene	-51.3	1.8	0.11	0.11	-0.74
***trans*-4-methylpent-2-ene**	-61.5	1.4	-4.89	-4.89	-4.34
hept-1-ene	-62.3	0.9	3.87	3.87	1.26
***cis*-but-2-ene**	-7.1	1	5.42	1.24	4.36
***cis*-pent-2-ene**	-27.6	0.9	5.07	0.88	3.50
2-methylbut-1-ene	-35.2	0.9	-1.86	-2.83	0.02

All units in kJ mol^{-1}	$\Delta_f H^o$ exp.	±	e	f	g
2-methylbut-2-ene	-41.7	1	5.43	-0.42	2.23
***cis*-hex-2-ene**	-52.3	1.3	1.40	-2.79	-1.78
***cis*-hex-3-ene**	-47.6	1.3	2.47	-1.71	0.95
2-methylpent-1-ene	-59.4	1.3	-4.07	-5.03	-1.75
2-methylpent-2-ene	-66.9	1.4	2.89	-2.97	3.46
***cis*-3-methylpent-2-ene**	-62.3	1.4	4.56	-3.72	1.54
***trans*-3-methylpent-2-ene**	-63.1	1.3	6.46	-1.83	5.87
***cis*-4-methylpent-2-ene**	-57.5	1.1	-1.42	-5.61	-4.40
2-ethylbut-1-ene	-56.0	1.4	-5.45	-7.38	-5.10
2,3-dimethylbut-1-ene	-62.4	1.3	-4.22	-8.82	-6.11
3,3-dimethylbut-1-ene	-60.3	1.2	0.79	-3.14	-3.40
2,3-dimethylbut-2-ene	-68.1	1.1	13.26	-5.15	-0.74
5-methylhex-1-ene	-65.7	1	7.38	6.00	4.07
***cis*-3-methylhex-3-ene**	-79.4	1.1	7.13	-1.16	3.52
***trans*-3-methylhex-3-ene**	-76.8	1.1	9.93	1.65	5.41
2,4-dimethylpent-1-ene	-83.8	1.3	-0.42	-1.38	-1.60
4,4-dimethylpent-1-ene	-81.6	1.9	3.84	2.71	1.56
***cis*-4,4-dimethylpent-2-ene**	-72.6	1.4	18.61	-6.49	-0.55
***trans*-4,4-dimethylpent-2-ene**	-88.8	1.1	1.52	-5.35	-3.13
2,4-dimethylpent-2-ene	-88.7	1.1	2.84	-3.02	0.50
2-ethyl-3-methylbut-1-ene	-79.5	1.4	5.38	-5.33	-4.06
2,3,3-trimethylbut-1-ene	-85.5	1.3	4.89	-10.59	-7.43
Standard deviation			5.36	4.00	3.18

3.1.6 Statistics

There are two standard deviations given in the following tables, s(40) and s(12). These refer to the standard deviations ascertained over the entire group of forty molecules and the twelve training group molecules respectively.

Table 3.12

Parameters for results in Tables 3.2 and 3.3, where labels a, b, c, d, e, f and g have the same meanings.

		k	**sp^{2}sp^3**	**E^o(C=C)**	**R^2**	**s(40)**	**s(12)**
Test group 1	a	88.842		608.280	0.9757	5.0	3.8
	b	92.897	2.293	606.273	0.9748	5.2	3.4
	c	92.897	2.293	606.273	0.9890	4.9	2.9
	d	96.188	4.192	605.849	0.9900	3.9	2.2
Test group 2	e	87.535		608.226	0.9760	5.1	3.5
	f	91.969	4.523	604.382	0.9714	6.9	2.1
	g	91.969	4.523	604.382	0.9904	3.1	2.1

Table 3.13

Parameters for results in Tables 3.4 and 3.5, where labels a, b, c, d, e, f and g have the same meanings.

		$K^{C\text{-}C}$	**$K^{C\text{-}C¨C=C}$**	**sp^{2}sp^3**	**E^o(C=C)**	**R^2**	**s(40)**	**s(12)**
Test group 1	a	88.838	85.000		588.505	0.9626	6.0	5.2
	b	92.846	38.635	7.594	589.456	0.9742	5.1	2.8
	c	92.846	38.635	7.594	589.456	0.9910	5.8	2.6
	d	91.931	25.718	10.123	589.496	0.9914	4.1	1.6
Test group 2	e	83.762	71.979		589.000	0.9626	6.0	5.4
	f	83.762	71.979		589.000	0.9440	8.8	5.4
	g	84.745	13.393	10.720	589.586	0.9908	2.9	1.6

Table 3.14

Parameters for results in Tables 3.6 and 3.7, where labels a, b, c, d, e, f and g have the same meanings.

		$K^{C\text{-}C}$	**$K^{C=C}$**	**sp^{2}sp^3**	**E^o(C=C)**	**R^2**	**s(40)**	**s(12)**
Test group 1	a	31.835	-273.013		553.247	0.9560	9.1	5.4
	b	90.857	69.507	9.833	602.821	0.9737	5.5	3.2
	c	90.857	69.507	9.833	602.821	0.9897	4.5	2.7
	d	88.795	32.584	11.412	596.333	0.9906	3.4	1.9
Test group 2	e	44.055	-248.076		559.975	0.9574	6.9	5.3

		$K^{C\text{-}C}$	$K^{C=C}$	sp^2sp^3	$E^o(C=C)$	R^2	s(40)	s(12)
	f	44.055	-248.076		559.975	0.9538	7.1	5.3
	g	89.605	58.238	12.607	599.108	0.9896	3.1	1.6

Table 3.15

Parameters for results in Tables 3.8 and 3.9, where labels a, b, c, d, e, f and g have the same meanings.

		$K^{\text{¨C-C}}$	$K^{\text{¨C=C}}$	sp^2sp^3	$E^o(C=C)$	R^2	s(40)	s(12)
Test group 1	a	95.531	156.499		617.504	0.9747	5.0	3.7
	b	90.453	58.195	2.912	600.850	0.9748	5.2	3.4
	c	90.453	58.195	2.912	600.850	0.9894	4.8	2.9
	d	92.845	48.717	5.039	598.431	0.9902	3.7	2.1
Test group 2	e	95.708	196.809		622.829	0.9746	5.5	3.1
	f	95.708	196.809		622.829	0.9824	4.7	3.1
	g	87.122	9.742	5.960	591.985	0.9896	3.1	1.9

Table 3.16

Parameters for results in Tables 3.10 and 3.11, where labels a, b, c, d, e, f and g have the same meanings.

		$K^{\text{¨C-C}}$	$K^{\text{¨}sp^2sp^3}$	$K^{\text{¨C=C}}$	sp^2sp^3	$E^o(C=C)$	R^2	s(40)	s(12)
Test group 1	a	79.813	51.037	30.347		595.284	0.9792	4.8	3.1
	b	79.813	51.037	30.347		595.284	0.9858	4.9	2.7
	c	83.224	49.527	71.087		601.228	0.9879	4.0	2.1
	d	88.805	94.543	70.296	5.123	600.991	0.9908	2.9	2.5
Test group 2	e	80.669	46.466	50.124		598.106	0.9788	5.4	2.0
	f	80.669	46.466	50.124		598.106	0.9872	4.0	2.0
	g	86.844	87.540	0.001	6.365	590.390	0.9892	3.2	1.9

Over the 12 molecules of a test group, parameterising only B(π), the standard deviation [s(12)], was best at 2.0 kJ mol^{-1}, when B(π) was ascertained with the ¨bond algorithm. However the bond¨bond (bond) and C-C¨bond algorithms had the best standard deviations, 1.6 kJ mol^{-1}, when the number of sp^2sp^3 bonds were also parameterised. Regardless of the algorithm chosen to sum B(π), if we parameterised using test group 2 and the number of sp^2sp^3 bonds over the range of alkenes the standard deviations [s(40)] were within in the range 3.0 ± 0.2 kJ mol^{-1}. It would appear that even though different specific B(π) were achieved by the different algorithms in a least squares regression analysis the parameters adjusted to accommodate most variations so that there was no particular algorithm vastly superior to any other. The bond¨bond and C-C¨bond algorithms might have a slight edge over other algorithms where the bond¨bond algorithm was the more versatile of the two as it did not require the presence of a C-C bond in a molecule.

3.2 STO-3G and 3-21G(*) comparison

Using the same algorithms and parameters of section 3.1 the 3-21G(*) basis set (Table 3.18) performed worse than the STO-3G basis set (Table 3.17). However when fitting a modern steric function of nonbonded atomic distance like the buffered 14 7 or similar function such as the algorithms of Section 4, the larger basis set gave improved results. It was this combination of extending the basis set and using a steric function of section 4 that one had to use to improve the fit of calculated to experimental values to around 2 kJ mol^{-1} standard deviations in alkenes.

The following tables use two different notations for single bonds next to double bonds, these being CC_1<CC_2>1 shortened to <CC_2> and sp^2sp^3. The difference is used to distinguish between an sp^2sp^3 B(π) parameter and an sp^2sp^3 bond energy term [$E(C\text{-}C)_{sp^2sp^3}$]. Thus <CC_2> would mean CC_1<CC_2>1¨CC_1<CC_2>1 and sp^2sp^3 represents a LSE for the number of sp^2sp^3 bonds. [S] represents a conventional steric correction term. An error was incurred in the standard deviation calculations by using the spreadsheet function standard deviation inappropriately. This caused standard deviations to be underestimated commensurate with the deviation from zero of the signed average of the residuals. We have not undertaken the task of recalculating correct standard deviations because ultimately we were looking for low standard deviations and the errant standard deviations here represent a lower bound. So when analysing the data it must be kept in mind that the correct standard deviation could not be smaller than the standard deviation of the spreadsheet function, but may be larger. Thus a large standard deviation can be immediately dismissed, but if it is small it may be correct or else erroneous from an unaccounted bias and absent regression constant in the model. This is only the case for Tables 3.17 and 3.18.

Table 3.17

Test group	**No. of Parameters and Associated Standard Deviation (kJ mol^{-1})**				**STO-3G geometry optimised B(π) and other parameters**
	2	**3**	**4**	**5**	
group 1	4.92				¨
		5.13			¨, sp^2sp^3
		3.44			¨, sp^2sp^3 - [S]
		3.19			¨, sp^2sp^3 with LSE for [S]
group 2	4.93				¨
		5.88			¨, sp^2sp^3
		2.99			¨, sp^2sp^3 - [S]
group 1		5.03			¨C-C, ¨C=C
			5.16		¨C-C, ¨C=C, sp^2sp^3
			3.35		¨C-C, ¨C=C, sp^2sp^3 - [S]
			3.13		¨C-C, ¨C=C, sp^2sp^3 with LSE for [S]
group 2		5.22			¨C-C, ¨C=C
		4.14			¨C-C, ¨C=C - [S]
			3.11		¨C-C, ¨C=C, sp^2sp^3 - [S]
group 1		5.91			C-C, C-C¨C=C
			5.03		C-C, C-C¨C=C, sp^2sp^3
			3.46		C-C, C-C¨C=C, sp^2sp^3 - [S]
			3.05		C-C, C-C¨C=C, sp^2sp^3 with LSE for [S]
group 2		5.97			C-C, C-C¨C=C
		7.28			C-C, C-C¨C=C - [S]
			2.92		C-C, C-C¨C=C, sp^2sp^3 - [S]
group 1		7.63			C-C, C=C
			5.39		C-C, C=C, sp^2sp^3
			3.22		C-C, C=C, sp^2sp^3 - [S]
			2.99		C-C, C=C, sp^2sp^3 with LSE for [S]
group 2		6.72			C-C, C=C

Test group	**No. of Parameters and Associated Standard Deviation (kJ mol^{-1})**				**STO-3G geometry optimised B(π) and other parameters**
	2	**3**	**4**	**5**	
		6.54			C-C, C=C, - [S]
			3.11		C-C, C=C, sp^2sp^3 - [S]
group 1	8.13				¡
		5.59			¡, sp^2sp^3
		6.23			¡, sp^2sp^3 - [S]
		5.89			¡, sp^2sp^3 with LSE for [S]
group 2	7.72				¡
		5.59			¡, sp^2sp^3
		6.41			¡, sp^2sp^3 - [S]
group 1		7.59			¡C-C, ¡C=C
			5.81		¡C-C, ¡C=C, sp^2sp^3
			5.83		¡C-C, ¡C=C, sp^2sp^3 - [S]
			5.84		¡C-C, ¡C=C, sp^2sp^3 with LSE for [S]
group 2		6.49			¡C-C, ¡C=C
			5.65		¡C-C, ¡C=C, sp^2sp^3
			6.04		¡C-C, ¡C=C, sp^2sp^3 - [S]
group 1			4.73		¨C-C, ¨<CC_2>1, ¨C=C
			3.77		¨C-C, ¨<CC_2>1, ¨C=C - [S]
			3.48		¨C-C, ¨<CC_2>1, ¨C=C with LSE for [S]
				2.92	¨C-C, ¨<CC_2>1, ¨C=C, sp^2sp^3 with LSE for [S]
group 2			4.84		¨C-C, ¨<CC_2>1, ¨C=C
			3.55		¨C-C, ¨<CC_2>1, ¨C=C - [S]
				3.17	¨C-C, ¨<CC_2>1, ¨C=C, sp^2sp^3 - [S]
group 1			6.44		C-C, C-C¨<CC_2>, C-C¨C=C
			4.41		C-C, C-C¨<CC_2>, C-C¨C=C - [S]
			3.99		C-C, C-C¨<CC_2>, C-C¨C=C with LSE for [S]
				4.03	C-C, C-C¨<CC_2>, C-C¨C=C, sp^2sp^3 with LSE for [S]

Test group	**No. of Parameters and Associated Standard Deviation (kJ mol^{-1})**				**STO-3G geometry optimised $B(\pi)$ and other parameters**
	2	**3**	**4**	**5**	
group 1			5.89		C-C, CC_1<CC_2>1, C=C
			7.35		C-C, CC_1<CC_2>1, C=C - [S]
			5.61		C-C, CC_1<CC_2>1, C=C with LSE for [S]
				5.58	C-C, CC_1<CC_2>1, C=C, sp^2sp^3 with LSE for [S]
group 1		4.95			C-C,sp^2sp^3¨C-C, C-C,sp^2sp^3¨sp^2sp^3 with LSE for [S]
			4.42		C-C,sp^2sp^3¨C-C, C-C,sp^2sp^3¨sp^2sp^3, C-C,sp^2sp^3¨C=C
			4.34		C-C,sp^2sp^3¨C-C, C-C,sp^2sp^3¨sp^2sp^3, C-C,sp^2sp^3¨C=C - [S]
			3.98		C-C,sp^2sp^3¨C-C, C-C,sp^2sp^3¨sp^2sp^3, C-C,sp^2sp^3¨C=C with LSE for [S]
				3.04	C-C,sp^2sp^3¨C-C, C-C,sp^2sp^3¨sp^2sp^3, C-C,sp^2sp^3¨C=C, sp^2sp^3 with LSE for [S]

Table 3.18

Test group	**No. of Parameters and Associated Standard Deviation**			**3-21G(*) geometry optimised $B(\pi)$ and other parameters**
	2	**3**	**4**	
group 1	5.24			¨
		5.21		¨, sp^2sp^3
		4.46		¨, sp^2sp^3 - [S]
		4.11		¨, sp^2sp^3 with LSE for [S]
group 2	5.17			¨
		6.20		¨, sp^2sp^3
		4.19		¨, sp^2sp^3 - [S]
group 1		4.96		¨C-C ¨C=C

Test group	No. of Parameters and Associated Standard Deviation			3-21G(*) geometry optimised B(π) and other parameters
	2	3	4	
			5.13	¨C-C ¨C=C sp^2sp^3
			4.47	¨C-C ¨C=C sp^2sp^3 - [S]
			4.30	¨C-C ¨C=C sp^2sp^3 with LSE for [S]
group 2		5.02		¨C-C ¨C=C
		5.57		¨C-C ¨C=C - [S]
			4.74	¨C-C ¨C=C sp^2sp^3 - [S]
group 1		6.18		C-C, C-C¨C=C
			5.16	C-C, C-C¨C=C, sp^2sp^3
			4.49	C-C, C-C¨C=C, sp^2sp^3 - [S]
			4.23	C-C, C-C¨C=C, sp^2sp^3 with LSE for [S]
group 2		6.18		C-C, C-C¨C=C
			7.90	C-C, C-C¨C=C - [S]
			4.34	C-C, C-C¨C=C sp^2sp^3 - [S]
group 1		8.66		C-C, C=C
			5.98	C-C, C=C, sp^2sp^3
			4.19	C-C, C=C, sp^2sp^3 - [S]
			4.25	C-C, C=C, sp^2sp^3 with LSE for [S]
group 2		7.01		C-C, C=C
		6.15		C-C, C=C - [S]
			4.76	C-C, C=C, sp^2sp^3 - [S]

4 Including Steric Strain

To obtain results with a standard deviation better than 3 kJ mol^{-1} for alkenes, a method of estimating the steric strain in a molecule needed to be incorporated. The different methods tested have been proposed by Skinner[5], Smith[18], Cao[122], the buffered 14 7 by Halgren[70], the MMFF94 steric energy [123] and a simple higher order inverse function of atomic distance (*r*) was also tested. Other common force fields are exp 6 functions and recently a modified Morse potential has been proposed[124].

In the previous section the conventional steric term was not included in the regression analysis. In this section a LSE is found for conventional steric terms. This is to make a fair comparison, though not having to parameterise may be the chief boast of the older methods and those based on them.

4.1.1 Conventional Steric Correction Terms

Skinner's steric scheme uses 21 parameters for alkanes and is less straightforward to calculate from the structural formula alone. Allen[4], and Benson and Buss[3] used a simpler method that counts the number of 1,4-*gauche* interactions, and propose the values 2.1 and 2.9 kJ mol^{-1} respectively. Cox and Pilcher[6] use the same approach using 2.5 kJ mol^{-1}, but include two further parameters for adjacent quaternary and tertiary carbon atoms, and two adjacent quaternary atoms. Smith uses the Cox and Pilcher value, but not the latter two steric correction terms. This is how the conventional steric correction term is calculated when not specified otherwise in this work.

In alkanes the number of 1,4-*gauche* interactions is arrived at by considering the secondary (S), tertiary (T) or quaternary (Q) structure of both carbons over each bond in a molecule. If the carbons in a bond are subscripted with its structure S, T or Q, then the following C-C bonds have the subsequent parenthetical number of 1,4-*gauche* interactions: C_S-C_T(1), C_S-C_Q(2), C_T-C_T(2), C_T-C_Q(4) and C_Q-C_Q(6). The number of 1,4-*gauche* interactions is then multiplied by 2.5 kJ mol^{-1}.

Table 4.1

CSE in alkanes.

molecule	Number of 1,4 gauche interactions	CSE kJ mol^{-1}
methane		0
ethane		0
propane		0
butane		0
2-methylpropane		0
pentane		0
2-methylbutane	C_S-C_T(1)	2.5

molecule	Number of 1,4 gauche interactions	CSE kJ mol^{-1}
2,2-dimethylpropane		0
hexane		0
2-methylpentane	C_S-C_T(1)	2.5
3-methylpentane	2 C_S-C_T(1)	5
2,2-dimethylbutane	C_S-C_Q(2)	5
2,3-dimethylbutane	C_T-C_T(2)	5
heptane		0
2-methylhexane	C_S-C_T(1)	2.5
3-methylhexane	2 C_S-C_T(1)	5
3-ethylpentane	3 C_S-C_T(1)	7.5
2,2-dimethylpentane	C_S-C_Q(2)	5
2,3-dimethylpentane	C_S-C_T(1), C_T-C_T(2)	7.5
2,4-dimethylpentane	2 C_S-C_T(1)	5
3,3-dimethylpentane	2 C_S-C_Q(2)	10
2,2,3-trimethylbutane	C_T-C_Q(4)	10
octane		0
2-methylheptane	C_S-C_T(1)	2.5
3-methylheptane	2 C_S-C_T(1)	5
4-methylheptane	2 C_S-C_T(1)	5
3-ethylhexane	3 C_S-C_T(1)	7.5
2,2-dimethylhexane	C_S-C_Q(2)	5
2,3-dimethylhexane	C_S-C_T(1), C_T-C_T(2)	7.5
2,4-dimethylhexane	3 C_S-C_T(1)	7.5
2,5-dimethylhexane	2 C_S-C_T(1)	5
3,3-dimethylhexane	2 C_S-C_Q(2)	10
3,4-dimethylhexane	2 C_S-C_T(1), C_T-C_T(2)	10
3-ethyl-2-methylpentane	2 C_S-C_T(1), C_T-C_T(2)	10
3-ethyl-3-methylpentane	3 C_S-C_Q(2)	15
2,2,3-trimethylpentane	C_S-C_T(1), C_T-C_Q(4)	12.5
2,2,4-trimethylpentane	C_S-C_T(1), C_S-C_Q(2)	7.5
2,3,3-trimethylpentane	C_S-C_Q(2), C_T-C_Q(4)	15
2,3,4-trimethylpentane	2 C_T-C_T(2)	10
2,2,3,3-tetramethylbutane	C_Q-C_Q(6)	15

Similar to a conventional method for strain is the method of Cao[122]. Neither is a function of interatomic distance. In Cao's method, molecular strain energy is related to a sum of squared products of vertex degrees V_i and V_j of the ith and jth carbon over each C_i-C_j bond, where $S_{ij} = (V_i V_j)^2$. The vertex degree is the number of carbon atoms bonded to a carbon. Thus for the second and third carbons in 2-methylbutane $S_{2,3} = (3 \times 2)^2$ and for the molecule $\Sigma\, S_{ij} = 58$.

4.1.2 The Simple Steric Repulsion Term

A purely destabilising steric parameter would neglect the dispersive term of a formula modelling the van der Waals potential. Thus in equation (36) *n* was varied to find the best fit, r_{ij} was the distance between non bonded H atoms and *A* was ascertained by least squares analysis for each value of *n*. We call this simple steric repulsion (SSR).

$$S_{ij} = A\sum \frac{1}{r_{ij}^{\ n}} \qquad \ldots (36)$$

This can be rearranged to take the form of the repulsive term of a Lennard-Jones potential. If $r^* = A^{1/n}$ then equation (36) has the form:

$$S_{ij} = \sum \left(\frac{r^*}{r_{ij}} \right)^n \qquad \ldots (37)$$

Parameterising against all 40 alkanes, the values of k¨ and *A* with associated *r** for each *n* with the statistics, standard deviation (s) and t-statistic: k¨/stderr and A/stderr for k¨ and *A* respectively are shown in Tables 4.2 and 4.3.

Table 4.2

B(π) is calculated on the STO-3G basis set and geometry optimised.

n	**k¨ (kJ mol^{-1})**	***A* (Å$^{1/n}$)**	**R²**	**s (kJ mol^{-1})**	**k¨/stderr**	***A*/stderr**	***r** (Å)**
1	58.485	-1.14	0.9929	3.41	17.01	9.36	1.136
2	53.834	-4.37	0.9945	2.99	16.41	11.22	2.090
3	48.804	-15.44	0.9955	2.69	15.02	12.87	2.490
4	44.448	-49.88	0.9965	2.40	14.21	14.76	2.658
5	42.741	-144.47	0.9971	2.18	14.72	16.49	**2.704**
6	44.448	-375.85	**0.9972**	**2.10**	16.51	**17.15**	2.686
7	48.474	-901.99	0.9970	2.13	19.37	16.85	2.643
8	53.295	-2059.76	0.9968	2.20	23.19	16.27	2.596

n	$\ddot{k}$ (kJ mol^{-1})	A (Å$^{1/n}$)	R^2	s (kJ mol^{-1})	$\ddot{k}$/stderr	A/stderr	r^* (Å)
9	57.982	-4575.40	0.9965	2.27	27.85	15.74	2.551
10	62.174	-10010.73	0.9964	2.33	33.18	15.28	2.512
11	65.801	-21714.23	0.9962	2.39	38.96	14.85	2.479
12	68.904	-46849.46	0.9960	2.45	45.00	14.40	2.450
13	71.553	-100715.68	0.9958	2.52	51.13	13.92	2.426
14	73.816	-215938.40	0.9955	2.60	57.22	13.41	2.404
15	75.756	-461996.76	0.9952	2.69	63.19	12.89	2.386
16	77.423	-986673.01	0.9949	2.78	68.97	12.37	2.369
17	78.860	-2103961.64	0.9946	2.88	74.52	11.85	2.355
18	80.102	-4480382.74	0.9943	2.97	79.80	11.36	2.342
19	81.178	-9529623.88	0.9939	3.07	84.78	10.88	2.330
20	82.112	-20248063.50	0.9936	3.17	89.43	10.44	2.319
21	82.925	-42983055.26	0.9932	3.26	93.72	10.03	2.309
22	83.633	-91174208.08	0.9929	3.36	97.65	9.65	2.300
23	84.251	-193267352.27	0.9926	3.44	101.19	9.29	2.292
24	84.792	-409452416.98	0.9922	3.53	**104.34**	8.97	2.285

Table 4.3

B(π) is calculated on the 3-21G(*) basis set and geometry optimised.

n	$\ddot{k}$ (kJ mol^{-1})	A (Å$^{1/n}$)	R^2	s (kJ mol^{-1})	$\ddot{k}$/stderr	A/stderr	r^* (Å)
1	83.268	-0.70	0.9938	3.10	19.02	5.36	0.701
2	78.560	-2.82	0.9945	2.93	16.99	6.09	1.678
3	73.087	-10.42	0.9951	2.75	14.91	6.86	2.184
4	67.868	-35.09	0.9959	2.52	13.86	7.94	2.434
5	65.490	-103.61	0.9967	2.29	14.59	9.19	2.530
6	67.004	-268.19	0.9971	2.12	17.25	10.23	**2.539**
7	70.939	-632.60	0.9974	2.03	21.54	10.88	2.513
8	75.522	-1415.88	0.9975	1.98	27.09	11.22	2.477
9	79.834	-3084.84	0.9975	**1.96**	33.57	11.36	2.442
10	83.580	-6631.04	0.9975	**1.96**	40.73	**11.37**	2.411
11	86.737	-14158.59	0.9974	1.98	48.36	11.27	2.384
12	89.372	-30132.55	0.9973	2.00	56.29	11.11	2.362
13	91.571	-64029.48	0.9973	2.03	64.40	10.90	2.343

14	93.411	-135963.15	0.9972	2.06	72.58	10.66	2.326
15	94.959	-288622.73	0.9970	2.10	80.75	10.40	2.312
16	96.268	-612600.33	0.9969	2.14	88.82	10.14	2.300
17	97.381	-1300111.24	0.9968	2.18	96.72	9.88	2.289
18	98.333	-2758907.20	0.9967	2.22	104.38	9.62	2.279
19	99.150	-5853768.17	0.9966	2.26	111.73	9.37	2.271
20	99.856	-12418332.34	0.9964	2.31	118.71	9.12	2.263
21	100.466	-26339647.78	0.9963	2.35	125.27	8.89	2.256
22	100.997	-55855806.27	0.9962	2.39	131.37	8.67	2.250
23	101.460	-118422511.38	0.9961	2.42	136.99	8.46	2.244
24	101.865	-251018855.45	0.9959	2.46	**142.11**	8.27	2.239

Using STO-3G B(π) the standard deviations in Table 4.2 are at a minimum around $n = 6$ and n = 7 with r^* 2.686 and 2.643 respectively. Using 3-21G(*) B(π) in Table 4.3 with $n = 7$ through to 13, with r^* ranging through 2.5 to 2.3, equivalent results are obtained.

4.1.3 The Buffered 14 7 Method

The Buffered 14 7 method used in molecular mechanics was modified and tested on alkanes. In the MMFF94 programme, the algorithm calculates the steric energy between any two nonbonded atoms no closer than the next nearest neighbour. It is our opinion that steric strain is mostly incurred between H···H nonbonded interactions. This was supported by improved results when only the H···H interactions were considered in our calculations. A further basis for excluding non H···H interactions is that the van der Waals energy profile given by the buffered 14 7 algorithm has an energy minimum where atoms attract. Between unlike atoms there can be a greater attraction than there is for like atoms, so the algorithm may perform more like a purely steric function between just the H···H nonbonded interactions. Another basis for modification arises from incorporating the buffered 14 7 in an oversimplified method. The following is a quote from Halgren[70] about the buffered 14 7 potential:

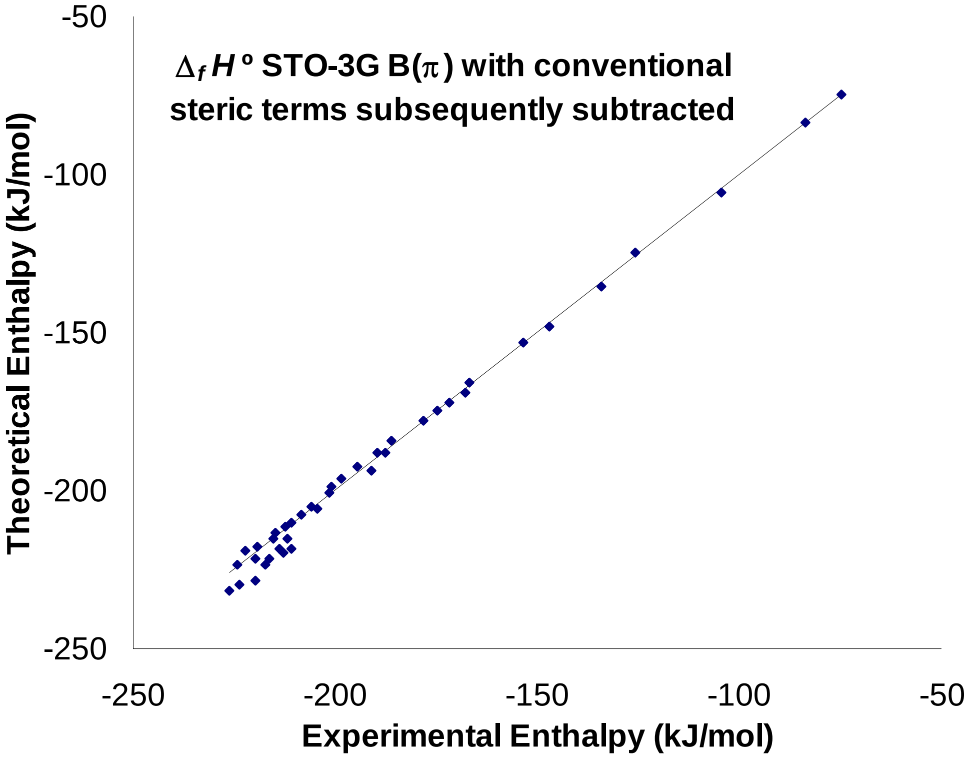
“…the vdW parameters… are ultimately intended for use in a more complex, physically superior force field, and indeed may perform optimally only in such a context.”

We used the parameters available from Halgren’s work[70]. The resultant value from the buffered 14 7 potential restricted to H···H interactions was then included in our least squares analysis, *c.f.* LSE in Tables 4.5 and 4.6.

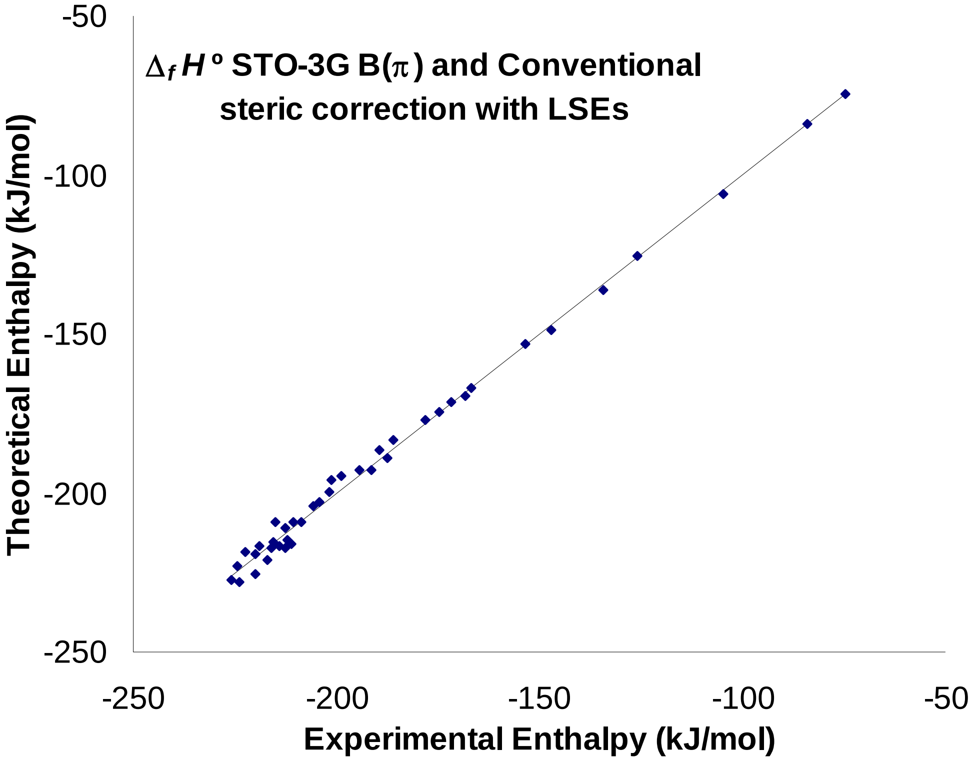
4.1.3.1 Buffered 14 7 Results for Alkanes

Graph 4.1 shows the $\Delta_f H^o$ calculated from STO-3G B(π) with the CSE subsequently subtracted, which can be compared with Graph 4.2 where CSE is included in the regression analysis. Graph 4.1 compared with Graph 4.2 or Graph 4.3 where the buffered 14 7 is evaluated with STO-3G B(π), seems to have a better correlation with B(π) when conventional steric correction is subtracted subsequent to regression analysis, though the standard deviation is worse. This is not the case for B(π) evaluated at the HF 3-21G(*) level where standard deviation as well as correlation is best parameterised with the buffered 14 7 method.

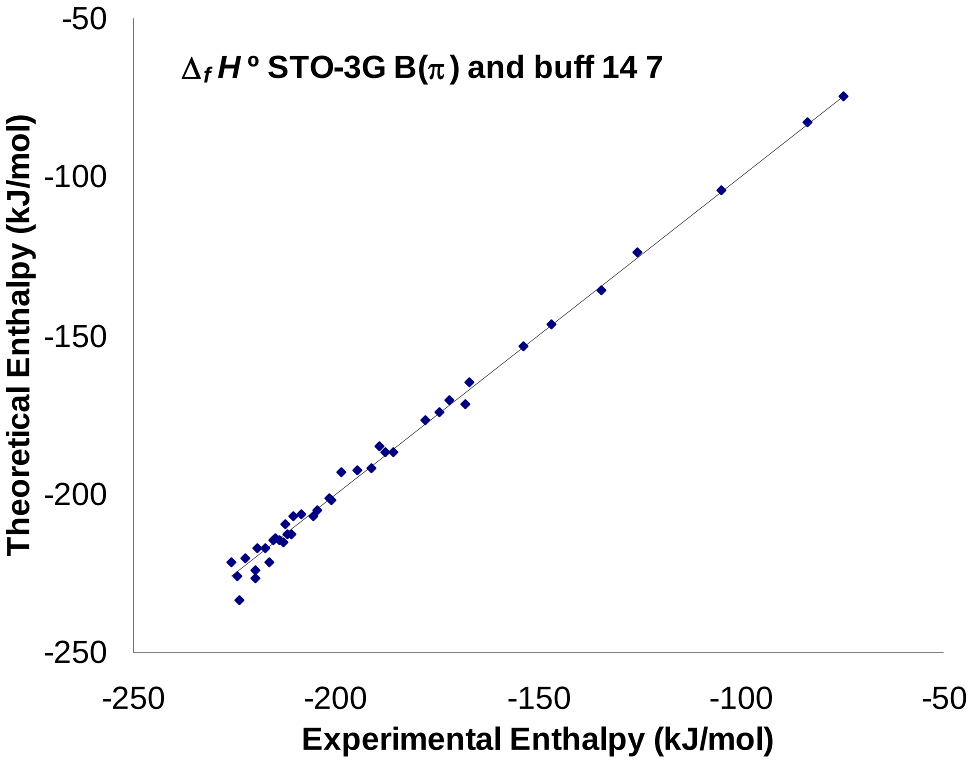
Graph 4.1



Graph 4.2



Graph 4.3



The statistics for Graph 4.3 are given in the first row of Table 4.4.

Graph 4.4

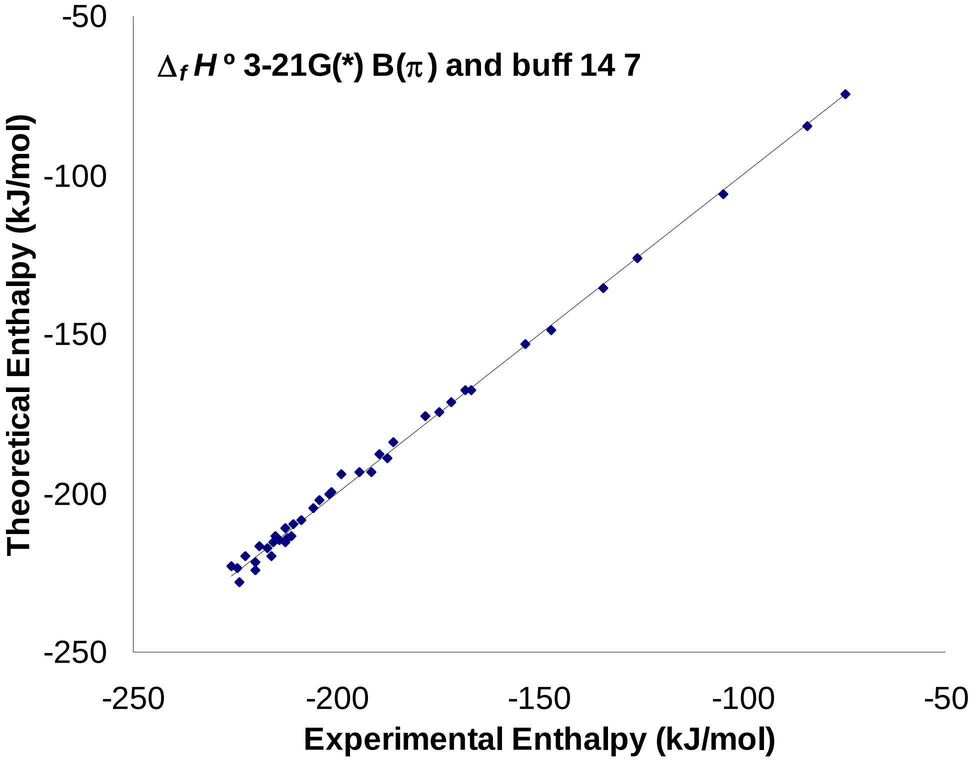


Table 4.4

Comparing buffered 14 7 and CSE using B(π) at the STO-3G and 3-21G(*) levels. The column headings have meanings: k¨ is the LSE for B(π) while the steric coefficient column is the LSE for the respective steric function and s is the standard deviation.

	K¨ (kJ mol^{-1})	**Steric coefficient**	**R²**	**s (kJ mol^{-1})**
STO-3G CSE*	83.856		0.9943	3.2
STO-3G CSE	82.848	-3.39	0.9952	2.7
STO-3G Buffered 14 7	82.572	-9.63	0.9944	3.0
3-21G(*) CSE*	102.520		0.9954	3.6
3-21G(*) CSE	101.291	-2.00	0.9955	2.6
3-21G(*) Buffered 14 7	99.895	-6.05	0.9972	2.0

* CSE term is subtracted subsequent to regression analysis.

4.2 Statistics and Steric Parameters

In the following tables the regression analysis is performed over the entire group of 40 alkanes, and each row has the parameters and statistics for each of the steric methods shown in the first column. The steric energy is from the MMFF94 molecular mechanics programme and r^{-n} is a partial form of the SSR which multiplied with the associated steric parameter is an equivalent form of *eq.* (36).

Table 4.5

With B(π) from a STO-3G basis set, the different steric methods are compared. The column headings have meanings: k¨ is the LSE for B(π) while the steric coefficient column is the LSE for the respective steric function and s is the standard deviation.

STO-3G	**K¨ (kJ mol^{-1})**	**Steric coefficient**	**R²**	**s (kJ/mol)**
r^{-6}	44.448	-375.849	0.9972	2.1
Buffered 14 7	82.572	-9.635	0.9944	3.0
CSE	82.848	-3.387	0.9952	2.7
Steric Energy	86.360	-0.624	0.9918	3.5
Cao	79.796	-0.086	0.9945	2.9

Table 4.6

With B(π) from a 3-21G(*) basis set, the different steric methods are compared. The column headings have meanings: k¨ is the LSE for B(π) while the steric coefficient column is the LSE for the respective steric function and s is the standard deviation.

3-21G(*)	**K¨ (kJ mol^{-1})**	**Steric coefficient**	**R²**	**s (kJ/mol)**
***r*$^{-9}$**	79.834	-3084.84	0.9975	2.0
Buffered 14 7	99.895	-6.05	0.9972	2.0
CSE	101.291	-2.00	0.9955	2.6
Steric energy	104.244	-0.30	0.9926	3.3
Cao	99.479	-0.05	0.9945	2.9

The two best methods are buffered 14 7 and SSR. Buffered 14 7 has the advantage that only one linear LSE needs to be determined for a model, but SSR requires a linear and nonlinear LSE for a model.

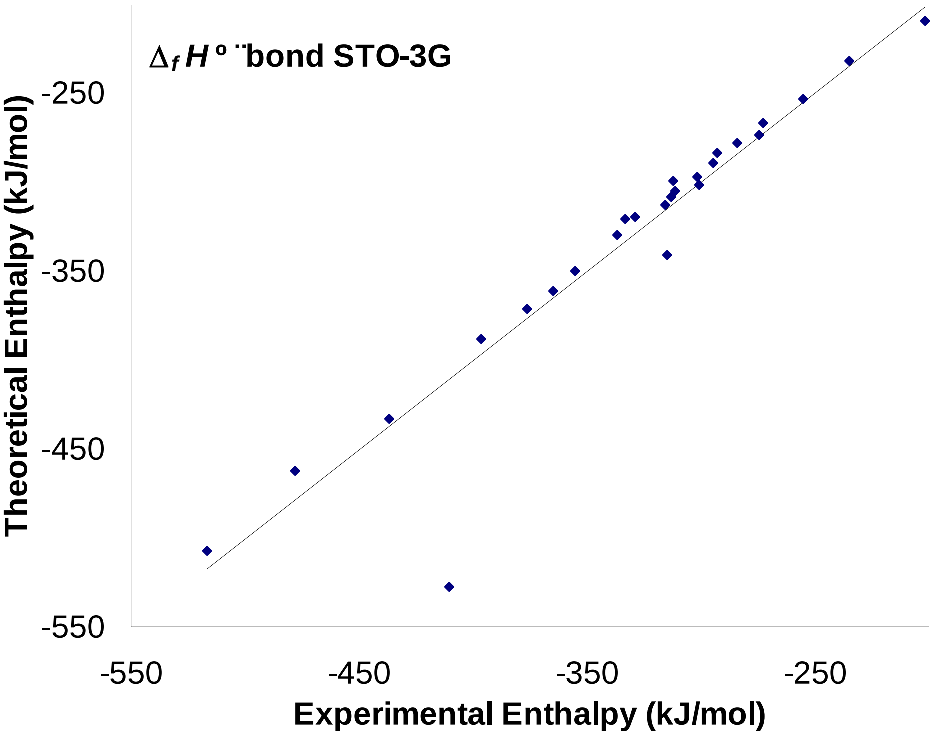
5 Alcohols

With alcohols we have another hetero atom to consider. This introduces a disparity of charge and electronegativity between oxygen and carbon. Thus either the electrostatic energy or an electronegativity scheme[125] may be necessary. Group methods have a different parameter for a C-O bond in primary, secondary and tertiary alcohols, or some equivalent parameters in the group method associated with the difference in these structures *mutatis mutandis*.

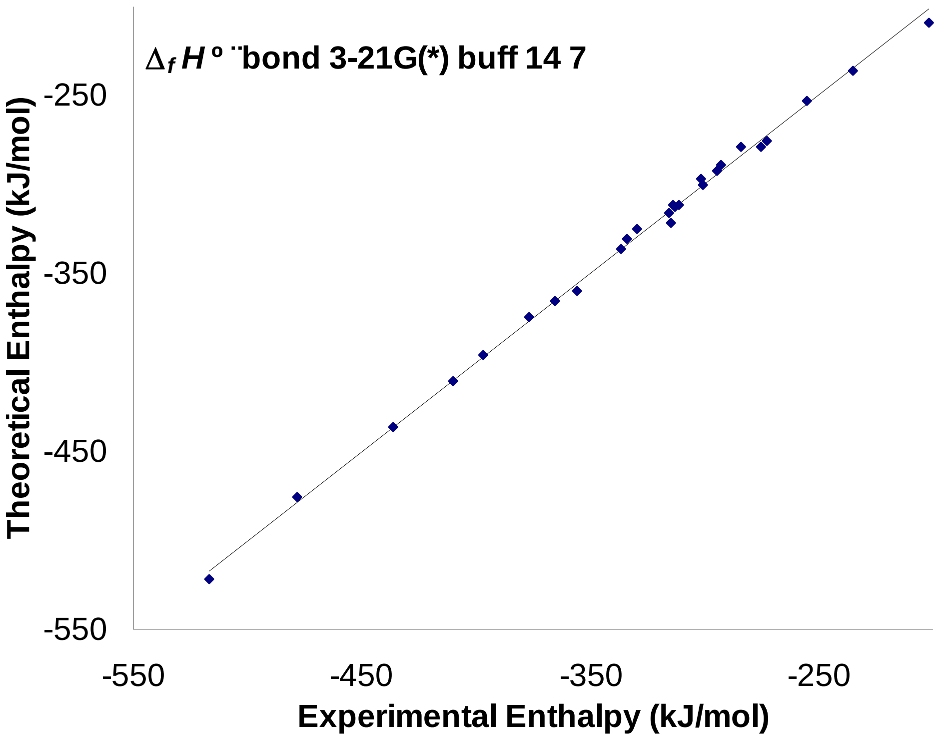
Using a STO-3G B(π) parameter, experimental values for molecules with steric crowding cannot be fitted. An additional steric or strain energy term is necessary. Graph 5.1 shows correlation between experimental and calculated values without a steric term.

The outlying data of Graph 5.1 are the calculated values for 3-pentanol and most seriously 3-tert-butyl-2,2,4,4-tetramethyl-3-pentanol. Without a steric term both these molecules have $\Delta_f H^o$ values calculated too low, by 25 and 94 kJ mol^{-1} respectively. If the SSR or modified buffered 14 7 algorithm (see section 4.1.3) is used, the graph is considerably improved. This is shown using buffered 14 7 energy in Graph 5.2.

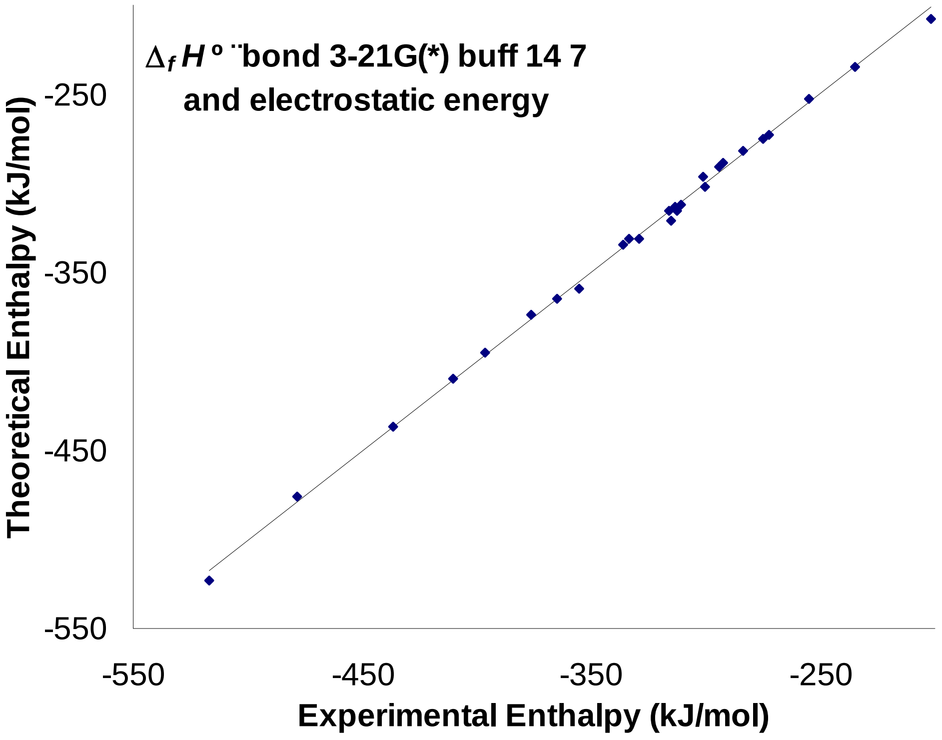
Graph 5.1



Graph 5.2



Graph 5.3



6 Conformational Analysis

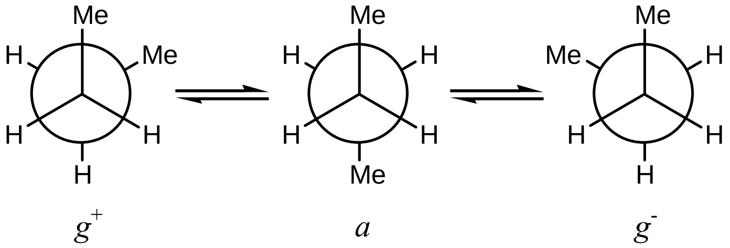
Our theoretical calculations with modern steric functions often calculate $\Delta_f H^o$ values too low for smaller alkanes. The experimental value is the average of all the different conformers in the test sample; this would then be higher than the actual value of the lowest energy conformer on which our calculations are based. The $\Delta_f H^o$ energies for alkanes are the most precisely known of the molecules tested in this work. Thus the adjustment for the lowest energy conformer relative to a higher experimental value due to population of higher energy conformers can exceed the experimental error for these molecules. For each conformer there is a distribution of molecules in higher energy vibrational states the average of which is $\Delta H_{vib}(T)$. In the calculations of this section the vibrational energies [$\Delta H_{vib}(T)$ and $H_{vib}(0)$] are approximated to be constant for each conformer and thus cancel when calculating relative energies. To calculate the relative energy for a set of conformers assuming vibrational energies cancel, the ΔH^o_{Tot} of the lowest energy conformer is subtracted from all the other conformers, so that energy relative to the lowest energy conformer, which is therefore zero, is obtained for each conformer. The smallest *n*-alkane to have a hairpin geometry lower in energy than an all-*trans* conformer probably occurs around the $C_{16} – C_{18}$ range, so the lowest energy conformer in this analysis will always be the all-*trans* conformer[126].

For alkanes, Eliel *et al.*[127] covers conformational analysis up to C5 and 2,3-dimethylbutane. Dale[128] also gives examples of calculations including hexane. An example calculation for butane is as follows:

There are three conformers shown in Figure 6.1. The plus *gauche* (g^+) and minus *gauche* (g^-) forms are equivalent in energy and are enantiomers, so an energy calculation only needs to be performed on one of them.

Figure 6.1

Butane *gauche* (g) and *anti* (a) conformers.



The experimental relative energy for the *gauche* conformers is 3.9 kJ mol^{-1}, the midrange value for the experimental range 3.7 – 4.1 kJ mol^{-1} given in Eliel *et al.*[127]. Alternatively, relative energies that include consideration for adjacent *gauche* stabilisation can be ascertained using the table of results from Klauda *et al.*, reproduced in Table 6.2 with only values for *n*-alkanes from butane to hexane (*cf.* Klauda *et al.* show results for *n*-alkanes from butane to decane). From their tables, a value based on the number and juxtaposition of *gauche* interactions can be ascertained. From Table 6.2, the relative energy for the g conformers is 2.64 kJ mol^{-1}. The relative proportions of each conformer can be determined from the equilibrium constant $K = [a]/[g]$ and $K = e^{-\Delta G^o/RT}$, where $\Delta G^o = \Delta H - T\Delta S$ is the free energy and depends on enthalpy (ΔH) as well as entropy (ΔS). Rotational and vibrational entropy differences between conformers are small and are not calculated. However two sources of entropy that must be considered are the entropy of mixing (S_{mix}) and entropy of symmetry (S_{sym}). The entropy of symmetry is based on the symmetry number (σ) which can be ascertained from Table 6.1, where $S_{sym} = -R\ln\sigma$. The entropy of mixing is ascertained by $S_{mix} = -R\Sigma\, n_i \ln n_i$, where n_i is the fraction of distinguishable types relative to each other. For a pair of enantiomers $S_{mix} = -R(0.5 \ln 0.5 + 0.5 \ln 0.5) = R\ln 2$. The S_{sym} and the S_{mix} values are added to give ΔS. The free energy can then be calculated with ΔS and relative energies either obtained from experiment, quantum chemical calculation or from the tables by Klauda *et al.*. A K_i for each i^{th} conformer can then be calculated. The mole fraction (M*f*) of each i^{th} conformer is then $K_i/\Sigma K_i$. These values are tabulated for butane in Table 6.3.

The weighted average energy (E_{exp}) in terms of the fraction (f_i) and energy (E_i) of the i^{th} conformer in a total of n confomers is

$$E_{exp} = \sum_{i=1}^{n} f_i E_i$$

If the relative energy to E_o is

$$\Delta_r E_i = E_i - E_o$$

then

$$E_i = \Delta_r E_i + E_o.$$

Therefore

$$\begin{aligned} E_{exp} &= \sum_{i=1}^{n} f_i(\Delta_r E_i + E_o) \\ &= \sum_{i=1}^{n} f_i \Delta_r E_i + \sum_{i=1}^{n} f_i E_o \\ &= \sum_{i=1}^{n} f_i \Delta_r E_i + E_o \end{aligned}$$

The difference between E_{exp} and the energy of the lowest energy conformer, for which $i = 1$, is then

$$E_{exp} - E_1 = \sum_{i=1}^{n} f_i \Delta_r E_i + E_o - E_1.$$

If $E_o = E_1$ then

$$E_{exp} - E_1 = \sum_{i=2}^{n} f_i \Delta_r E_i \ .$$

If $\Delta_f H^o$ is substituted for E_{exp} then the equation is in one unknown: E_1, which corresponds to the corrected energy of a conformer relative to the experimental $\Delta_f H^o$. If the experimental value for the *g* conformers of butane is used, then the results of Table 6.3 are ascertained with each step in the process proceeding by column from left to right.

Table 6.1

Symmetry number and order in point group

Group	C_1	C_n	D_n	C_s	S_n	C_{nv}; C_{nh}	$C_{\infty v}$	D_{nd}; D_{nh}	$D_{\infty h}$	T_d	O_h	I_h
Order	1	*n*	2*n*	2	*n*	2*n*	∞	4*n*	∞	24	48	120
σ	1	*n*	2*n*	1	*n*/2	*n*	1	2*n*	2	12	24	60

Table 6.2

ΔE in kJ mol^{-1} , where $(g^+)_m$ is for the first alkane in a column and $a(g^+)_m a_{n-m-4}$ for *n*-alkane carbons.

alkane	***m* = 1**	**2**	**3**
***n* = 4**	2.64		
5	2.59	4.14	
6	2.51	3.89	5.31

We have updated results for pentane with recent values calculated at the CCSD(T) cc-pVDZ level by Klauda *et al.*[129]. The remaining molecules hexane, 2-methylpentane and 3-methylpentane we have calculated at the MP2 cc-pVDZ level while 2-methylbutane was calculated at the MP2 cc-pVTZ level. The calculations are as follows:

At the bottom of each graph the correction for the conformer on which calculations are performed in the other sections of this work is shown with its relation to the experimental $\Delta_f H^o$ and the $\Delta_f H^o$ of the conformer.

Table 6.3

The butane conformers.

	ΔH kJ mol^{-1}	**σ**	**d,l**	**ΔS J mol^{-1}**	**G_{298} kJ mol^{-1}**	**ΔG kJ mol^{-1}**	**M*f***
a	0	2	1	- Rln2	1.72	0	0.293
g	3.9	2	2	0	3.9	2.18	0.707

$\Delta_f H^o = \Delta_f H^o\ a + 1.17$ kJ mol^{-1}

Table 6.4

The pentane conformers.

	ΔH kJ mol^{-1}	**σ**	**d,l**	**ΔS J mol^{-1}**	**G_{298} kJ mol^{-1}**	**ΔG kJ mol^{-1}**	**M*f***
aa	0	2	1	- Rln2	1.72	0	0.358
ag	2.59	1	2	+ Rln2	0.87	-0.85	0.505
gg	4.14	2	2	0	4.14	2.42	0.136

$\Delta_f H^o = \Delta_f H^o aa + 1.87$ kJ mol^{-1}

Table 6.5

The 2-methylbutane conformers.

	ΔH kJ mol^{-1}	**σ**	**d,l**	**ΔS J mol^{-1}**	**G_{298} kJ mol^{-1}**	**ΔG kJ mol^{-1}**	**M*f***
(*ga*)	0	1	2	+ Rln2	1.72	0	0.645

(g^-g)	3.2	1	1	0	3.2	1.48	0.355

$\Delta_f H^o = \Delta_f H^o\ (ga) + 1.12\ \text{kJ mol}^{-1}$

Table 6.6

The hexane conformers.

	ΔH kJ mol^{-1}	**σ**	**d,l**	**ΔS J mol^{-1}**	**G$_{298}$ kJ mol^{-1}**	**ΔG kJ mol^{-1}**	**M*f***
aaa	0	2	1	- Rln2	1.72	0	0.203
aag	2.51	1	2	+ Rln2	0.79	-0.93	0.303
aga	2.51	2	2	0	2.51	0.79	0.151
agg	3.89	1	2	+ Rln2	2.51	0.45	0.172
gag	4.94	2	2	0	4.94	3.22	0.056
g^+ag^-	4.94	1	1	0	4.94	3.22	0.056
ggg	5.31	2	2	0	5.31	3.59	0.050

$\Delta_f H^o = \Delta_f H^o aaa + 2.63\ \text{kJ mol}^{-1}$

Table 6.7

The 2-methylpentane conformers.

	ΔH kJ mol^{-1}	**σ**	**d,l**	**ΔS J mol^{-1}**	**G$_{298}$ kJ mol^{-1}**	**ΔG kJ mol^{-1}**	**M*f***
(*ga*)*a*	0	1	2	+ Rln2	-1.717	0	0.571
(*ga*)*g*	1.212	1	2	+ Rln2	-0.506	1.212	0.350
(g^-g)*a*	3.186	1	1	0	3.186	4.903	0.079

$\Delta_f H^o = \Delta_f H^o\ (ga)a + 0.68\ \text{kJ mol}^{-1}$

Table 6.8

The 3-methylpentane conformers.

	ΔH kJ mol^{-1}	**σ**	**d,l**	**ΔS J mol^{-1}**	**G$_{298}$ kJ mol^{-1}**	**ΔG kJ mol^{-1}**	**M*f***
(*ga*)(g^-a)	0	1	1	0	0	0	0.206
(*ga*)(*ga*)	0.269	1	2	+ Rln2	-1.449	-1.449	0.370
(g^-g)(g^-a)	1.098	1	2	+ Rln2	-0.620	-0.620	0.265

(*g*$^-$*g*)(*ga*)	2.360	1	2	+ Rln2	0.642	0.642	0.159

$\Delta_f H^o = \Delta_f H^o$ (*ga*)(*g*$^-$*a*) + 0.77 kJ mol^{-1}

Table 6.9

The heptane conformers.

	ΔH **kJ mol^{-1}**	**σ**	**d,l**	**ΔS** **J mol^{-1}**	**G$_{298}$** **kJ mol^{-1}**	**ΔG** **kJ mol^{-1}**	**M*f***
aaaa	0	2	1	- Rln2	1.718	0.000	0.100
aaag	2.242	1	2	+ Rln2	0.524	-1.195	0.162
aaga	1.990	1	2	+ Rln2	0.272	-1.447	0.180
aagg	3.121	1	2	+ Rln2	1.403	-0.316	0.114
agga	2.853	2	2	0	2.853	1.135	0.063
agag	4.061	1	2	+ Rln2	2.343	0.624	0.078
aggg	4.156	1	2	+ Rln2	2.438	0.719	0.075
agag$^-$	5.006	1	2	+ Rln2	3.288	1.569	0.053
gaag	4.429	2	2	0	4.429	2.711	0.034
gaag$^-$	4.539	1	1	0	4.539	2.821	0.032
gagg	5.014	1	2	+ Rln2	3.296	1.577	0.053
g$^-$***agg***	6.142	1	2	+ Rln2	4.424	2.705	0.034
gggg	5.618	2	2	0	5.618	3.900	0.021

$\Delta_f H^o = \Delta_f H^o$ *aaaa* + 3.04 kJ mol^{-1}

Table 6.10

The octane conformers.

	ΔH **kJ mol^{-1}**	**σ**	**d,l**	**ΔS** **J mol^{-1}**	**G$_{298}$** **kJ mol^{-1}**	**ΔG** **kJ mol^{-1}**	**M*f***
aaaaa	0.000	2	1	- Rln2	1.718	0.000	0.052
aaaag	2.228	1	2	+ Rln2	0.509	-1.209	0.085
aaaga	1.921	1	2	+ Rln2	0.202	-1.516	0.096
aaagg	3.019	1	2	+ Rln2	1.300	-0.418	0.061
aagaa	1.759	2	2	0	1.759	0.041	0.051
aagag	3.818	1	2	+ Rln2	2.100	0.381	0.045
aagag$^-$	4.781	1	2	+ Rln2	3.062	1.344	0.030

	ΔH kJ mol^{-1}	**σ**	**d,l**	**ΔS J mol^{-1}**	**G$_{298}$ kJ mol^{-1}**	**ΔG kJ mol^{-1}**	**M*f***
aagga	2.486	1	2	+ Rln2	0.768	-0.951	0.076
aaggg	3.785	1	2	+ Rln2	2.067	0.348	0.045
agaag	4.080	1	2	+ Rln2	2.362	0.644	0.040
$agaag^{-}$	4.193	1	2	+ Rln2	2.475	0.756	0.038
agaga	3.671	2	2	0	3.671	1.952	0.024
agagg	4.549	1	2	+ Rln2	2.831	1.112	0.033
$agag^{-}a$	4.705	1	1	0	4.705	2.986	0.016
$agag^{-}g^{-}$	5.808	1	2	+ Rln2	4.090	2.371	0.020
aggag	4.337	1	2	+ Rln2	2.619	0.901	0.036
$aggag^{-}$	5.518	1	2	+ Rln2	3.800	2.082	0.022
aggga	3.422	2	2	0	3.422	1.704	0.026
agggg	4.905	1	2	+ Rln2	3.187	1.468	0.029
gaaag	4.429	2	2	0	4.429	2.710	0.017
$gaaag^{-}$	4.488	1	1	0	4.488	2.769	0.017
gaagg	5.150	1	2	+ Rln2	3.432	1.713	0.026
$gaag^{-}g^{-}$	5.195	1	2	+ Rln2	3.477	1.759	0.026
gagag	5.895	2	2	0	5.895	4.177	0.010
$gagag^{-}$	6.829	1	2	+ Rln2	5.111	3.393	0.013
gaggg	5.585	1	2	+ Rln2	3.867	2.149	0.022
$gag^{-}ag$	7.828	2	2	0	7.828	6.110	0.004
$gag^{-}g^{-}g^{-}$	6.784	1	2	+ Rln2	5.065	3.347	0.013
ggagg	5.251	2	2	0	5.251	3.533	0.012
$ggag^{-}g^{-}$	6.905	1	1	0	6.905	5.187	0.006
ggggg	6.366	2	2	0	6.366	4.648	0.008

$\Delta_f H^{\circ} = \Delta_f H^{\circ}$ *aaaaa* + 3.56 kJ mol^{-1}

Table 6.11

Adjusted $\Delta_f H^{\circ}$ due to conformational mixing.

	$\Delta_f H^{\circ}$ experimental	**$\Delta_f H^{\circ}$ adjusted**
butane	-125.7	-126.9
2-methylpropane	-134.2	-134.2

pentane	-146.9	-148.8
2-methylbutane	-153.6	-154.7
hexane	-166.9	-169.5
2-methylpentane	-174.6	-175.3
3-methylpentane	-171.9	-172.7
heptane	-187.6	-190.6
octane	-208.5	-212.1

These calculations assume that contributions from the most stable conformers dominate the population analysis of the sample, and the more transient or less stable intermediaries are ignored. For instance the g^+g^- conformer of pentane is ignored. In the g^+g^- conformer, if diamond angles and bond lengths are maintained, the hydrogens of the terminal carbons would be positioned closer than a normal H-H bond. For instance if a g^+g^- conformer is constructed with typical bond lengths and angles in the Spartan build programme, prior to geometry optimisation of any kind, this H-H bond distance is 0.71 Å. In this orientation the molecule asymmetrically deforms so that two energy minima are then associated with the g^+g^- conformer. Osawa *et al.*[130] developed an exhaustive conformation search algorithm which ascertained a large number of conformers for population analysis, where asymmetrical deformations in high energy g^+g^- interactions were also included in the conformational search. Saunders'[131] method is used in MM4 for conformational searching.

7 Developing B(π) with nonbonded steric functions

What minimum number of parameters for specific B(π) and steric terms are required for a model that calculates $\Delta_f H^o$ values, with errors comparable with current group methods *viz.* Pedley's[15] method?

A test group of 346 molecules (see appendix A.3.1) was initially chosen to parameterise against $\Delta_f H^o$ values. To compare with Pedley's scheme we used the experimental $\Delta_f H^o$ (g) values of his work[15]. We anticipated developing a model that might eventually be suitable for calculating $\Delta_f H^o$ (g) for amino acids and peptides, and so molecules that exhibited functional groups common in amino acids were chosen, where there were at least a few $\Delta_f H^o$ (g) values in a group according to the grouping in Pedley's work. If the 346 molecules are grouped in the same way as Pedley, then 17 sub groups of the 346 molecule test can be distinguished. The 17 groups that the molecules of the test group belong to are: alkanes, alkenes, alkynes, alcohols, ethers, aldehydes, ketones, carboxylic acids, esters, amines, amides, diazenes, nitriles, nitroalkanes, nitrates, thiols and benzenoids. Of the 346 molecules eight extra molecules not available in Pedley's compilation were also included. The $\Delta_f H^o$ values for the following were obtained from the JANAF tables[110]: pent-1-yne, pent-2-yne, 3-methylbut-1-yne, 3,3-dimethylbut-1-yne, 3-tert-butyl-2,2,4,4-tetramethyl-3-pentanol, hydrogen cyanide, ethenetricarbonitrile and ethenetetracarbonitrile. Over the test set of 346 molecules, excepting the latter list, using around 107 parameters, Pedley's method has a standard deviation of 7 kJ mol^{-1}.This was our benchmark. A programme was written to ascertain all the molecular fragments required for a group method like Pedley's. The number of molecular fragments based on a bond and its next nearest neighbours for the test set of 346 molecules was 238, for which Pedley only supplies values for about 107.

The bond energies of Table 7.1 are fixed from the atomisation enthalpies of methane, diamond, water, ammonia and hydrogen sulfide obtained from the JANAF tables[110].

Table 7.1

Bond energies obtained directly from CH_4, diamond, H_2O, NH_3 and H_2S.

	Bond energy (kJ mol^{-1})
E(C-H)	415.87
E(C-C)	357.4
E(O-H)	463.50
E(N-H)	390.86
E(S-H)	366.74

Trying combinations of different algorithms for B(π), buffered 14 7 or SSR and electrostatic energy, our method developed in the following way.

SSR was parameterised for each type of H···H interaction. We started by only discerning two kinds of H environments as is done in MMFF94, *i.e.* H attached to C and H attached to anything else. Following Halgren's[70] classification for the Buffered 14 7 method, these are called type 5 and type 21 hydrogens. MMFF94 actually distinguishes many other types for hydrogen but they all have the same parameters as a type 21 hydrogen. Hardness parameters were allowed to vary from 1 – 24.

The chemical environment algorithm distinguished the following 26 chemical environments to parameterise B(π) in:

CC_1
CC_1<CC_2>1
CC_1<CC_2>1<CN_3>1
CC_1<CC_2>1<CO_1>1<CO_2>1
CC_1<CC_5>2
CC_1<CN_1>1
CC_1<CN_1>1<CO_2>1
CC_1<CN_1>2
CC_1<CN_3>1
CC_1<CN_3>2
CC_1<CO_1>1
CC_1<CO_1>1<CO_2>1
CC_1<CO_1>2
CC_1<CO_2>1
CC_1<CS_1>1
CC_2
CC_3
CC_5<CC_5>2
CN_1
CN_3
CO_1
CO_2
CS_1
NN_2
NO_1
NO_5

From the work done on alkenes, the method chosen for summation of B(π) did not seem to matter. However we retested and compared algorithms that returned ¨bond and bond¨bond (bond) specific B(π) (see notation in section 1.7). It was found the bond algorithm had an improved standard deviation over the ¨bond algorithm by around 1 kJ mol^{-1}.

The bond algorithm gave a standard deviation of 10.1 kJ mol^{-1}. This was much larger than Pedley's standard deviation of 7.0 kJ mol^{-1}, though the method used only 42 parameters. The algorithm was modified to ascertain a greater diversity of chemical environments for the H···H nonbonded steric interactions. Thus a type of H was distinguished by what it was immediately attached to. The subsequent number referenced the atomic number of the atom attached. If H types (steric environments) are distinguished this way, there are four different types of nonbonded steric interaction in the test group of 346 molecules: H16···H6, H6···H6, H6···H7 and H6···H8.

A notation for the steric interactions used by our computer programmes is as follows: The atomic symbols of the two atoms of the steric interaction are written in alphabetical order without any separation. This part is followed by a space and the two atoms along with their type of steric environment, which is specified with a number are written in alphabetical order separated by a dash. The first part of the notation is redundant and can be ascertained easily from the second part, however its presence allows a programme (as well as humans) to easily sort and categorise steric interactions hierarchically by atoms involved and then by differences in steric environments. By this notation the nonbonded interactions H16···H6, H6···H6, H6···H7 and H6···H8 are written HH H16-H6, HH H6-H6, HH H6-H7 and HH H6-H8. This notation is used in the tables of this section.

Distinguishing nonbonded H···H did not give any improvement on the standard deviation, so SSR was extended to include nonbonded interactions between H, N and O, where we distinguish, O attached to one or two atoms as type O6 and O7 respectively and N attached to one, two or three atoms as types N8, N9 and N61 respectively. This distinguished 18 types of steric interaction. However by this algorithm too many steric environments were created, so all atomic numbers greater than 6 were made 21, where the number 21 is an analogy with a type 21 H in a buffered 14 7 calculation. This reduced the steric interaction types to 15.

If the number of hardness parameters (*h*) gives a base number and an exponent is given by the number of types of steric interactions (*S*) distinguishable by the different types of H, N or O (steric environments) then the number of comparisons to ascertain the hardness parameter for each steric environment is shown in *eq.* (38).

$$comparisons = h^S \qquad \ldots(38)$$

If 15 types of steric interaction are distinguished. Testing hardness parameters from 1 – 24 would require 24^{15} comparisons. With 346 molecules it takes roughly a second for each regression analysis calculation, thus it would take 16,000 billion years using one of today's computers. The test was limited to hardness parameters 6 and 12. By distinguishing these 15 steric interactions with either hardness parameter 6 or 12, a standard deviation of 7.6 kJ mol^{-1} was achieved. This required 68 parameters. However 1,2,3-propanetriol trinitrate had four steric interactions peculiar to itself resulting in four collinear parameters in the regression analysis. Removing it from the test group resulted in 11 different steric interactions, and a test group of 345 molecules, with 58 parameters (see column d of Table 7.2).

Table 7.2

The hardness parameters tested for all the models were 6 and 12, except the column with a ‘c’ label. The labels a, b, c, d and e have the following meanings: ‘a’ 346 molecules in test group with SSR ascertained for H, N and O interactions with an extra parameter for electrostatic energy; ‘b’ 346 molecules in test group with SSR ascertained for H, N and O interactions with no extra parameter; ‘c’ 346 molecules in test group with SSR ascertained for H interactions only, reducing nonbonded interactions to 8, so hardness parameters 3, 6, 9 and 12 were tested, as well as an extra parameter for electrostatic energy; ‘d’ 345 molecules in test group with SSR ascertained for H, N and O interactions, with no extra parameter; ‘e’ 345 molecules in test group with SSR ascertained over H, N and O interactions with no extra parameter and with CC_1<CN_1>2 included with CC_1<CN_1>1.

All units in kJ mol^{-1} unless specified parenthetically	**A1 bond a**	**A1 bond b**	**A1 bond c**	**A1 bond d**	**A3 bond e**
Number of parameters	**68**	**67**	**54**	**58**	**57**
HH H21-H6 (hardness)	6	6	9	6	6
HH H6-H6 (hardness)	6	6	6	6	6
HN H6-N61 (hardness)	12	12	12	12	12
HN H6-N8 (hardness)	12	12	3	12	12
HN H6-N9 (hardness)	12	6	3	6	6
HO H21-O7 (hardness)	6	6	3	6	6
HO H6-O6 (hardness)	12	6	12	12	12
HO H6-O7 (hardness)	6	12	3	6	6
NN N61-N61 (hardness)	6	6		6	6
NN N8-N8 (hardness)	6	6			
NO N8-O6 (hardness)	6	6			
NO N8-O7 (hardness)	6	6		6	6
OO O6-O6 (hardness)	6	6			

All units in kJ mol^{-1} unless specified parenthetically	**A1 bond a**	**A1 ¨bond b**	**A1 bond c**	**A1 bond d**	**A3 bond e**
Number of parameters	**68**	**67**	**54**	**58**	**57**
OO O6-O7 (hardness)	6	6			
OO O7-O7 (hardness)	6	12		6	6
E(CC_2)	600.7	622.9	600.7	602.0	602.7
E(CC_3)	813.7	815.8	815.4	815.2	833.9
E(CC_5<CC_5>2)	508.5	509.6	508.4	508.4	508.4
E(CN_1)	297.7	320.5	296.0	295.6	295.1
E(CN_3)	864.5	875.9	841.9	864.3	876.0
E(CO_1)	370.7	423.6	371.9	370.3	372.7
E(CO_2)	795.5	737.9	786.0	797.5	798.2
E(CS_1)	653.8	288.8	651.4	285.6	285.1
E(NN_2)	494.5	426.7	459.2	479.1	488.0
E(NO_1)	294.9	262.0	264.0	295.1	274.1
E(NO_5)	396.2	429.4	417.9	396.5	414.3
Electrostatic (Å e^{-2})	-2.6		-2.0		
CC_1	6.8	32.2	14.5	8.4	8.7
CC_1<CC_2>1	-10.4	34.2	-6.2	-13.2	-11.8
CC_1<CC_2>1<CN_3>1	-196.2	792.1	-294.7	-196.2	-194.0
CC_1<CC_2>1<CO_1…*	-137.5	5.7	-157.5	-119.2	-117.8
CC_1<CC_5>2	1.0	56.5	5.6	-0.8	-0.6
CC_1<CN_1>1	5.4	45.9	20.5	6.8	6.6
CC_1<CN_1>1<CO_2>1	110.7	100.6	212.3	109.6	150.2
CC_1<CN_1>2	1.1×10^3	178.8	349.7	1.1×10^3	
CC_1<CN_3>1	-38.7	52.8	-89.2	-42.1	-62.6
CC_1<CN_3>2	-81.3	0.0	-124.8	-83.2	-121.3
CC_1<CO_1>1	20.5	166.7	18.2	12.3	9.1
CC_1<CO_1>1<CO_2>1	-35.2	210.3	1.2	-40.2	-31.5
CC_1<CO_1>2	1.2×10^3	791.1	1.4×10^3		
CC_1<CO_2>1	37.7	287.4	47.6	34.1	23.5
CC_1<CS_1>1	2.7	38.7	8.5	-4.3	-4.1
CC_2	8.0	155.5	7.1	18.6	21.8
CC_3	-1.7×10^3	-31.0	-1.9×10^3	-1.7×10^3	-37.7

All units in kJ mol^{-1} unless specified parenthetically	**A1 bond a**	**A1 ¨bond b**	**A1 bond c**	**A1 bond d**	**A3 bond e**
Number of parameters	**68**	**67**	**54**	**58**	**57**
CC_5<CC_5>2	8.9	26.5	9.6	6.5	6.3
CN_1	-4.6	150.9	-23.0	-14.7	-22.6
CN_3	6.7×10^{5}	4.2	7.4×10^{5}	6.4×10^{5}	-32.1
CO_1	116.2	351.4	117.5	117.6	130.6
CO_2	1.1×10^{3}	357.8	1.1×10^{3}	1.1×10^{3}	1.1×10^{3}
CS_1	113.7	276.2	96.7	98.9	94.1
NN_2	142.3	185.1	39.8	114.1	145.0
NO_1	14.5	-28.7	41.9	22.9	31.1
NO_5	-272.5	90.3	23.0	-267.9	-79.2
HH H21-H6 (Å$^{1/hardness}$)	-1.2×10^{3}	-510.5	-1.7×10^{4}	-1.1×10^{3}	-1.1×10^{3}
HH H6-H6 (Å$^{1/hardness}$)	-711.5	-531.8	-656.2	-681.4	-677.9
HN H6-N61 (Å$^{1/hardness}$)	2.3×10^{6}	1.6×10^{5}	5.9×10^{6}	2.3×10^{6}	8.3×10^{5}
HN H6-N8 (Å$^{1/hardness}$)	3.6×10^{5}	7.3×10^{4}	35.3	4.5×10^{5}	4.0×10^{5}
HN H6-N9 (Å$^{1/hardness}$)	-3.8×10^{4}	1.2×10^{3}	12.1	233.1	159.4
HO H21-O7 (Å$^{1/hardness}$)	707.0	719.6	156.0	915.7	1.4×10^{3}
HO H6-O6 (Å$^{1/hardness}$)	3.9×10^{5}	1.4×10^{3}	3.2×10^{5}	4.0×10^{5}	3.8×10^{5}
HO H6-O7 (Å$^{1/hardness}$)	1.4×10^{3}	2.9×10^{5}	93.5	1.5×10^{3}	1.3×10^{3}
NN N61-N61 (Å$^{1/hardness}$)	-8.1×10^{4}	-1.1×10^{5}		-8.1×10^{4}	-8.4×10^{4}
NN N8-N8 (Å$^{1/hardness}$)	0.0	0.0			
NO N8-O6 (Å$^{1/hardness}$)	0.0	0.0			
NO N8-O7 (Å$^{1/hardness}$)	4.2×10^{4}	-5.0×10^{3}		4.3×10^{4}	1.7×10^{4}
OO O6-O6 (Å$^{1/hardness}$)	0.0	0.0			
OO O6-O7 (Å$^{1/hardness}$)	0.0	0.0			
OO O7-O7 (Å$^{1/hardness}$)	-7.9×10^{4}	1.1×10^{6}		-7.9×10^{4}	-3.4×10^{4}
Standard deviation (kJ mol^{-1})	7.6	9.0	8.0	7.7	8.1

* CC_1<CC_2>1<CO_1>1<CO_2>1

Table 7.2 shows a number of anomalous parameters for some B(π) environments, namely CC_1<CN_1>2, CC_3, CN_3 and CO_2. CC_1<CN_1>2 is unique to 1,1-dinitropropane so we bundled it into CC_1<CN_1>1-2. The remaining chemical environments by their π bonding nature gave small $B(\pi)$ values. Here a third algorithm (A3) was used that ascertains positive $B(\pi)$ over bonds with bond order greater than some specified value. The results in column e of Table 7.2, were when A3 sumed positive $B(\pi)$ in triple bonds.

The algorithm that ascertained the chemical environment ignored extra chemical information for bonds between hetero atoms. Modifying the algorithm to ascertain this information over the test group of 345 molecules, the following 37 chemical environments were ascertained:

CC_1
CC_1<CC_2>1
CC_1<CC_2>1<CN_3>1
CC_1<CC_2>1<CO_1>1<CO_2>1
CC_1<CC_5>2
CC_1<CN_1>1-2
CC_1<CN_1>1<CO_2>1
CC_1<CN_3>1
CC_1<CN_3>2
CC_1<CO_1>1-2
CC_1<CO_1>1<CO_2>1
CC_1<CO_2>1
CC_1<CS_1>1
CC_2
CC_3
CC_5<CC_5>2
CN_1
CN_1<CN_1>1-3
CN_1<CN_1>1-3<CO_2>1
CN_1<CN_1>1-3<NO_5>2
CN_1<CO_2>1
CN_1<NN_2>1
CN_1<NO_5>2
CN_3
CO_1
CO_1<CO_1>1
CO_1<CO_1>1<CO_2>1
CO_1<CO_2>1
CO_1<NO_1>1
CO_2
CO_2<CN_1>1
CO_2<CO_1>1
CS_1
NN_2
NO_1<CO_1>1<NO_5>2
NO_5<CN_1>1<NO_5>1
NO_5<NO_1>1<NO_5>1

Distinguishing these chemical environments for specific $B(\pi)$ and the formerly mentioned 11 nonbonded steric interaction types and varying a bond order value above which A3 sums positive $B(\pi)$ the parameters of Table 7.3 were obtained.

Table 7.3

The A3 algorithm with different bond order switch values, where: TDAr sp^2sp^3 sums positive B(π) for bond orders above one; TDAr sums positive B(π) for bond orders 1.5 and above; TD sums B(π) for triple and double bonds; T sums positive B(π) in triple bonds.

345 molecules. All units in kJ mol^{-1} unless specified parenthetically	**A3 Bond TDAr sp^2sp^3**	**A3 Bond TDAr**	**A3 Bond TD**	**A3 Bond T**	**A3 Bond T adjusted**
Number of parameters	**70**	**70**	**70**	**70**	**70**
HH H21-H6 (hardness)	6	12	12	12	12
HH H6-H6 (hardness)	6	6	6	6	6
HN H6-N61 (hardness)	12	6	6	6	6
HN H6-N8 (hardness)	6	6	6	6	6
HN H6-N9 (hardness)	12	12	12	6	6
HO H21-O7 (hardness)	6	6	6	6	6
HO H6-O6 (hardness)	12	12	12	12	12
HO H6-O7 (hardness)	6	12	12	6	6
NN N61-N61 (hardness)	6	6	6	6	6
NO N8-O7 (hardness)	6	6	12	6	6
OO O7-O7 (hardness)	6	6	12	12	12
E(CC_2)	710.6	609.0	607.9	603.4	603.3
E(CC_3)	831.4	844.4	843.5	845.5	845.4
E(CC_5)	509.6	509.8	508.5	508.1	508.1
E(CN_1)	318.4	318.0	319.2	301.8	301.7
E(CN_3)	894.9	875.4	874.7	876.7	876.7
E(CO_1)	393.3	369.2	369.0	348.4	348.3
E(CO_2)	732.6	730.8	731.4	787.1	787.4
E(CS_1)	285.1	282.1	282.0	279.6	279.5
E(NN_2)	443.4	443.2	441.5	471.4	471.5
E(NO_1)	161.2	197.1	233.3	237.6	237.4
E(NO_5)	417.1	418.2	422.2	431.6	431.8
CC_1	10.1	28.3	27.1	25.8	25.5
CC_1<CC_2>1	12.2	7.1	6.3	1.4	1.3
CC_1<CC_2>1<CN_3>1	20.7	-212.2	-211.9	-212.6	-212.6

345 molecules. All units in kJ mol^{-1} unless specified parenthetically	**A3 Bond TDAr sp^{2}sp^3**	**A3 Bond TDAr**	**A3 Bond TD**	**A3 Bond T**	**A3 Bond T adjusted**
Number of parameters	**70**	**70**	**70**	**70**	**70**
CC_1<CC_2>1<CO_1…*	1168.0	-180.4	-182.9	-270.7	-270.9
CC_1<CC_5>2	-1.2	10.5	11.9	11.0	10.8
CC_1<CN_1>1-2	-4.3	14.4	12.0	11.5	11.3
CC_1<CN_1>1<CO_2>1	-12.6	51.3	45.6	65.8	65.7
CC_1<CN_3>1	1.0	-103.8	-102.3	-95.8	-95.4
CC_1<CN_3>2	75.6	-147.5	-146.5	-140.8	-140.5
CC_1<CO_1>1-2	-8.4	17.4	16.6	2.0	1.8
CC_1<CO_1>1<CO_2>1	-130.7	-28.9	-27.4	33.3	33.1
CC_1<CO_2>1	-26.6	46.2	45.4	-24.7	-25.0
CC_1<CS_1>1	0.9	13.2	12.5	9.1	8.8
CC_2	-471.3	-26.9	-21.8	21.6	21.1
CC_3	-31.0	-51.4	-49.9	-55.3	-55.3
CC_5<CC_5>2	-14.4	-22.4	10.5	6.9	6.7
CN_1	164.7	162.4	171.4	81.2	80.7
CN_1<CN_1>1-3	127.3	126.4	132.6	38.2	37.4
CN_1<CN_1>1-3<CO_2>1	119.6	-93.1	49.1	-125.8	-124.7
CN_1<CN_1>1-3<NO_5>2	-21.3	-1.2	49.9	-40.8	-40.9
CN_1<CO_2>1	92.2	145.2	195.0	152.0	151.8
CN_1<NN_2>1	0.3	-27.7	-22.7	13.9	13.5
CN_1<NO_5>2	7.5	-5.8	23.1	4.5	4.5
CN_3	-54.1	-31.0	-29.5	-32.5	-32.4
CO_1	254.6	124.6	126.0	45.2	45.2
CO_1<CO_1>1	243.5	102.5	101.7	12.3	11.9
CO_1<CO_1>1<CO_2>1	160.5	-535.6	-533.1	-216.9	-216.3
CO_1<CO_2>1	20.9	111.6	115.8	95.6	95.3
CO_1<NO_1>1	-248.9	-308.6	140.5	53.8	53.4
CO_2	-76.3	-48.7	-49.7	1.1×10^3	1.1×10^3
CO_2<CN_1>1	50.4	38.4	45.0	4.6×10^3	4.6×10^3
CO_2<CO_1>1	36.4	21.0	20.8	-14.3	-11.0

345 molecules. All units in kJ mol^{-1} unless specified parenthetically	A3 Bond TDAr sp^{2}sp^3	A3 Bond TDAr	A3 Bond TD	A3 Bond T	A3 Bond T adjusted
Number of parameters	**70**	**70**	**70**	**70**	**70**
CS_1	138.2	97.8	99.1	48.7	46.3
NN_2	-39.0	-39.6	-39.3	116.3	116.1
NO_1<CO_1>1<NO_5>2	89.9	-14.1	10.9	7.7	7.8
NO_5<CN_1>1<NO_5>1	-12.7	-18.5	69.4	96.5	96.6
NO_5<NO_1>1<NO_5>1	-36.0	-47.5	-706.5	-265.2	-265.0
HH H21-H6 (Å$^{1/hardness}$)	-756.0	-1.8×10^{5}	-1.7×10^{5}	-1.6×10^{5}	-1.6×10^{5}
HH H6-H6 (Å$^{1/hardness}$)	-688.6	-551.9	-561.2	-553.2	-554.6
HN H6-N61 (Å$^{1/hardness}$)	2.5×10^{6}	-1.5×10^{3}	-1.4×10^{3}	-1.3×10^{3}	-1.3×10^{3}
HN H6-N8 (Å$^{1/hardness}$)	241.7	149.7	67.2	354.4	355.9
HN H6-N9 (Å$^{1/hardness}$)	-1.7×10^{5}	-2.0×10^{5}	-2.0×10^{5}	223.1	223.1
HO H21-O7 (Å$^{1/hardness}$)	3.2×10^{3}	1.0×10^{4}	1.0×10^{4}	4.7×10^{3}	4.7×10^{3}
HO H6-O6 (Å$^{1/hardness}$)	3.8×10^{5}	2.7×10^{5}	2.8×10^{5}	2.5×10^{5}	2.5×10^{5}
HO H6-O7 (Å$^{1/hardness}$)	1.3×10^{3}	4.0×10^{5}	3.9×10^{5}	1.3×10^{3}	1.3×10^{3}
NN N61-N61 (Å$^{1/hardness}$)	-9.8×10^{4}	-9.9×10^{4}	-9.8×10^{4}	-9.7×10^{4}	-9.7×10^{4}
NO N8-O7 (Å$^{1/hardness}$)	9.5×10^{3}	8.5×10^{3}	-2.3×10^{6}	-6.3×10^{3}	-6.3×10^{3}
OO O7-O7 (Å$^{1/hardness}$)	-2.3×10^{4}	-2.2×10^{4}	2.3×10^{6}	3.6×10^{6}	3.6×10^{6}
Standard deviation	8.876	8.173	8.200	6.348	6.331

* CC_1<CC_2>1<CO_1>1<CO_2>1

Without compromising our constant transferable bond energies, the necessary parameters required to achieve a standard deviation comparable or better than Pedley's value is achieved at the expense of the alkanes.

Empirical schemes are parameterised such that parameters for one group of molecules are not achieved at the expense of other groups, by allowing a sort of regression constant for each grouping of molecules. This is perhaps most obvious in the Allen scheme[4], where a parameter for two adjacent C-C bonds (Γ_{CCC}) is altered between some groups. This scheme parameterised by Skinner[10] uses Γ_{CCC} = 11.21 kJ mol^{-1} for alkanes but the value is altered to 10.7 kJ mol^{-1} for alcohols. The equivalence of the empirical schemes has been shown by Cox and Pilcher[6]. By not incorporating a "pseudo regression constant" for each grouping of molecules we attempt to describe such chemical phenomena by the success of our scheme. In our method so far we have overlooked adjustments for zero-point energy (ZPE) and population analysis of the conformers. Both corrections are computationally expensive.

Unfortunately the methods of the present work give large errors for small alkanes e.g. methane, ethane and propane. (*ca.* 7-8 kJ mol^{-1}).

Parameterised over the test group of 345 molecules Table 7.4 shows the experimental minus calculated values of the alkanes for all structural isomers up to C6 as well as heptane and octane. The focus for the remainder of this section is this set of molecules, which will be referred to as the small alkane test group. These are listed in Table 7.4 as a sample of the test set of 340 molecules.

Table 7.4

Results for the small alkane test group, with columns corresponding to the columns of Table 7.2. s(15) and s(346), are the standard deviations for the small alkane test group and the 346 test group respectively.

Calculated error in kJ mol^{-1}	**A1 bond a**	**A1 ¨bond b**	**A1 bond c**	**A1 bond d***	**A3 bond e***
Parameter numbers	**68**	**67**	**54**	**58**	**57**
methane	1.2	0.0	1.0	0.0	0.0
ethane	7.0	4.9	6.4	7.1	7.1
propane	8.0	4.9	7.2	7.5	7.5
butane	8.0	4.2	7.1	7.6	7.6
2-methylpropane	6.8	2.3	5.8	4.7	4.7
pentane	7.2	4.2	6.5	7.8	7.9
2-methylbutane	3.4	-0.3	2.7	2.1	2.2
2,2-dimethylpropane	6.7	-0.1	5.2	3.0	3.0
hexane	8.3	4.0	7.3	8.7	8.8
2-methylpentane	2.7	-0.8	2.0	1.9	2.0
3-methylpentane	-1.0	-2.2	-1.1	-0.3	-0.2

Calculated error in kJ mol^{-1}	A1 bond a	A1 ¨bond b	A1 bond c	A1 bond d*	A3 bond e*
Parameter numbers	**68**	**67**	**54**	**58**	**57**
2,2-dimethylbutane	0.0	-4.5	-0.8	-2.4	-2.3
2,3-dimethylbutane	-2.3	-5.2	-2.7	-3.9	-3.8
heptane	8.1	3.8	7.1	8.9	9.0
octane	7.6	3.0	6.5	9.0	9.1
s(15)	6.2	3.6	5.4	6.1	6.1
s(346)	7.6	9.0	8.0	7.7*	8.1*

* values and standard deviations are for the 345 test group.

Table 7.5 shows the difference between the experimental and calculated $\Delta_f H^\circ$ values as calculated by MM4[36] and Pedley[15], as well as our estimate for the number of parameters in the respective methods, which is probably conservative for MM4.

Table 7.5

Experimental minus calculated $\Delta_f H^\circ$ for MM4[36] and Pedley's group method[15].

Calculated error in kJ mol^{-1}	MM4	Pedley
Number of parameters	**68**	**107**
methane	0.0	0.0
ethane	-2.1	0.0
propane	0.7	0.0
butane	-0.8	-0.2
2-methylpropane	0.9	-0.1
pentane	1.3	-0.7
2-methylbutane	-1.0	-0.7
2,2-dimethylpropane	1.7	0.0
hexane	0.7	0.0
2-methylpentane		-0.9
3-methylpentane		-0.3

Calculated error in kJ mol^{-1}	MM4	Pedley
Number of parameters	**68**	**107**
2,2-dimethylbutane		-2.1
2,3-dimethylbutane	-1.4	-0.7
heptane	1.1	0.2
octane	1.6	0.0
Standard deviation	1.3	0.7

7.1 Testing Zero-Point Energy and Adjustment for Population Analysis

The more accurate schemes must incorporate zero-point and thermal energies, and values must be fitted to the conformer's energy relative to the experimental $\Delta_f H^o$ which is not necessarily the value of only one conformer, but takes into account conformational mixing (see section 6). Considering *n*-alkanes the value of these computationally expensive factors is examined in combination with B(π) and steric functions.

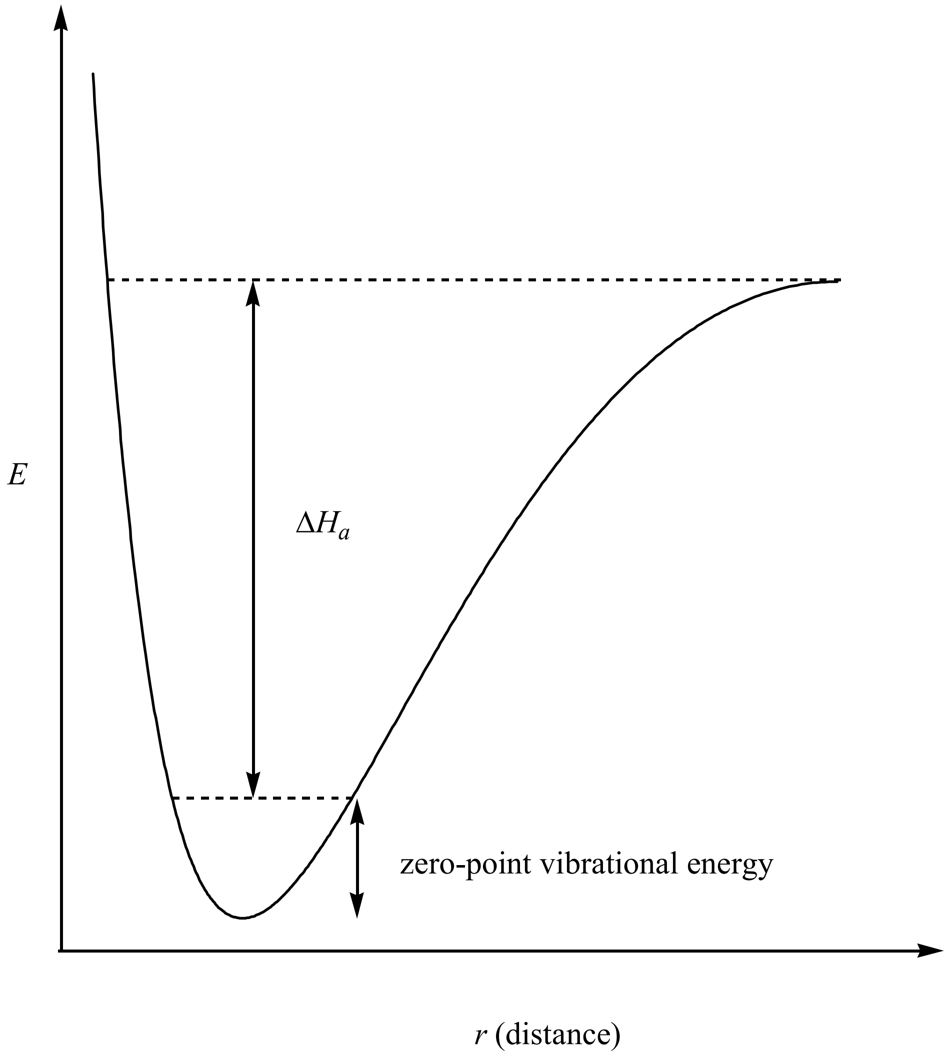
7.1.1 Zero-Point Vibrational Energy

The zero-point vibrational energy [$H_{vib}(0)$] is calculated at the HF 6-31G(*d*) and B3LYP 6-31G(*2df,p*) level in G3 and G4 calculations respectively. However to accommodate a test group containing larger molecules we obtained vibrational information at the HF 3-21G(*) level. Even so at this level we had to exclude some of the larger molecules formerly tested in the test group of 346, reducing the test group to 340 (1,2,3-propanetriol trinitrate was eliminated for reasons mentioned previously). These molecules were 1-hexadecanol, decylbenzene, 1,2,4,5-tetramethylbenzene, pentamethylbenzene and hexamethylbenzene.

The atomisation energy (ΔH_a) for a diatomic molecule is the bond energy. As can be seen in Figure 7.1 bond energies based on experimental ΔH_a do not include $H_{vib}(0)$, *i.e.* bond energies are not arrived at based on an energy at the bottom of the potential energy well. Therefore a bond energy to be used with an explicit calculation of $H_{vib}(0)$ needs to be commensurately larger to calculate an energy at the bottom of a potential energy well.

Figure 7.1

Morse curve for interatomic distance (*r*) with energy (*E*) for a diatomic molecule.



The $\Delta_f H^o$ was calculated with the explicit calculation of $H_{vib}(0)$ at the HF 3-21G(*) level {H_{vib}[0, 3-21G(*)]} with the scale factor 0.9207 recommended for H_{vib}[0, 3-21G(*)][113, 132], and with E(C-H), E(O-H), E(N-H) and E(S-H) values adjusted to the values shown in Table 7.6, by dividing the scaled $H_{vib}(0)$ for methane, water, ammonia and hydrogen sulfide by the number of C-H, O-H, N-H and S-H bonds respectively, and adding these values to the respective bond energy given in Table 7.1. The E(C-C) value was treated slightly differently. Fujimoto and Shingu's[133] additivity scheme for alkanes is given by *eq.* (39).

$$H_{vib}(0) = 30.46n_1 + 13.14n_2 - 3.35n_3 \qquad \ldots (39)$$

where n_1, n_3 and n_2 are respectively the number of C-H bonds, C-C bonds and chain ends, and $H_{vib}(0)$ is for the gaseous paraffin in kJ mol^{-1}. This is a Cottrell[1] constant increment per methylene group as fitted to a scheme that includes corrections for branching. Attributing two C-C bonds per carbon atom in diamond, we have $n_1 = n_2 = 0$ and $n_3 = 2$. Accordingly 13.14 kJ mol^{-1} is added per E(C-C). In this way these bond energies are fixed for use in a model that explicitly calculates $H_{vib}(0)$; LSEs are obtained for the remaining bond energies. This method having some bond energy terms fixed is compared with a similar model where H_{vib}[0, 3-21G(*)] is added to ΔH_a and LSEs are ascertained for all bond energies. These results fitting with $H_{vib}(0)$ for the small alkanes test group are shown in columns a and b of Table 7.9. In both methods B(π) was ascertained with the A3 bond algorithm switching for triple bonds only, and hardness parameters were restricted to 6 or 12 for the SSR.

In the same way that bond energies were adjusted for use with the explicit calculation of $H_{vib}(0)$, adjustments were made to bond energies for use with the explicit calculation of vibrational and thermal energies [$\Delta H_{VT} = H_{vib}$[0, 3-21G(*)] - $\Delta H(T)$]. An experimental ΔH_{VT} can be calculated from the fundamental frequencies using *eq.* (10) modified by subtracting the second parenthetical term from the first [the first term determines $H_{vib}(0)$ and the second $H_{vib}(298)$] and subtracting $4RT$ (for nonlinear molecules). The ΔH_{VT} calculated at the HF 3-21G(*) level [ΔH_{VT} 3-21G(*)], H_{vib}[0, 3-21G(*)] and H_{vib}[298, 3-21G(*)] energies are scaled by 0.9, the scaling factor recommended for vibrational frequencies by Hehre *et al.*[71]. The experimental ΔH_{VT}, ΔH_{VT} 3-21G(*) and scaled ΔH_{VT} 3-21G(*) are shown in Table 7.8.

The two regression analyses of the previous paragraph were repeated substituting ΔH_{VT} for $H_{vib}(0)$. Also no fixed value for E(C-C) is ascertained. If ΔH_{VT} is calculated at the HF 3-21G(*) level and multiplied by a factor of 0.9, then 1 kJ mol^{-1} must be subtracted. The subtraction of 1 kJ mol^{-1} is to compensate for scaling the translational and rotational energies and ideal gas value that are part of the ΔH_{VT} that do not require scaling. i.e. $(1 - 0.9)\ (2.4789 + 2 \times 3.7184)$. For linear molecules, of which there are three in the 340 test set (ethyne, hydrogen cyanide and ethane dinitrile), the value that must be added is 0.87 kJ mol^{-1} (1 –

0.9) (2 × 2.4789 + 3.7184). However as this is a small difference we have not distinguished linear from nonlinear molecules in this instance. Fitting a scaled ΔH_{VT} 3-21G(*) using the fixed bond energies of Table 7.6 for use with ΔH_{VT}, gives the results for the small alkanes test group in column c of Table 7.9. Fitting ΔH_a plus ΔH_{VT} without any scaling, and ascertaining LSEs for all bond energies to calculate $\Delta_f H^o$ gives the results of column d for the small alkanes test group in Table 7.9. In both methods B(π) was ascertained with the A3 bond algorithm switching for triple bonds only, and hardness parameters were restricted to 6 or 12 for the SSR.

Table 7.6

Bond energy (kJ mol^{-1})	**Use with ZPE correction**	**Use with ΔH_{VT} correction**
E(C-C)	370.54	
E(C-H)	444.8699	441.7268
E(O-H)	489.8154	484.2600
E(N-H)	434.4290	428.4109
E(S-H)	386.5497	381.1327

The experimental $H_{vib}(0)$ can be calculated from the fundamental frequencies from the JANAF tables[110] where $H_{vib}(0) = 50hc\Sigma\overline{\nu}_i$ (if $\overline{\nu}_i$ is in cm^{-1}). The experimental $H_{vib}(0)$ is compared with $H_{vib}(0)$ calculated at the HF 3-21G(*) level (H_{vib}[0, 3-21G(*)]) as well as the scaled calculated $H_{vib}(0)$.

Table 7.7

Experimental and calculated zero-point energies. Experimental $H_{vib}(0)$ are determined from fundamental frequencies from the JANAF tables.

Energies are in kJ mol^{-1}	**$H_{vib}(0)$ experimental**	**$H_{vib}[0, 3\text{-}21G(*)]$**	**0.9207 × $H_{vib}[0, 3\text{-}21G(*)]$**
CH_4	113.4	126.0	116.0
H_2O	53.8	57.2	52.6
NH_3	90.1	94.6	87.1
H_2S	38.4	43.0	39.6

Table 7.8

Experimental and calculated total vibrational and thermal energies (ΔH_{VT}). Experimental ΔH_{VT} are determined from fundamental frequencies from the JANAF tables.

Energies are in kJ mol^{-1}	**ΔH_{VT} experimental**	**ΔH_{VT} 3-21G(*)**	**0.9ΔH_{VT} 3-21G(*) - 1**
CH_4	103.4	116.0	103.4
H_2O	43.9	47.2	41.5
NH_3	80.1	84.6	75.1
H_2S	28.5	33.1	28.8

Table 7.9

Calculations for the small alkanes test group using parameters ascertained over the 340 test group, with bondingness ascertained using the A3 algorithm changing only for triple bonds and hardness parameters restricted to 6 or 12.

All values in kJ mol^{-1}	**Exper-imental error**	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**	**$\Delta\Delta_f H^o$ d**
parameters		**70**	**75**	**71**	**75**
methane	0.4	0.0	1.1	0.0	4.0
ethane	0.3	4.6	5.4	3.2	6.1
propane	0.5	4.6	5.3	2.8	5.3
butane	0.6	4.9	5.6	3.3	5.4
2-methylpropane	0.6	1.9	2.6	0.0	2.3
pentane	0.8	5.6	6.2	4.2	6.1
2-methylbutane	0.9	0.5	1.1	-0.9	0.9
2,2-dimethylpropane	0.8	-0.3	0.4	-2.4	0.1
hexane	0.8	6.8	7.5	5.8	7.3
2-methylpentane	0.9	0.8	1.4	-0.3	1.2
3-methylpentane	0.9	-0.2	0.2	-1.1	0.2
2,2-dimethylbutane	0.9	-3.5	-2.9	-5.0	-3.1
2,3-dimethylbutane	0.9	-3.8	-3.3	-5.0	-3.6
heptane	1.3	7.5	8.0	6.9	8.1

All values in kJ mol^{-1}	Experimental error	$\Delta\Delta_f H^o$ a	$\Delta\Delta_f H^o$ b	$\Delta\Delta_f H^o$ c	$\Delta\Delta_f H^o$ d
parameters		**70**	**75**	**71**	**75**
octane	1.3	8.0	8.5	7.8	8.7
s(15)		4.6	5.0	4.2	5.2
s(340)		6.5	6.4	6.7	6.6

It can be seen in Table 7.9 that the largest deviations in calculated $\Delta_f H^o$ from experimental $\Delta_f H^o$ are the *n*-alkanes. These also have the greatest number of rotamers. The $\Delta_f H^o$ of the *n*-alkanes adjusted by energy corrections based on population of high energy rotamers is examined in section 7.1.2.

7.1.2 Adjustment from Population Analysis

Fitting *n*-alkane $\Delta_f H^o$ experimental or $\Delta_f H^o$ adjusted values of Table 6.11 against B(π) and SSR allowing hardness parameters 2 – 24, the results of Table 7.10 are obtained.

Table 7.10

Allowing hardness parameters to vary from 2 to 24 in SSR the labels a,b,c,d and e have the following meanings: ‘a’ fitted to experimental $\Delta_f H^o$ values with conformational adjustment, with LSEs found for B(π) and SSR; ‘b’ fitted to experimental $\Delta_f H^o$ values without conformational adjustment, with LSEs found for B(π) and SSR; ‘c’ fitted to experimental $\Delta_f H^o$ values adjusted with scaled H_{vib}[0, 3-21G(*)], while E(C-C) and E(C-H) are fixed at 357.4 and 386.8701 kJ mol^{-1} respectively, with LSEs found for B(π) and SSR; ‘d’ fitted to experimental $\Delta_f H^o$ values adjusted with scaled H_{vib}[0, 3-21G(*)], while E(C-H) is fixed at 386.8701 kJ mol^{-1}, with LSEs found for B(π), E(C-C) and SSR; ‘e’ fitted to experimental $\Delta_f H^o$ values adjusted with scaled H_{vib}[0, 3-21G(*)], while LSEs are found for B(π), E(C-C), E(C-H) and SSR. The experimental error is given in the ± column.

	±	**$\Delta\Delta_f H^o$ a**	**$\Delta\Delta_f H^o$ b**	**$\Delta\Delta_f H^o$ c**	**$\Delta\Delta_f H^o$ d**	**$\Delta\Delta_f H^o$ e**
methane	0.4	0.01	0.01	0.010	0.01	0
ethane	0.3	-0.281	-0.267	0.417	0.164	0.164
propane	0.5	0.947	0.466	0.085	-0.049	-0.049
butane	0.6	-0.769	-0.603	-0.575	-0.361	-0.361
pentane	0.8	0.685	0.851	0.214	-0.025	-0.026
hexane	0.8	-0.852	-0.524	-0.082	0.306	0.305
heptane	1.3	0.77	0.737	0.424	0.175	0.176
octane	1.3	-0.333	-0.564	-0.276	-0.209	-0.208
hardness		2	2	12	5	5
E(C-C)		357.4*	357.4*	370.54*	385.5	385.5
E(C-H)		415.87*	415.87*	444.9*	444.9*	444.9
CC_1		117.9	113.5	59.9	24.4	24.4
A†		2.233	1.343	-55121.8	-426.7	-426.0
Standard deviation		0.704	0.599	0.342	0.217	0.216

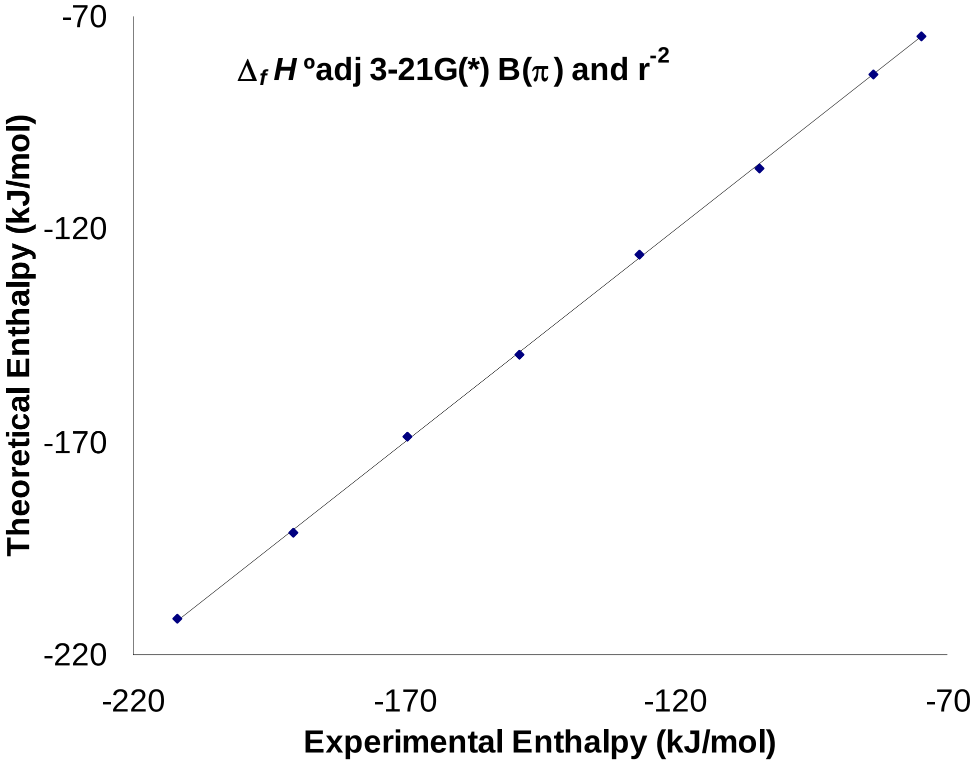
* Not a LSE.

† see equation (36) section 4.1.2.

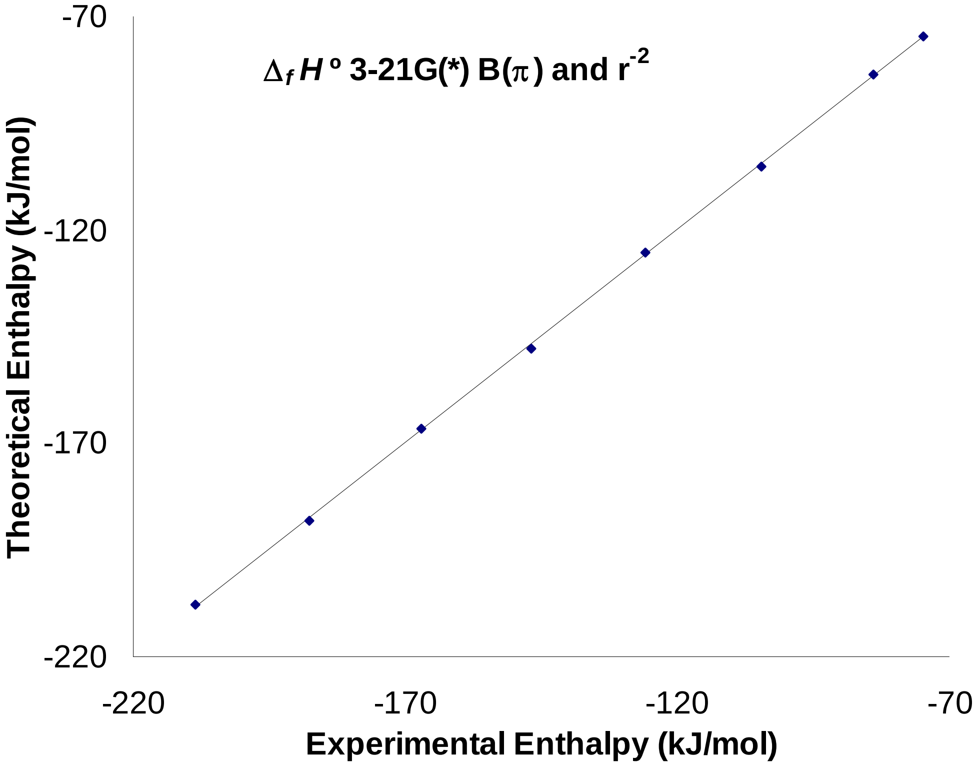
Columns a and b of Table 7.10 are shown in Graphs 7.1 and 7.2. In both cases, whether $\Delta_f H^o$ values are adjusted or not, there is an alternation above and below the 1:1 line, or put another way, having positive residuals for an odd number of carbons and negative for an even carbon number. This would seem to be an anisotropic problem. However before an anisotropic method is to be applied, a few simpler methods should be investigated, namely how the buffered 14 7 function performs, how a Lennard-Jones potential might perform and what $H_{vib}(0)$ adjustment achieves.

The $H_{vib}(0)$ of column c uses the scaled ZPE of methane to adjust E(C-H) to the value shown in Table 7.6. Because there are only eight observables in the form of experimental $\Delta_f H^o$, the degrees of freedom are quickly consumed, so regression analysis was performed using the E(C-C) value also from Table 7.6.

Graph 7.1



Graph 7.2



The $H_{vib}(0)$ adjustment seems to account for the alternation in the *n*-alkanes with just the SSR. However this is still uncertain with such a small test group, and the small gain in standard deviation does not recommend such a labour intensive calculation. The conformational adjustment gave no improvement.

7.1.3 Lennard-Jones Potential

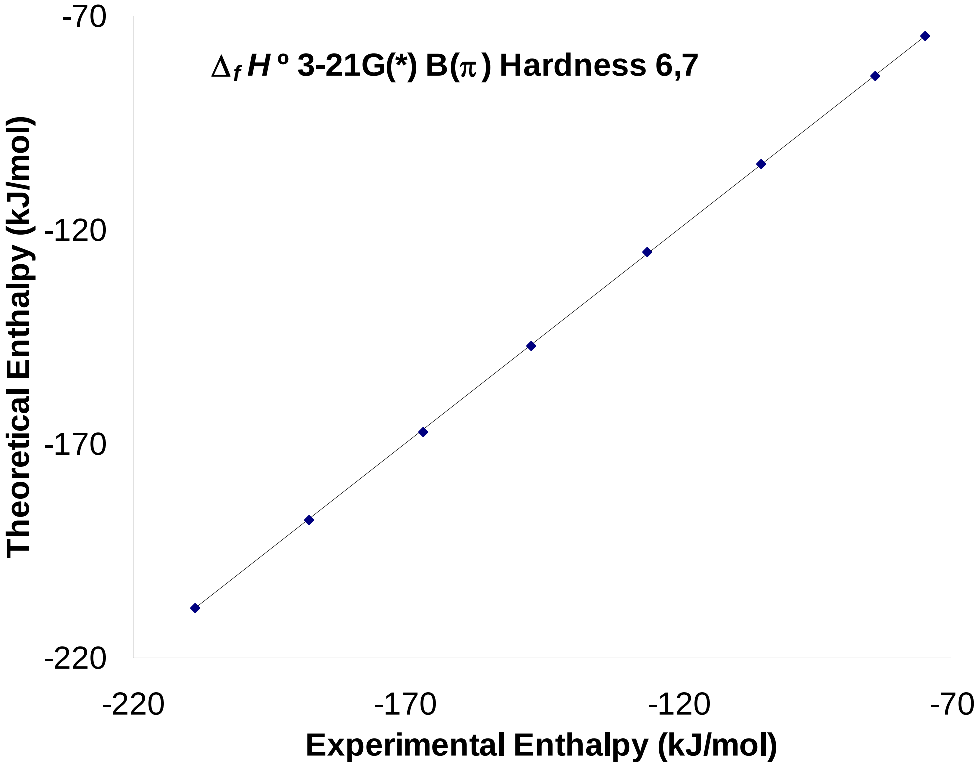
So far we have been neglecting the dispersive term of the Lennard-Jones potential [see *eqs.* (35), (40) or (41)]. The regression analysis is performed in the following way. Two terms for every type of steric interaction are subject to linear least squares analysis to ascertain LSEs for Σr^{-n} and Σr^{-m}, where $m < n$. Every allowable m for every n over the *n*-alkanes (only one type of steric interaction) where m and n vary between 2 and 24 is tested. Strictly speaking $m = 6$ for a dispersive term. We have allowed m for this term to vary, so the term is called the attractive term henceforth. The analysis was repeated for the *n*-alkanes with conformational adjustment. The results are shown in Table 7.11.

Table 7.11

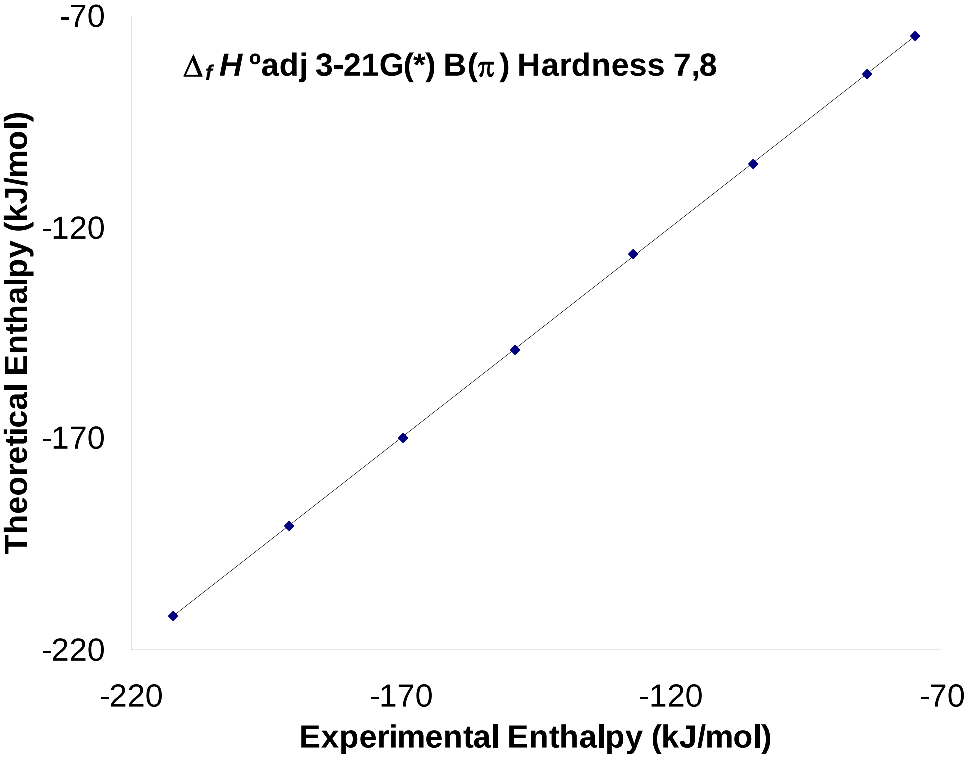
The labels a and b have the following meanings: ‘a’ no conformational adjustment and ‘b’ with conformational adjustment.

All values in kJ mol^{-1}	Exper-imental error	$\Delta\Delta_f H^o$ a	$\Delta\Delta_f H^o$ b
methane	0.4	0.01	0.01
ethane	0.3	0.202	0.145
propane	0.5	-0.013	0.221
butane	0.6	-0.477	-0.6
pentane	0.8	0.096	0.026
hexane	0.8	0.276	0.238
heptane	1.3	0.203	0.124
octane	1.3	-0.248	-0.119
Standard deviation		0.255	0.272

Graph 7.3



Graph 7.4



If the Lennard-Jones potential is written:

$$V_{vdW} = \frac{n\varepsilon}{n-m}\left[\frac{m}{n}\left(\frac{r_o}{r}\right)^n - \left(\frac{r_o}{r}\right)^m\right] \qquad \ldots (40)$$

where ε is the well depth, r_o is interaction distance minimising V_{vdW}, m and n are integers where $m < n$ and are the hardness parameters for the attractive and repulsive terms respectively, and r is the interatomic distance. We parameterise V_{vdW} with the form shown in *eq.* (41).

$$S_{ij} = \sum_r Ar^{-n} - Br^{-m} \qquad \ldots (41)$$

The A and B variables of *eq.* (41) can be converted to ε and r_o of *eq.* (40) in the following way.

$$V_{dvW} = Ar^{-n} - Br^{-m}$$

$$\frac{dV_{vdW}}{dr} = -nAr^{-n-1} + mBr^{-m-1}$$

$$\text{When } \frac{dV_{vdW}}{dr} = 0: \quad mBr^{-1-m} = nAr^{1-n} \text{ and } r = r_o$$

$$\frac{nA}{mB} = \frac{r^{1-m}}{r^{1-n}}$$

$$ln\frac{nA}{mB} = (1\text{-}m)\ ln\ r - (1\text{-}n) ln\ r$$

$$= n\text{-}m\ ln\ r$$

$$\frac{nA}{mB} = r^{n\text{-}m}$$

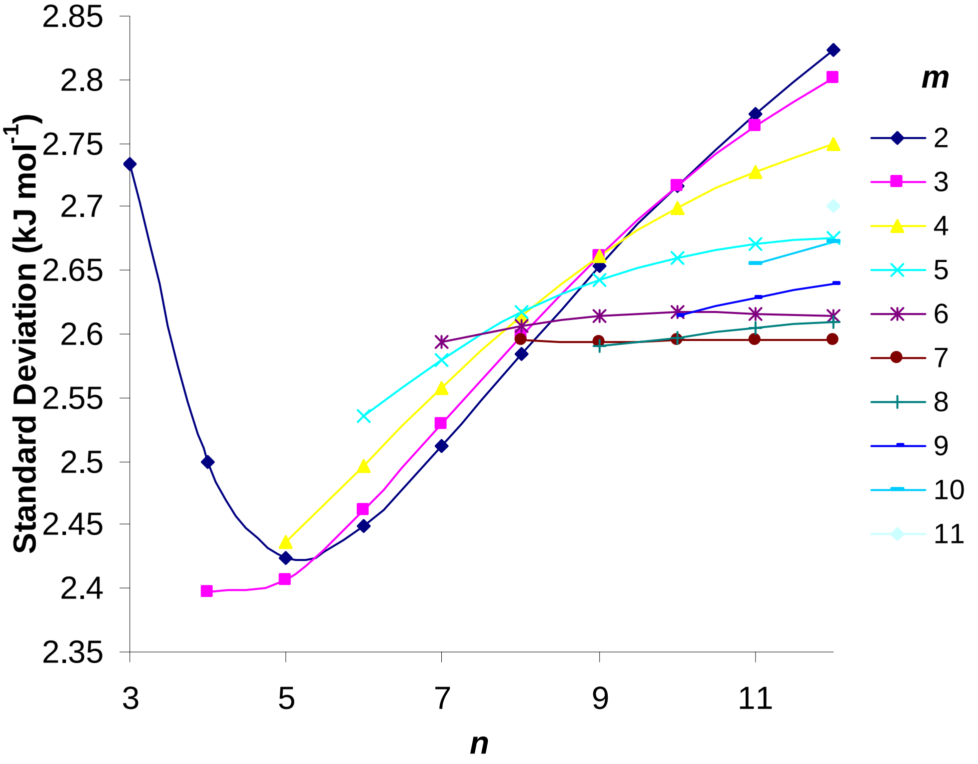
Therefore:

$$r_o = \left(\frac{nA}{mB}\right)^{\frac{1}{n-m}} \qquad \ldots (42)$$

Substituting r_o for r into *eq.* (41) gives ε.

The test group was expanded to include alkanes and alkenes, making a test group of 80. This introduced another steric interaction. Allowing m to vary from 2 to 11 and n from 3 to 12, this required 55 combinations. If the standard deviation is graphed against n for every m, as seen in Graph 7.5, for all instances of m, the value of n that minimises the standard deviation is $n = m + 1$, except for $m = 2$ and $m = 7$.

Graph 7.5



If $m = n - 1$ *eqs.* (40), (41) and (42) become:

$$V_{vdW} = n\varepsilon\left[\frac{n-1}{n}\left(\frac{r_o}{r}\right)^{n} - \left(\frac{r_o}{r}\right)^{1-n}\right] \qquad \ldots (43)$$

$$S_{ij} = \sum_r Ar^{-n} - Br^{1-n} \qquad \ldots (44)$$

$$r_o = \frac{nA}{(n-1)B} \qquad \ldots (45)$$

By the algorithm that distinguishes chemical environment for steric interactions, it is only the test group of alkanes and alkenes that has only two distinguishable chemical environments for steric interactions. The test group of 25 primary, secondary and tertiary alcohols (see alcohols in Appendix A.3.1 excluding 1-hexadecanol) have three distinct chemical environments by this algorithm: $O\cdots H_C$ $H_C\cdots H_C$ and $H_O\cdots H_C$ (H_C: H attached to C and H_O: H attached to O). Allowing *m* to vary from 3 to 11, and *n* from 4 to 12, requires 91,125 combinations to be computed, which is not so easily graphed. However the *m* and *n* for each steric interaction of a chemical environment that minimised the standard deviation of the alcohols are shown in Table 7.12, these being 3, 4; 3, 4 and 11, 12.

Table 7.12

Alcohols. All units in kJ mol^{-1} unless specified parenthetically	**A1 bond parameters**
No. parameters	**17**
m HH H21-H6 (hardness)	11
n HH H21-H6 (hardness)	12
m HH H6-H6 (hardness)	3
n HH H6-H6 (hardness)	4
m HO H6-O6 (hardness)	3
n HO H6-O6 (hardness)	4
E(CO_1)	371.6
electrostatic	7.7
CC_1	40.9

Alcohols. All units in kJ mol^{-1} unless specified parenthetically	**A1 bond parameters**
No. parameters	**17**
CC_1<CO_1>1-2	-27.3
CO_1	-3.8
B HH H21-H6 (Å$^{1/hardness}$)	-2255995.7
A HH H21-H6 (Å$^{1/hardness}$)	4771523.3
B HH H6-H6 (Å$^{1/hardness}$)	188.9
A HH H6-H6 (Å$^{1/hardness}$)	-638.7
B HO H6-O6 (Å$^{1/hardness}$)	1395.9
A HO H6-O6 (Å$^{1/hardness}$)	-3680.3
Standard deviation kJ mol^{-1}	1.179

It would seem that the approximation $m = n - 1$ is a global minimum for the test group of alcohols when an electrostatic function and B(π) are parameterised with the steric function.

Over the test group of 345 molecules there are 11 steric environments. If we allow n to vary from 4 to 12, this would require 9^{11} comparisons. So the approximation still does not allow us to ascertain the most transferable constant bond energy parameters over our largest test group. However it does allow a larger test group for which we can ascertain a global minimum. To ascertain a global minimum with 6 steric environments is 9^6 comparisons, which might take a day or two and 7 steric environments (9^7) might take about a month.

Using the commonly used hardness parameters where $m = 6$ and $n = 12$ of *eq.* (41) the parameters of Table 7.13 are found for the test group of 345 molecules.

Table 7.13

345 molecules. All units in kJ mol^{-1} unless specified parenthetically	**A1 bond**
parameters	**70**
E(CC_2)	604.5
E(CC_3)	818.0
E(CC_5)	508.2
E(CN_1)	303.6
E(CN_3)	854.4
E(CO_1)	348.5
E(CO_2)	787.0
E(CS_1)	280.6
E(NN_2)	-20.4
E(NO_1)	245.0
E(NO_5)	427.2
CC_1	23.8
CC_1<CC_2>1	-1.6
CC_1<CC_2>1<CN_3>1	-257.6
CC_1<CC_2>1<CO_...*	-216.6
CC_1<CC_5>2	9.5
CC_1<CN_1>1-2	11.6
CC_1<CN_1>1<CO_2>1	88.8
CC_1<CN_3>1	-157.6
CC_1<CN_3>2	1171.9
CC_1<CO_1>1-2	-12.3
CC_1<CO_1>1<CO_2>1	30.8
CC_1<CO_2>1	-22.7
CC_1<CS_1>1	7.7
CC_2	37.3
CC_3	-2033.3
CC_5<CC_5>2	7.0
CN_1	77.6
CN_1<CN_1>1-3	45.3

345 molecules. All units in kJ mol^{-1} unless specified parenthetically	A1 bond
parameters	**70**
CN_1<CN_1>1-3<CO_2>1	112.9
CN_1<CN_1>1-3<NO_5>2	-493.9
CN_1<CO_2>1	501.1
CN_1<NN_2>1	-2.4
CN_1<NO_5>2	1.6
CN_3	31199.3
CO_1	39.8
CO_1<CO_1>1	12.7
CO_1<CO_1>1<CO_2>1	-213.4
CO_1<CO_2>1	111.0
CO_1<NO_1>1	88.8
CO_2	1056.5
CO_2<CN_1>1	2030.8
CO_2<CO_1>1	-33.8
CS_1	58.1
NN_2	-2010.9
NO_1<CO_1>1<NO_5>2	16.7
NO_5<CN_1>1<NO_5>1	63.5
NO_5<NO_1>1<NO_5>1	-427.3
B HH H21-H6 (Å$^{1/6}$)	-97.9
A HH H21-H6 (Å$^{1/12}$)	-167960.3
B HH H6-H6 (Å$^{1/6}$)	-566.7
A HH H6-H6 (Å$^{1/12}$)	-646.9
B HN H6-N61 (Å$^{1/6}$)	-6749.2
A HN H6-N61 (Å$^{1/12}$)	5466088.0
B HN H6-N8 (Å$^{1/6}$)	1364.2
A HN H6-N8 (Å$^{1/12}$)	-392779.2
B HN H6-N9 (Å$^{1/6}$)	2455.1
A HN H6-N9 (Å$^{1/12}$)	-880624.8

345 molecules. All units in kJ mol^{-1} unless specified parenthetically	**A1 bond**
parameters	**70**
B HO H21-O7 (Å$^{1/6}$)	19364.4
A HO H21-O7 (Å$^{1/12}$)	-2474828.1
B HO H6-O6 (Å$^{1/6}$)	-1277.8
A HO H6-O6 (Å$^{1/12}$)	621225.7
B HO H6-O7 (Å$^{1/6}$)	1306.6
A HO H6-O7 (Å$^{1/12}$)	16739.1
B NN N61-N61 (Å$^{1/6}$)	-624137.2
A NN N61-N61 (Å$^{1/12}$)	3069541073.5
B NO N8-O7 (Å$^{1/6}$)	-56869.3
A NO N8-O7 (Å$^{1/12}$)	23495175.3
B OO O7-O7 (Å$^{1/6}$)	-22789.7
A OO O7-O7 (Å$^{1/12}$)	17250495.7
Standard deviation kJ mol^{-1}	5.893

* CC_1<CC_2>1<CO_1>1<CO_2>1

The values for the small alkane test group are:

Table 7.14

All values in kJ mol^{-1}	**Experimental error**	**$\Delta\Delta_f H^o$**
methane	0.4	0.01
ethane	0.3	6.108
propane	0.5	6.55
butane	0.6	5.288
2-methylpropane	0.6	4.065
pentane	0.8	4.949
2-methylbutane	0.9	0.812
2,2-dimethylpropane	0.8	2.042
hexane	0.8	4.726
2-methylpentane	0.9	1.178

3-methylpentane	0.9	-0.516
2,2-dimethylbutane	0.9	-2.194
2,3-dimethylbutane	0.9	-3.072
heptane	1.3	7.659
octane	1.3	7.466
s(15)		4.680
s(345)		5.893

7.2 Bondingness as an anisotropic substitute for a Urey-Bradley force

Some functional groups have a non-spherical arrangement of hydrogen atoms about a heavy atom, for instance amino and hydroxyl groups. These groups have different interactions with neighbouring groups or atoms, depending on which direction the hydrogens are oriented relative to the other group or atom. The phenomenon where the spatial direction of something has a bearing on an outcome is anisotropy. Anisotropy in a functional group is usually accounted for by creating a pseudo atom to represent the lone pair of electrons usually present in such a functional group. The anisotropic problem may to some degree be accounted for by bondingness between nonbonded atoms in a molecule. Urey-Bradley forces[134] are between nearest neighbour atoms i.e. atoms not bonded to each other, but bonded to a common atom. $B(\pi)$ is directional, as is $B(\sigma)$ when p or higher *l* quantum number atomic orbitals are involved. Nonbonded $B(\pi)$ [nb-$B(\pi)$] is only distinguished by the two atoms involved, and no further chemical environment information. It therefore does not accrue parameters as quickly as the algorithm for bonded $B(\pi)$ values. Over the test group of 345 molecules the algorithm distinguishes the following nb-$B(\pi)$: CC, CN, CO, CS, NN, NO and OO. An nb-$B(\pi)$ value is arrived at by summing all $B(\pi)$ over all MOs for common interaction types. This much simpler algorithm works just as well if not better than an A1 or A3 type of algorithm.

7.3 Comparison between 340 and 345 test groups

The main difference between the 345 and 340 test groups, aside from there being 5 less molecules in the 340 test group, is vibrational data *e.g.* $H_{vib}(0)$ can be more easily ascertained with the 340 test group. The desktop computer in our laboratory could not calculate vibrational data for the five extra molecules in the 345 test group not present in the 340 test group. The calculation of $H_{vib}(0)$ is required to reduce the standard deviation over the small alkane test group. Parameterisation of the 340 test group without explicit calculation of $H_{vib}(0)$ gave a s(15) of 5.2 kJ mol^{-1} and a s(340) of 5.3 kJ mol^{-1}.

Table 7.15

All units in kJ mol^{-1} unless specified parenthetically	**A3 Bond 345 molecules**	**A3 Bond 340 molecules**
No. parameters	81	81
s(15)	5.5	3.8
s(345) and s(340)	5.2	5.3
E(CC_1)	360.9	376.4
E(CC_2)	602.9	620.0
E(CC_3)	862.5	882.5
E(CC_5)	509.7	523.9
E(CH_1)	415.9†	444.9†
E(CN_1)	309.6	325.7
E(CN_3)	832.6	851.1
E(CO_1)	350.3	366.8
E(CO_2)	740.3	761.3
E(CS_1)	273.3	285.3
E(HN_1)	372.9	403.2
E(HO_1)	442.7	472.7
E(HS_1)	366.7†	386.5†
E(NN_2)	-101.3	-54.8
E(NO_1)	148.5	190.7
E(NO_5)	379.5	388.1
$H_{vib}(0)$		0.9207†
electrostatic (Å e^{-2})	-1.2	-1.2

All units in kJ mol^{-1} unless specified parenthetically	A3 Bond 345 molecules	A3 Bond 340 molecules
CC_1	22.6	17.1
CC_1<CC_2>1	-11.1	-14.7
CC_1<CC_2>1<CN_3>1	-284.7	-282.4
CC_1<CC_2>1<CO_...*	-262.0	-252.8
CC_1<CC_5>2	5.1	15.8
CC_1<CN_1>1-2	2.7	1.2
CC_1<CN_1>1<CO_2>1	65.7	54.5
CC_1<CN_3>1	-156.4	-160.2
CC_1<CN_3>2	1547.0	1669.2
CC_1<CO_1>1-2	-15.7	-8.8
CC_1<CO_1>1<CO_2>1	46.6	46.8
CC_1<CO_2>1	-21.5	-23.0
CC_1<CS_1>1	14.2	11.9
CC_2	34.3	35.3
CC_3	-88.8	-94.3
CC_5<CC_5>2	3.6	-11.1
CN_1	-21.8	-28.3
CN_1<CN_1>1-3	58.4	50.6
CN_1<CN_1>1-3<CO_2>1	-13.4	-157.1
CN_1<CN_1>1-3<NO_5>2	-91156.9	11068.5
CN_1<CO_2>1	152.8	127.3
CN_1<NN_2>1	3.0	-0.7
CN_1<NO_5>2	8.1	8.2
CN_3	49.3	43.9
CO_1	-9.2	-11.2
CO_1<CO_1>1	14.4	9.7
CO_1<CO_1>1<CO_2>1	-81.6	-75.4
CO_1<CO_2>1	168.3	159.6
CO_1<NO_1>1	-30.3	-38.6
CO_2	533.8	620.4
CO_2<CN_1>1	-24733.9	-2752.3
CO_2<CO_1>1	188.7	200.4

All units in kJ mol^{-1} unless specified parenthetically	**A3 Bond 345 molecules**	**A3 Bond 340 molecules**
CS_1	-11.3	-50.8
NN_2	-2366.7	-2224.1
NO_1<CO_1>1<NO_5>2	-7.8	-13.6
NO_5<CN_1>1<NO_5>1	49.6	-59.8
NO_5<NO_1>1<NO_5>1	146.9	152.0
CC	51.3	69.0
CN	40.0	26.0
CO	63.3	54.9
CS	127.4	128.3
NN	-193543.9	24198.1
NO	-200.8	-186.7
OO	-292.7	-253.0

* CC_1<CC_2>1<CO_1>1<CO_2>1

† Not a LSE.

The remaining parameters for steric interactions for the test groups of 345 and 340 molecules are shown in Tables 7.16 and 7.17 respectively.

Table 7.16

Lennard-Jones potentials for the 345 test group.

345 molecules		**A and B values**	**Lennard-Jones equivalent**	
HH H21-H6	B (Å$^{1/6}$)	414.6	r_o (Å)	3.000
HH H21-H6	A (Å$^{1/12}$)	-151004.5	ε (kJ mol^{-1})	-0.285
HH H6-H6	B (Å$^{1/6}$)	-679.4	r_o (Å)	*1.587
HH H6-H6	A (Å$^{1/12}$)	-5429.9	ε (kJ mol^{-1})	63.758
HN H6-N61	B (Å$^{1/6}$)	-10817.8	r_o (Å)	3.423
HN H6-N61	A (Å$^{1/12}$)	8699956.5	ε (kJ mol^{-1})	3.363
HN H6-N8	B (Å$^{1/6}$)	-1674.9	r_o (Å)	3.055
HN H6-N8	A (Å$^{1/12}$)	681439.2	ε (kJ mol^{-1})	1.029
HN H6-N9	B (Å$^{1/6}$)	2218.2	r_o (Å)	3.096
HN H6-N9	A (Å$^{1/12}$)	-975986.8	ε (kJ mol^{-1})	-1.260

345 molecules	A and B values	Lennard-Jones equivalent	
HO H21-O7 B ($\text{Å}^{1/6}$)	17525.0	r_o (Å)	2.480
HO H21-O7 A ($\text{Å}^{1/12}$)	-2036830.9	ε (kJ mol^{-1})	-37.697
HO H6-O6 B ($\text{Å}^{1/6}$)	-2639.7	r_o (Å)	3.030
HO H6-O6 A ($\text{Å}^{1/12}$)	1020683.7	ε (kJ mol^{-1})	1.707
HO H6-O7 B ($\text{Å}^{1/6}$)	1435.9	r_o (Å)	2.306
HO H6-O7 A ($\text{Å}^{1/12}$)	-108037.7	ε (kJ mol^{-1})	-4.771
NN N61-N61 B ($\text{Å}^{1/6}$)	-798222.8	r_o (Å)	4.636
NN N61-N61 A ($\text{Å}^{1/12}$)	3963280853.4	ε (kJ mol^{-1})	40.191
NO N8-O7 B ($\text{Å}^{1/6}$)	-9656064.3	r_o (Å)	3.084
NO N8-O7 A ($\text{Å}^{1/12}$)	4157779608.2	ε (kJ mol^{-1})	5606.332
OO O7-O7 B ($\text{Å}^{1/6}$)	-1242779.5	r_o (Å)	3.827
OO O7-O7 A ($\text{Å}^{1/12}$)	1953435277.2	ε (kJ mol^{-1})	197.665

* Both *A* and *B* parameters are of the same sign. Therefore there is no minimum or maximum for r at r_o. All values were calculated using the absolute values of *A* and *B* in Equation (42).

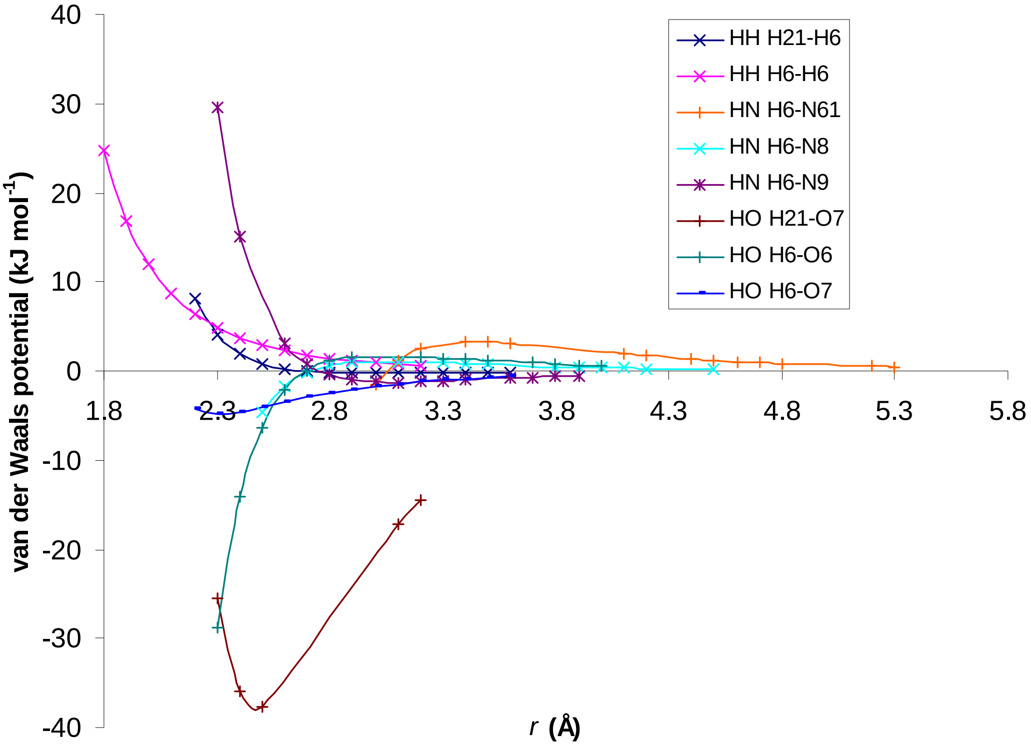
The usual shape of a van der Waals potential has an energy minimum when $r = r_o$ (when $r = r_o$ the interaction is its most attractive) steeply increasing and quickly becoming repulsive at shorter distances. The energy of a steric function is subtracted from a calculated ΔH_a value, as an increase in steric energy should reduce ΔH_a. In the regression analysis, the parameters of Tables 7.16 and 7.17 represent the energy that is added. Since the potential that is added is to be compared with what is normally subtracted, the negative of these parameters should be taken. If the potential is in the form of *eq.* (41) then a negative A value and positive B value in Tables 7.16 and 7.17 will give a potential with the aforementioned behaviour of a van der Waals potential.

From Table 7.16 the H···H purely aliphatic interactions (HH H6-H6) are purely repulsive with no attractive part in the function, while aliphatic H interactions with H on electronegative atoms (HH H21-H6) have a potential with the usual shape. H···N and H···O interactions have the usual shape only for interactions with monovalent O (O7) and divalent N (N9). The remaining H···N and H···O interactions are aliphatic H (H6) with divalent O (O6) and monovalent or trivalent N (N61 or N8). These, as well as the interactions not involving H, have a form the negative of a van der Waals potential, such that *r* at r_o represents a potential maximum instead of a minimum. There are six potentials with maxima for *r* at r_o. Graphs 7.6 and 7.7 show the potentials ascertained by linear least squares analysis for the 345 test group.

A potential's domain in Graphs 7.6, 7.7, 7.8 and 7.9 is limited to the distances of Table 7.18 for the steric interactions of that table. The other steric interactions represented are truncated after 15 points. *viz.* HH H21-H6, HH H6-H6, HN H6-N61, HN H6-N8, HN H6-N9, HO H6-O6, HO H6-O7.

Graph 7.6

Lennard-Jones potentials involving H for the 345 test group.



Graph 7.7

Lennard-Jones potentials not involving H for the 345 test group.

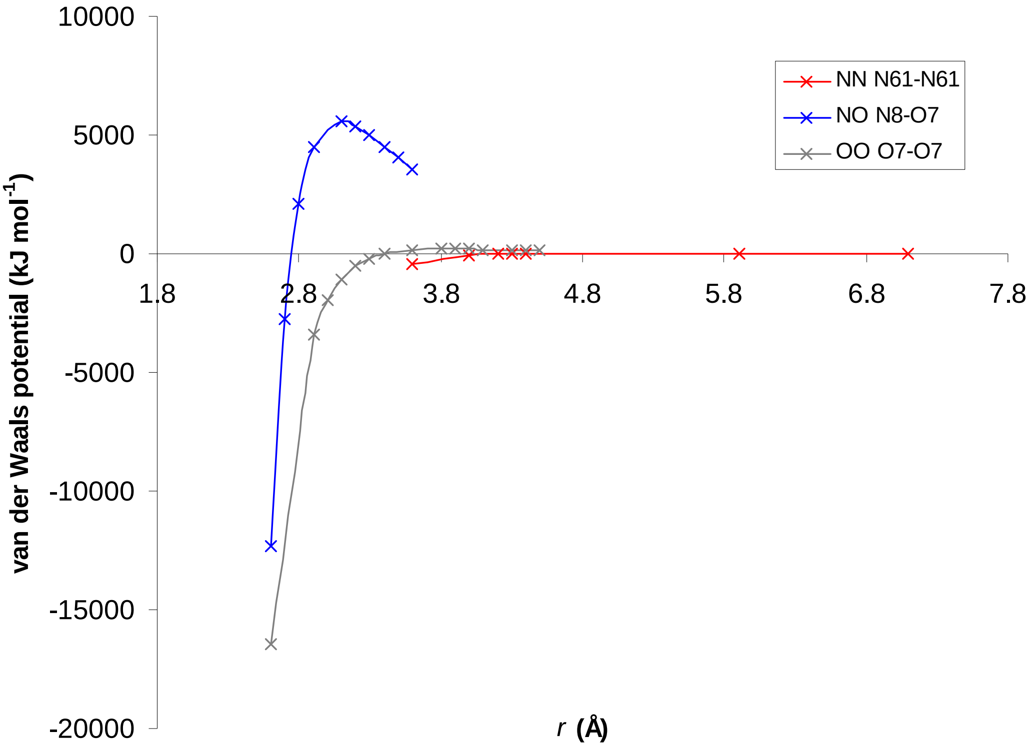


Table 7.17

Lennard-Jones potentials for the 340 test group.

340 molecules		**A and B values**	**Lennard-Jones equivalent**	
HH H21-H6	B (Å$^{1/6}$)	394.2	r_o (Å)	2.992
HH H21-H6	A (Å$^{1/12}$)	-141293.2	ε (kJ mol^{-1})	-0.275
HH H6-H6	B (Å$^{1/6}$)	-668.7	r_o (Å)	*1.485
HH H6-H6	A (Å$^{1/12}$)	-3590.3	ε (kJ mol^{-1})	93.411
HN H6-N61	B (Å$^{1/6}$)	-12607.6	r_o (Å)	3.421
HN H6-N61	A (Å$^{1/12}$)	10102744.3	ε (kJ mol^{-1})	3.933
HN H6-N8	B (Å$^{1/6}$)	-2104.1	r_o (Å)	3.037
HN H6-N8	A (Å$^{1/12}$)	824925.7	ε (kJ mol^{-1})	1.342
HN H6-N9	B (Å$^{1/6}$)	2128.8	r_o (Å)	3.116
HN H6-N9	A (Å$^{1/12}$)	-974541.8	ε (kJ mol^{-1})	-1.163

340 molecules	**A and B values**	**Lennard-Jones equivalent**	
HO H21-O7 B (Å$^{1/6}$)	17662.9	r_o (Å)	2.491
HO H21-O7 A (Å$^{1/12}$)	-2107687.3	ε (kJ mol^{-1})	-37.005
HO H6-O6 B (Å$^{1/6}$)	-1919.8	r_o (Å)	3.042
HO H6-O6 A (Å$^{1/12}$)	761017.9	ε (kJ mol^{-1})	1.211
HO H6-O7 B (Å$^{1/6}$)	1232.6	r_o (Å)	2.146
HO H6-O7 A (Å$^{1/12}$)	-60238.1	ε (kJ mol^{-1})	-6.306
NN N61-N61 B (Å$^{1/6}$)	-871652.4	r_o (Å)	4.629
NN N61-N61 A (Å$^{1/12}$)	4287575576.9	ε (kJ mol^{-1})	44.301
NO N8-O7 B (Å$^{1/6}$)	1252359.7	r_o (Å)	3.061
NO N8-O7 A (Å$^{1/12}$)	-514610365.6	ε (kJ mol^{-1})	-761.938
OO O7-O7 B (Å$^{1/6}$)	-58344.0	r_o (Å)	*4.315
OO O7-O7 A (Å$^{1/12}$)	-188386492.9	ε (kJ mol^{-1})	13.552

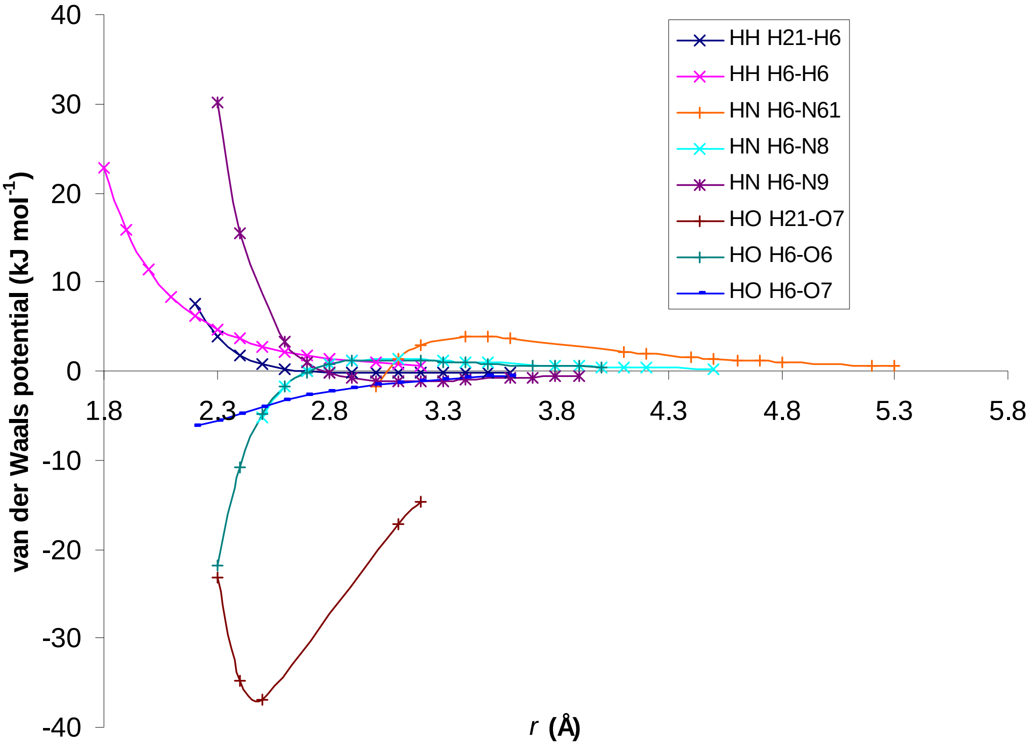
* Both *A* and *B* parameters are the same sign. Therefore there is no minimum or maximum for r at r_o. All values were calculated using the absolute values of *A* and *B* in *eq.* (42).

The results are slightly more sensible for the 340 test group. The potentials of the interactions involving at least one H have maxima or minima for r at r_o for the same types of interactions.

Of the interactions between NN, NO and OO, only the NN interaction has a potential maximum for r at r_o, compared with all three having maxima for r at r_o for the same potentials used in the 345 test group. The OO interactions which are monovalent with monovalent O interactions are purely repulsive. The slight improvement in the form of the potentials in the 340 test group over those of the 345 test group might have been due to explicit calculation of $H_{vib}(0)$ in the 340 test group. However ascertaining the same parameters without explicitly calculating $H_{vib}(0)$ in the 340 test group had no affect on the form of the potentials, so the different forms of the potentials in the two groups is most likely due to the inclusion of a further five molecules in the 345 test group.

Graph 7.8

Lennard-Jones potentials involving H for the 340 test group.



Graph 7.9

Lennard-Jones potentials not involving H for the 340 test group.

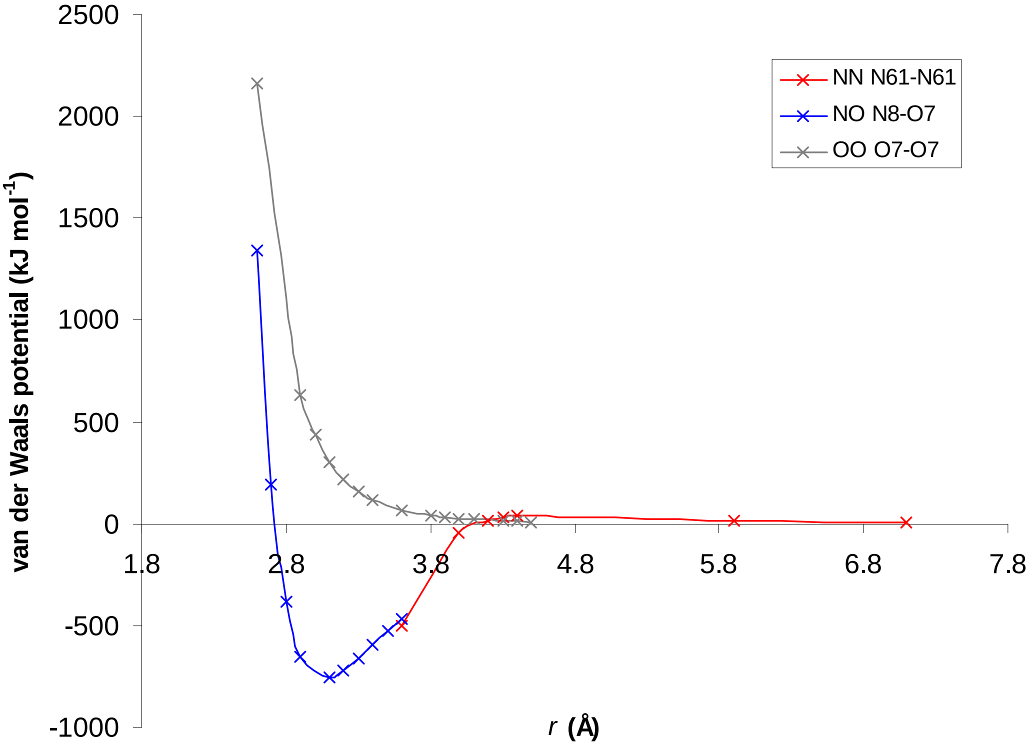


Table 7.18 shows the steric interactions with the most limited data for values of *r* (Å). For instance HO H21-O7 is parameterised over a 0.9 Å range from 2.3 to 3.2. Not much greater is the range for NO N8-O7, which is parameterised over a 1.0 Å range. Interactions of the types in Table 7.18 have a smaller statistical weighting than the other types, as the reduced range is commensurate with the paucity of the interaction.

Table 7.18

Nonbonded interaction distances in Å.

HO **H21-O7**	**NN** **N61-N61**	**NO** **N8-O7**	**OO** **O7-O7**
2.3	3.6	2.6	2.6
2.4	4.0	2.7	2.9
2.5	4.2	2.8	3.0
3.1	4.3	2.9	3.1
3.2	4.4	3.1	3.2
	5.9	3.2	3.3
	7.1	3.3	3.4
		3.4	3.6
		3.5	3.8
		3.6	3.9
			4.0
			4.1
			4.3
			4.4
			4.5

Table 7.19 shows the shortest distance for each steric environment for steric interactions, used in the test group of 345 molecules.

Table 7.19

Over test group of 345 molecules	**Closest nonbonded distance (Å)**	**Potential has a maximum at r_o for 345 test set**	**Potential has a maximum at r_o for 340 test set**
HH H21-H6	2.2	NO	NO
HH H6-H6	1.8	NO	NO
HN H6-N61	3	YES	YES
HN H6-N8	2.5	YES	YES
HN H6-N9	2.3	NO	NO
HO H21-O7	2.3	NO	NO
HO H6-O6	2.3	YES	YES
HO H6-O7	2.2	NO	NO
NN N61-N61	3.6	YES	YES
NO N8-O7	2.6	YES	NO
OO O7-O7	2.6	YES	NO

If the potentials with parameters ascertained from a test group are used to calculate a $\Delta_f H^o$ value for a molecule not in the test set, then a doubtful result might be expected if there are nonbonded interactions at shorter distances than the closest nonbonded distance of Table 7.19 for a potential with a maximum at $r = r_o$.

The form of the Lennard-Jones potential in *eq.* (41) allows parameterisation by linear least squares analysis. In a least squares analysis the *A* and *B* parameters are allowed to have positive or negative signs and this does not preserve the relationship between the dispersive and repulsive terms. Keeping this relationship requires a restricted least squares analysis or nonlinear least squares analysis on a function preserving the repulsive and dispersive terms, using a Gauss-Newton or modified Newton method. Performing a nonlinear least squares Gauss-Newton analysis failed to converge. The problem at hand is a difficult one as we have one observable value per molecule ($\Delta_f H^o$) which has many interatomic distances. Probably for this reason it is difficult to get these methods to converge. Also the Lennard-Jones potential has weak local minima for hardness parameters. Alternatively the SSR does not permit a negative function of the van der Waals potential. Its simpler form makes it more obvious what is an attractive or repulsive steric interaction, and can thus be kept in check more easily. The SSR also has weak local minima for hardness parameters. This can be seen for SSR in Tables 4.2 and 4.3.

8 Conclusion

The systematic corrections to a ΔE_{Tot} HF-SCF or $\Delta_f H^o$ PM3 calculation show some promise for further development. However it is likely a system that determines a parameter specific to a group or type of molecule may be necessary. We have avoided the use of such parameters, and have attempted to develop constant transferable bond energies that vary according to B(π). However the trend obtainable for a homologous series is diminished for the same molecules, when fitting these molecules with others in a polyfunctional test set. This can be seen for the small alkane test set in Tables 7.4, 7.9 and 7.14 where $\Delta_f H^o$ values are calculated with parameters ascertained over a much larger polyfunctional test set. These standard deviations for the set of 15 alkanes are greater than the lower standard deviations achievable over the entire set of 40 alkanes when parameters are ascertained by parameterisation of only the 40 alkanes. The results of Table 7.5 show the $\Delta_f H^o$ values for small alkanes calculated by MM4 as well as Pedley's results. Not all the MM4 values were available for the small alkane test group from Allinger *et al.*[36]. The standard deviations of the small alkane test set [s(15)] for MM4 and Pedley's method, were 1.3 and 0.7 kJ mol^{-1} respectively. The best s(15) value using bondingness was 3.6 kJ mol^{-1} which had a 9.0 kJ mol^{-1} s(346) value (see Table 7.4). The s(15) is not much worse at 3.8 kJ mol^{-1}, but with a much better s(340) of 5.3 kJ mol^{-1} for the method using B(π), $H_{vib}(0)$ and the Lennard-Jones potential (see Table 7.15). Our overall standard deviation is better than Pedley's for a similar test group, and this is because not all parameters can be ascertained for all molecular fragments that are required in a group additivity method *i.e.* there are 238 different molecular fragments determined by next nearest neighbours in the 346 test group, for which Pedley only determines about 107. MM4 has a better standard deviation for the small alkanes test group, but without knowing the values for most of the 346 test group, no overall comparison can be made.

The Lennard-Jones potential and the $H_{vib}(0)$ have different limitations. The $H_{vib}(0)$ is computationally expensive and not feasible for large molecules. The Lennard-Jones potential is not so computationally expensive but with the form and method of parameterisation we have used, it is difficult to ascertain parameters for polyfunctional test groups. One may decide to use either $H_{vib}(0)$ or the Lennard – Jones potential. The Lennard-Jones potential is probably the method to choose when a closely related test group of molecules is available with polyfunctionality, giving a maximum of six or seven steric environments, in which case a reasonable approximation can be made using Equation (43) or (44). This method would be limited by the polyfunctionality of the test group and less by molecular size. However if molecules are sufficiently small enough that vibrational data can be ascertained, parameterisation with $H_{vib}(0)$ and SSR could give better predictions for molecules, when due to polyfunctionality the number of steric environments exceeds seven. However more work needs to be done for the Lennard-Jones potential using restricted regression analysis, so that potentials that are the negative of the usual potential are restricted. These potentials have energy maxima for r at r_o. It is likely that $\Delta_f H^o$ calculations for molecules that require Lennard –Jones potentials with maxima for r at r_o will give spurious results if the nonbonded distances are shorter than those of Table 7.19.

Molecular mechanics (MM) is considered by many to be the method of choice for $\Delta_f H^o$ calculations, and the most advanced force fields employ the most plenary knowledge of the current understanding of molecular forces. For theoreticians, the inability of a MM force field to account for some experimental observations is a pointer to new or misunderstood chemical phenomena. When our methods were too simple, we looked to MM and what potentials are used by some of the more popular models, for instance the buffered 14 7 potential. Future development in bondingness algorithms may be most fruitful in ascertaining hyperconjugation information, a chemical phenomenon used in class 3 force fields. Hyperconjugation might be feasible to parameterise as an algorithm, presumably of B(σ) or an interaction of B(π) with B(σ) on adjacent bonds. Also more work can be done in refining the chemical environment labelling algorithm to incorporate similar chemical environments into the same label, or remove redundant B(π) chemical environments, as well as testing more options for bondingness over 1-3 nonbonded interactions.

9 References

1. Cottrell, T.L., *Binding Energies in Hydrocarbons.* Journal of the Chemical Society, **1948**: p. 1448-1453.
2. Laidler, K.J., *A System of Molecular Thermochemistry for Organic Gases and Liquids.* Canadian Journal of Chemistry, **1956**. **34**(5): p. 626-648.
3. Benson, S.W. and J.H. Buss, *Additivity Rules for the Estimation of Molecular Properties. Thermodynamic Properties.* Journal of Chemical Physics, **1958**. **29**(3): p. 546-572.
4. Allen, T.L., *Bond Energies and the Interactions between Next-Nearest Neighbors. I. Saturated Hydrocarbons, Diamond, Sulfanes, S8, and Organic Sulfur Compounds.* Journal of Chemical Physics, **1959**. **31**(4): p. 1039-1049.
5. Skinner, H.A., *An Examination of Allen's Empirical Bond-Energy Scheme, and its Application to Paraffins and Cycloalkanes, Olefins, Alkyl Alcohols and Bromides, and Amines.* Journal of the Chemical Society, **1962**: p. 4396-4408.
6. Cox, J.D. and G. Pilcher, *Thermochemistry of Organic and Organometallic Compounds*. **1970**, Academic Press Inc.: London & New York.
7. Benson, S.W. and M. Luria, *Electrostatics and the Chemical Bond. II. Unsaturated Hydrocarbons.* Journal of the American Chemical Society, **1975**. **97**(12): p. 3337-3342.
8. Benson, S.W. and M. Luria, *Electrostatics and the Chemical Bond.*[1] *I. Saturated Hydrocarbons.* Journal of the American Chemical Society, **1975**. **97**(4): p. 704-709.
9. Benson, S.W. and M. Luria, *Electrostatics and the Chemical Bond. III. Free Radicals.* Journal of the American Chemical Society, **1975**. **97**(12): p. 3342-3345.
10. Skinner, H.A., *Thermochemistry and Its Applications to Chemical and Biochemical Systems*. NATO ASI series. Series C: Mathematical and Physical Sciences, ed. M.A.V.R. da Silva. Vol. 119. **1984**, D. Reidel: Dordrecht.
11. Pihlaja, K., K. Rossi, and P. Vainiotalo, *Bond-Bond Interactions in Alkanes and Their Heteroanalogues. Allen-Type Group Increments for Estimating Enthalpies of Formation of Alkanes and Their Oxygen, Sulfur, and Nitrogen Analogues and Aliphatic Ketones.* Journal of Chemical and Engineering Data, **1985**. **30**(4): p. 387-394.
12. Garbalena, M. and W.C. Herndon, *Optimum Graph-Theoretical Models for Enthalpic Properties of Alkanes.* Journal of Chemical Information and Computer Sciences, **1992**. **32**(1): p. 37-42.
13. Cohen, N. and S.W. Benson, *Estimation of Heats of Formation of Organic Compounds by Additivity Methods.* Chemical Reviews, **1993**. **93**(7): p. 2419-2438.
14. Laurencelle, N. and P.D. Pacey, *Simple Electrostatic Model for Enthalpies of Formation of Singly Substituted Alkanes.* Journal of the American Chemical Society, **1993**. **115**(2): p. 625-631.
15. Pedley, J.B., *Thermochemical Data and Structures of Organic Compounds*. Vol. Volume I. **1994**, Thermodynamics Research Center: Texas.

16. Fishtik, I., I. Gutman, and I. Nagypal, *Response Reactions in Chemical Thermodynamics.* Journal of the Chemical Society, Faraday Transactions, **1996**. **92**(19): p. 3525-3532.
17. Smith, D.W., *Additive Bond Energy Scheme for the Calculation of Enthalpies of Formation and Bond Dissociation Energies for Alkyl Radicals.* Journal of the Chemical Society, Faraday Transactions, **1996**. **92**(22): p. 4415-4417.
18. Smith, D.W., *Additive Bond-Energy Scheme for the Calculation of Enthalpies of Formation of Hydrocarbons Including Geminal H-H Terms.* Journal of the Chemical Society, Faraday Transactions, **1996**. **92**(7): p. 1141-1147.
19. Smith, D.W., *Additive Bond-Energy Scheme with Geminal H–H Terms. Applications to the Enthalpies of Formation of Alkane Derivatives.* Journal of the Chemical Society, Faraday Transactions, **1997**. **93**(11): p. 2037-2041.
20. Smith, D.W., *Is Overlap Repulsion between C-H Bonds Responsible for Differences in Enthalpies of Formation Among Isomers in Alkanes, Alkenes and Simple Derivatives? An Empirical Inquiry.* Physical Chemistry Chemical Physics, **1999**. **1**: p. 4065-4068.
21. Smith, D.W., *Carbon-Carbon π Antibonding Effects on the Thermochemistry of Alkanes, Elucidated by Angular Overlap and MO Calculations.* Physical Chemistry Chemical Physics, **2001**. **3**(17): p. 3562-3568.
22. Fishtik, I., R. Datta, and J.F. Liebman, *Group Additivity Methods in Terms of Responce Reactions.* Journal of Physical Chemistry A, **2003**. **107**(13): p. 2334-2342.
23. Fishtik, I. and R. Datta, *Group Additivity vs Ab Initio.* Journal of Physical Chemistry A, **2003**. **107**(34): p. 6698-6707.
24. Smith, D.W. and M.C.J. Lee, *Correlation between MO Eigenvectors and Enthalpies of Formation for Alkanes.* Thermochimica Acta, **2005**. **435**: p. 43-47.
25. Bhattacharya, A. and S. Shivalkar, *Re-tooling Benson's Group Additivity Method for Estimation of the Enthalpy of Formation of Free Radicals: C/H and C/H/O Groups.* Journal of Chemical and Engineering Data, **2006**. **51**(4): p. 1169-1181.
26. Fishtik, I., *Group Additivity Methods without Group Values.* Journal of Physical Chemistry A, **2006**. **110**(49): p. 13264-13269.
27. Fishtik, I., D. Urban, and J. Wilcox, *The Effect of Stoichiometry on Ab Initio-Based Thermochemistry Predictions.* Chemical Physics Letters, **2006**. **417**(1-3): p. 185-189.
28. Hayes, M.Y., L. Baiqing, and H. Rabitz, *Estimation of Molecular Properties by High-Dimensional Model Representation.* Journal of Physical Chemistry A, **2006**. **110**(1): p. 264-272.
29. Leal, J.P., *Additive Methods for Prediction of Thermochemical Properties. The Laidler Method Revisited. 1. Hydrocarbons.* Journal of Physical and Chemical Reference Data, **2006**. **35**(1): p. 55-76.
30. Salmon, A. and D. Dalmazzone, *Prediction of Enthalpy of Formation in the Solid State (at 298.15 k) using Second-Order Group Contributions. Part 1. Carbon-Hydrogen and Carbon-Hydrogen-Oxygen Compounds.* Journal of Physical and Chemical Reference Data, **2006**. **35**(3): p. 1443-1457.

31. Wodrich, M. and P. Schleyer, *New Additivity Schemes for Hydrocarbon Energies.* Organic Letters, **2006**. **8**(10): p. 2135-2138.
32. Salmon, A. and D. Dalmazzone, *Prediction of Enthalpy of Formation in the Solid State (at 298.15 K) Using Second-Order Group Contributions—Part 2: Carbon-Hydrogen, Carbon-Hydrogen-Oxygen, and Carbon-Hydrogen-Nitrogen-Oxygen Compounds.* Journal of Physical and Chemical Reference Data, **2007**. **36**(1): p. 19-58.
33. Burkert, U. and N.L. Allinger, *Molecular Mechanics*. ACS Monographs, ed. M.C. Caserio. Vol. 177. **1982**, American Chemical Society: Washington, D.C.
34. Schachtschneider, J.H. and R.G. Snyder, *Vibrational Analysis of the n-Paraffins--II : Normal Co-ordinate Calculations.* Spectrochimica Acta, **1963**. **19**(1): p. 117-168.
35. Hwang, M.J., T.P. Stockfisch, and A.T. Hagler, *Derivation of Class II Force Fields. 2. Derivation and Characterization of a Class II Force Field, CFF93, for the Alkyl Functional Group and Alkane Molecules.* Journal of the American Chemical Society, **1994**. **116**(6): p. 2515-2525.
36. Allinger, N.L., K.-H. Chen, and J.-H. Lii, *An Improved Force Field (MM4) for Saturated Hydrocarbons.* Journal of Computational Chemistry, **1996**. **17**(5 & 6): p. 642-668.
37. Allinger, N.L., *Conformational Analysis. 130. MM2. A Hydrocarbon Force Field Utilizing V1 and V2 Torsional Terms.* Journal of the American Chemical Society, **1977**. **99**(25): p. 8127-8134.
38. Allinger, N.L., Y.H. Yuh, and J.-H. Lii, *Molecular Mechanics. The MM3 Force Field for Hydrocarbons. 1.* Journal of the American Chemical Society, **1989**. **111**(23): p. 8551-8566.
39. Lii, J.-H. and N.L. Allinger, *Molecular Mechanics. The MM3 Force Field for Hydrocarbons. 2. Vibrational Frequencies and Thermodynamics.* Journal of the American Chemical Society, **1989**. **111**(23): p. 8566-8575.
40. Lii, J.-H. and N.L. Allinger, *Molecular Mechanics. The MM3 Force Field for Hydrocarbons. 3. The van der Waals' Potentials and Crystal Data for Aliphatic and Aromatic Hydrocarbons.* Journal of the American Chemical Society, **1989**. **111**(23): p. 8576-8582.
41. Allinger, N.L., et al., *Hyperconjugative Effects on Carbon - Carbon Bond Lengths in Molecular Mechanics (MM4).* Journal of Computational Chemistry, **1996**. **17**(5 & 6): p. 747-755.
42. Nevins, N., K. Chen, and N.L. Allinger, *Molecular Mechanics (MM4) Calculations on Alkenes.* Journal of Computational Chemistry, **1996**. **17**(5 & 6): p. 669-694.
43. Nevins, N., J.-H. Lii, and N.L. Allinger, *Molecular Mechanics (MM4) Calculations on Conjugated Hydrocarbons.* Journal of Computational Chemistry, **1996**. **17**(5 & 6): p. 695-729.
44. Nevins, N. and N.L. Allinger, *Molecular Mechanics (MM4) Vibrational Frequency Calculations for Alkenes and Conjugated Hydrocarbons.* Journal of Computational Chemistry, **1996**. **17**(5 & 6): p. 730-746.
45. Allinger, N.L. and Y. Fan, *Molecular Mechanics Studies (MM4) of Sulfides and Mercaptans.* Journal of Computational Chemistry, **1997**. **18**(15): p. 1827-1847.
46. Allinger, N.L. and K.A. Durkin, *Van der Waals Effects between Hydrogen and First-Row Atoms in Molecular Mechanics (MM3/MM4).* Journal of Computational Chemistry, **2000**. **21**(14): p. 1229-1242.

47. Langley, C.H., J.-H. Lii, and N.L. Allinger, *Molecular Mechanics (MM4) Calculations on Carbonyl Compounds Part I: Aldehydes.* Journal of Computational Chemistry, **2001**. **22**(13): p. 1396-1425.
48. Langley, C.H., J.-H. Lii, and N.L. Allinger, *Molecular Mechanics Calculations on Carbonyl Compounds. II. Open-Chain Ketones.* Journal of Computational Chemistry, **2001**. **22**(13): p. 1426-1450.
49. Langley, C.H., J.-H. Lii, and N.L. Allinger, *Molecular Mechanics Calculations on Carbonyl Compounds. III. Cycloketones.* Journal of Computational Chemistry, **2001**. **22**(13): p. 1451-1475.
50. Langley, C.H., J.-H. Lii, and N.L. Allinger, *Molecular Mechanics Calculations on Carbonyl Compounds. IV. Heats of Formation.* Journal of Computational Chemistry, **2001**. **22**(13): p. 1476-1483.
51. Mastryukov, V.S., K.-H. Chen, and N.L. Allinger, *Pyramidalized Cycloalkenes (Cyclohexene, Cycloheptene, and cis-Cyclooctene): An MM4 and ab Initio Study.* J. Phys. Chem. A, **2001**. **105**(37): p. 8562-8566.
52. Chen, K.-H. and N.L. Allinger, *Molecular Mechanics (MM4) Study of Saturated Four-Membered Ring Hydrocarbons.* Journal of Molecular Structure: THEOCHEM, **2002**. **581**(1-3): p. 215-237.
53. Langley, C.H. and N.L. Allinger, *Molecular Mechanics (MM4) Calculations on Amides.* J. Phys. Chem. A, **2002**. **106**(23): p. 5638-5652.
54. Lii, J.-H., *Molecular Mechanics (MM4) Studies of Carboxylic Acids, Esters, and Lactones.* J. Phys. Chem. A, **2002**. **106**(37): p. 8667-8679.
55. Allinger, N.L., et al., *Alcohols, Ethers, Carbohydrates, and Related Compounds. I. The MM4 Force Field for Simple Compounds.* Journal of Computational Chemistry, **2003**. **24**(12): p. 1447-1472.
56. Langley, C.H. and N.L. Allinger, *Molecular Mechanics (MM4) and ab Initio Study of Amide-Amide and Amide-Water Dimers.* J. Phys. Chem. A, **2003**. **107**(26): p. 5208-5216.
57. Lii, J.-H., et al., *Alcohols, ethers, carbohydrates, and related compounds. II. The anomeric effect.* Journal of Computational Chemistry, **2003**. **24**(12): p. 1473-1489.
58. Lii, J.-H., et al., *Alcohols, ethers, carbohydrates, and related compounds. III. The 1,2-dimethoxyethane system.* Journal of Computational Chemistry, **2003**. **24**(12): p. 1490-1503.
59. Lii, J.-H., K.-H. Chen, and N.L. Allinger, *Alcohols, ethers, carbohydrates, and related compounds. IV. carbohydrates.* Journal of Computational Chemistry, **2003**. **24**(12): p. 1504-1513.
60. Sakakibara, K., et al., *Molecular Mechanics (MM4) Calculations on [3.3]- and [4.4]Orthoparacyclophanes.* J. Phys. Chem. A, **2004**. **108**(15): p. 3048-3055.
61. Lii, J.-H., K.-H. Chen, and N.L. Allinger, *Alcohols, Ethers, Carbohydrates, and Related Compounds Part V. The Bohlmann Torsional Effect.* J. Phys. Chem. A, **2004**. **108**(15): p. 3006-3015.
62. Lii, J.-H., et al., *The External-Anomeric Torsional Effect.* Carbohydrate Research: Conformations of Oligo- and Poly-saccharides, **2005**. **340**(5): p. 853-862.
63. Chen, K.-H., et al., *Molecular Mechanics (MM4) Study of Fluorinated Hydrocarbons.* Journal of Physical Chemistry A, **2006**. **110**: p. 7202-7227.
64. Clark, M., R.D. Cramer III, and N. van Opdenbosch, *Validation of the general purpose tripos 5.2 force field.* Journal of Computational Chemistry, **1989**. **10**(8): p. 982-1012.

65. Halgren, T.A., *Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94*.* Journal of Computational Chemistry, **1996**. **17**(5-6): p. 490-519.
66. Halgren, T.A., *Merck Molecular Force Field. V. Extension of MMFF94 Using Experimental Data, Additional Computational Data, and Empirical Rules.* Journal of Computational Chemistry, **1996**. **17**(5-6): p. 616-641.
67. Halgren, T.A., *Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions.* Journal of Computational Chemistry, **1996**. **17**(5-6): p. 520-552.
68. Halgren, T.A. and R.B. Nachbar, *Merck Molecular Force Field. IV. Conformational Energies and Geometries for MMFF94.* Journal of Computational Chemistry, **1996**. **17**(5-6): p. 587-615.
69. Halgren, T.A., *Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94.* Journal of Computational Chemistry, **1996**. **17**(5-6): p. 553-586.
70. Halgren, T.A., *Representation of van der Waals (vdW) Interactions in Molecular Mechanics Force Fields: Potential Form, Combination Rules, and vdW Parameters.* Journal of the American Chemical Society, **1992**. **114**(20): p. 7827-7843.
71. Hehre, W.J., et al., *Ab Initio Molecular Orbital Theory*. **1986**, Wiley: New York.
72. Kong, J., et al., *Q-Chem 2.1: A High-Performance Ab Initio Electronic Structure Program Package.* Journal of Computational Chemistry, **2000**. **21**(16): p. 1532-1548.
73. Hehre, W.J., *A Guide to Molecular Mechanics and Quantum Chemical Calculations*. **2003**, Wavefunction Inc.: Irvine.
74. Atkins, P. and R. Friedman, *Molecular Quantum Mechanics*. Fourth ed. **2005**, Oxford University Press: Oxford.
75. Lowe, J.P. and K.A. Peterson, *Quantum Chemistry*. Third ed. **2006**, Elsevier Acdemic Press: Burlington.
76. Roothaan, C.C.J., *New Developments in Molecular Orbital Theory.* Reviews of Modern Physics, **1951**. **23**(2): p. 69-89.
77. Hall, G.G., *The Molecular Orbital Theory of Chemical Valency. VIII. A Method of Calculating Ionization Potentials.* Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences, **1951**. **205**(1083): p. 541-552.
78. Møller, C. and M.S. Plesset, *Note on an Approximation Treatment for Many-Electron Systems.* Physical Review, **1934**. **46**(7): p. 618-622.
79. Binkley, J.S. and J.A. Pople, *Møller-Plesset Theory for Atomic Ground State Energies.* International Journal of Quantum Chemistry, **1975**. **9**(2): p. 229-236.
80. Hohenberg, P. and W. Kohn, *Inhomogeneous Electron Gas.* Physical Review, **1964**. **136**(3B): p. B864-B871.
81. Kohn, W. and L.J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects.* Physical Review, **1965**. **140**(4A): p. A1133-A1138.
82. Becke, A.D., *Density-functional exchange-energy approximation with correct asymptotic behavior.* Physical Review A, **1988**. **38**(6): p. 3098-3100.
83. Lee, C., W. Yang, and R.G. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density.* Physical Review B, **1988**. **37**(2): p. 785-789.

84. Pople, J.A., et al., *Gaussian-1 theory: A general procedure for prediction of molecular energies.* The Journal of Chemical Physics, **1989**. **90**(10): p. 5622-5629.

85. Curtiss, L.A., et al., *Gaussian-2 Theory for Molecular Energies of First- and Second-Row Compounds.* Journal of Chemical Physics, **1991**. **94**(11): p. 7221-7230.

86. Curtiss, L.A., et al., *Gaussian-3 (G3) Theory for Molecules Containing First and Second-Row Atoms.* Journal of Chemical Physics, **1998**. **109**: p. 7764.

87. Curtiss, L.A., K. Raghavachari, and P.C. Redfern, *Gaussian-4 Theory.* Journal of Chemical Physics, **2007**. **126**(8): p. 084108.

88. Karton, A., et al., *W4 Theory for Computational Thermochemistry: In Pursuit of Confident Sub-kJ/mol Predictions.* Journal of Chemical Physics, **2006**. **125**(14): p. 144108/1-144108/17.

89. Tasi, G., et al., *The Origin of Systematic Error in the Standard Enthalpies of Formation of Hydrocarbons Computed via Atomization Schemes.* ChemPhysChem, **2006**. **7**: p. 1664-1667.

90. Dewar, M.J.S., et al., *AM1: A New General Purpose Quantum Mechanical Molecular Model.* Journal of the American Chemical Society, **1985**. **107**(13): p. 3902.

91. Dewar, M.J.S. and W. Thiel, *Ground States of Molecules. 38. The MNDO Method. Approximations and Parameters.* Journal of the American Chemical Society, **1977**. **99**(15): p. 4899.

92. Thiel, W. and A.A. Voityuk, *Extension of MNDO to d Orbitals: Parameters and Results for the Halogens.* International Journal of Quantum Chemistry, **1992**. **44**(5): p. 807-829.

93. Stewart, J.J.P., *Optimization of Parameters for Semiempirical Methods I. Method.* Journal of Computational Chemistry, **1989**. **10**(2): p. 209-220.

94. Tubert-Brohman, I., C.R.W. Guimarães, and W.L. Jorgensen, *Extension of the PDDG/PM3 Semiempirical Molecular Orbital Method to Sulfur, Silicon, and Phosphorus.* J. Chem. Theory Comput., **2005**. **1**(5): p. 817-823.

95. Tubert-Brohman, I., et al., *Extension of the PDDG/PM3 and PDDG/MNDO Semiempirical Molecular Orbital Methods to the Halogens.* Journal of Computational Chemistry, **2004**. **25**(1): p. 138-150.

96. Repasky, M.P., J. Chandrasekhar, and W.L. Jorgensen, *PDDG/PM3 and PDDG/MNDO: Improved Semiempirical Methods.* Journal of Computational Chemistry, **2002**. **23**(16): p. 1601-1622.

97. Rocha, G.B., et al., *RM1: A Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I.* Journal of Computational Chemistry, **2006**. **27**(10): p. 1101-1111.

98. Sattelmeyer, K.W., J. Tirado-Rives, and W.L. Jorgensen, *Comparison of SCC-DFTB and NDDO-Based Semiempirical Molecular Orbital Methods for Organic Molecules.* Journal of Physical Chemistry A, **2006**. **110**(50): p. 13551-13559.

99. Wiberg, K.B., *Group Equivalents for Converting Ab Initio Energies to Enthalpies of Formation.* Journal of Computational Chemistry, **1984**. **5**(2): p. 197-199.

100. Wiberg, K.B., *Structures and Energies of the Tricyclo[4.1.0.0^{1,3}]heptanes and the Tetracyclo[4.2.1.0^{2,9}0^{5,9}]nonanes. Extended Group Equivalents for Converting Ab Initio Energies to Heats of Formation.* Journal of Organic Chemistry, **1985**. **50**: p. 5285-5291.

101. Ibrahim, M.R. and P.v.R. Schleyer, *Atom Equivalents for Relating Ab Initio Energies to Enthalpies of Formation.* Journal of Computational Chemistry, **1985**. **6**(3): p. 157.
102. Herndon, W.C., *Hydrocarbon Enthalpies of Formation and Ab Initio Calculations.* Chemical Physics Letters, **1995**. **234**(1-3): p. 82-86.
103. Liu, M.H. and C. Chen, *Modified Calculations of Hydrocarbon Thermodynamic Properties.* Journal of Computational Chemistry, **2006**. **27**(5): p. 537-544.
104. Habibollahzadeh, D., et al., *Nonlocal Density Functional Calculation of Gas Phase Heats of Formation.* Journal of Computational Chemistry, **1995**. **16**(5): p. 654-658.
105. Mole, S.J., X. Zhou, and R. Liu, *Density Functional Theory (DFT) Study of Enthalpy of Formation. 1. Consistency of DFT Energies and Atom Equivalents for Converting DFT Energies into Enthalpies of Formation.* Journal of Physical Chemistry, **1996**. **100**(35): p. 14665-1467.
106. Repasky, M.P., J. Chandrasekhar, and W.L. Jorgensen, *Improved Semiempirical Heats of Formation through the Use of Bond and Group Equivalents.* Journal of Computational Chemistry, **2002**. **23**(4): p. 498-510.
107. Delley, B., *Ground-State Enthalpies: Evaluation of Electronic Structure Approaches with Emphasis on the Density Functional Method.* Journal of Physical Chemistry A, **2006**. **110**(50): p. 13632-13639.
108. Smith, D.W., *A Simple Molecular Orbital Treatment of the Barrier to Internal Rotation in the Ethane Molecule.* Journal of Chemical Education, **1998**. **75**(7): p. 907-909.
109. Smith, D.W., *The Antibonding Effect.* Journal of Chemical Education, **2000**. **77**(6): p. 780-784.
110. Chase, M.W., et al., *JANAF Thermochemical Tables Third Edition.* Journal of Physical and Chemical Reference Data, **1985**. **14**(Supplement No. 1).
111. Personal Communication: Smith, D.W. 9/12 **2004**.
112. *Spartan*. **2004**, Wavefunction, Inc.: 18401 Von Karman Avenue, Suite 370, Irvine, CA 92612 U.S.A.
113. Scott, A.P. and L. Radom, *Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Møller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors.* Journal of Physical Chemistry, **1996**. **100**(41): p. 16502-16513.
114. Wiberg, K.B. and D.J. Wasserman, *Enthalpies of Hydration of Alkenes. 1. The n-Hexenes.* Journal of the American Chemical Society, **1981**. **103**(22): p. 6563-6566.
115. Wiberg, K.B., D.J. Wasserman, and E. Martin, *Enthalpies of Hydration of Alkenes. 2. The n-Heptenes and n-Pentenes.* Journal of Physical Chemistry, **1984**. **88**(16): p. 3684-3688.
116. Liebman, J.F. and A. Greenberg, eds. *Physical Measurements*. Molecular Structure and Energetics, ed. J.F. Liebman and A. Greenberg. Vol. 2. **1987**, VCH Publishers Inc.: New York.
117. Rogers, D.W. and E.L. Crooks, *Enthalpies of Hydrogenation of the Isomers of n-Hexene.* The Journal of Chemical Thermodynamics, **1983**. **15**(11): p. 1087-1092.
118. Rogers, D.W., E. Crooks, and K. Dejroongruang, *Enthalpies of Hydrogenation of the Hexenes.* The Journal of Chemical Thermodynamics, **1987**. **19**(11): p. 1209-1215.

119. Rogers, D.W. and K. Dejroongruang, *Enthalpies of Hydrogenation of the n-Heptenes and the Methylhexenes.* The Journal of Chemical Thermodynamics, **1988**. **20**(6): p. 675-680.
120. Rogers, D.W. and K. Dejroongruang, *Enthalpies of hydrogenation of the Dimethylpentenes, Ethylpentenes, Methylethylbutene, and Trimethylbutene.* The Journal of Chemical Thermodynamics, **1989**. **21**(11): p. 1115-1120.
121. Rogers, D.W., et al., *Enthalpies of Hydrogenation of the Octenes and the Methylheptenes.* The Journal of Chemical Thermodynamics, **1992**. **24**(6): p. 561-565.
122. Cao, C. and H. Yuan, *A New Approach of Evaluating Bond Dissociation Energy from Eigenvalue of Bonding Orbital-Connection Matrix for C-C and C-H Bonds in Alkane.* Journal of Chemical Information and Computer Sciences, **2003**. **43**: p. 600-608.
123. *Spartan*, Wavefunction, Inc.: 18401 Von Karman Avenue, Suite 370, Irvine, CA 92612 U.S.A.
124. Hayes, J.M., J.C. Greer, and D.A. Morton-Blake, *A Force-Field Description of Short-Range Repulsions for High Density Alkane Molecular Dynamics Simulations.* Journal of Computational Chemistry, **2004**. **25**: p. 1953-1966.
125. Smith, D.W., *Group Electronegativities from Electronegativity Equilibration.* Journal of the Chemical Society, Faraday Transactions, **1998**. **94**(2): p. 201-205.
126. Thomas, L.L., T.J. Christakis, and W.L. Jorgensen, *Conformation of Alkanes in the Gas Phase and Pure Liquids.* Journal of Physical Chemistry B, **2006**. **110**(42): p. 21198-21204.
127. Eliel, E.L., S.H. Wilen, and L.N. Mander, *Stereochemistry of Organic Compounds*. **1994**, Wiley: New York.
128. Dale, J., *Stereochemistry and Conformational Analysis*. **1978**, Universitetsforlaget, Verlag Chemie: Oslo, New York-Weinheim.
129. Klauda, J.B., R.W. Pastor, and B.R. Brooks, *Adjacent Gauche Stabilization in Linear Alkanes: Implications for Polymer Models and Conformational Analysis.* Journal of Physical Chemistry B Letters, **2005**. **109**: p. 15684-15686.
130. Goto, H., E. Osawa, and M. Yamoto, *How Many Conformers Are There for Small n-Alkanes? Consequences of Asymmetric Deformation in GG' Segment.* Tetrahedron, **1993**. **49**(2): p. 387-396.
131. Saunders, M., *Stochastic exploration of molecular mechanics energy surfaces. Hunting for the global minimum.* J. Am. Chem. Soc., **1987**. **109**(10): p. 3150-3152.
132. Lynch, B.J., Y. Zhao, and D.G. Truhlar, http://comp.chem.umn.edu/database/freq_scale.htm site last updated 28 Oct. 2003.
133. Shingu, H. and T. Fujimoto, *Empirical Correlation between Molecular Structure and Zero-Point Energies in Hydrocarbons.* Journal of Chemical Physics, **1959**. **31**(2): p. 556-557.
134. Urey, H.C. and C.A. Bradley, *The Vibrations of Pentatonic Tetrahedral Molecules.* Physical Review, **1931**. **38**(11): p. 1969-1978.
135. Lide, D.R., ed. *CRC Handbook of Chemistry and Physics*. 87th ed. **2006**, CRC Press. Taylor & Francis Group: Boca Raton, London, New York.

A Appendix

A.1 Mathematical conventions in this work

Scalars are represented by italic non-boldface letters of upper or lowercase. A matrix is symbolised by boldface uppercase letters, while its elements which are scalar, are the same letter in lowercase, italicised with two subscripts representing the matrix element position. Vectors are lowercase boldface letters with their elements represented by lowercase italicised letters with one subscript representing the position in the vector.

Algebra of matrices:

$$\mathbf{A}(\mathbf{B} \pm \mathbf{C})\mathbf{D} = \mathbf{ABD} \pm \mathbf{ACD}$$

$$(\mathbf{ABC}\ldots)' = (\ldots\mathbf{C}'\mathbf{B}'\mathbf{A}')$$

$$(\mathbf{ABC}\ldots)^{-1} = (\ldots\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1})$$

A.2 Conversion from atomic units

To calculate a precise conversion factor from the Rydberg constant the speed of light, Planck’s constant and Avogadro’s number are used.

Constants are from CRC[135].

Molar gas constant	R (J mol^{-1} K^{-1})	8.314 472(15)
Permeability of a vacuum	μ_o (J s^{2} C^{-2} m^{-1})	$4\pi \times 10^{-7}$
Mass of electron	μ_e (kg)	$9.109\,382\,6\,(16) \times 10^{-31}$
Mass of proton	μ_p (kg)	$1.672\,621\,71\,(29) \times 10^{-27}$
Reduced mass of electron	μ_m ($\mu_e\mu_p/(\mu_e + \mu_p)$)	$9.104\,575\,49 \times 10^{-31}$
Charge on proton	e (C)	$1.602\,176\,53\,(14) \times 10^{-19}$
Speed of light	c (m s^{-1})	$2.997\,924\,58\,(1.2) \times 10^{8}$
Planck constant	h (J s)	$6.626\,069\,3\,(11) \times 10^{-34}$
Avogadro number	N_A (mol^{-1})	$6.022\,141\,5\,(10) \times 10^{23}$

The Rydberg constant can be related to less precisely known physical constants in the following way:

$$R_\infty = \frac{\mu_o^2 \mu_e e^4 c^3}{8h^3} = 1.097\ 373\ 156\ 852\ 5\ (83) \times 10^7\ \mathrm{m}^{-1}$$

Using the speed of light, Avogadro's number and Planck's constant the conversion from atomic units to S.I. units can be calculated in the following way:

$$\frac{2hcN_A R_\infty}{1000} = 2624.49963\ (97)\ \mathrm{kJ\ mol^{-1}\ a.u.^{-1}}$$

A.3 Test groups

The spartan files of the 346 and 340 test groups are included on the accompanying CD-ROM.

A.3.1 346 test molecules

Alkanes: methane, ethane, propane, butane, 2-methylpropane, pentane, 2-methylbutane, 2,2-dimethylpropane, hexane, 2-methylpentane, 3-methylpentane, 2,2-dimethylbutane, *gauche* 2,3-dimethylbutane, heptane, 2-methylhexane, 3-methylhexane, 3-ethylpentane, 2,2-dimethylpentane, (R) 2,3-dimethylpentane, 2,4-dimethylpentane, 3,3-dimethylpentane, 2,2,3-trimethylbutane, octane, 2-methylheptane, 3-methylheptane, 4-methylheptane, 3-ethylhexane, 2,2-dimethylhexane, 2,3-dimethylhexane, 2,4-dimethylhexane, 2,5-dimethylhexane, 3,3-dimethylhexane, 3,4-dimethylhexane, 3-ethyl-2-methylpentane, 3-ethyl-3-methylpentane, 2,2,3-trimethylpentane, 2,2,4-trimethylpentane, 2,3,3-trimethylpentane, 2,3,4-trimethylpentane, 2,2,3,3-tetramethylbutane.

Alkenes: ethene, propene, but-1-ene, *cis*-but-2-ene, *trans*-but-2-ene, 2-methylpropene, pent-1-ene, *cis*-pent-2-ene, *trans*-pent-2-ene, 2-methyl-but-1-ene, 3-methyl-but-1-ene, 2-methyl-but-2-ene, hex-1-ene, *cis*-hex-2-ene, *trans*-hex-2-ene, *cis*-hex-3-ene, *trans*-hex-3-ene, 2-methylpent-1-ene, 3-methylpent-1-ene, 4-methylpent-1-ene, 2-methylpent-2-ene, *cis*-3-methylpent-2-ene, *trans*-3-methylpent-2-ene, *cis*-4-methylpent-2-ene, *trans*-4-methylpent-2-ene, 2-ethylbut-1-ene, 2,3-dimethylbut-1-ene, 3,3-dimethylbut-1-ene, 2,3-dimethylbut-2-ene, hept-1-ene, 5-methylhex-1-ene, *cis*-3-methylhex-3-ene, *trans*-3-methylhex-3-ene, 2,4-dimethylpent-1-ene, 4,4-dimethylpent-1-ene, *cis*-4,4-dimethylpent-2-ene, *trans*-4,4-

dimethylpent-2-ene, 2,4-dimethylpent-2-ene, 2-ethyl-3-methylbut-1-ene, 2,3,3-trimethylbut-1-ene.

Alkynes: ethyne, propyne, but-1-yne, but-2-yne, pent-1-yne, pent-2-yne, 3-methylbut-1-yne, 3,3-dimethylbut-1-yne.

Alcohols: methanol, ethanol, 1-propanol, 1-butanol, 1-pentanol, 1-hexanol, 1-heptanol, 1-octanol, 1-nonanol, 1-decanol, 1-dodecanol, 1-tetradecanol, 1-hexadecanol, 2-methyl-1-propanol, 2-methyl-1-butanol, 2-ethyl-1-hexanol, 3-methyl-1-butanol, 2-propanol, 2-butanol, 2-pentanol, 3-pentanol, 2-hexanol, 3-methyl-2-butanol, 2-methyl-2-propanol, 2-methyl-2-butanol, 3-tert-butyl-2,2,4,4-tetramethyl-3-pentanol.

Ethers: dimethyl ether, methyl ethyl ether, methyl *n*-propyl ether, diethyl ether, methyl *n*-butyl ether, ethyl *n*-propyl ether, di-*n*-propyl ether, di-*n*-butyl ether, isopropyl methyl ether, methyl tert-butyl ether, ethyl tert-butyl ether, methyl tert-pentyl ether, diisopropyl ether, tert-butyl isopropyl ether, di-sec-butyl ether, butyl tert-butyl ether, isobutyl tert-butyl ether, sec-butyl tert-butyl ether, di-tert-butyl ether, pentyl tert-butyl ether.

Aldehydes: formaldehyde, ethanal, propanal, butanal, pentanal, heptanal, 2-methylpropanal, 2-ethylhexanal.

Ketones: propanone, 2-butanone, 2-pentanone, 3-pentanone, 3-methyl-2-butanone, 2-hexanone, 3-hexanone, 2-methyl-3-pentanone, 3,3-dimethyl-2-butanone, 2,4-dimethyl-3-pentanone, 2,2-dimethyl-3-pentanone, 2,2,4-trimethyl-3-pentanone, 2-nonanone, 5-nonanone, 2,6-dimethyl-4-heptanone, 2,2,4,4-tetramethyl-3-pentanone.

Carboxylic Acids: formic acid, acetic acid, propanoic acid, butanoic acid, pentanoic acid, 3-methylbutanoic acid, 2,2-dimethylpropanoic acid, hexanoic acid, heptanoic acid, octanoic acid, 2-ethylhexanoic acid, nonanoic acid, decanoic acid.

Esters: methyl methanoate, propyl methanoate, methyl ethanoate, ethyl ethanoate, 1-methylethyl ethanoate, butyl ethanoate, ethyl propanoate, 1-methylpropyl butanoate, methyl pentanoate, ethyl pentanoate, propyl pentanoate, 1-methylethyl pentanoate, butyl pentanoate, 1-methylpropyl pentanoate, 2-methylpropyl pentanoate, methyl 2-methylbutanoate, ethyl 2-methylbutanoate, methyl 3-methylbutanoate, ethyl 3-methylbutanoate, methyl 2,2-dimethylpropanoate, ethyl 2,2-dimethylpropanoate, methyl hexanoate, methyl heptanoate, methyl octanoate.

Esters (Alkenoates): methyl 2-propenoate, methyl *trans*-2-butenoate, ethyl *trans*-2-butenoate, propyl *trans*-2-butenoate, 1-methylethyl *trans*-2-butenoate, butyl *trans*-2-butenoate, 1-methylpropyl *trans*-2-butenoate, 3-methylbutyl *trans*-2-butenoate, 1,1-dimethylethyl *trans*-2-butenoate, ethyl *cis*-2-pentenoate, ethyl *trans*-2-pentenoate, propyl *trans*-2-pentenoate, 1-methylethyl *trans*-2-pentenoate, butyl *trans*-2-pentenoate, 1-methylpropyl *trans*-2-pentenoate, 2-methylpropyl *trans*-2-pentenoate, ethyl *cis*-3-pentenoate, ethyl *trans*-3-pentenoate, propyl *trans*-3-pentenoate, 1-methylethyl *trans*-3-pentenoate, butyl *trans*-3-pentenoate, 1-methylpropyl *trans*-3-pentenoate, 2-methylpropyl *trans*-3-pentenoate, ethyl 4-pentenoate, propyl 4-pentenoate, 1-methylethyl 4-pentenoate, butyl 4-pentenoate, 1-methylpropyl 4-pentenoate, 2-methylpropyl 4-pentenoate.

Amines: methylamine, ethylamine, 1-propylamine, 2-propylamine, 1-butylamine, 2-butylamine, 2-methyl-2-propylamine, 2-methylpropylamine, dimethylamine, diethylamine, methyl-tert-butylamine, di-*n*-propylamine, diisopropylamine, isopropyl-tert-butylamine, di-*n*-butylamine, butylisobutylamine, diisobutylamine, di-tert-butylamine, trimethylamine, triethylamine, tri-*n*-propylamine.

Amides: methanamide, ethanamide, propanamide, butanamide, pentanamide, hexanamide, octanamide, 2-methylpropanamide, 2,2-dimethylpropanamide, butylethanamide, dimethylmethanamide, dimethylacetamide, *n*-butyldiacetamide, triacetamide.

Diazenes: *trans*-butylmethyldiazene, *trans*-dipropyldiazene, *trans*-dibutyldiazene, *trans*-1,2-diisopropyldiazene, di-tert-butyldiazene.

Nitriles: hydrogen cyanide, ethanenitrile, propanenitrile, butanenitrile, 2-methylpropanenitrile, pentanenitrile, 2,2-dimethylpropanenitrile, heptanenitrile, octanenitrile, decanenitrile, ethanedinitrile, propanedinitrile, butanedinitrile, hexanedinitrile, tetramethylbutanedinitrile, methanetetracarbonitrile, propenenitrile, *cis*-2-butenenitrile, *trans*-2-butenenitrile, 3-butenenitrile, *cis*-2-pentenenitrile, *trans*-2-pentenenitrile, *trans*-3-pentenenitrile, *trans*-butenedinitrile, ethenetricarbonitrile, ethenetetracarbonitrile.

Nitroalkanes: nitromethane, nitroethane, 1-nitropropane, 2-nitropropane, 1-nitrobutane, 2-nitrobutane, 2-methyl-2-nitropropane, dinitromethane, 1,1-dinitropropane, trinitromethane, tetranitromethane.

Nitrates: methyl nitrate, ethyl nitrate, propyl nitrate, 1-methylethyl nitrate.

Thiols: methanethiol, ethanethiol, 1-propanethiol, 2-propanethiol, 1-butanethiol, 2-butanethiol, 2-methyl-1-propanethiol, 2-methyl-2-propanethiol, 1-pentanethiol, 2-methyl-1-butanethiol, 3-methyl-1-butanethiol, 2-methyl-2-butanethiol, 3-methyl-2-butanethiol, 2,2-dimethyl-1-propanethiol, 1-hexanethiol, 2-methyl-2-pentanethiol, 2,3-dimethyl-2-butanethiol, 1-heptanethiol.

Benzenoids: benzene, toluene, ethylbenzene, *n*-propylbenzene, isopropylbenzene, *n*-butylbenzene, isobutylbenzene, sec-butylbenzene, tert-butylbenzene, *n*-decylbenzene, 1,2-dimethylbenzene, 1-ethyl-2-methylbenzene, 1,3-dimethylbenzene, 1-ethyl-3-methylbenzene, 1,4-dimethylbenzene, 1-ethyl-4-methylbenzene, 1,2,3-trimethylbenzene, 1,2,4-trimethylbenzene, 1,3,5-trimethylbenzene, 1,2,4,5-tetramethylbenzene, pentamethylbenzene, hexamethylbenzene.

A.4 Computer Programmes

This section gives the file name of a computer programme accompanied with a brief description of what the file does. All files assimilate data from the OLE storage files, with extension “spartan” created by the Spartan 04© quantum chemical software package. Any number of files with the spartan extension and with calculations pertaining to one molecule per file (files from the archive folder named M0001 are retrieved within the object) are copied to a working directory. From each of these files the quantum chemical output is assimilated. This is mostly geometry and AO coefficients for B(π) calculation, but atomic charges and molecular energies are also commonly assimilated. These files are stored in the accompanying CD-ROM, as well as the necessary modules, which must be copied to an appropriate directory to match the “use lib” command in the perl scripts. The directory expected without modifying the scripts is “c:/perl/lib/bss”. As well as these script specific modules, the OLE/Storage_Lite.pm and Math/Matrix.pm modules must be installed from the perl repository.

Scripts that perform regression analysis use a modified version of the Statistics/Regression.pm module that returns the required modified standard deviation. To run the scripts without modifying them, either, install the statistics regression module and replace regression.pm with the modified regression.pm file or copy the modified regression.pm file to a sub directory of site or lib called Statistics.

A.4.1 The energy.pl Script

This file uses the parameters ascertained from the file parameters.edb, and generates a tab delimited file with extension xls of calculated energies for all the spartan files in the working directory. It also uses enthalpy.edb to ascertain experimental $\Delta_f H^o$ values, and prints these on the same row if the molecule is suitably named to match the name in the enthalpy.edb file. This file does not fit molecules to data, its purpose is only to calculate energies based on a parameter file.

A.4.2 The VdW_par_bss.pl Script

The vdW_par_bss.pl script generates parameters by least squares analysis for a van der Waals potential as well as other extra parameters and B(π) parameters.

This script generates the following files for debugging purposes: bssdump.wri, calcdump.wri, connect.wri, datadump.wri, mol_list.wri, nb_bssdump.wri and reg_input.wri. The following files are output for analysis: dispersion.xls, exp_calc.xls and parameters.xls.

The exp_calc.xls file contains the calculated $\Delta_f H^o$ compared with the experimental $\Delta_f H^o$ as well as the $\Delta_f H^o$ experimental error on the same row for each spartan file. There are several files from which experimental data can be accessed depending on the options chosen in the switches section in the vdW_par_bss.pl script. These are: enthalpy.edb, enthalpy_adjusted.edb, enthalpy_ex.edb, enthalpy_no_alk.edb, enthalpy_therm.edb or enthalpy_ZPVE.edb.

A.4.3 The Steric_par_bss.pl Script

The steric_par_bss.pl script generates parameters by least squares analysis for the SSR potential as well as other extra parameters and B(π) parameters. It generates and uses the same files as vdW_par_bss.pl *mutatis mutandis* (see A.4.2).

A.4.4 The Par_bss.pl Script

The par_bss.pl script ascertains LSEs for the keyed parameters of the complex data structure: $mol_list->{$molecule_name}{'parameters'}{'parameter to ascertain LSE'}. This script generates and uses the same files as vdW_par_bss.pl *mutatis mutandis* (see A.4.2).

A.4.5 The Steric_tot_en.pl Script

The steric_tot_en.pl script fits SSR, B(π) and extra parameters to quantum chemical total energies or heats of formation in the semiempirical case. Parameters are ascertained according to *eq.* (34). It may also assign a value for each group or type of molecules based on the molecular empirical formula. The script generates the usual files (see A.4.2), but as it requires total ionisation energies, requires a different edb file: formation.edb.

A.4.6 The Steric_tot_en_bak.pl Script

The steric_tot_en_bak.pl script does the same as the steric_tot_en.pl script except for the use of calculated $H_{vib}(0)$. The $H_{vib}(0)$ is scaled with a LSE, but using $H_{vib}(0)$ in this way is *ad hoc*, and is not being used in the calculation as a zero-point energy.

This script generates the following files for debugging purposes: bssdump.wri, calcdump.wri, connect.wri, datadump.wri, mol_list.wri and reg_input.wri. The following files are output for analysis: parameters.xls and vib_sort.edb. The most important file is the parameters.xls file which contains the parameters and associated standard deviation (labelled sigma). A sort on the sigma column ascertains the parameters with the lowest standard deviation for the list of hardness parameters. Alternatively the best_sd.pl script can be executed with the parameters.xls file in the working directory for very large files. This returns the index, which is related to the iteration number in the loop, from which can be ascertained the hardness parameters and the standard deviation.

A.4.7 The Pedley.pl Script

The pedley.pl script ascertains the molecular fragments and bond energies according to the nearest neighbours that are found over all the spartan files placed in a working directory. The fragments and bond energies are written to the connect.wri file which may be opened as a tab delimited file as a spreadsheet.

A.4.8 The Create_conf_list.pl Script

The create_conf_list.pl script generates a list of the conformers not in the excluded volume for an *n*-alkane. The list is written to the conf_list.wri file.