J Mol Model (2007) 13:1173–1213
DOI 10.1007/s00894-007-0233-4

ORIGINAL PAPER

Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements

James J. P. Stewart

Received: 9 May 2007 / Accepted: 10 July 2007 / Published online: 9 September 2007
© Springer-Verlag 2007

Abstract Several modifications that have been made to the NDDO core-core interaction term and to the method of parameter optimization are described. These changes have resulted in a more complete parameter optimization, called PM6, which has, in turn, allowed 70 elements to be parameterized. The average unsigned error (AUE) between calculated and reference heats of formation for 4,492 species was 8.0 kcal mol^{-1}. For the subset of 1,373 compounds involving only the elements H, C, N, O, F, P, S, Cl, and Br, the PM6 AUE was 4.4 kcal mol^{-1}. The equivalent AUE for other methods were: RM1: 5.0, B3LYP 6–31G*: 5.2, PM5: 5.7, PM3: 6.3, HF 6–31G*: 7.4, and AM1: 10.0 kcal mol^{-1}. Several long-standing faults in AM1 and PM3 have been corrected and significant improvements have been made in the prediction of geometries.

Keywords NDDO · Parameterization · PM6 · Transition metals

Introduction

Over the past 30 years, NDDO-type [1, 2] semiempirical methods have evolved steadily. The earliest of these methods was MNDO [3, 4], which itself was a major advance over even earlier non-NDDO methods such as MINDO/3 [5]. The main advantage of MNDO over earlier methods was that the values of the parameters were optimized to reproduce molecular rather than atomic properties. When it first appeared, MNDO was immediately popular because of its increased accuracy, but, with the passage of time, various limitations were found, among the most important of which was the almost total absence of a hydrogen bond. As hydrogen bonding is essential to life, this particular fault essentially precluded MNDO being used in modeling biochemistry.

In 1985 an attempt, AM1 [6], was made to improve MNDO by adding a stabilizing Gaussian function to the core-core interaction to represent the hydrogen bond. Despite the fact that this was an over-simplification of a very complicated phenomenon, the overall effect was similar, and for the first time NDDO methods gave a good, albeit limited, model of hydrogen bonding.

In the course of the next several years, improvements were made to the method of parameter optimization. The result of this was the PM3 method [7–10], which culminated in the parameterization of all the elements in the main group in 2004 [11]. At the same time, various changes to the original set of approximations used in MNDO were proposed, the most important of which were the addition of *d*-orbitals to main-group elements [12, 13] and the introduction of diatomic parameters. Work started on the transition metals, and parameters for some of these have been reported [14, 15]. More recently, parameter sets tailored to reproduce specific phenomena such as the binding energy of nucleic acid base pairs [16], iron complex catalyzed hydrogen abstraction [17], phosphatase-catalyzed reaction barriers [18], and the redox properties of iron containing proteins [19] have been developed.

Because of the way advances in NDDO developments occurred, in terms of the modifications of the approximations

Electronic supplementary material The online version of this article (doi:10.1007/s00894-007-0233-4) contains supplementary material, which is available to authorized users.

J. J. P. Stewart (✉)
Stewart Computational Chemistry,
15210 Paddington Circle,
Colorado Springs, CO 80921, USA
e-mail: MrMOPAC@OpenMOPAC.net

and the extensions to specific elements or groups of elements, there has been an inevitable lack of consistency. The aim of the current work was three-fold: to investigate the incorporation of some of the reported modifications to the core-core approximations into the NDDO methodology; to carry out a systematic global parameter optimization of all the main group elements, with emphasis on compounds of interest in biochemistry; and to extend the methodology by performing a restricted optimization of parameters for the transition metals. This resulted in the development of a new method, consisting of the final set of approximations used and the optimized parameters. This method will be referred to as parametric method number 6, or PM6. The name PM6 was chosen to avoid any confusion with two other unpublished methods, PM4 and PM5.

Theory

Despite the apparent complexity of semiempirical methods, there are only three possible sources of error: reference data may be inaccurate or inadequate, the set of approximations may include unrealistic assumptions or be too inflexible, and the parameter optimization process may be incomplete. In order for a method to be accurate, all three potential sources of error must be carefully examined, and, where faults are found, appropriate corrective action taken.

Reference data

In contrast to earlier methods, in which reference data was assembled by painstakingly searching the original literature, the current work relies heavily on the large compendia of data that have been developed in recent years. The most important of these are the WebBook [20], for thermochemistry, and the Cambridge Structural Database [21] (CSD), for molecular geometries.

During the early stages of the current work, consistency checks were performed to ensure that erroneous data were not used. These checks revealed many cases in which the calculated heats of formation were inconsistent with the reference heats of formation reported in the NIST database. On further checking, many of these reference data were also found [22, 23] to be inconsistent with other data in the WebBook. In those cases where there was strong evidence of error in the reference data, the offending data were deleted, and the webbook updated [24].

For molecular geometries, gas phase reference data are preferred, but in many instances such data were unavailable, and recourse was made to condensed-phase data. Provided that care was taken to exclude those species whose geometries were likely to be significantly distorted by crystal forces, or which carried a large formal charge, condensed-phase data of the type found in the CSD were regarded as being suitable as reference data.

Because earlier methods used only a limited number of reference data, most of the cases where the method gave bad results were not discovered until after the method was published. In an attempt to minimize the occurrence of such unpleasant surprises, the set of reference data used was made as large as practical. To this end, where there was a dearth or even a complete absence of experimental reference data, recourse was made to high level calculations. Thus, for the Group VIII elements, there are relatively few stable compounds, and the main phenomena of interest involve rare gas atoms colliding with other atoms or molecules, so reference data representing the mechanics of rare gas atoms colliding with other atoms was generated from the results of ab-initio calculations. Additionally, there is an almost complete lack of thermochemical data for many types of complexes involving transition metals, so augmenting what little data there was with the results of ab-initio calculations was essential.

Use of Ab-Initio results

Ab-initio calculations provide a convenient source of reference data; for this work, extensive use has been made of results of Hartree Fock and B3LYP density functional [25, 26] methods (DFT), both with the 6–31G(d) basis set for elements in the periodic table up to argon. For systems involving heavier elements, the B88–PW91 functional [27, 28] was used with the DZVP basis set. Within the spectrum of ab-initio methods these methods are not particularly accurate; many methods with larger basis sets and with post-Hartree-Fock corrections are more accurate. However, the methods used in this work were chosen because they were regarded as robust, practical methods, allowing many systems to be modeled in a reasonable amount of time, a condition that could not be achieved with the more sophisticated ab-initio methods.

Procedure used in deriving ΔH_f

Reference heats of formation, ΔH_f, for compounds and ions of elements for which there was a paucity of data were derived from DFT total energies in two stages. In the first stage, a basic set of ~1,400 well-behaved compounds, for which reliable reference values of experimental ΔH_f were available, was assembled. Only compounds containing one or more of the elements H, C, N, O, F, P, S, Cl, Br, and I were used. For this set, a root-mean-square fit was made to the reference ΔH_f using the calculated total energies, E_{tot} and the atom counts. Thus, the error function, S, in Eq. (1) was minimized.

$$S = \sum_j \left(\Delta H_j(Ref.) - 627.51 \left(E_{Tot} + \sum_i C_i n_i \right) \right)_j^2 \quad (1)$$

In this expression, the C_i are constants for each atom of type i, and the n_i are the number of atoms of that type.

In the second stage, the contribution to the total energy of compounds containing element X arising from the elements in the first stage was removed using the coefficients from Equation (1). A second RMS fit was then performed. In this, the function minimized, *S*, was the RMS difference between the reference ΔH_f of compound X and the values predicted from the DFT energy, Eq. (2).

$$S = \sum_j \left(\Delta H_j(\text{Ref.}) - 627.51 \left(E_{\text{Tot}} + \sum_i C_i n_i + C_x n_x \right) \right)_j^2 \quad (2)$$

In this expression, the only unknown is the multiplier coefficient C_x. After solving for C_x, the ΔH_f of any compound of X could then be predicted as soon as its DFT total energy was evaluated.

Training set reference data

The training set of reference data used was considerably larger than that used in parameterizing PM3 [7, 8], where approximately 800 discrete species were used. In optimizing the parameters for PM6, somewhat over 9,000 separate species were used, of which about 7,500 were well-behaved stable molecules. The remainder consisted of reference data that were tailored to help define the values of individual parameters or sets of parameters.

Use of rules in parameter optimization

Most reference data can be expressed as simple facts. Indeed, all the earlier NDDO methods were parameterized using precisely four types of reference data: ΔH_f, molecular geometries, dipole moments, and ionization potentials. During the development of PM6, however, the use of other types of reference data was found to be necessary. Because of their behavior, these new data are best described as "rules." In this context, a rule can therefore be regarded as a reference datum that is a function of one or more other data. To illustrate the use of a rule, consider the binding energy of a hydrogen bond in the water dimer. By default, the weighting factor for ΔH_f for normal compounds is 1.0 kcal mol^{-1}. With this weighting factor, average unsigned errors in the predicted ΔH_f of the order of 3–5 kcal mol^{-1} would be acceptable, particularly as the spectrum of values of ΔH_f spans several hundreds of kilocalories per mole. However, the binding energy of a hydrogen bond in a water dimer is only 5 kcal mol^{-1}. To have an average unsigned error (AUE) of 4 kcal mol^{-1} in the prediction of hydrogen bond energies would render such a method almost useless for modeling such phenomena.

One way to increase the importance of the hydrogen bond in water would be to increase the weight for the ΔH_f of the water molecule, −57.8 kcal mol^{-1}, and the water dimer system, ca. −120.6 kcal mol^{-1}. While this would have the intended effect of increasing the weight of the hydrogen bond energy, it would also have the undesired effect of increasing the weight of the ΔH_f of water.

An alternative would be to express the ΔH_f of the water dimer in terms of the ΔH_f of two individual water molecules. The difference between the two ΔH_f, that of water dimer and that of two isolated water molecules, would be the energy of the hydrogen bond. If the weight assigned to this quantity were then increased, it would increase the weight for the hydrogen bond energy without also increasing the weight for the ΔH_f of water. Such a reference datum is referred to here as a rule. That is, rules relate the ΔH_f of a moiety to that of one or more other moieties. Thus, in the above example, the simple reference datum H, representing the ΔH_f of an isolated water molecule, could be expressed as:

$$H = -57.8$$

Using a rule-based reference datum to represent the strength of the hydrogen bond, and giving a weight of 10 to the hydrogen bond energy, the ΔH_f of the water dimer would then be defined as

$$H = 10(-5 + H_{H2O} + H_{H2O})$$

In this expression, H_{H2O} was the calculated ΔH_f, in kcal mol^{-1}, of an isolated water molecule. This rule could be interpreted as "The calculated strength of the hydrogen bond formed when two water molecules form the dimer should be 5 kcal mol^{-1}, and the importance should be 100 times that of ordinary heats of formation."

Rules are very useful in defining the parameter hypersurface. Examples of such tailoring are as follows:

Correcting qualitatively incorrect predictions

During the parameterization of transition metals, some systems were predicted to have qualitatively the wrong structure. For example, $[Cu^{II}Cl_4]^{2-}$ was initially predicted to have a tetrahedral structure, instead of the D_{2d} geometry observed. To induce the parameters to change so as to make the D_{2d} geometry more stable than the T_d geometry, a rule was added to the set of reference data for copper compounds. This rule was constructed using the results of B3LYP calculations on $[Cu^{II}Cl_4]^{2-}$. First, the total energies of the optimized B3LYP structure and that of the structure resulting from the semiempirical calculation were evaluated. The difference between these energies was then used in constructing the rule. In this case, the rule was that "The ΔH_f of the geometry predicted by the faulty semiempirical method should be *n.n* kcal mol^{-1} more than that of the B3LYP geometry." When such a rule was included in the parameter optimization, with an appropriate large weight, any tendency of the parameters to predict the incorrect geometry resulted in a large contribution

to the error function. That is, with the new rule in place, there was a strong disincentive to prediction of the incorrect structure. Usually one rule was sufficient to correct most qualitative errors, but for a few complicated structures more than one rule was needed. The commonest need for multiple rules occurred when, initially, one rule was used to correct a faulty prediction and, after re-optimizing the parameters, the geometry optimized to a new structure that was distinctly different from either the correct structure or the incorrect structure covered by the rule. When that happened, the procedure just described was repeated, and a new rule added to the set of reference data to address the new incorrect structure. In extreme cases, several such rules might be needed, each one defining a geometry that was incorrect and should therefore be avoided.

Rare gas atoms at sub-equilibrium distances

For some elements, specifically those of Group VIII, there is an understandable shortage of useful experimental reference data. In addition, most simulations involving these elements are likely to involve a rare-gas atom dynamically interacting with another atom or with a molecule at distances significantly less than the equilibrium distance. This makes determining the potential energy surface at sub-equilibrium distances important. As with hydrogen bond energies, the energies involved in this domain are likely to be in the order of a few kcal mol^{-1}. The shape of the potential energy surface (PES) can readily be mapped using DFT methods. By selecting two or three representative points on this PES, reference data rules can be constructed that describe the mechanical properties of the interactions. As with hydrogen bonding, a large weight can be assigned to these rules.

Use of rules to restrain parameter values

In general, uncharged atoms that are separated by a distance sufficiently large so that all overlaps between orbitals on the two atoms are vanishingly small will not interact significantly, and what interaction energy exists would arise from VDW terms: of their nature, these are mildly stabilizing. Although statements of this type are obviously true, when they are expressed as rules and added to the training set of reference data they can help define the parameter values. For a pair of atoms, A and B, a simple diatomic system would be constructed in which the interatomic separation was the minimum distance at which any overlaps of the atomic orbitals would still be insignificant. The electronic state of such a system would then be the sum of the states of the two isolated atoms. Thus, if both A and B were silicon, then, since the ground state of an isolated silicon atom is a triplet, the combined state would be a quintet. Because the two atoms do not interact significantly, a rule could then be constructed that said "The energy of the diatomic system is equal to the addition of energies of the two individual systems." By giving this rule a large weight, any tendency of the method to generate a spurious attraction or repulsion between the atoms would be prevented.

Atomic energy levels

In keeping with the philosophy that a large amount of reference data should be used in the parameter optimization, spin-free atomic energy levels were used for most elements. The exceptions were carbon, nitrogen, and oxygen, where there were enough conventional reference data that the addition of atomic energy levels would not significantly improve the definition of the parameter surface.

NDDO approximations do not allow for spin-orbit coupling. Therefore, spin-free levels were needed. For a few elements, there were insufficient spin states to allow the spin-free energy levels to be calculated. For all the remaining elements, spin-free energy levels were calculated.

In Moore's compendia [29–31] of atomic energy levels, observed emission spectra were used in determining the energy levels of the various states of neutral and ionized atoms. Most of these energy levels were characterized by three quantum numbers: the spin and orbital angular momenta, and the "J" or spin-orbit quantum number. The starting point for determining the spin-free atomic energy levels for a given element consisted of identifying each complete manifold of atomic energy levels for that element, that is, each set of levels split by spin-orbit coupling. If all members of the set were present, i.e., all energy levels from L+S to |L−S|, then the weighted barycenter of energy could be calculated. The spin-free energy level, E, was derived from the spin-split levels $E(S,L,J)$ using Eq. (3).

$$E = \frac{1}{(2S+1)(2L+1)} \sum_{J=|L-S|}^{L+S} (2J+1)E(S,L,J) \qquad (3)$$

In those cases where the ground state of an atom was itself a member of a spin-split manifold, the barycenter of the ground state manifold was calculated and used in re-defining the spin-free ground state. For all elements except tungsten, this change in definition was benign. There is a $^7S_3$ level present in tungsten that is located only 8.4 kcal mol^{-1} above the ground state. This puts it inside the ^5D_J, manifold, which has a barycenter at 12.7 kcal mol^{-1}. The effect of this was that, on going from a spin-split to a spin-free ground state, the ground state changed from $6d^25d^4$ or ^5D to $6d^15d^5$ or ^7S, and the ^5D state now became an excited state with an energy of 4.4 kcal mol^{-1}. To allow for this, a corresponding change was made to the ground state configuration in the PM6 definition of tungsten.

Where there were relatively few other reference data, the singly-ionized, and, in rare cases, the doubly-ionized, spin-free states were also evaluated and used as reference data.

Each energy level contributed one reference datum to the training set. Most atoms have a large number of atomic energy levels, so in order to minimize the probability that a level might be incorrectly assigned, each level was labeled with three quantum numbers: the total spin momentum, the total angular momentum, and the principal quantum number for these two quantum numbers. These were compared with the corresponding values calculated from the state functions. Since each set of three quantum numbers is unique, the potential for miss-assignment was minimized. In rare cases, particularly during the early stages of parameter optimization, two states with the same total spin and angular quantum numbers would be interchanged, with the result that the calculated principal quantum number would also be interchanged. All such cases always involved the ground state, and were quickly identified and corrected.

Approximations

Most of the approximations used in PM6 are identical to those in AM1 and PM3. The differences are:

Core-core interactions

In the original MNDO set of approximations, two changes were made to the simple point-charge expression for the core-core repulsion term. Beyond about five Ångstroms, there should be no significant interaction of two neutral atoms. However, in MNDO, the two-electron, two-center $\langle s_A s_A | s_B s_B \rangle$ integrals and the electron-core interactions do not converge to the exact point charge expression; instead, they are always slightly smaller. To prevent there being a small net repulsion between two uncharged atoms, the core-core expression is modified by the exact $1/R_{AB}$ term being replaced by the term used in the $\langle s_A s_A | s_B s_B \rangle$ integrals. An additional term is needed to represent the increased core-core repulsion at small distances due to the unpolarizable core. These two changes can be expressed as the MNDO core-core repulsion term as shown in Eq. (4).

$$E_n(A,B) = Z_A Z_B \langle s_A s_A | s_B s_B \rangle \left(1 + e^{-\alpha_A R_{AB}} + e^{-\alpha_B R_{AB}}\right) \quad (4)$$

This approximation works well for most main-group elements, but when molybdenum was being parameterized, Voityuk [14] found that the errors in heats of formation and geometries were unacceptably large, and good results were achieved only when a diatomic term was added to the core-core approximation, as shown in Eq. (5).

$$E_n(A,B) = Z_A Z_B \langle s_A s_A | s_B s_B \rangle \left(1 + x_{AB} e^{-\alpha_{AB} R_{AB}}\right) \quad (5)$$

When PM3 parameters for elements of Groups IA were being optimized, the MNDO approximation to the core-core expression was found to be unsuitable. In these elements there is only one valence electron so the core charge is the same as that of hydrogen. A consequence of this was that the apparent size of these elements was also approximately that of a hydrogen atom, in marked contrast with observation. For these elements, diatomic core-core parameters were also found to be essential.

Further examination showed that when diatomic parameters were used, there was always an increase in accuracy; therefore, in the current work, Eq. (4) was replaced systematically by Eq. (5).

As the interatomic separation increased, Voityuk's equation converged to the exact point-charge interaction, as expected. However, for rare gas interactions, an increase in accuracy was found when the rate of convergence was increased by the addition of a small perturbation. Subsequently, the perturbed function was found to be generally beneficial. Because of this, the general form of the core-core interaction used in PM6 is that given in Eq. (6).

$$E_n(A,B) = Z_A Z_B \langle s_A s_A | s_B s_B \rangle \left(1 + x_{AB} e^{-\alpha_{AB}\left(R_{AB} + 0.0003 R_{AB}^6\right)}\right) \quad (6)$$

At normal chemical bonding distances, Eqs. (5) and (6) have essentially similar behavior, but at distances of greater than about 3 Å the effect of the perturbation is to make the PM6 function significantly smaller than the Voityuk approximation.

d-orbitals on main-group elements

Thiel and Voityuk have shown [13] that a large increase in accuracy results when *d*-orbitals are added to main-group elements that have the potential to be hypervalent. During preliminary stages of this work, *d*-orbitals were excluded from main-group elements, and the parameters were optimized. This work was then repeated but with *d*-orbitals on various main-group elements. The results were in accordance with Thiel's observation: the accuracy of the method increased significantly. Because of this, *d*-orbitals were added to several main-group elements: the value of the increased accuracy far outweighs the extra computational cost.

The effect of the addition of *d*-orbitals was fundamentally different between main-group elements and transition metals. For main-group elements, the effect of *d*-orbitals is merely a perturbation: to a large degree the chemistry of these elements is determined by the *s* and *p* atomic orbitals. This is not the case with transition metals, where the *d*-orbitals are of paramount importance and the *s* and *p* orbitals are of only very minor significance. In recognition of the importance of the *s* and *p* shells in main-group chemistry, specific parameters are used for the five one-center two-electron integrals. Conversely, for the transition metals, the values of these integrals are derived directly from the internal orbital exponents.

Unpolarizable core

As noted earlier, the NDDO core-core interaction is a function of the number of valence electrons. For elements on the left of the periodic table these numbers are small and can cause the elements to appear to be too small. This was part of the rationale behind the adoption of Voityuk's diatomic core-core parameters. However, even the Voityuk approximation failed during parameter optimization when, in rare cases, a pair of atoms would approach each other very closely. Examination of these catastrophes indicated that the cause was the complete neglect of the unpolarizable core of the atoms involved. To allow for its presence, the core-core interaction for all element pairs was modified by the addition of a simple function, f_{AB}, based on the first term of the Lennard-Jones potential [32]. A candidate function was constructed, Eq. (7), using the fact that, to a first approximation, the size of an atom increases as the third power of its atomic number.

$$f_{AB} = c\left(\frac{\left(Z_A^{1/3} + Z_B^{1/3}\right)}{R_{AB}}\right)^{12} \tag{7}$$

The value of c was set to 10^{-8}, this being the best compromise between the requirements that the function should have a vanishingly small value at normal chemical distances. That is, under normal conditions the value of the function should be negligible, and at small interatomic separations the function should be highly repulsive, i.e., that it should represent the unpolarizable core.

Individual core-core corrections

For a small number of diatomic interactions, the general expression for the core-core interaction was modified in order to correct a specific fault. Because it is desirable to keep the methodology as simple as possible, modifications of the approximations were made only after determining that the existing approximations were inadequate. The diatomic specific modifications were:

O–H and N–H

In the original MNDO formalism, the general core-core interaction, Eq. (4), was replaced in the cases of O–H and N–H pairs with Eq. (8).

$$E_n(A,B) = Z_A Z_B \langle s_A s_A | s_B s_B \rangle \left(1 + R_{AB} e^{-\alpha_A R_{AB}} + R_{AB} e^{-\alpha_B R_{AB}}\right) \tag{8}$$

An unintended effect of this change was that at distances where hydrogen-bonding interactions are important, the diatomic contribution to the ΔH_f is greater than if the general approximation, Eq. (4), had been used. This contributed to a reduced hydrogen-bonding interaction in MNDO, and was a contributor to the need for modified core-core interactions in AM1 and PM3.

In PM6, the MNDO core-core approximation is replaced by Voityuk's diatomic expression, but even with that modification, the resulting hydrogen bond interaction energy was too small. In an attempt to increase it, the Voityuk approximation was replaced by Eq. (9).

$$E_n(A,B) = Z_A Z_B \langle s_A s_A | s_B s_B \rangle \left(1 + x_{AB} e^{-\alpha_{AB} R_{AB}^2}\right) \tag{9}$$

At normal O–H and N–H separations, approximately 1 Å, Eqs. (5) and (9) have similar values, but at hydrogen bonding distances, ~2 Å, the contribution arising from the exponential term is significantly reduced, resulting in a corresponding increased hydrogen bond interaction energy.

C–C

After optimizing all parameters, it was found that compounds containing yne groups, -C≡C-, were predicted to be too stable by about 10 kcal mol^{-1} per yne group. This error was unique to compounds with extremely short C–C distances, and in light of the increased emphasis on accurately reproducing the properties of organic compounds, the C–C core-core term was perturbed by the addition of a repulsive term. This term was optimized to correct the error in the yne groups and to have a negligible effect on all other C–C interactions. The optimized form of the C–C core-core interaction is given in Eq. (10).

$$E_n(A,B) = Z_A Z_B \langle s_A s_A | s_B s_B \rangle \left(1 + x_{AB} e^{-\alpha_{AB}\left(R_{AB} + 0.0003 R_{AB}^6\right)} + 9.28 e^{-5.98 R_{AB}}\right) \tag{10}$$

Si–O

During testing of PM6, neutral silicate layers of the type found in talc, $H_2Mg_3Si_4O_{12}$, were found to be slightly repulsive instead of being slightly bound. An attempt was made to correct for this error by adding a weak perturbation to the Si–O interaction, illustrated by Eq. (11).

$$E_n(A,B) = Z_A Z_B \langle s_A s_A | s_B s_B \rangle \left(1 + x_{AB} e^{-\alpha_{AB}\left(R_{AB} + 0.0003 R_{AB}^6\right)} - 0.0007 e^{-(R_{AB} - 2.9)^2}\right) \tag{11}$$

Nitrogen sp^2 pyramidalization

Although PM6 predicted the degree of pyramidalization of primary amines correctly, it overestimated the pyramidalization of secondary and tertiary amines. The degree of pyramidalization of these amines was decreased by adding a function to make the calculated ΔH_f more negative as the nitrogen became more planar, as shown in Eq. (12).

$$\Delta H'_f = \Delta H_f - 0.5e^{-10\phi} \quad (12)$$

In this equation, the angle ϕ is a measure of the non-planarity of the nitrogen environment, and is given by 2π minus the sum of the three contained angles about the nitrogen atom. For planar sp^2 secondary and tertiary amines, this correction amounted to 0.5 kcal mol^{-1} per nitrogen atom.

More elements

The NDDO basis sets of many of the elements parameterized in PM6 have not previously been described. For all elements except hydrogen, which has only an *s* orbital, the basis set consists of an *s* orbital, three *p* orbitals, and, for most elements, a set of five *d* orbitals. Slater atomic orbitals are used exclusively; these are of form:

$$\varphi = \frac{(2\xi)^{n+1/2}}{((2n)!)^{1/2}} r^{n-1} e^{-\xi r} Y_l^m(\theta, \phi)$$

Where ξ is the orbital exponent, n is the principal quantum number (PQN), and the $Y_l^m(\theta, \phi)$ are the normalized real spherical harmonics. The PQN are those of the valence shell, i.e., the set of atomic orbitals most important in forming chemical bonds. For PM6, the PQN used are shown in Table 1. For most main-group elements, the *s* and *p* PQN are the same, and, when *d* orbitals are present, all three PQN are the same: that is, the PQN are (n*s*, n*p*, n*d*). For transition metals, the *d* PQN is one less than that of the *s* and *p* shells, i.e., (n*s*, n*p*, (n–1)*d*). An exception to this generalization occurs in the elements of Group VIII. Here, the valence shell is completely filled, so in all chemical interactions that could occur between an atom of a Group VIII element and any other atom, electron density could only migrate from the Group VIII element to the other atom. That is, when a rare gas element forms any type of chemical bond it would necessarily become slightly positive. This is an unrealistic result. In order to allow rare gas atoms to have the potential of being slightly negative, the set of valence orbitals was changed from (n*s*, n*p*) to (n*p*, (n+1)*s*), for the elements Ne, Ar, Kr, and Xe. Helium is the only exception to this change, because it does not have a "1*p*" valence shell. For helium, the valence shell used was (1*s*, 2*p*), this being considered the best compromise.

Table 1 Principal quantum numbers for atomic orbitals

	s	p	d		s	p	d
H	1			Kr	5	4	
He	1	2		Rb	5	5	
Li	2	2		Sr	5	5	
Be	2	2		Y	5	5	4
B	2	2		Zr	5	5	4
C	2	2		Nb	5	5	4
N	2	2		Mo	5	5	4
O	2	2		Tc	5	5	4
F	2	2		Ru	5	5	4
Ne	3	2		Rh	5	5	4
Na	3	3		Pd	5	5	4
Mg	3	3		Ag	5	5	4
Al	3	3	3	Cd	5	5	
Si	3	3	3	In	5	5	
P	3	3	3	Sn	5	5	
S	3	3	3	Sb	5	5	5
Cl	3	3	3	Te	5	5	5
Ar	4	3		I	5	5	5
K	4	4		Xe	6	5	
Ca	4	4		Cs	6	6	
Sc	4	4	3	Ba	6	6	
Ti	4	4	3	La	6	6	5
V	4	4	3	Lu	6	6	5
Cr	4	4	3	Hf	6	6	5
Mn	4	4	3	Ta	6	6	5
Fe	4	4	3	W	6	6	5
Co	4	4	3	Re	6	6	5
Ni	4	4	3	Os	6	6	5
Cu	4	4	3	Ir	6	6	5
Zn	4	4		Pt	6	6	5
Ga	4	4		Au	6	6	5
Ge	4	4		Hg	6	6	
As	4	4	4	Tl	6	6	
Se	4	4	4	Pb	6	6	
Br	4	4	4	Bi	6	6	

Parameter optimization

Background

The objective of parameter optimization is to modify the values of the parameters so as to minimize the error function, *S*, Eq. (13), representing the square of the differences between the values of reference data, $Q_{ref}(i)$, and the values calculated using the semiempirical method, $Q_{calc}(i)$, with appropriate weighting factors, g_i.

$$S = \sum_i \left(g_i\left(Q_{calc}(i) - Q_{ref}(i)\right)\right)^2 \quad (13)$$

This process is initiated by rendering the reference data in the training set dimensionless. The default conversion factors are given in Table 2, with weighting factors for

Table 2 Default weighting factors for reference data

Reference data	Weight
ΔH_f	1.0 mol.kcal^{-1}
Bond length	0.7 Å^{-1}
Angle	0.7 degrees^{-1}
Dipole	20 Debye^{-1}
I.P.	10 eV^{-1}
Elements	Multiplier
Core	1.0
Organic	0.9
Main group	0.8
Transition metals	0.7

reference data represented by rules being much larger, typically in the order of 5–20 kcal mol^{-1}.

The elements were divided into four sets: core elements, (H, C, N, and O), other elements important in organic chemistry (F, Na, P, S, Cl, K, Br, I), the rest of the main group, and the transitions metals. Elements were assigned to the different sets based on their presumed degree of importance in biochemistry, and this importance was converted into a weighting factor to be used in the parameterization optimization procedure. Reference data representing species consisting only of core elements were given their default weight. When other elements were present, the weight was set to the default weight times the smallest multiplier shown in Table 2. Thus the default weight for a reference datum involving tetramethyllead, $Pb(CH_3)_4$, would be multiplied by 0.8 reflecting the fact that this species contains an element in the main group set.

For a given set of parameters, ***P***, optimization proceeds by calculating the values of all the $Q_{calc}(i)$, their first derivatives with respect to each parameter, $P(j)$, and the second derivatives with respect to every pair of parameters. Evaluating these quantities is time-consuming, and considerable effort was expended in minimizing the need for explicit evaluation of these functions. The most efficient strategy developed [7] involved assuming that, in the region of parameter space near to the current values of the parameters, the values of the first derivatives of the $Q_{calc}(i)$ with respect to ***P*** were, at least to a first approximation, constant. By making this assumption the values of the parameters could then be updated using perturbation theory. Because the assumption is only valid in the region of the starting point in parameter space, periodically the focus was moved to the new point in parameter space and a complete explicit re-evaluations of all the functions performed. The parameter optimization process terminated when the scalar of the first derivatives dropped below a preset limit. This process was fully automated, and for given sets of reference data and parameters, parameter optimization could be performed rapidly, easily, and reliably.

Sequence of optimization of parameters

Notwithstanding the reliability of the parameter optimization procedure, a simple global optimization of all the parameters for all 70 elements involving about over 9,000 discrete species was found to be impractical because of the large number of derivatives involved. Such an optimization would involve over 2,000 parameters and over 10,000 reference data. The set of second derivatives alone would consist of 2×10^{10} terms. With more powerful computers, evaluating such large sets of derivatives might be practical some day, but even then, one faulty reference datum or one faulty initial parameter value would ruin an optimization run. The strategy of parameter optimization was approached with great caution, and the procedure finally adopted was as follows:

Because the elements H, C, N, and O are of paramount importance in biochemistry, and because large amounts of reference data are available, the starting point for parameter optimization involved the simultaneous optimization of parameters for these four elements. For the purposes of discussion, this set of four elements will be called the "core elements".

Once stable parameters had been obtained, parameters for other elements important in organic chemistry were optimized in two stages. First, the parameters for the core elements were held constant, and parameters for the elements F, P, S, Cl, Br, and I were optimized one at a time. Then all parameters for all ten elements were simultaneously optimized. This set (the organic elements) was then used as the starting point for parameterizing the rest of the main group.

The same sequence was followed for the rest of the main-group elements. That is, parameters for each element were optimized while freezing the parameters for the organic elements. Then, once all the elements had been processed, all parameters for all of the 39 main-group elements, plus zinc, cadmium, and mercury, were optimized simultaneously.

When parameters for the transition metals were being optimized, all parameters for the main group elements were held constant. There were several reasons for this. Most importantly, the reference data for the transition metals, particularly the thermochemical data, was of lower quality, so one consideration was to prevent the transition metals from having a deleterious effect on the main-group elements. Another important consideration was that most compounds involving transition metals also involved only elements of the organic set. Since parameters for these elements had been optimized using a training set consisting of all the main-group elements, the values of the optimized parameters would likely be relatively insensitive to the influence of the small number of additional reference data involving transition metals.

In general, all parameters for a given element were optimized simultaneously; this was both efficient and convenient. In some optimizations, specifically those involving a new element, only sub-sets of parameters were used. Three main sub-sets were used:

Parameters that determine atomic electronic properties

For most elements, atomic energy levels are determined by six parameters: the one-electron one-center integrals U_{ss}, U_{pp}, U_{dd}, and the internal orbital exponents ζ_{sn}, ζ_{pn} and ζ_{dn}. If the heat of ionization and sufficient atomic energy level data were available, these quantities could be uniquely defined; there would be no need for the use of molecular reference data. These parameters were the first to be optimized whenever an optimization was started for an element that had not previously been parameterized

Parameters that determine molecular electronic properties

Two of the more important electronic molecular properties are the dipole moment, which indicates the degree of polarization within a molecule, and the ionization potential. These properties are determined primarily by 12 parameters: the six parameters that determine atomic electronic properties and six additional parameters: β_s, β_p, and β_d and the Slater orbital exponents ζ_s, ζ_p, and ζ_d. In the second stage of parameter optimization, the first six parameters were held constant at the values defined using atomic data and the second set optimized. During this operation, all geometries were fixed at their reference values.

Parameters that determine geometries

As soon as an initial optimized set of electronic parameters was available, the diatomic and other core-core parameters could be optimized. The most efficient process was to optimize these parameters initially without allowing the electronic parameters or the molecular geometries to optimize. If geometries were allowed to optimize, optimization of the core-core parameters would be slowed considerably, because of the tight dependency of the optimized geometries on the values of the core-core parameters, and vice versa.

As soon as all parameters had been optimized using fixed geometries, the geometries were allowed to relax and the parameters that determine geometry re-optimized. After that there would be three sets of incompletely optimized parameters: the six atomic electronic parameters, the six molecular electronic parameters and the core-core parameters. The only remaining operation was the simultaneous optimization of all the parameters. If the training set of reference data was insufficient to unambiguously define the values of all the parameters, then, at that stage, the potential existed for the parameters to become ill-defined. An example of this would be where there were too few atomic energy levels to allow all six parameters in the first set to be defined. To allow for this, a penalty function was added to each parameter. If the values of a parameter exceeded pre-defined limits, the error function *S* was incremented by a constant times the square of the excess. No penalty was applied if the value of a parameter was between the pre-defined limits; that is, no bias was applied to the numerical value of a parameter. During the early stages of simultaneous optimization of all the parameters for a given element the penalty function was used frequently. In the later stages the penalty function was invoked rarely, and then only when there was a distinct shortage of reference data.

Results

Parameters for PM6

PM6 atomic and diatomic parameters for the 70 elements are presented in Tables 3 and 4, respectively. Not all elements have all parameters: where monatomic parameters are missing, the associated approximations were not used. For diatomic parameters, where an atom-pair is missing, no representatives of that type of bond were used.

Accuracy

Comparison with other semiempirical methods

Using the program MOPAC2007 [33], an extensive comparison was made between the results obtained using PM6 and those from PM5, PM3, and AM1. This comparison was started by generating tables of reference data (that is, ΔH_f, geometries, ionization potentials (I.P.s), and dipole moments) and differences between the calculated and reference values, using each of the four methods presented here. Because of their size they are provided in the supplementary material. To simplify navigating within the tables, all species are listed in the order of their empirical formula.

Average unsigned errors (AUE) for ΔH_f for each element parameterized at the PM6 level are shown in Table 5, together with AUE for PM5, PM3, and AM1. The number of data used in each average varies depending on the elements available in each method. AM1 boron [34] uses a different core-core interaction expression from the other elements and was not used. AUE for bond-lengths are shown in Table 6. In those cases where a calculated bond-length was very large, indicating that the bond had broken, the bond-length was not used in the analysis. If such data had been used, the resulting statistics would have been misleading. AUE for angles are shown in Table 7. Errors in angles for many elements that form very ionic, i.e., labile, bonds are of less importance than errors involving elements that form strong

Table 3 PM6 parameters for 70 elements

Element	Parameters	Element	Parameters	Element	Parameters
Hydrogen		Iron		Indium	
U_{ss} [eV]	-11.246958	U_{ss} [eV]	-70.515047	U_{ss} [eV]	-28.339246
β_s [eV]	-8.352984	U_{pp} [eV]	-62.963069	U_{pp} [eV]	-23.373875
ζ_s [bohr^{-1}]	1.268641	U_{dd} [eV]	-103.631790	β_s [eV]	-1.982376
g_{ss} [eV]	14.448686	β_s [eV]	8.027621	β_p [eV]	-3.330294
a [none]	0.024184	β_p [eV]	-1.125760	ζ_s [bohr^{-1}]	2.023087
b [Å^{-2}]	3.055953	β_d [eV]	-3.507531	ζ_p [bohr^{-1}]	2.106618
c [Å]	1.786011	ζ_s [bohr^{-1}]	1.479150	g_{ss} [eV]	9.906091
		ζ_p [bohr^{-1}]	6.002246	g_{sp} [eV]	10.520060
Helium		ζ_d [bohr^{-1}]	1.080747	g_{pp} [eV]	4.826006
		z_{sn} [bohr^{-1}]	1.459152	g_{p2} [eV]	7.906563
U_{ss} [eV]	-31.770969	z_{pn} [bohr^{-1}]	1.392614	h_{sp} [eV]	3.500299
U_{pp} [eV]	-5.856382	z_{dn} [bohr^{-1}]	2.161909		
β_s [eV]	-58.903774	g_{ss} [eV]	7.977036	Tin	
β_p [eV]	-37.039974	g_{sp} [eV]	7.786867		
ζ_s [bohr^{-1}]	3.313204	g_{pp} [eV]	8.295758	U_{ss} [eV]	-29.888217
ζ_p [bohr^{-1}]	3.657133	g_{p2} [eV]	7.272041	U_{pp} [eV]	-22.156954
g_{ss} [eV]	9.445299	h_{sp} [eV]	1.880189	β_s [eV]	-8.621087
g_{sp} [eV]	11.201419	ρ(core) [bohr^{-1}]	1.272092	β_p [eV]	-4.989752
g_{pp} [eV]	9.214548	F^0_{sd} [eV]	9.300165	ζ_s [bohr^{-1}]	2.383941
g_{p2} [eV]	13.046115	G^2_{sd} [eV]	1.601345	ζ_p [bohr^{-1}]	2.057908
h_{sp} [eV]	0.299954			g_{ss} [eV]	8.269655
		Cobalt		g_{sp} [eV]	5.013349
Lithium				g_{pp} [eV]	6.584874
		U_{ss} [eV]	-21.039413	g_{p2} [eV]	5.855159
U_{ss} [eV]	-4.709912	U_{pp} [eV]	10.000000	h_{sp} [eV]	0.531212
U_{pp} [eV]	-2.722581	U_{dd} [eV]	-28.068971	a [none]	-1.004587
β_s [eV]	-2.283946	β_s [eV]	-8.992062	b [Å^{-2}]	4.706252
β_p [eV]	-7.535573	β_p [eV]	-0.100000	c [Å]	1.180218
ζ_s [bohr^{-1}]	0.981041	β_d [eV]	-2.481509		
ζ_p [bohr^{-1}]	2.953445	ζ_s [bohr^{-1}]	1.166613	Antimony	
g_{ss} [eV]	11.035907	ζ_p [bohr^{-1}]	3.000000		
g_{sp} [eV]	19.998647	ζ_d [bohr^{-1}]	1.860218	U_{ss} [eV]	-41.688879
g_{pp} [eV]	11.543650	z_{sn} [bohr^{-1}]	0.519518	U_{pp} [eV]	-39.541180
g_{p2} [eV]	9.059036	z_{pn} [bohr^{-1}]	1.000000	U_{dd} [eV]	-6.581663
h_{sp} [eV]	1.641886	z_{dn} [bohr^{-1}]	0.352115	β_s [eV]	-7.472322
		g_{ss} [eV]	2.840152	β_p [eV]	-5.940750
Beryllium		g_{sp} [eV]	3.425933	β_d [eV]	-3.979108
		g_{pp} [eV]	5.956968	ζ_s [bohr^{-1}]	2.391178
U_{ss} [eV]	-16.360315	g_{p2} [eV]	5.221864	ζ_p [bohr^{-1}]	1.773006
U_{pp} [eV]	-16.339216	h_{sp} [eV]	0.390087	ζ_d [bohr^{-1}]	2.465590
β_s [eV]	-3.199549	F^0_{sd} [eV]	1.446283	z_{sn} [bohr^{-1}]	5.993591
β_p [eV]	-4.451920	G^2_{sd} [eV]	1.680225	z_{pn} [bohr^{-1}]	6.145086
ζ_s [bohr^{-1}]	1.212539			z_{dn} [bohr^{-1}]	5.704031
ζ_p [bohr^{-1}]	1.276487	Nickel		g_{ss} [eV]	10.588832
g_{ss} [eV]	7.552804			g_{sp} [eV]	7.310023
g_{sp} [eV]	10.203146	U_{ss} [eV]	-47.620247	g_{pp} [eV]	9.281609
g_{pp} [eV]	12.862153	U_{pp} [eV]	-32.878408	g_{p2} [eV]	8.954081
g_{p2} [eV]	13.602858	U_{dd} [eV]	-93.026395	h_{sp} [eV]	0.779112
h_{sp} [eV]	1.501452	β_s [eV]	-9.151521		
a [none]	0.164180	β_p [eV]	-8.086696	Tellurium	

Table 3 (continued)

Element	Parameters	Element	Parameters	Element	Parameters
b [Å^{-2}]	1.704828	β_d [eV]	-8.655910		
c [Å]	1.785591	ζ_s [bohr^{-1}]	1.591828	U_{ss} [eV]	-114.733316
		ζ_p [bohr^{-1}]	2.304739	U_{pp} [eV]	-50.096389
Boron		ζ_d [bohr^{-1}]	2.514761	β_s [eV]	-70.001062
		z_{sn} [bohr^{-1}]	0.746470	β_p [eV]	-6.151642
U_{ss} [eV]	-25.967679	z_{pn} [bohr^{-1}]	0.753327	ζ_s [bohr^{-1}]	2.769862
U_{pp} [eV]	-19.115864	z_{dn} [bohr^{-1}]	1.461345	ζ_p [bohr^{-1}]	1.731319
β_s [eV]	-4.959706	g_{ss} [eV]	4.080876	g_{ss} [eV]	7.030626
β_p [eV]	-4.656753	g_{sp} [eV]	4.099452	g_{sp} [eV]	12.601389
ζ_s [bohr^{-1}]	1.634174	g_{pp} [eV]	4.487545	g_{pp} [eV]	7.883479
ζ_p [bohr^{-1}]	1.479195	g_{p2} [eV]	3.933771	g_{p2} [eV]	6.973163
g_{ss} [eV]	8.179341	h_{sp} [eV]	0.993498	h_{sp} [eV]	5.000826
g_{sp} [eV]	7.294021	ρ(core) [bohr^{-1}]	1.586979		
g_{pp} [eV]	7.829395	F^0_{sd} [eV]	4.651664	Iodine	
g_{p2} [eV]	6.401072	G^2_{sd} [eV]	1.880502		
h_{sp} [eV]	1.252845			U_{ss} [eV]	-59.973232
		Copper		U_{pp} [eV]	-56.459835
Carbon				U_{dd} [eV]	-28.822603
		U_{ss} [eV]	-97.002205	β_s [eV]	-30.522481
U_{ss} [eV]	-51.089653	U_{pp} [eV]	-1.000000	β_p [eV]	-5.942120
U_{pp} [eV]	-39.937920	U_{dd} [eV]	-110.442592	β_d [eV]	-7.676107
β_s [eV]	-15.385236	β_s [eV]	-9.369508	ζ_s [bohr^{-1}]	4.498653
β_p [eV]	-7.471929	β_p [eV]	-0.100000	ζ_p [bohr^{-1}]	1.917072
ζ_s [bohr^{-1}]	2.047558	β_d [eV]	-16.982092	ζ_d [bohr^{-1}]	1.875175
ζ_p [bohr^{-1}]	1.702841	ζ_s [bohr^{-1}]	1.669096	z_{sn} [bohr^{-1}]	9.135244
g_{ss} [eV]	13.335519	ζ_p [bohr^{-1}]	3.000000	z_{pn} [bohr^{-1}]	6.888191
g_{sp} [eV]	11.528134	ζ_d [bohr^{-1}]	2.734990	z_{dn} [bohr^{-1}]	3.791523
g_{pp} [eV]	10.778326	z_{sn} [bohr^{-1}]	1.899598	g_{ss} [eV]	7.234759
g_{p2} [eV]	9.486212	z_{pn} [bohr^{-1}]	3.000000	g_{sp} [eV]	9.154406
h_{sp} [eV]	0.717322	z_{dn} [bohr^{-1}]	1.484317	g_{pp} [eV]	9.877466
a [none]	0.046302	g_{ss} [eV]	10.384910	g_{p2} [eV]	8.035916
b [Å^{-2}]	2.100206	g_{sp} [eV]	12.145361	h_{sp} [eV]	5.004215
c [Å]	1.333959	g_{pp} [eV]	17.870905	a [none]	-0.035519
		g_{p2} [eV]	15.665592	b [Å^{-2}]	1.744389
Nitrogen		h_{sp} [eV]	2.037394	c [Å]	1.223844
		F^0_{sd} [eV]	9.848807		
U_{ss} [eV]	-57.784823	G^2_{sd} [eV]	9.847577	Xenon	
U_{pp} [eV]	-49.893036				
β_s [eV]	-17.979377	Zinc		U_{ss} [eV]	-18.270227
β_p [eV]	-15.055017			U_{pp} [eV]	-167.163063
ζ_s [bohr^{-1}]	2.380406	U_{ss} [eV]	-18.040862	β_s [eV]	-3.980622
ζ_p [bohr^{-1}]	1.999246	U_{pp} [eV]	-7.834895	β_p [eV]	-38.822792
g_{ss} [eV]	12.357026	β_s [eV]	-13.276583	ζ_s [bohr^{-1}]	2.759787
g_{sp} [eV]	9.636190	β_p [eV]	1.479642	ζ_p [bohr^{-1}]	1.977446
g_{pp} [eV]	12.570756	ζ_s [bohr^{-1}]	1.512875	g_{ss} [eV]	20.000252
g_{p2} [eV]	10.576425	ζ_p [bohr^{-1}]	1.789482	g_{sp} [eV]	4.175902
h_{sp} [eV]	2.871545	g_{ss} [eV]	8.707424	g_{pp} [eV]	2.305787
a [none]	-0.001436	g_{sp} [eV]	3.436116	g_{p2} [eV]	4.063220
b [Å^{-2}]	0.495196	g_{pp} [eV]	20.000041	h_{sp} [eV]	4.418843
c [Å]	1.704857	g_{p2} [eV]	6.782785		
		h_{sp} [eV]	0.662036	Cesium	
Oxygen					

Table 3 (continued)

Element	Parameters	Element	Parameters	Element	Parameters
		Gallium		U_{ss} [eV]	-3.748609
U_{ss} [eV]	-91.678761			U_{pp} [eV]	-2.348109
U_{pp} [eV]	-70.460949	U_{ss} [eV]	-30.600226	β_s [eV]	2.287838
β_s [eV]	-65.635137	U_{pp} [eV]	-21.032425	β_p [eV]	-5.908071
β_p [eV]	-21.622604	β_s [eV]	-10.808320	ζ_s [bohr^{-1}]	5.956008
ζ_s [bohr^{-1}]	5.421751	β_p [eV]	-4.185500	ζ_p [bohr^{-1}]	1.619485
ζ_p [bohr^{-1}]	2.270960	ζ_s [bohr^{-1}]	2.339067	g_{ss} [eV]	6.464751
g_{ss} [eV]	11.304042	ζ_p [bohr^{-1}]	1.729592	g_{sp} [eV]	4.004501
g_{sp} [eV]	15.807424	g_{ss} [eV]	10.354885	g_{pp} [eV]	13.775390
g_{pp} [eV]	13.618205	g_{sp} [eV]	7.993674	g_{p2} [eV]	12.912537
g_{p2} [eV]	10.332765	g_{pp} [eV]	6.090184	h_{sp} [eV]	1.026928
h_{sp} [eV]	5.010801	g_{p2} [eV]	6.299226		
a [none]	-0.017771	h_{sp} [eV]	1.295974	Barium	
b [Å^{-2}]	3.058310				
c [Å]	1.896435	Germanium		U_{ss} [eV]	-9.306985
				U_{pp} [eV]	-8.826713
Fluorine		U_{ss} [eV]	-32.747338	β_s [eV]	10.003125
		U_{pp} [eV]	-24.709016	β_p [eV]	-6.335160
U_{ss} [eV]	-140.225626	β_s [eV]	-14.854297	ζ_s [bohr^{-1}]	1.395379
U_{pp} [eV]	-98.778044	β_p [eV]	-2.591260	ζ_p [bohr^{-1}]	1.430139
β_s [eV]	-69.922593	ζ_s [bohr^{-1}]	2.546073	g_{ss} [eV]	3.600823
β_p [eV]	-30.448165	ζ_p [bohr^{-1}]	1.709130	g_{sp} [eV]	4.740579
ζ_s [bohr^{-1}]	6.043849	g_{ss} [eV]	7.518301	g_{pp} [eV]	3.345166
ζ_p [bohr^{-1}]	2.906722	g_{sp} [eV]	6.594443	g_{p2} [eV]	3.142783
g_{ss} [eV]	12.446818	g_{pp} [eV]	6.066801	h_{sp} [eV]	0.929429
g_{sp} [eV]	18.496082	g_{p2} [eV]	5.305947		
g_{pp} [eV]	8.417366	h_{sp} [eV]	0.290742	Lanthanum	
g_{p2} [eV]	12.179816				
h_{sp} [eV]	2.604382	Arsenic		U_{ss} [eV]	-19.641953
a [none]	-0.010792			U_{pp} [eV]	-22.059431
b [Å^{-2}]	6.004648	U_{ss} [eV]	-37.956965	U_{dd} [eV]	-22.638986
c [Å]	1.847724	U_{pp} [eV]	-38.453701	β_s [eV]	0.796727
		U_{dd} [eV]	-30.282658	β_p [eV]	-10.856056
Neon		β_s [eV]	-11.963725	β_d [eV]	-0.484922
		β_p [eV]	-7.340073	ζ_s [bohr^{-1}]	2.673780
U_{ss} [eV]	-2.978729	β_d [eV]	3.753005	ζ_p [bohr^{-1}]	1.248192
U_{pp} [eV]	-85.441118	ζ_s [bohr^{-1}]	2.926171	ζ_d [bohr^{-1}]	1.688562
β_s [eV]	-69.793475	ζ_p [bohr^{-1}]	1.765191	z_{sn} [bohr^{-1}]	1.617784
β_p [eV]	-33.261962	ζ_d [bohr^{-1}]	1.392142	z_{pn} [bohr^{-1}]	4.331620
ζ_s [bohr^{-1}]	6.000148	z_{sn} [bohr^{-1}]	2.006543	z_{dn} [bohr^{-1}]	2.285738
ζ_p [bohr^{-1}]	3.834528	z_{pn} [bohr^{-1}]	3.316832	g_{ss} [eV]	6.154440
g_{ss} [eV]	19.999574	z_{dn} [bohr^{-1}]	4.653440	g_{sp} [eV]	7.322704
g_{sp} [eV]	16.896951	g_{ss} [eV]	6.665030	g_{pp} [eV]	18.077465
g_{pp} [eV]	8.963560	g_{sp} [eV]	6.213867	g_{p2} [eV]	15.679057
g_{p2} [eV]	16.027799	g_{pp} [eV]	9.310836	h_{sp} [eV]	0.138601
h_{sp} [eV]	1.779280	g_{p2} [eV]	8.712542	ρ(core) [bohr^{-1}]	2.511701
		h_{sp} [eV]	0.280662	F^0_{sd} [eV]	8.856858
Sodium				G^2_{sd} [eV]	7.925585
		Selenium			
U_{ss} [eV]	-4.537153			Lutetium	
U_{pp} [eV]	-2.433015	U_{ss} [eV]	-32.671088		
β_s [eV]	0.244853	U_{pp} [eV]	-32.522220	U_{ss} [eV]	-15.954994
β_p [eV]	0.491998	β_s [eV]	2.636001	U_{pp} [eV]	-11.606213

Table 3 (continued)

Element	Parameters	Element	Parameters	Element	Parameters
ζ_s [bohr^{-1}]	0.686327	β_p [eV]	-9.557700	U_{dd} [eV]	-13.050056
ζ_p [bohr^{-1}]	0.950068	ζ_s [bohr^{-1}]	2.512366	β_s [eV]	-5.590778
g_{ss} [eV]	4.059972	ζ_p [bohr^{-1}]	2.007576	β_p [eV]	-0.937679
g_{sp} [eV]	7.061183	g_{ss} [eV]	5.522356	β_d [eV]	-7.737752
g_{pp} [eV]	9.283540	g_{sp} [eV]	2.907562	ζ_s [bohr^{-1}]	5.471741
g_{p2} [eV]	17.034978	g_{pp} [eV]	8.042391	ζ_p [bohr^{-1}]	1.712296
h_{sp} [eV]	0.640715	g_{p2} [eV]	6.735106	ζ_d [bohr^{-1}]	2.225892
a [none]	-1.026036	h_{sp} [eV]	3.095789	z_{sn} [bohr^{-1}]	1.632335
b [Å^{-2}]	2.014506			z_{pn} [bohr^{-1}]	4.033128
c [Å]	1.271202	Bromine		z_{dn} [bohr^{-1}]	0.921999
				g_{ss} [eV]	6.209796
Magnesium		U_{ss} [eV]	-45.834364	g_{sp} [eV]	7.379102
		U_{pp} [eV]	-50.293675	g_{pp} [eV]	16.831746
U_{ss} [eV]	-14.574226	U_{dd} [eV]	7.086738	g_{p2} [eV]	14.598613
U_{pp} [eV]	-7.583850	β_s [eV]	-32.131665	h_{sp} [eV]	0.209008
β_s [eV]	-9.604932	β_p [eV]	-9.514484	ρ(core) [bohr^{-1}]	2.743262
β_p [eV]	3.416908	β_d [eV]	-9.839124	F^0_{sd} [eV]	3.924927
ζ_s [bohr^{-1}]	1.310830	ζ_s [bohr^{-1}]	4.670684	G^2_{sd} [eV]	1.000946
ζ_p [bohr^{-1}]	1.388897	ζ_p [bohr^{-1}]	2.035626		
g_{ss} [eV]	7.115328	ζ_d [bohr^{-1}]	1.521031	Hafnium	
g_{sp} [eV]	3.253024	z_{sn} [bohr^{-1}]	3.094777		
g_{pp} [eV]	4.737311	z_{pn} [bohr^{-1}]	3.065764	U_{ss} [eV]	-22.375140
g_{p2} [eV]	8.428485	z_{dn} [bohr^{-1}]	2.820003	U_{pp} [eV]	-13.081670
h_{sp} [eV]	0.877379	g_{ss} [eV]	7.616791	U_{dd} [eV]	-20.637741
		g_{sp} [eV]	5.010425	β_s [eV]	-5.366351
Aluminium		g_{pp} [eV]	9.649216	β_p [eV]	-21.550119
		g_{p2} [eV]	8.343792	β_d [eV]	-3.884443
U_{ss} [eV]	-24.546778	h_{sp} [eV]	4.996553	ζ_s [bohr^{-1}]	3.085344
U_{pp} [eV]	-20.104434	a [none]	-0.004996	ζ_p [bohr^{-1}]	1.575819
U_{dd} [eV]	8.004394	b [Å^{-2}]	6.001292	ζ_d [bohr^{-1}]	1.840840
β_s [eV]	-18.375229	c [Å]	2.895153	z_{sn} [bohr^{-1}]	0.946927
β_p [eV]	-9.382700			z_{pn} [bohr^{-1}]	3.538911
β_d [eV]	-20.840474	Krypton		z_{dn} [bohr^{-1}]	0.940283
ζ_s [bohr^{-1}]	2.364264			g_{ss} [eV]	3.602338
ζ_p [bohr^{-1}]	1.749102	U_{ss} [eV]	8.535384	g_{sp} [eV]	4.293729
ζ_d [bohr^{-1}]	1.269384	U_{pp} [eV]	-80.484321	g_{pp} [eV]	14.769194
z_{sn} [bohr^{-1}]	4.742341	β_s [eV]	-2.727088	g_{p2} [eV]	12.809708
z_{pn} [bohr^{-1}]	4.669626	β_p [eV]	-16.142951	h_{sp} [eV]	0.011028
z_{dn} [bohr^{-1}]	7.131138	ζ_s [bohr^{-1}]	1.312248	F^0_{sd} [eV]	4.842900
g_{ss} [eV]	6.652155	ζ_p [bohr^{-1}]	4.491371	G^2_{sd} [eV]	4.386101
g_{sp} [eV]	7.459435	g_{ss} [eV]	19.999857		
g_{pp} [eV]	7.668857	g_{sp} [eV]	1.175304	Tantalum	
g_{p2} [eV]	6.673299	g_{pp} [eV]	9.174784		
h_{sp} [eV]	0.435060	g_{p2} [eV]	14.926948	U_{ss} [eV]	-39.009984
a [none]	1.002222	h_{sp} [eV]	0.299867	U_{pp} [eV]	1.163975
b [Å^{-2}]	1.517400			U_{dd} [eV]	-43.266315
c [Å]	0.659101	Rubidium		β_s [eV]	-17.199605
				β_p [eV]	-5.818839
Silicon		U_{ss} [eV]	-3.636505	β_d [eV]	-9.816794
		U_{pp} [eV]	-2.500671	ζ_s [bohr^{-1}]	4.578087
U_{ss} [eV]	-27.358058	β_s [eV]	9.998744	ζ_p [bohr^{-1}]	4.841244
U_{pp} [eV]	-20.490578	β_p [eV]	1.343004	ζ_d [bohr^{-1}]	1.838249

Table 3 (continued)

Element	Parameters	Element	Parameters	Element	Parameters
U_{dd} [eV]	-22.751900	ζ_s [bohr^{-1}]	5.510145	z_{sn} [bohr^{-1}]	1.741367
β_s [eV]	-8.686909	ζ_p [bohr^{-1}]	1.335170	z_{pn} [bohr^{-1}]	3.430157
β_p [eV]	-1.856482	g_{ss} [eV]	6.680824	z_{dn} [bohr^{-1}]	2.311198
β_d [eV]	-6.360627	g_{sp} [eV]	20.001098	g_{ss} [eV]	6.624580
ζ_s [bohr^{-1}]	1.752741	g_{pp} [eV]	5.068874	g_{sp} [eV]	7.805321
ζ_p [bohr^{-1}]	1.198413	g_{p2} [eV]	2.747860	g_{pp} [eV]	14.315323
ζ_d [bohr^{-1}]	2.128593	h_{sp} [eV]	3.602834	g_{p2} [eV]	12.416054
z_{sn} [bohr^{-1}]	8.388111			h_{sp} [eV]	0.577263
z_{pn} [bohr^{-1}]	1.843048	Strontium		F^0_{sd} [eV]	8.544427
z_{dn} [bohr^{-1}]	0.708600			G^2_{sd} [eV]	2.074254
g_{ss} [eV]	5.194805	U_{ss} [eV]	-10.427671		
g_{sp} [eV]	5.090534	U_{pp} [eV]	-9.943751	Tungsten	
g_{pp} [eV]	5.185150	β_s [eV]	-6.253108		
g_{p2} [eV]	4.769775	β_p [eV]	-9.844498	U_{ss} [eV]	-44.524950
h_{sp} [eV]	1.425012	ζ_s [bohr^{-1}]	2.197303	U_{pp} [eV]	-40.011500
a [none]	0.208571	ζ_p [bohr^{-1}]	1.730137	U_{dd} [eV]	-46.490410
b [Å^{-2}]	6.000483	g_{ss} [eV]	4.603664	β_s [eV]	-16.946460
c [Å]	1.185245	g_{sp} [eV]	5.716069	β_p [eV]	5.623170
		g_{pp} [eV]	7.334620	β_d [eV]	-2.947340
Phosphorus		g_{p2} [eV]	7.443088	ζ_s [bohr^{-1}]	2.664560
		h_{sp} [eV]	0.831527	ζ_p [bohr^{-1}]	1.624010
U_{ss} [eV]	-48.729905	a [none]	-0.012948	ζ_d [bohr^{-1}]	1.794400
U_{pp} [eV]	-40.354689	b [Å^{-2}]	6.000126	z_{sn} [bohr^{-1}]	1.498860
U_{dd} [eV]	-7.349246	c [Å]	3.011964	z_{pn} [bohr^{-1}]	1.965900
β_s [eV]	-14.583780			z_{dn} [bohr^{-1}]	1.876450
β_p [eV]	-11.744725	Yttrium		g_{ss} [eV]	5.702025
β_d [eV]	-20.099893			g_{sp} [eV]	6.323145
ζ_s [bohr^{-1}]	2.158033	U_{ss} [eV]	-14.247809	g_{pp} [eV]	8.204433
ζ_p [bohr^{-1}]	1.805343	U_{pp} [eV]	-14.817140	g_{p2} [eV]	7.115919
ζ_d [bohr^{-1}]	1.230358	U_{dd} [eV]	-16.394302	h_{sp} [eV]	1.319912
z_{sn} [bohr^{-1}]	6.042706	β_s [eV]	0.343336	F^0_{sd} [eV]	7.788180
z_{pn} [bohr^{-1}]	2.376473	β_p [eV]	-3.180807	G^2_{sd} [eV]	1.684940
z_{dn} [bohr^{-1}]	7.147750	β_d [eV]	-4.508957		
g_{ss} [eV]	8.758856	ζ_s [bohr^{-1}]	0.593368	Rhenium	
g_{sp} [eV]	8.483679	ζ_p [bohr^{-1}]	1.490422		
g_{pp} [eV]	8.662754	ζ_d [bohr^{-1}]	1.650893	U_{ss} [eV]	-41.291342
g_{p2} [eV]	7.734264	z_{sn} [bohr^{-1}]	0.902611	U_{pp} [eV]	-35.089592
h_{sp} [eV]	0.871681	z_{pn} [bohr^{-1}]	1.484400	U_{dd} [eV]	-44.178985
a [none]	-0.034320	z_{dn} [bohr^{-1}]	1.384238	β_s [eV]	3.830075
b [Å^{-2}]	6.001394	g_{ss} [eV]	4.046733	β_p [eV]	-1.638530
c [Å]	2.296737	g_{sp} [eV]	4.726277	β_d [eV]	-1.414411
		g_{pp} [eV]	7.278752	ζ_s [bohr^{-1}]	2.411839
Sulfur		g_{p2} [eV]	6.343281	ζ_p [bohr^{-1}]	1.815351
		h_{sp} [eV]	0.679228	ζ_d [bohr^{-1}]	2.522766
U_{ss} [eV]	-47.530706	ρ(core) [bohr^{-1}]	2.773703	z_{sn} [bohr^{-1}]	1.680823
U_{pp} [eV]	-39.191045	F^0_{sd} [eV]	4.972716	z_{pn} [bohr^{-1}]	1.331218
U_{dd} [eV]	-46.306944	G^2_{sd} [eV]	5.016364	z_{dn} [bohr^{-1}]	1.490623
β_s [eV]	-13.827440			g_{ss} [eV]	6.394256
β_p [eV]	-7.664613	Zirconium		g_{sp} [eV]	5.555571
β_d [eV]	-9.986172			g_{pp} [eV]	5.555669
ζ_s [bohr^{-1}]	2.192844	U_{ss} [eV]	-20.008884	g_{p2} [eV]	4.818577
ζ_p [bohr^{-1}]	1.841078	U_{pp} [eV]	-14.559692	h_{sp} [eV]	1.220913

Table 3 (continued)

Element	Parameters	Element	Parameters	Element	Parameters
ζ_d [bohr^{-1}]	3.109401	U_{dd} [eV]	-21.302657	F^0_{sd} [eV]	5.442818
z_{sn} [bohr^{-1}]	0.479722	β_s [eV]	9.551952	G^2_{sd} [eV]	2.376279
z_{pn} [bohr^{-1}]	1.015507	β_p [eV]	-4.551915		
z_{dn} [bohr^{-1}]	4.317470	β_d [eV]	-3.213274	Osmium	
g_{ss} [eV]	9.170350	ζ_s [bohr^{-1}]	1.692590		
g_{sp} [eV]	5.944296	ζ_p [bohr^{-1}]	1.694916	U_{ss} [eV]	-26.434080
g_{pp} [eV]	8.165473	ζ_d [bohr^{-1}]	1.567392	U_{pp} [eV]	-48.739500
g_{p2} [eV]	7.301878	z_{sn} [bohr^{-1}]	1.189109	U_{dd} [eV]	-55.837880
h_{sp} [eV]	5.005404	z_{pn} [bohr^{-1}]	0.809092	β_s [eV]	-12.508730
a [none]	-0.036928	z_{dn} [bohr^{-1}]	1.190249	β_p [eV]	0.846880
b [Å^{-2}]	1.795067	g_{ss} [eV]	5.331208	β_d [eV]	5.164360
c [Å]	2.082618	g_{sp} [eV]	4.150579	ζ_s [bohr^{-1}]	3.031000
		g_{pp} [eV]	3.967381	ζ_p [bohr^{-1}]	1.593960
Chlorine		g_{p2} [eV]	3.457490	ζ_d [bohr^{-1}]	1.775570
		h_{sp} [eV]	0.743676	z_{sn} [bohr^{-1}]	1.844700
U_{ss} [eV]	-61.389930	F^0_{sd} [eV]	5.010704	z_{pn} [bohr^{-1}]	1.564220
U_{pp} [eV]	-54.482801	G^2_{sd} [eV]	2.943652	z_{dn} [bohr^{-1}]	1.770010
U_{dd} [eV]	-38.258155			g_{ss} [eV]	7.017683
β_s [eV]	-2.367988	Niobium		g_{sp} [eV]	6.384200
β_p [eV]	-13.802139			g_{pp} [eV]	6.528073
β_d [eV]	-4.037751	U_{ss} [eV]	-31.269298	g_{p2} [eV]	5.661968
ζ_s [bohr^{-1}]	2.637050	U_{pp} [eV]	-20.151277	h_{sp} [eV]	1.508926
ζ_p [bohr^{-1}]	2.118146	U_{dd} [eV]	-35.893116	F^0_{sd} [eV]	2.021170
ζ_d [bohr^{-1}]	1.324033	β_s [eV]	-12.045244	G^2_{sd} [eV]	1.392130
z_{sn} [bohr^{-1}]	0.956297	β_p [eV]	1.465762		
z_{pn} [bohr^{-1}]	2.464067	β_d [eV]	-5.920160	Iridium	
z_{dn} [bohr^{-1}]	6.410325	ζ_s [bohr^{-1}]	2.355562		
g_{ss} [eV]	11.142654	ζ_p [bohr^{-1}]	1.386907	U_{ss} [eV]	-29.703974
g_{sp} [eV]	7.487881	ζ_d [bohr^{-1}]	1.977324	U_{pp} [eV]	-38.210924
g_{pp} [eV]	9.551886	z_{sn} [bohr^{-1}]	1.490754	U_{dd} [eV]	-32.538202
g_{p2} [eV]	8.128436	z_{pn} [bohr^{-1}]	0.892760	β_s [eV]	-10.943427
h_{sp} [eV]	5.004267	z_{dn} [bohr^{-1}]	1.443837	β_p [eV]	2.908880
a [none]	-0.013213	g_{ss} [eV]	6.683592	β_d [eV]	-3.791731
b [Å^{-2}]	3.687022	g_{sp} [eV]	4.685339	ζ_s [bohr^{-1}]	1.500907
c [Å]	2.544635	g_{pp} [eV]	4.377647	ζ_p [bohr^{-1}]	4.106373
		g_{p2} [eV]	3.815028	ζ_d [bohr^{-1}]	2.676047
Argon		h_{sp} [eV]	0.650679	z_{sn} [bohr^{-1}]	0.927246
		F^0_{sd} [eV]	6.550674	z_{pn} [bohr^{-1}]	3.191892
U_{ss} [eV]	-7.797931	G^2_{sd} [eV]	1.065577	z_{dn} [bohr^{-1}]	0.662007
U_{pp} [eV]	-83.211487			g_{ss} [eV]	3.527467
β_s [eV]	-8.839842	Molybdenum		g_{sp} [eV]	4.203820
β_p [eV]	-28.427303			g_{pp} [eV]	13.320955
ζ_s [bohr^{-1}]	6.000272	U_{ss} [eV]	-53.467728	g_{p2} [eV]	11.553612
ζ_p [bohr^{-1}]	5.949170	U_{pp} [eV]	-35.291951	h_{sp} [eV]	0.018501
g_{ss} [eV]	17.858776	U_{dd} [eV]	-55.836977	F^0_{sd} [eV]	2.627170
g_{sp} [eV]	4.168451	β_s [eV]	-0.189344	G^2_{sd} [eV]	2.996029
g_{pp} [eV]	11.852500	β_p [eV]	7.017762		
g_{p2} [eV]	15.669543	β_d [eV]	-10.941126	Platinum	
h_{sp} [eV]	4.574549	ζ_s [bohr^{-1}]	1.060429		
		ζ_p [bohr^{-1}]	1.350412	U_{ss} [eV]	-73.516173
Potassium		ζ_d [bohr^{-1}]	1.827152	U_{pp} [eV]	-68.320056
		z_{sn} [bohr^{-1}]	1.912995	U_{dd} [eV]	-76.598873

Table 3 (continued)

Element	Parameters	Element	Parameters	Element	Parameters
U_{ss} [eV]	-3.801108	z_{pn} [bohr^{-1}]	1.355055	β_s [eV]	1.151418
U_{pp} [eV]	-3.339656	z_{dn} [bohr^{-1}]	1.876231	β_p [eV]	3.298694
β_s [eV]	-8.755195	g_{ss} [eV]	8.576652	β_d [eV]	-18.044737
β_p [eV]	-1.788061	g_{sp} [eV]	6.888293	ζ_s [bohr^{-1}]	2.301264
ζ_s [bohr^{-1}]	6.000478	g_{pp} [eV]	6.644509	ζ_p [bohr^{-1}]	1.662404
ζ_p [bohr^{-1}]	1.127503	g_{p2} [eV]	5.790552	ζ_d [bohr^{-1}]	3.168852
g_{ss} [eV]	3.369251	h_{sp} [eV]	1.317368	z_{sn} [bohr^{-1}]	2.270699
g_{sp} [eV]	6.129351	F^0_{sd} [eV]	10.000608	z_{pn} [bohr^{-1}]	1.949896
g_{pp} [eV]	0.999505	G^2_{sd} [eV]	1.216752	z_{dn} [bohr^{-1}]	1.713856
g_{p2} [eV]	18.999148			g_{ss} [eV]	8.638286
h_{sp} [eV]	0.300325	Technetium		g_{sp} [eV]	7.922254
a [none]	0.157519			g_{pp} [eV]	8.137643
b [Å^{-2}]	6.000566	U_{ss} [eV]	-41.850292	g_{p2} [eV]	7.057990
c [Å]	2.047539	U_{pp} [eV]	-34.910293	h_{sp} [eV]	1.892617
		U_{dd} [eV]	-45.530412	F^0_{sd} [eV]	7.098591
Calcium		β_s [eV]	-2.791024	G^2_{sd} [eV]	4.484183
		β_p [eV]	-8.086697		
U_{ss} [eV]	-10.770058	β_d [eV]	-5.724335	Gold	
U_{pp} [eV]	-9.754177	ζ_s [bohr^{-1}]	1.956245		
β_s [eV]	-4.343881	ζ_p [bohr^{-1}]	6.006299	U_{ss} [eV]	-95.041846
β_p [eV]	-1.296612	ζ_d [bohr^{-1}]	1.767360	U_{pp} [eV]	-63.890158
ζ_s [bohr^{-1}]	1.528258	z_{sn} [bohr^{-1}]	1.411033	U_{dd} [eV]	-88.066087
ζ_p [bohr^{-1}]	2.060094	z_{pn} [bohr^{-1}]	1.141313	β_s [eV]	-7.479625
g_{ss} [eV]	5.725773	z_{dn} [bohr^{-1}]	1.159312	β_p [eV]	3.664356
g_{sp} [eV]	4.781065	g_{ss} [eV]	6.326174	β_d [eV]	-61.715468
g_{pp} [eV]	7.172103	g_{sp} [eV]	5.587138	ζ_s [bohr^{-1}]	1.814169
g_{p2} [eV]	7.431876	g_{pp} [eV]	5.596426	ζ_p [bohr^{-1}]	1.618657
h_{sp} [eV]	1.240572	g_{p2} [eV]	4.877169	ζ_d [bohr^{-1}]	5.053167
a [none]	-0.025275	h_{sp} [eV]	1.258989	z_{sn} [bohr^{-1}]	2.444680
b [Å^{-2}]	0.500017	F^0_{sd} [eV]	5.434886	z_{pn} [bohr^{-1}]	7.014990
c [Å]	2.329051	G^2_{sd} [eV]	1.106875	z_{dn} [bohr^{-1}]	1.777089
				g_{ss} [eV]	9.300152
Scandium		Ruthenium		g_{sp} [eV]	11.073443
				g_{pp} [eV]	29.276168
U_{ss} [eV]	-15.544461	U_{ss} [eV]	-44.901521	g_{p2} [eV]	25.391984
U_{pp} [eV]	-18.646295	U_{pp} [eV]	-41.424409	h_{sp} [eV]	0.144384
U_{dd} [eV]	-16.069444	U_{dd} [eV]	-37.934514	F^0_{sd} [eV]	8.827257
β_s [eV]	-8.620944	β_s [eV]	-12.859508	G^2_{sd} [eV]	4.915625
β_p [eV]	3.075948	β_p [eV]	-8.475518		
β_d [eV]	-9.768661	β_d [eV]	-3.830797	Mercury	
ζ_s [bohr^{-1}]	1.402469	ζ_s [bohr^{-1}]	1.459195		
ζ_p [bohr^{-1}]	1.345196	ζ_p [bohr^{-1}]	5.537201	U_{ss} [eV]	-17.608732
ζ_d [bohr^{-1}]	1.859012	ζ_d [bohr^{-1}]	2.093164	U_{pp} [eV]	-18.369417
z_{sn} [bohr^{-1}]	0.848418	z_{sn} [bohr^{-1}]	0.984449	β_s [eV]	-3.045239
z_{pn} [bohr^{-1}]	2.451729	z_{pn} [bohr^{-1}]	4.586613	β_p [eV]	-5.693556
z_{dn} [bohr^{-1}]	0.789372	z_{dn} [bohr^{-1}]	0.765332	ζ_s [bohr^{-1}]	2.104896
g_{ss} [eV]	4.638215	g_{ss} [eV]	4.413643	ζ_p [bohr^{-1}]	1.516293
g_{sp} [eV]	5.739164	g_{sp} [eV]	5.356996	g_{ss} [eV]	6.372822
g_{pp} [eV]	14.604872	g_{pp} [eV]	22.490448	g_{sp} [eV]	10.143176
g_{p2} [eV]	12.802595	g_{p2} [eV]	19.599957	g_{pp} [eV]	10.397393
h_{sp} [eV]	0.193835	h_{sp} [eV]	0.008058	g_{p2} [eV]	14.794056
ρ(core) [bohr^{-1}]	3.173734	F^0_{sd} [eV]	5.917404	h_{sp} [eV]	0.926128

Table 3 (continued)

Element	Parameters	Element	Parameters	Element	Parameters
F^0_{sd} [eV]	4.798313	G^2_{sd} [eV]	5.859738		
G^2_{sd} [eV]	5.380136			Thallium	
		Rhodium			
Titanium				U_{ss} [eV]	-29.518621
		U_{ss} [eV]	-20.513756	U_{pp} [eV]	-29.826907
U_{ss} [eV]	-25.507973	U_{pp} [eV]	-40.045431	β_s [eV]	-7.230170
U_{pp} [eV]	-17.260909	U_{dd} [eV]	-35.818492	β_p [eV]	-7.575544
U_{dd} [eV]	-23.809486	β_s [eV]	-8.222141	ζ_s [bohr^{-1}]	3.335883
β_s [eV]	3.389142	β_p [eV]	-15.556691	ζ_p [bohr^{-1}]	1.766141
β_p [eV]	-3.355350	β_d [eV]	-13.396182	g_{ss} [eV]	5.015118
β_d [eV]	-1.842829	ζ_s [bohr^{-1}]	1.324919	g_{sp} [eV]	13.932049
ζ_s [bohr^{-1}]	5.324777	ζ_p [bohr^{-1}]	4.306111	g_{pp} [eV]	10.495551
ζ_p [bohr^{-1}]	1.164068	ζ_d [bohr^{-1}]	2.901406	g_{p2} [eV]	10.526198
ζ_d [bohr^{-1}]	1.418280	z_{sn} [bohr^{-1}]	0.809923	h_{sp} [eV]	0.293760
z_{sn} [bohr^{-1}]	1.045904	z_{pn} [bohr^{-1}]	6.898259		
z_{pn} [bohr^{-1}]	1.076844	z_{dn} [bohr^{-1}]	0.643134	Lead	
z_{dn} [bohr^{-1}]	0.717945	g_{ss} [eV]	3.631179		
g_{ss} [eV]	5.717851	g_{sp} [eV]	4.407820	U_{ss} [eV]	-35.038145
g_{sp} [eV]	5.800015	g_{pp} [eV]	33.825599	U_{pp} [eV]	-25.413401
g_{pp} [eV]	6.414726	g_{p2} [eV]	29.478305	β_s [eV]	-8.323792
g_{p2} [eV]	5.623133	h_{sp} [eV]	0.000092	β_p [eV]	-2.237891
h_{sp} [eV]	1.403732	F^0_{sd} [eV]	1.775497	ζ_s [bohr^{-1}]	2.368901
F^0_{sd} [eV]	6.560562	G^2_{sd} [eV]	1.851571	ζ_p [bohr^{-1}]	1.685246
G^2_{sd} [eV]	3.396235			g_{ss} [eV]	5.254128
		Palladium		g_{sp} [eV]	7.061016
Vanadium				g_{pp} [eV]	6.818551
		U_{ss} [eV]	-76.140196	g_{p2} [eV]	5.603019
U_{ss} [eV]	-32.162276	U_{pp} [eV]	-21.073362	h_{sp} [eV]	1.018819
U_{pp} [eV]	-21.572501	U_{dd} [eV]	-85.325301	a [none]	-0.239463
U_{dd} [eV]	-34.506245	β_s [eV]	-8.038245	b [Å^{-2}]	5.444338
β_s [eV]	-1.211330	β_p [eV]	0.740037	c [Å]	1.613682
β_p [eV]	0.740746	β_d [eV]	-2.394498		
β_d [eV]	3.153669	ζ_s [bohr^{-1}]	1.658503	Bismuth	
ζ_s [bohr^{-1}]	1.974330	ζ_p [bohr^{-1}]	1.156718		
ζ_p [bohr^{-1}]	1.063106	ζ_d [bohr^{-1}]	2.219861	U_{ss} [eV]	-42.409177
ζ_d [bohr^{-1}]	1.394806	z_{sn} [bohr^{-1}]	1.794085	U_{pp} [eV]	-36.393746
z_{sn} [bohr^{-1}]	1.094426	z_{pn} [bohr^{-1}]	6.158778	β_s [eV]	-34.951578
z_{pn} [bohr^{-1}]	0.755378	z_{dn} [bohr^{-1}]	1.630913	β_p [eV]	-7.359060
z_{dn} [bohr^{-1}]	1.099367	g_{ss} [eV]	8.043535	ζ_s [bohr^{-1}]	3.702377
g_{ss} [eV]	5.983116	g_{sp} [eV]	9.755042	ζ_p [bohr^{-1}]	1.872327
g_{sp} [eV]	4.736769	g_{pp} [eV]	30.199556	g_{ss} [eV]	5.851803
g_{pp} [eV]	4.499763	g_{p2} [eV]	26.318284	g_{sp} [eV]	6.790583
g_{p2} [eV]	3.944481	h_{sp} [eV]	0.086121	g_{pp} [eV]	8.389442
h_{sp} [eV]	0.901105	F^0_{sd} [eV]	8.004447	g_{p2} [eV]	7.724219
F^0_{sd} [eV]	6.810021	G^2_{sd} [eV]	2.613148	h_{sp} [eV]	0.295606
G^2_{sd} [eV]	1.831407				
		Silver			
Chromium					
		U_{ss} [eV]	-25.484137		
U_{ss} [eV]	-34.864339	U_{pp} [eV]	-36.116023		
U_{pp} [eV]	-26.978615	U_{dd} [eV]	-35.668272		
U_{dd} [eV]	-54.431036	β_s [eV]	-6.129623		
β_s [eV]	-5.122615	β_p [eV]	1.004115		

Table 3 (continued)

Element	Parameters	Element	Parameters
β_p [eV]	3.926711	β_d [eV]	-69.238347
β_d [eV]	-4.230550	ζ_s [bohr^{-1}]	1.994004
ζ_s [bohr^{-1}]	3.283460	ζ_p [bohr^{-1}]	0.681817
ζ_p [bohr^{-1}]	1.029394	ζ_d [bohr^{-1}]	6.007328
ζ_d [bohr^{-1}]	1.623119	z_{sn} [bohr^{-1}]	0.695514
z_{sn} [bohr^{-1}]	1.619853	z_{pn} [bohr^{-1}]	4.729949
z_{pn} [bohr^{-1}]	0.848266	z_{dn} [bohr^{-1}]	0.506522
z_{dn} [bohr^{-1}]	1.405015	g_{ss} [eV]	3.118242
g_{ss} [eV]	8.855572	g_{sp} [eV]	3.785152
g_{sp} [eV]	5.588631	g_{pp} [eV]	23.193295
g_{pp} [eV]	5.053094	g_{p2} [eV]	20.212474
g_{p2} [eV]	4.429530	h_{sp} [eV]	0.000432
h_{sp} [eV]	0.648039	F^0_{sd} [eV]	1.938327
F^0_{sd} [eV]	6.150136	G^2_{sd} [eV]	1.071901
G^2_{sd} [eV]	2.000300		
		Cadmium	
Manganese			
		U_{ss} [eV]	-14.645792
U_{ss} [eV]	-51.460000	U_{pp} [eV]	-9.318664
U_{pp} [eV]	-37.543990	β_s [eV]	-11.613183
U_{dd} [eV]	-47.655370	β_p [eV]	1.663178
β_s [eV]	-4.185290	ζ_s [bohr^{-1}]	1.384108
β_p [eV]	-3.479630	ζ_p [bohr^{-1}]	1.957413
β_d [eV]	-13.473190	g_{ss} [eV]	6.677284
ζ_s [bohr^{-1}]	2.131680	g_{sp} [eV]	5.953373
ζ_p [bohr^{-1}]	1.525880	g_{pp} [eV]	18.729843
ζ_d [bohr^{-1}]	2.607800	g_{p2} [eV]	9.917452
z_{sn} [bohr^{-1}]	1.132450	h_{sp} [eV]	0.825192
z_{pn} [bohr^{-1}]	1.390740		
z_{dn} [bohr^{-1}]	0.962550		
g_{ss} [eV]	6.190990		
g_{sp} [eV]	6.757427		
g_{pp} [eV]	8.284594		
g_{p2} [eV]	7.262255		
h_{sp} [eV]	1.520518		
F^0_{sd} [eV]	7.690920		
G^2_{sd} [eV]	1.105330		

covalent bonds. The angles subtended by such bonds are often determined largely by the electronic structure of the atom. Information on the accuracy of prediction of molecular electronic structure can also be inferred from the AUE of dipole moments, Table 8, and ionization potentials, Table 9.

Comparison of the accuracy of PM6 with the other NDDO methods PM5, PM3, and AM1, was made more complicated by the fact that different sets of elements were available in each method. To allow a simple comparison, therefore, average unsigned errors (AUE) for the four common properties for various subsets are presented in Tables 10, 11, 12, 13 and 14. To ensure a valid comparison the same number of data were used in each method, except for AM1 in "whole of main group", where data for cadmium and boron were not used.

*Comparison with AM1**

Winget, et al. [15], developed AM1* parameters for P, S, and Cl, in which Voityuk's diatomic parameters were used for all atom-pairs involving P, S, and Cl with H, C, N, O, F, P, S, Cl and Mo. In the AM1* method, all parameters for elements other than the ones being optimized are held constant at the AM1 values. As such, AM1* could be regarded as a hybrid method: parameters for a few individual elements are re-optimized, in this case with some changes in the set of approximations, while holding the parameters for the other methods constant at their AM1 values. Tables comparing individual P, S, and Cl species calculated with AM1* and PM6 are given in the supplementary material. A summary of the statistical analysis is given in Table 15. Winget et al. also reported AM1* parameters for titanium and zirconium [15]. These parameters were not used in the comparison given here because the set of approximations used was incompatible with the set used in PM6.

Comparison with RM1

In 2006, ten elements, H, C, N, O, F, P, S, Cl, Br, and I, that had been parameterized at the AM1 level were re-parameterized [35]; the result was a new method, RM1. No changes were made to the set of approximations used, so that, for example, P, S, Cl, Br, and I used only the *s-p* basis set. That is, RM1 was functionally identical to AM1. A statistical analysis showed that RM1 was more accurate than any of the other NDDO methods, and therefore was the method of choice for modeling organic compounds. An indication of the effect of the current changes to the set of approximations can be obtained by comparing the AUE for PM6 and RM1 in Tables 10, 11, 12, 13 and 14.

Voityuk reported the parameterization of molybdenum [14] at the AM1* level. These parameters were added to the standard AM1 parameters and were used in the analysis.

Comparison with high-level methods

A comparison of PM6, HF 6–31G(d) and B3LYP 6–31G(d) errors in predicted ΔH_f for 1373 compounds is given in the supplementary material. Only compounds containing the elements H, C, N, O, F, P, S, Cl, and Br were considered, these being the more important elements in biochemistry. Ab-initio ΔH_f were obtained from the calculated total energies by the addition of a simple atomic correction and conversion from atomic units to kcal mol^{-1}. No allowance was made for thermal population effects, zero point energies, etc., the assumption being made that such effects could be absorbed into the atomic corrections.

A statistical analysis of errors in thermochemical predictions for the three methods is given in Table 16. A

Table 4 Diatomic core–core parameters

		α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}
H	H	3.540942	2.243587	Cu	Cl	2.776531	0.139065	Te	P	1.453718	1.109289
				Cu	Cu	3.616846	5.184376	Te	S	1.830170	0.943925
He	H	2.989881	2.371199					Te	Cl	1.300260	0.285478
He	He	3.783559	3.450900	Zn	H	1.987891	3.109193	Te	Zn	1.218929	1.756070
				Zn	C	1.802327	0.991465	Te	Ge	2.342372	7.019049
Li	H	2.136265	2.191985	Zn	N	1.844579	0.952476	Te	As	1.189253	0.685774
Li	He	3.112403	9.273676	Zn	O	2.335054	2.265313	Te	Se	1.566008	1.187826
Li	Li	4.714674	16.116384	Zn	F	2.410021	1.225545	Te	Br	1.250940	0.394202
				Zn	Si	1.832058	3.783905	Te	Cd	1.307262	1.085919
Be	H	2.475418	2.562831	Zn	P	1.220480	0.581530	Te	In	1.540988	2.039582
Be	He	3.306702	12.544878	Zn	S	1.455000	0.648000	Te	Sn	1.763941	2.951976
Be	Li	2.236728	3.287165	Zn	Cl	1.625176	0.721351	Te	Te	1.164978	0.642486
Be	Be	1.499907	0.238633	Zn	Ca	1.119180	1.240290				
				Zn	Zn	0.929000	0.465000	I	H	2.139913	0.981898
B	H	2.615231	1.321394					I	He	2.172984	1.630721
B	He	3.163140	1.974170	Ga	H	1.847350	1.386652	I	Li	2.121251	4.168599
B	Li	3.759397	7.886018	Ga	C	2.325410	1.962990	I	Be	2.288023	2.351898
B	Be	1.888998	1.151792	Ga	N	2.121820	1.188338	I	B	2.667605	3.161385
B	B	3.318624	3.593619	Ga	O	2.348347	1.523644	I	C	2.068710	0.810156
				Ga	F	2.679869	1.416942	I	N	1.677518	0.264903
C	H	1.027806	0.216506	Ga	Si	1.913780	1.002290	I	O	2.288919	0.866204
C	He	3.042705	3.213971	Ga	P	2.979650	0.500000	I	F	2.203580	0.392425
C	Li	3.241874	16.180002	Ga	S	2.232108	2.456284	I	Ne	2.414415	1.503568
C	Be	4.212882	25.035879	Ga	Cl	2.024710	1.186661	I	Na	1.403090	1.986112
C	B	2.919007	1.874859	Ga	Ga	1.334643	1.198394	I	Mg	2.045137	3.276914
C	C	2.613713	0.813510					I	Al	1.816068	2.929080
				Ge	H	2.206793	1.733226	I	Si	1.559579	0.700299
N	H	0.969406	0.175506	Ge	C	2.257469	1.297510	I	P	2.131593	3.047207
N	He	2.814339	1.077861	Ge	N	1.988226	0.637506	I	S	1.855110	0.709929
N	Li	2.640623	2.823403	Ge	O	2.139413	0.826964	I	Cl	1.574161	0.310474
N	Be	2.580895	1.740605	Ge	F	2.384777	0.651977	I	Ar	1.576587	0.305367
N	B	2.477004	0.952882	Ge	Si	0.299721	0.178680	I	K	1.539714	4.824353
N	C	2.686108	0.859949	Ge	P	2.469291	5.616349	I	Ca	2.196490	7.689921
N	N	2.574502	0.675313	Ge	S	2.024588	1.160957	I	Sc	1.814884	3.114282
				Ge	Cl	1.771228	0.545239	I	Ti	1.933469	2.426747
O	H	1.260942	0.192295	Ge	Mn	2.382834	2.255151	I	V	2.683520	6.198112
O	He	3.653775	6.684525	Ge	Co	2.852610	2.151850	I	Cr	2.634224	2.598590
O	Li	2.584442	1.968598	Ge	Ge	2.019000	3.023000	I	Mn	2.266600	1.193410
O	Be	3.051867	3.218155					I	Fe	1.912829	0.532622
O	B	2.695351	1.269801	As	H	1.993527	1.090589	I	Co	3.235204	1.105239
O	C	2.889607	0.990211	As	C	1.855069	0.579098	I	Ni	1.085343	0.017459
O	N	2.784292	0.764756	As	N	1.496543	0.273337	I	Cu	0.834305	0.006781
O	O	2.623998	0.535112	As	O	2.003950	0.701614	I	Zn	1.394762	0.976607

Table 4 (continued)

		α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}
				As	F	2.012583	0.402628	I	Ga	1.671729	1.252168
F	H	3.136740	0.815802	As	Al	1.152786	1.003580	I	Ge	1.817425	1.323267
F	He	2.856543	0.745107	As	Si	1.915600	1.430706	I	As	1.245262	0.310824
F	Li	3.043901	1.975985	As	S	1.954368	1.033784	I	Br	1.579376	0.483054
F	Be	3.726923	3.882993	As	Cl	1.691070	0.454433	I	Kr	1.238574	0.201136
F	B	2.823837	0.862761	As	Ti	1.932911	1.581317	I	Rb	1.432675	4.092446
F	C	3.027600	0.732968	As	Co	3.368140	1.675240	I	Sr	1.262042	2.103941
F	N	2.856646	0.635854	As	Zn	1.459130	3.156571	I	Y	1.279110	1.021402
F	O	3.015444	0.674251	As	Ga	1.730977	1.686298	I	Zr	1.995182	4.513943
F	F	3.175759	0.681343	As	As	1.588264	0.737307	I	Nb	1.967251	2.399298
								I	Mo	0.948461	0.124695
Ne	H	5.999680	5.535021	Se	H	2.035068	0.847998	I	Tc	1.292312	0.110594
Ne	He	3.677758	1.960924	Se	C	2.387118	1.114787	I	Ru	3.953203	7.837710
Ne	Li	2.193666	0.704958	Se	N	1.937764	0.482840	I	Rh	3.708170	2.357944
Ne	Be	1.316588	0.392628	Se	O	2.484263	0.955161	I	Pd	5.144544	3.522017
Ne	B	2.756190	2.764140	Se	F	2.302180	0.444806	I	Ag	2.593161	0.048904
Ne	C	3.441188	5.468780	Se	Si	1.529817	0.518227	I	Cd	0.996238	0.396784
Ne	N	4.426370	29.999609	Se	P	1.048183	0.292052	I	In	2.351758	5.947821
Ne	O	2.889587	0.763899	Se	S	1.479606	0.391721	I	Sn	1.855633	1.783163
Ne	F	3.675611	2.706754	Se	Cl	2.128861	0.981067	I	Sb	1.155315	0.318190
Ne	Ne	3.974567	2.794830	Se	Mn	2.648038	2.180720	I	Te	1.493951	1.101116
				Se	Co	2.523450	2.202410	I	I	1.519925	0.510542
Na	H	0.500326	0.207831	Se	Zn	1.186242	0.511594				
Na	He	1.703029	4.282517	Se	Ge	2.669057	5.872051	Xe	H	1.356861	0.701016
Na	Li	1.267299	0.881482	Se	As	1.665280	0.711261	Xe	He	2.497832	2.599471
Na	Be	1.255480	3.121620	Se	Se	1.795894	0.821823	Xe	Li	2.466895	4.582081
Na	B	1.569961	3.188608					Xe	Be	6.000003	0.660525
Na	C	2.196050	4.520429	Br	H	2.192803	0.850378	Xe	B	5.051957	1.100612
Na	N	2.494384	8.586387	Br	He	2.128275	1.062043	Xe	C	1.704440	0.826727
Na	O	1.981449	3.270079	Br	Li	2.074441	1.858866	Xe	N	1.932952	0.925624
Na	F	2.619551	7.047351	Br	Be	2.367146	1.940933	Xe	O	0.839233	0.035356
Na	Ne	1.774236	1.343037	Br	B	2.307890	1.226420	Xe	F	1.128812	0.065011
Na	Na	0.446435	0.287137	Br	C	2.015086	0.570686	Xe	Ne	1.330202	0.293862
				Br	N	4.224901	30.000133	Xe	Na	2.103003	8.368204
Mg	H	2.651594	7.758237	Br	O	2.283046	0.706584	Xe	Mg	2.698414	9.723572
Mg	He	2.210603	3.725850	Br	F	2.031765	0.293500	Xe	Al	2.412039	7.404465
Mg	Li	1.184380	2.490250	Br	Ne	2.464172	1.006159	Xe	Si	3.087060	16.092000
Mg	Be	1.557591	2.066392	Br	Na	1.622218	1.752937	Xe	Cl	1.546396	0.463758
Mg	B	2.527441	6.146701	Br	Mg	2.195697	2.916280	Xe	Ar	0.591520	0.049266
Mg	C	3.040946	10.517690	Br	Al	1.894141	2.357130	Xe	K	1.171250	1.224889
Mg	N	2.079125	1.208075	Br	Si	1.570825	0.589511	Xe	Ca	1.510653	1.717121
Mg	O	2.251520	1.535734	Br	P	1.402139	0.456521	Xe	Br	1.439618	0.475116
Mg	F	3.362208	5.859023	Br	S	1.509874	0.286688	Xe	Kr	0.551561	0.049793
Mg	Ne	2.031676	1.214859	Br	Cl	1.710331	0.389238	Xe	Rb	1.087823	0.974965

Table 4 (continued)

		α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}
Mg	Na	1.506773	8.675619	Br	Ar	2.450801	3.262668	Xe	I	0.799155	0.112090
Mg	Mg	1.093573	0.465645	Br	K	1.616093	3.322795	Xe	Xe	1.244762	0.344474
				Br	Ca	2.078405	4.052910				
Al	H	2.025996	2.958379	Br	Sc	1.793486	2.098251	Cs	H	0.264882	0.096901
Al	He	2.255830	2.701400	Br	Ti	1.674847	0.883434	Cs	B	1.487110	10.392610
Al	Li	1.581593	1.106819	Br	V	1.902904	0.612698	Cs	C	2.147104	24.514623
Al	Be	1.938237	5.037214	Br	Cr	1.566028	0.217853	Cs	N	2.446532	29.711077
Al	B	2.059569	2.741479	Br	Mn	2.283820	1.183580	Cs	O	2.085139	8.176843
Al	C	2.267440	2.928056	Br	Fe	3.641782	6.061921	Cs	F	2.834100	22.233416
Al	N	2.009754	1.345202	Br	Co	2.632688	0.425148	Cs	P	2.924953	0.506512
Al	O	2.498660	2.131396	Br	Ni	2.772136	0.632145	Cs	S	0.289412	0.091743
Al	F	3.084258	1.975635	Br	Cu	5.826407	0.768517	Cs	Cl	1.673663	4.531965
Al	Ne	2.447869	1.709200	Br	Zn	1.416120	0.747027	Cs	Br	1.167189	1.658427
Al	Na	1.202871	2.071847	Br	Ga	1.819105	1.261036	Cs	I	0.919562	1.072178
Al	Mg	1.972530	13.472443	Br	Ge	1.602366	0.627737	Cs	Cs	1.170843	25.320055
Al	Al	1.387714	2.139200	Br	As	1.520170	0.514153				
				Br	Se	1.483713	0.319342	Ba	H	6.000135	2.040004
Si	H	1.896950	0.924196	Br	Br	1.758146	0.615308	Ba	C	0.770626	0.119793
Si	He	2.040498	1.853583					Ba	N	1.148233	0.207934
Si	Li	1.789609	3.090791	Kr	H	3.770453	5.125897	Ba	O	1.283018	0.348945
Si	Be	1.263132	0.623433	Kr	He	1.996943	0.627701	Ba	F	3.000618	5.575255
Si	B	1.982653	1.028287	Kr	Li	3.314562	8.758697	Ba	Al	2.105924	9.539099
Si	C	1.984498	0.785745	Kr	Be	3.253048	10.237796	Ba	Si	1.240420	1.212660
Si	N	1.818988	0.592972	Kr	B	2.363169	2.946781	Ba	S	0.705188	0.215386
Si	O	1.923600	0.751095	Kr	C	2.076738	0.652623	Ba	Cl	1.071044	0.160177
Si	F	2.131028	0.543516	Kr	N	1.644052	0.199606	Ba	Ti	2.176040	9.493530
Si	Ne	2.867784	14.378676	Kr	O	0.292300	0.006733	Ba	Br	1.190346	0.828794
Si	Na	2.007615	9.237644	Kr	F	3.452321	4.134407	Ba	I	0.982528	0.835597
Si	Mg	3.139749	29.994520	Kr	Ne	2.813679	1.433722	Ba	Ba	0.339269	0.356186
Si	Al	1.900000	2.000000	Kr	Na	2.480598	8.354448				
Si	Si	1.329000	0.273477	Kr	Mg	1.391487	0.888436	La	H	0.833667	0.623501
				Kr	Al	2.467131	5.091716	La	C	0.604869	0.108649
P	H	1.926537	1.234986	Kr	Si	1.764100	0.554250	La	N	0.758881	0.104778
P	He	2.093158	1.490218	Kr	Cl	1.884974	0.520217	La	O	1.318333	0.557957
P	Li	1.394544	1.122950	Kr	Ar	1.995125	0.554874	La	F	2.379335	2.401903
P	Be	1.800070	1.684831	Kr	K	2.182487	8.609782	La	Al	1.003510	0.500540
P	B	1.923168	1.450886	Kr	Ca	1.305197	0.878891	La	Si	2.016820	3.219030
P	C	1.994653	0.979512	Kr	Br	1.529006	0.308098	La	P	0.954450	0.541660
P	N	2.147042	0.972154	Kr	Kr	1.135319	0.052099	La	S	1.834129	2.682412
P	O	2.220768	0.878705					La	Cl	0.993753	0.230203
P	F	2.234356	0.514575	Rb	H	2.443556	29.861632	La	Br	0.758184	0.238582
P	Ne	2.219036	0.774954	Rb	He	1.270741	1.862585	La	I	0.592666	0.226883
P	Na	1.500320	2.837095	Rb	B	5.532239	9.040493	La	La	4.248067	5.175162
P	Mg	1.383773	1.177881	Rb	C	2.765830	29.974031				

Table 4 (continued)

		α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}
P	Al	1.980727	5.050816	Rb	N	0.761047	0.024636	Gd	H	0.390870	0.135810
P	Si	3.313466	13.239121	Rb	O	1.334908	1.125350	Gd	C	0.446870	0.053040
P	P	1.505792	0.902501	Rb	F	3.638122	28.815278	Gd	N	1.159410	0.205050
				Rb	Ne	2.267591	7.736563	Gd	O	0.862040	0.175800
S	H	2.215975	0.849712	Rb	Al	0.798774	2.992457	Gd	F	1.497980	0.334630
S	He	1.959149	0.437618	Rb	S	1.303184	0.964411	Gd	Al	1.003510	0.500540
S	Li	2.294275	2.642502	Rb	Cl	2.274411	10.384486	Gd	Si	2.016820	3.219030
S	Be	2.781736	3.791565	Rb	Ar	2.510977	18.433329	Gd	P	0.954450	0.541660
S	B	2.403696	1.125394	Rb	Br	1.797766	5.176214	Gd	S	2.003930	2.655400
S	C	2.210305	0.666849	Rb	Kr	2.268753	15.307503	Gd	Cl	0.806810	0.089970
S	N	2.289990	0.738710	Rb	Rb	1.180818	20.147610	Gd	Br	0.715810	0.240740
S	O	2.383289	0.747215					Gd	I	0.585360	0.278240
S	F	2.187186	0.375251	Sr	H	2.105914	12.973316	Gd	Gd	3.348180	2.670400
S	Ne	2.787058	3.296160	Sr	C	1.986688	6.654657				
S	Na	1.400850	0.852434	Sr	N	2.183629	6.853866	Lu	H	1.415790	0.787920
S	Mg	1.500163	0.500748	Sr	O	2.138399	3.561396	Lu	C	2.312813	4.453825
S	Al	1.976705	2.347384	Sr	F	3.050666	10.971705	Lu	N	2.141302	2.860828
S	Si	1.885916	0.876658	Sr	Si	2.969780	2.764750	Lu	O	2.192486	2.917076
S	P	1.595325	0.562266	Sr	P	2.789150	2.552100	Lu	P	5.618820	0.500000
S	S	1.794556	0.473856	Sr	S	1.598106	3.129603	Lu	Cl	2.753636	12.757099
				Sr	Cl	1.854190	3.783955	Lu	Br	2.322618	8.648274
Cl	H	2.402886	0.754831	Sr	Ti	2.880030	2.817250	Lu	I	2.248348	10.082315
Cl	He	1.671677	0.272964	Sr	Br	1.524316	2.766567				
Cl	Li	2.783001	4.227794	Sr	Sr	1.000040	5.372120	Hf	H	1.423788	3.427312
Cl	Be	2.822676	2.507275					Hf	B	1.633500	0.659270
Cl	B	2.259323	0.822129	Y	H	1.189053	0.612399	Hf	C	1.002194	0.378579
Cl	C	2.162197	0.515787	Y	C	1.336094	0.504306	Hf	N	1.332410	0.655795
Cl	N	2.172134	0.520745	Y	N	1.778796	1.627903	Hf	O	1.633289	1.034718
Cl	O	2.323236	0.585510	Y	O	1.851030	1.742922	Hf	F	2.290803	1.679335
Cl	F	2.313270	0.411124	Y	F	2.648046	4.433809	Hf	Mg	1.911350	4.330250
Cl	Ne	1.703151	0.125133	Y	Al	1.003500	0.500670	Hf	Al	0.949150	0.622520
Cl	Na	1.816429	1.357894	Y	Si	2.016820	3.219030	Hf	Si	2.189300	3.382300
Cl	Mg	2.391806	2.430856	Y	P	0.954450	0.541660	Hf	P	1.231220	0.505530
Cl	Al	2.125939	2.153451	Y	S	0.971688	0.318222	Hf	S	2.327110	1.666760
Cl	Si	1.684978	0.513000	Y	Cl	1.630152	1.154959	Hf	Cl	1.297117	0.706421
Cl	P	1.468306	0.352361	Y	Br	1.401208	1.054316	Hf	Ca	2.054500	4.319510
Cl	S	1.715435	0.356971	Y	Y	1.012681	1.691725	Hf	As	1.799500	1.280820
Cl	Cl	1.823239	0.332919					Hf	Br	1.090759	0.692456
				Zr	H	1.379703	0.593732	Hf	I	1.014096	0.820948
Ar	H	4.056167	3.933445	Zr	C	2.029427	1.999182	Hf	Ba	2.264830	9.022520
Ar	He	2.716562	1.177211	Zr	N	1.707083	0.995045	Hf	Hf	0.544144	1.058911
Ar	Li	3.122895	3.362910	Zr	O	1.709570	1.057525				
Ar	Be	3.044007	2.755492	Zr	F	1.900925	0.861142	Ta	H	2.288014	2.827669

Table 4 (continued)

		α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}
Ar	B	2.415471	1.931586	Zr	Al	1.270620	0.874060	Ta	C	1.838949	0.847439
Ar	C	1.471309	0.122309	Zr	Si	1.750833	1.723343	Ta	N	2.053679	1.015461
Ar	N	2.326805	0.562581	Zr	P	1.091858	0.748376	Ta	O	2.412629	1.751083
Ar	O	2.240673	0.355795	Zr	S	2.129761	2.429324	Ta	F	3.107390	3.146520
Ar	F	3.920658	9.269715	Zr	Cl	1.328835	0.443099	Ta	Na	2.551120	8.276130
Ar	Ne	2.963747	1.304697	Zr	Br	1.446868	0.858909	Ta	P	2.513800	6.261880
Ar	Na	2.167677	3.398138	Zr	Zr	3.865968	3.077773	Ta	S	2.246723	2.975980
Ar	Mg	2.092664	1.970638					Ta	Cl	1.608805	0.516413
Ar	Al	2.645165	1.852009	Nb	H	2.505912	3.603779	Ta	K	4.521470	2.026700
Ar	Si	1.780350	1.067890	Nb	C	2.621012	4.575481	Ta	Br	1.640376	0.791445
Ar	P	4.372516	0.171014	Nb	N	2.023863	1.213587	Ta	I	2.401053	6.551551
Ar	S	2.049398	0.653769	Nb	O	2.049489	1.184719	Ta	Ta	2.082863	10.987053
Ar	Cl	2.554449	2.256094	Nb	F	3.003157	3.663682				
Ar	Ar	2.306432	0.972699	Nb	Na	2.551010	8.276020	W	H	2.130880	1.832270
				Nb	P	2.221608	6.201507	W	C	2.097480	1.160770
K	H	0.648173	0.369340	Nb	S	2.249482	2.460020	W	N	1.596040	0.478350
K	He	1.418501	2.895045	Nb	Cl	2.215275	1.891557	W	O	1.359020	0.349010
K	Li	1.036487	4.374567	Nb	K	4.521360	2.026590	W	F	1.446050	0.213890
K	Be	1.931888	6.732221	Nb	Br	2.006678	1.921269	W	Na	2.551030	8.276040
K	B	2.031768	8.900541	Nb	Nb	1.727941	2.122388	W	P	2.338060	5.953860
K	C	2.241757	10.317987					W	S	1.542570	0.488630
K	N	2.325859	7.977707	Mo	H	2.035748	0.934686	W	Cl	1.310690	0.278000
K	O	1.508571	1.012275	Mo	C	2.198672	1.190742	W	K	4.521380	2.026610
K	F	3.182817	6.592971	Mo	N	1.869475	0.608268	W	Br	1.293260	0.372390
K	Ne	1.138021	0.233995	Mo	O	1.755424	0.511267	W	I	1.573570	1.077370
K	Na	0.884307	5.563027	Mo	F	2.202593	0.610429	W	W	2.940870	7.471390
K	Mg	0.884810	3.290502	Mo	Na	2.440770	8.286550				
K	Al	1.976076	29.944708	Mo	P	1.850441	1.522846	Re	H	1.634500	0.345894
K	Si	1.675930	8.279200	Mo	S	1.939658	0.830428	Re	C	2.306285	0.690687
K	P	1.443738	4.475384	Mo	Cl	1.783362	0.474325	Re	N	1.918332	0.445213
K	S	2.512156	29.528951	Mo	K	3.939420	2.142390	Re	O	1.967747	0.635960
K	Cl	1.622163	1.231481	Mo	Cr	2.674616	1.741943	Re	F	2.154219	0.535966
K	Ar	2.302803	9.710508	Mo	Br	1.283334	0.225918	Re	Si	2.775930	0.849450
K	K	1.435514	5.934329	Mo	Mo	2.034254	0.626462	Re	P	1.804168	0.966942
								Re	S	1.083919	0.068874
Ca	H	2.141859	7.728606	Tc	H	2.830345	6.310334	Re	Cl	1.433875	0.146319
Ca	He	1.719847	2.913852	Tc	C	3.198326	3.972439	Re	Ge	2.852340	2.151580
Ca	B	1.700010	1.700010	Tc	N	2.315417	0.727130	Re	Se	2.523170	2.202140
Ca	C	1.035305	0.148450	Tc	O	2.405190	1.024616	Re	Br	1.603060	0.287528
Ca	N	2.386600	2.988074	Tc	F	3.604815	5.811784	Re	Sb	2.204360	2.275780
Ca	O	3.263897	17.028946	Tc	S	2.463401	1.496502	Re	I	2.610119	3.559286
Ca	F	2.645053	3.482821	Tc	Cl	2.572043	1.651583	Re	Re	6.000258	4.488852
Ca	Ne	0.954530	0.332586	Tc	Ge	2.852820	2.152060				
Ca	Na	3.107104	9.657509	Tc	Se	2.523660	2.202620	Os	H	3.404180	4.393870

Table 4 (continued)

		α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}
Ca	Mg	2.299800	8.599800	Tc	Br	2.828264	3.820130	Os	C	2.336500	0.498410
Ca	Al	1.612565	4.188555					Os	N	1.143090	0.080870
Ca	Si	1.218788	0.336233	Ru	H	2.892899	7.137976	Os	O	1.350360	0.184300
Ca	P	1.024142	0.410840	Ru	C	2.784833	1.134936	Os	F	1.507620	0.140050
Ca	S	0.958171	0.325739	Ru	N	3.055504	2.334094	Os	Na	2.550740	8.275750
Ca	Cl	2.383391	5.956144	Ru	O	3.134940	2.976279	Os	P	2.836090	6.058300
Ca	Ar	1.034881	0.291072	Ru	F	3.878711	6.947128	Os	S	2.809500	4.186050
Ca	K	1.119200	1.240320	Ru	Si	2.775910	0.849430	Os	Cl	1.833070	0.327920
Ca	Ca	1.889674	30.003591	Ru	P	0.298916	0.056974	Os	K	4.521090	2.026320
				Ru	S	2.508076	1.006683	Os	Br	1.766880	0.382430
Sc	H	1.179485	0.351199	Ru	Cl	1.759883	0.126586	Os	I	2.203760	2.199190
Sc	C	2.630490	8.608052	Ru	Ge	2.852320	2.151560	Os	Os	2.021630	0.830440
Sc	N	2.270004	3.231881	Ru	Se	2.523160	2.202120				
Sc	O	2.256516	3.058672	Ru	Br	2.584735	0.659881	Ir	H	1.033900	0.058047
Sc	F	3.107985	7.252347	Ru	Ru	0.572056	0.097805	Ir	C	1.690295	0.115047
Sc	Al	1.003550	0.500620					Ir	N	3.934508	8.518640
Sc	Si	2.016870	3.219070	Rh	H	3.104165	2.306107	Ir	O	3.748272	9.625402
Sc	P	0.868165	0.626749	Rh	C	3.415991	3.488079	Ir	F	2.982799	1.499639
Sc	S	0.422939	0.211850	Rh	N	3.585462	4.000947	Ir	Na	2.550820	8.275830
Sc	Cl	2.141474	2.996129	Rh	O	3.927830	10.298676	Ir	P	2.714060	6.284670
Sc	Sc	1.132838	2.598166	Rh	F	4.051654	9.065384	Ir	S	3.204834	4.135732
				Rh	Si	2.776490	0.850010	Ir	Cl	2.009770	0.258916
Ti	H	0.832669	0.143722	Rh	P	2.334607	1.038141	Ir	K	4.521170	2.026400
Ti	B	1.628710	0.649360	Rh	S	3.154006	4.816410	Ir	Br	2.038142	0.171879
Ti	C	1.597973	0.416706	Rh	Cl	3.300130	3.586865	Ir	I	3.410914	1.497148
Ti	N	1.678686	0.545461	Rh	Ge	2.852900	2.152140	Ir	Ir	5.771663	11.175193
Ti	O	1.789118	0.799486	Rh	Se	2.523740	2.202700				
Ti	F	2.307087	1.085742	Rh	Br	2.928082	1.510149	Pt	H	4.001198	8.924015
Ti	Mg	1.911340	4.330240	Rh	Rh	2.497328	2.070114	Pt	C	3.306722	3.493403
Ti	Al	1.369486	2.091841					Pt	N	2.307923	0.540730
Ti	Si	2.856038	6.773815	Pd	H	2.183761	0.443269	Pt	O	2.110563	0.487756
Ti	P	2.151929	4.150500	Pd	C	4.777192	9.853715	Pt	F	3.714441	5.617014
Ti	S	1.846439	0.943784	Pd	N	2.328046	0.249703	Pt	Al	1.572360	1.056930
Ti	Cl	1.461034	0.333297	Pd	O	2.154867	0.216403	Pt	Si	0.999990	0.099990
Ti	Ca	2.000000	4.109141	Pd	F	4.237312	6.945312	Pt	P	1.403239	0.233712
Ti	Ti	2.648597	2.000000	Pd	Al	1.572720	1.057290	Pt	S	2.791500	2.224263
				Pd	Si	2.948200	2.225104	Pt	Cl	2.108526	0.341001
V	H	1.280133	0.105204	Pd	P	0.803630	0.045017	Pt	Br	2.185307	0.520361
V	C	2.789855	1.938760	Pd	S	2.177801	0.255229	Pt	I	3.077338	4.601248
V	N	1.607540	0.276725	Pd	Cl	3.871243	2.969891	Pt	Pt	3.404276	9.010252
V	O	1.623973	0.415312	Pd	Br	5.994879	4.638051				
V	F	1.825160	0.342815	Pd	Pd	1.064375	0.051956	Au	H	3.369041	2.605283
V	Na	2.551010	8.276020					Au	C	4.580016	21.485634
V	P	2.549154	6.250624	Ag	H	2.895936	1.995168	Au	N	2.138095	0.222059

Table 4 (continued)

		α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}
V	S	2.704124	2.035039	Ag	C	4.404336	11.335456	Au	O	1.548763	0.077192
V	Cl	1.688529	0.243657	Ag	N	4.659871	19.803710	Au	F	4.453145	9.594384
V	K	4.521360	2.026590	Ag	O	1.893874	0.165661	Au	Al	1.572570	1.057140
V	V	4.832391	10.779892	Ag	F	4.628423	12.695884	Au	P	1.618713	0.067001
				Ag	Al	1.928800	0.896514	Au	S	4.306238	21.619145
Cr	H	0.882661	0.044469	Ag	P	6.000006	0.049932	Au	Cl	3.539414	2.257702
Cr	C	3.656754	6.110187	Ag	S	3.653121	11.188022	Au	Br	0.581911	0.004237
Cr	N	3.029186	1.920324	Ag	Cl	4.441176	23.765459	Au	I	0.577916	0.008816
Cr	O	2.500000	1.055511	Ag	Br	3.677491	1.714369	Au	Au	0.903162	0.013091
Cr	F	2.716521	0.737607	Ag	Ag	2.127645	0.557742				
Cr	Na	2.295056	8.364274					Hg	H	1.136587	0.799399
Cr	Si	1.860760	1.029110	Cd	H	2.628748	11.914201	Hg	C	0.795816	0.147128
Cr	P	1.695383	0.600177	Cd	C	1.425678	0.603441	Hg	N	0.332152	0.050240
Cr	S	2.260978	0.550334	Cd	N	0.970423	0.180663	Hg	O	1.052145	0.240720
Cr	Cl	2.152618	0.369073	Cd	O	1.696673	0.926146	Hg	F	1.240572	0.113827
Cr	K	2.000000	2.000000	Cd	F	2.312135	1.353665	Hg	Si	2.770860	3.680740
Cr	Cr	4.655419	10.318607	Cd	Si	1.371225	2.253346	Hg	P	0.608604	0.214951
				Cd	S	1.182202	0.361389	Hg	S	1.041682	0.347383
Mn	H	2.309940	1.269210	Cd	Cl	0.943547	0.140424	Hg	Cl	0.430731	0.053660
Mn	C	3.000750	2.583110	Cd	Br	1.001451	0.272267	Hg	Ti	3.414630	2.957200
Mn	N	2.921470	1.956750	Cd	Cd	1.564044	18.617999	Hg	Br	0.638717	0.172363
Mn	O	2.577540	1.285620					Hg	Te	0.291500	0.212732
Mn	F	2.791950	1.113070	In	H	3.064144	14.975293	Hg	I	0.758162	0.342058
Mn	Al	1.768360	1.040790	In	C	2.189272	2.187385	Hg	Hg	0.474413	0.423276
Mn	Si	1.937959	0.950580	In	N	2.469868	3.369993				
Mn	P	1.947020	1.130320	In	O	2.662095	4.128583	Tl	H	0.673658	0.138205
Mn	S	2.482510	1.612650	In	F	2.948797	3.701016	Tl	B	1.528347	10.504338
Mn	Cl	1.657010	0.201850	In	S	2.542131	6.341105	Tl	C	1.390345	0.582895
Mn	Ca	1.491440	0.620180	In	Cl	2.233405	2.388552	Tl	N	0.982335	0.158812
Mn	Mn	2.665420	2.460040	In	Ga	1.628870	2.421987	Tl	O	1.550068	0.636906
				In	As	2.299552	6.208350	Tl	F	1.469516	0.226166
Fe	H	0.854488	0.025195	In	Se	1.906572	2.319323	Tl	S	0.994851	0.303426
Fe	C	3.991343	0.366835	In	Br	2.257957	3.728598	Tl	Cl	0.846193	0.162037
Fe	N	2.500486	0.155342	In	In	2.073241	8.063491	Tl	Br	0.874419	0.296836
Fe	O	1.726313	0.136422					Tl	I	0.902012	0.430033
Fe	F	4.294707	3.657350	Sn	H	2.648910	6.535162	Tl	Tl	1.191684	9.535127
Fe	P	2.567534	0.431291	Sn	C	2.440538	3.374355				
Fe	S	0.988991	0.033478	Sn	N	2.085589	1.391900	Pb	H	1.522676	0.840096
Fe	Cl	1.229793	0.019473	Sn	O	2.727260	4.374017	Pb	Li	1.001810	1.285064
Fe	K	2.000000	6.000000	Sn	F	3.724286	18.598664	Pb	B	0.911197	1.138157
Fe	Fe	2.720785	1.846890	Sn	S	2.131542	2.314870	Pb	C	1.525593	0.404656
				Sn	Cl	1.771522	0.807782	Pb	N	1.317394	0.335787
Co	H	2.966518	2.472465	Sn	Ge	2.524633	12.343411	Pb	O	1.763210	0.782506
Co	C	3.716233	2.123930	Sn	Se	2.127377	3.061885	Pb	F	3.288902	8.368562

Table 4 (continued)

		α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}			α_{ij} [Å^{-1}]	x_{ij}
Co	N	3.618638	2.653836	Sn	Br	1.535089	0.668798	Pb	P	4.516800	5.033200
Co	O	3.726911	5.252022	Sn	Sn	0.921000	0.287000	Pb	S	1.027519	0.175150
Co	F	3.956347	4.585030					Pb	Cl	1.094123	0.164814
Co	Si	2.469805	1.090240	Sb	H	1.571272	0.795343	Pb	V	1.500000	1.000000
Co	P	1.152505	0.105936	Sb	C	1.696206	0.579212	Pb	Cr	1.860760	1.029110
Co	S	2.429255	0.436707	Sb	N	0.676115	0.082065	Pb	Zn	1.500000	1.000000
Co	Cl	3.217497	1.033414	Sb	O	1.846384	0.634234	Pb	Se	2.000000	0.111195
Co	Co	3.288166	3.919618	Sb	F	2.182922	0.650277	Pb	Br	0.865550	0.148229
				Sb	Al	1.422641	1.616690	Pb	Nb	1.500000	1.000000
Ni	H	2.635280	1.763124	Sb	Si	2.686590	8.713749	Pb	Mo	2.000000	5.000000
Ni	C	4.285513	7.133324	Sb	S	1.418837	0.396969	Pb	Te	1.002559	0.809042
Ni	N	3.845215	4.286800	Sb	Cl	1.117287	0.156475	Pb	I	0.983474	0.267426
Ni	O	2.937232	0.885942	Sb	Mn	2.400320	2.236710	Pb	Pb	1.881764	2.362343
Ni	F	3.440241	1.088208	Sb	Co	2.204630	2.276050				
Ni	Si	2.068881	0.938646	Sb	Br	1.063916	0.198044	Bi	H	1.679905	1.397462
Ni	P	3.260283	5.059727	Sb	Tc	2.204850	2.276260	Bi	Li	0.340140	0.695320
Ni	S	2.002752	0.274852	Sb	Ru	2.204350	2.275760	Bi	C	1.534025	0.576179
Ni	Cl	2.200512	0.202313	Sb	Rh	2.204930	2.276340	Bi	N	1.143876	0.152738
Ni	Ni	1.097960	0.035474	Sb	In	2.141933	6.660801	Bi	O	1.553297	0.333042
				Sb	Sb	1.348535	0.724885	Bi	F	2.355400	1.035324
Cu	H	2.335359	0.603591					Bi	S	1.466879	0.620997
Cu	C	4.638773	7.067794	Te	H	2.039130	1.807679	Bi	Cl	1.272975	0.326871
Cu	N	4.214337	3.228667	Te	C	1.992816	0.970494	Bi	Se	1.344746	0.651208
Cu	O	3.959951	2.000000	Te	N	1.722269	0.358593	Bi	Br	1.146233	0.381170
Cu	F	4.478832	1.282108	Te	O	1.853064	0.382926	Bi	I	1.302171	0.862377
Cu	P	0.210640	0.020126	Te	F	1.998576	0.200822	Bi	Bi	1.074064	1.168214
Cu	S	0.273112	0.005248	Te	Al	1.387541	2.106812				

check was also done to verify that the error distribution was approximately Gaussian. The resulting histogram, shown in Fig. 1, shows that the distribution is indeed Gaussian.

Hydrogen bonding

One of the commonest forms of hydrogen bonding involves a hydrogen atom attached to an oxygen atom and forming a weak bond to a distant oxygen atom. The simplest, well-characterized case is that of the water dimer. In an exhaustive analysis of this system, Tschumper, et. al. [36], characterized this system using CCSD(T) and a large basis set. They identified and characterized ten stationary points on the 12-dimensional potential energy surface of the dimer and determined that the lowest energy conformer of the water dimer was 5.00 kcal mol^{-1} more stable than two isolated water molecules. A comparison of the relative heats of formation of these points calculated using NDDO methods is shown in Table 17. The AUE for the various methods are as follows: PM6: 1.35 kcal mol^{-1}, PM5: 3.35, PM3: 2.16, and AM1: 1.67.

The energies of various different types of hydrogen bonds were estimated from the energy released when the two small molecules involved associate to form a hydrogen-bonded system. Table 18 lists the values predicted using B3LYP and the NDDO methods.

Nitrogen pyramidalization

A well-documented fault in PM3 nitrogen was its exaggerated degree of pyramidalization when in the sp^2 configuration. This is dramatically evident in N-methylacetamide, where the H-N-C–C torsion angle should be 180 °, but is predicted by PM3 to be 136 °. That is, the nitrogen, instead of being in a planar environment, is predicted to be highly

Table 5 Average unsigned errors in calculated heats of formation (kcal mol^{-1})

Element	PM6	No.	PM5	No.	PM3	No.	AM1	No.
Hydrogen	7.29	3039	13.89	2340	17.09	2340	21.12	2270
Helium	0.00	1	0.00	1	0.00	1	0.00	1
Lithium	7.98	83	15.31	83	18.02	83	18.84	82
Beryllium	5.92	34	29.06	34	29.58	34	18.51	34
Boron	6.44	122	10.81	120	11.84	120	–	–
Carbon	7.31	2828	13.03	2155	15.06	2155	19.42	2123
Nitrogen	8.22	1067	16.45	761	20.96	761	24.23	744
Oxygen	8.42	1758	16.59	1243	20.13	1244	27.68	1229
Fluorine	8.49	497	22.31	350	21.25	350	37.40	334
Neon	0.00	1	0.00	1	0.00	1	0.00	1
Sodium	5.72	40	8.57	39	9.47	39	10.77	38
Magnesium	9.84	66	12.07	66	17.94	66	18.71	66
Aluminum	7.61	75	17.49	75	19.15	75	18.99	75
Silicon	6.51	98	9.28	96	12.80	96	17.00	95
Phosphorus	8.20	110	16.01	98	17.36	98	20.06	95
Sulfur	8.81	427	15.40	330	18.44	330	26.38	323
Chlorine	8.28	670	16.69	390	18.71	390	23.06	383
Argon	0.00	1	0.00	1	0.00	1	0.00	1
Potassium	6.53	43	12.33	42	9.36	42	28.38	41
Calcium	11.87	43	28.68	43	43.44	43	63.20	43
Scandium	10.33	52	–	–	–	–	–	–
Titanium	10.20	85	–	–	–	–	–	–
Vanadium	14.29	59	–	–	–	–	–	–
Chromium	14.09	60	–	–	–	–	–	–
Manganese	12.77	44	–	–	–	–	–	–
Iron	18.31	76	–	–	–	–	–	–
Cobalt	15.51	42	–	–	–	–	–	–
Nickel	15.10	51	–	–	–	–	–	–
Copper	13.00	47	–	–	–	–	–	–
Zinc	5.56	54	17.84	54	32.93	54	37.06	54
Gallium	7.51	47	29.12	47	37.58	47	46.87	47
Germanium	9.83	67	12.20	67	15.86	67	19.12	67
Arsenic	6.94	49	15.22	49	16.68	49	17.34	49
Selenium	4.40	25	39.58	25	39.71	25	32.00	25
Bromine	7.37	330	17.20	199	25.04	199	28.22	199
Krypton	0.00	1	0.00	1	0.00	1	0.00	1
Rubidium	10.91	24	16.57	24	21.47	24	29.33	23
Strontium	7.72	38	52.46	38	103.16	38	57.21	38
Yttrium	13.28	51	–	–	–	–	–	–
Zirconium	11.18	46	–	–	–	–	–	–
Niobium	8.57	51	–	–	–	–	–	–
Molybdenum	13.41	70	–	–	–	–	35.77	69
Technetium	15.14	50	–	–	–	–	–	–
Ruthenium	13.87	56	–	–	–	–	–	–
Rhodium	20.92	32	–	–	–	–	–	–
Palladium	11.65	47	–	–	–	–	–	–
Silver	4.67	14	–	–	–	–	–	–
Cadmium	3.49	38	34.66	38	61.92	38	–	–
Indium	7.33	54	31.53	54	29.83	54	32.16	54
Tin	7.14	77	16.83	77	17.10	77	20.21	77
Antimony	5.41	58	30.98	58	34.61	58	35.00	58
Tellurium	8.20	45	35.66	45	46.80	45	22.91	45
Iodine	7.23	279	23.77	176	25.90	176	36.55	175
Xenon	0.00	1	0.00	1	0.00	1	0.00	1
Cesium	6.89	40	37.01	40	35.22	40	55.33	39
Barium	12.12	37	98.20	37	154.65	37	161.09	37

Table 5 (continued)

Element	PM6	No.	PM5	No.	PM3	No.	AM1	No.
Lanthanum	10.37	37	–	–	–	–	–	–
Lutetium	7.68	24	–	–	–	–	–	–
Hafnium	8.52	37	–	–	–	–	–	–
Tantalum	14.37	36	–	–	–	–	–	–
Tungsten	7.38	28	–	–	–	–	–	–
Rhenium	10.40	57	–	–	–	–	–	–
Osmium	6.46	19	–	–	–	–	–	–
Iridium	10.21	25	–	–	–	–	–	–
Platinum	11.61	77	–	–	–	–	–	–
Gold	12.82	32	–	–	–	–	–	–
Mercury	5.94	51	16.39	51	17.67	51	19.75	51
Thallium	10.42	44	32.63	44	73.96	45	73.18	45
Lead	7.92	44	18.08	44	14.18	44	16.71	44
Bismuth	7.74	53	99.88	53	28.95	53	119.23	53

Table 6 Average unsigned errors in bond lengths (Å)

Element	PM6	No.	PM5	No.	PM3	No.	AM1	No.
Hydrogen	0.044	238	0.056	219	0.032	217	0.035	181
Helium	0.251	6	0.459	6	0.182	4	0.655	5
Lithium	0.175	111	0.191	110	0.167	110	0.171	105
Beryllium	0.076	42	0.131	42	0.067	42	0.085	42
Boron	0.027	116	0.043	116	0.066	122	–	–
Carbon	0.057	1191	0.066	693	0.051	634	0.063	628
Nitrogen	0.090	663	0.145	309	0.124	259	0.163	253
Oxygen	0.095	1163	0.122	625	0.103	577	0.117	571
Fluorine	0.063	396	0.096	246	0.069	251	0.101	228
Neon	0.353	5	0.182	2	0.062	1	0.030	1
Sodium	0.229	33	0.200	33	0.208	30	0.140	29
Magnesium	0.089	106	0.067	106	0.167	105	0.073	106
Aluminium	0.045	77	0.120	72	0.098	70	0.138	70
Silicon	0.039	97	0.056	94	0.074	95	0.077	90
Phosphorus	0.039	141	0.078	92	0.073	92	0.083	87
Sulfur	0.094	359	0.107	216	0.091	207	0.134	200
Chlorine	0.069	672	0.098	283	0.095	284	0.130	285
Argon	0.258	4	0.303	1	–	–	–	–
Potassium	0.139	46	0.135	47	0.148	47	0.281	46
Calcium	0.133	67	0.177	69	0.151	67	0.102	60
Scandium	0.053	90	–	–	–	–	–	–
Titanium	0.078	140	–	–	–	–	–	–
Vanadium	0.090	168	–	–	–	–	–	–
Chromium	0.080	89	–	–	–	–	–	–
Manganese	0.083	107	–	–	–	–	–	–
Iron	0.102	117	–	–	–	–	–	–
Cobalt	0.107	100	–	–	–	–	–	–
Nickel	0.065	133	–	–	–	–	–	–
Copper	0.174	130	–	–	–	–	–	–
Zinc	0.076	77	0.084	77	0.098	77	0.142	76
Gallium	0.048	80	0.105	81	0.192	81	0.135	81
Germanium	0.038	131	0.045	131	0.056	133	0.068	133
Arsenic	0.073	72	0.069	70	0.080	72	0.099	72
Selenium	0.056	56	0.094	55	0.071	54	0.061	54
Bromine	0.104	358	0.106	184	0.146	182	0.136	184
Krypton	0.059	6	0.417	3	0.623	3	0.602	3
Rubidium	0.413	36	0.498	37	0.176	34	0.230	36

Table 6 (continued)

Element	PM6	No.	PM5	No.	PM3	No.	AM1	No.
Strontium	0.087	56	0.199	55	0.128	32	0.242	47
Yttrium	0.132	69	–	–	–	–	–	–
Zirconium	0.063	65	–	–	–	–	–	–
Niobium	0.060	88	–	–	–	–	–	–
Molybdenum	0.104	89	–	–	–	–	0.095	84
Technetium	0.078	84	–	–	–	–	–	–
Ruthenium	0.073	113	–	–	–	–	–	–
Rhodium	0.162	68	–	–	–	–	–	–
Palladium	0.080	120	–	–	–	–	–	–
Silver	0.151	41	–	–	–	–	–	–
Cadmium	0.159	54	0.179	55	0.121	50	–	–
Indium	0.039	77	0.085	77	0.155	77	0.102	77
Tin	0.073	96	0.065	96	0.078	96	0.087	94
Antimony	0.060	92	0.169	91	0.083	91	0.135	92
Tellurium	0.070	80	0.162	79	0.123	77	0.122	79
Iodine	0.144	286	0.137	147	0.146	145	0.175	141
Xenon	0.620	8	0.584	4	0.472	2	0.793	6
Cesium	0.258	40	0.335	43	0.372	25	0.358	43
Barium	0.202	51	0.228	47	0.207	48	0.261	51
Lanthanum	0.253	47	–	–	–	–	–	–
Lutetium	0.050	60	–	–	–	–	–	–
Hafnium	0.071	42	–	–	–	–	–	–
Tantalum	0.074	59	–	–	–	–	–	–
Tungsten	0.141	57	–	–	–	–	–	–
Rhenium	0.068	108	–	–	–	–	–	–
Osmium	0.072	50	–	–	–	–	–	–
Iridium	0.169	71	–	–	–	–	–	–
Platinum	0.057	140	–	–	–	–	–	–
Gold	0.158	84	–	–	–	–	–	–
Mercury	0.143	64	0.110	64	0.135	63	0.139	64
Thallium	0.202	59	0.248	55	0.208	45	0.268	43
Lead	0.140	53	0.167	53	0.121	53	0.125	51
Bismuth	0.142	81	0.616	75	0.225	82	0.682	75

Table 7 Average unsigned errors in bond angles (Degrees)

Element	PM6	No. in set	PM5	No. in set	PM3	No. in set	AM1	No. in set
Lithium	7.79	28	6.82	28	3.53	28	9.48	28
Beryllium	6.61	14	6.44	14	6.94	14	5.98	14
Boron	3.27	31	4.41	31	4.61	31	–	–
Carbon	2.50	134	2.79	134	2.75	131	2.25	131
Nitrogen	7.32	37	8.01	37	6.75	35	7.94	31
Oxygen	12.14	59	11.12	58	10.17	53	9.57	42
Fluorine	8.32	3	16.18	3	26.34	3	24.67	2
Sodium	21.00	4	2.87	4	3.43	4	5.32	4
Magnesium	8.44	24	7.28	24	14.23	24	7.10	24
Aluminum	4.05	20	5.26	20	7.21	19	4.33	19
Silicon	5.25	35	3.37	35	2.81	34	2.88	34
Phosphorus	3.24	35	4.40	35	6.01	35	5.07	35
Sulfur	5.23	46	5.64	45	5.42	41	5.05	41
Chlorine	3.65	5	19.47	5	10.31	5	14.80	5
Potassium	17.90	11	10.27	11	12.93	11	12.75	11
Calcium	14.99	16	11.35	16	16.81	16	18.06	15
Scandium	7.98	32	–	–	–	–	–	–
Titanium	7.86	39	–	–	–	–	–	–
Vanadium	7.46	44	–	–	–	–	–	–

Table 7 (continued)

Element	PM6	No. in set	PM5	No. in set	PM3	No. in set	AM1	No. in set
Chromium	3.77	19	–	–	–	–	–	–
Manganese	6.02	26	–	–	–	–	–	–
Iron	11.21	30	–	–	–	–	–	–
Cobalt	10.68	29	–	–	–	–	–	–
Nickel	10.44	48	–	–	–	–	–	–
Copper	10.77	44	–	–	–	–	–	–
Zinc	10.92	27	14.41	27	8.16	27	13.34	27
Gallium	4.43	18	10.86	18	14.43	18	13.84	18
Germanium	4.58	52	5.37	52	8.95	52	5.71	52
Arsenic	6.29	36	6.52	36	6.48	36	5.03	36
Selenium	7.27	24	16.16	24	12.37	23	5.46	23
Bromine	12.64	4	20.03	4	19.21	3	3.27	3
Rubidium	9.69	11	10.20	11	21.03	11	6.68	11
Strontium	18.16	25	32.91	25	32.92	25	31.00	25
Yttrium	12.29	34	–	–	–	–	–	–
Zirconium	10.36	12	–	–	–	–	–	–
Niobium	6.54	23	–	–	–	–	–	–
Molybdenum	8.15	27	–	–	–	–	8.73	27
Technetium	4.96	22	–	–	–	–	–	–
Ruthenium	6.93	34	–	–	–	–	–	–
Rhodium	10.66	22	–	–	–	–	–	–
Palladium	9.19	46	–	–	–	–	–	–
Silver	23.36	9	–	–	–	–	–	–
Cadmium	15.23	10	13.52	10	20.09	10	–	–
Indium	4.47	17	7.21	17	5.30	17	4.94	17
Tin	3.06	34	4.09	34	3.74	34	11.81	34
Antimony	6.49	41	12.24	41	6.84	41	7.40	41
Tellurium	4.85	25	7.00	25	5.33	25	7.87	25
Iodine	8.33	1	12.55	1	20.66	1	4.53	1
Cesium	15.50	12	8.52	12	19.38	12	11.75	12
Barium	28.65	10	28.43	10	37.04	10	36.17	10
Lanthanum	9.25	14	–	–	–	–	–	–
Lutetium	7.08	26	–	–	–	–	–	–
Hafnium	5.64	10	–	–	–	–	–	–
Tantalum	9.88	15	–	–	–	–	–	–
Tungsten	10.90	9	–	–	–	–	–	–
Rhenium	7.39	32	–	–	–	–	–	–
Osmium	12.67	10	–	–	–	–	–	–
Iridium	7.86	18	–	–	–	–	–	–
Platinum	5.92	72	–	–	–	–	–	–
Gold	13.59	16	–	–	–	–	–	–
Mercury	20.20	15	20.99	15	18.47	15	21.49	15
Thallium	5.73	10	10.28	10	19.95	10	25.38	10
Lead	4.33	20	5.24	20	4.61	20	3.57	19
Bismuth	8.01	25	21.74	25	8.28	25	33.99	25

pyramidal. The results of a survey of 19 molecules that contain sp^2 nitrogen are presented in Table 19.

Transition metals

Optimizing parameters for transition metals was not as straightforward as for the main group elements. As with the main group compounds, there is a wealth of structural reference data on transition metal complexes. However, unlike main group compounds, there is a distinct shortage of reliable thermochemical data. To alleviate this shortage, the thermochemical data that was available was augmented by the results of DFT calculations. It was recognized, however, that these derived reference data were likely to be of a lower accuracy than the experimental data. Many transition metal complexes are also highly labile; a consequence of this was that some moieties that are known to exist in the solid phase were predicted to be

Table 8 Average unsigned errors in dipole moments (D)

Element	PM6	No.	PM5	No.	PM3	No.	AM1	No.
Hydrogen	0.62	266	0.80	265	0.64	222	0.50	204
Lithium	0.78	16	0.95	16	0.79	16	0.52	16
Beryllium	1.63	1	1.49	1	0.27	1	0.53	1
Boron	0.66	17	0.66	17	0.73	17	–	–
Carbon	0.51	219	0.62	218	0.41	176	0.42	165
Nitrogen	0.61	48	0.66	48	0.46	40	0.55	39
Oxygen	0.99	198	1.27	196	1.05	74	0.74	75
Fluorine	0.80	124	1.11	121	0.59	63	0.69	59
Sodium	1.34	6	0.80	6	1.97	6	1.26	6
Aluminium	0.33	1	1.50	1	1.76	1	0.53	1
Silicon	0.21	11	1.09	11	0.72	11	0.29	11
Phosphorus	0.83	14	0.79	14	0.37	10	0.87	10
Sulfur	0.62	28	1.01	28	0.74	21	0.70	21
Chlorine	0.99	103	1.27	100	0.77	47	0.84	43
Potassium	0.44	4	0.34	4	1.30	4	0.58	4
Calcium	0.73	4	1.12	4	1.23	4	0.33	4
Scandium	1.11	9	–	–	–	–	–	–
Titanium	1.02	8	–	–	–	–	–	–
Vanadium	0.82	8	–	–	–	–	–	–
Chromium	1.98	9	–	–	–	–	–	–
Manganese	1.06	11	–	–	–	–	–	–
Iron	1.61	14	–	–	–	–	–	–
Cobalt	1.04	6	–	–	–	–	–	–
Nickel	1.40	15	–	–	–	–	–	–
Copper	1.11	10	–	–	–	–	–	–
Zinc	0.21	4	0.18	4	0.16	4	0.16	4
Gallium	0.20	1	1.81	1	1.35	1	0.64	1
Germanium	0.63	23	0.63	23	0.55	23	0.59	23
Arsenic	0.37	6	0.99	6	0.35	6	0.37	6
Selenium	0.66	10	0.94	10	0.61	10	0.80	10
Bromine	0.90	88	1.34	87	1.01	37	0.50	39
Rubidium	1.84	6	2.43	6	1.65	6	0.44	6
Strontium	1.64	6	1.31	6	2.55	6	1.51	6
Yttrium	1.70	8	–	–	–	–	–	–
Zirconium	0.94	8	–	–	–	–	–	–
Niobium	0.91	10	–	–	–	–	–	–
Molybdenum	1.09	8	–	–	–	–	1.48	8
Technetium	1.74	13	–	–	–	–	–	–
Ruthenium	1.13	12	–	–	–	–	–	–
Rhodium	1.09	6	–	–	–	–	–	–
Palladium	0.97	8	–	–	–	–	–	–
Silver	1.98	9	–	–	–	–	–	–
Cadmium	0.42	2	2.22	2	0.67	2	–	–
Indium	0.47	3	0.78	3	0.75	3	1.36	3
Tin	0.28	13	0.41	13	0.88	13	0.81	13
Antimony	0.55	5	0.77	5	0.48	5	0.61	5
Tellurium	0.47	2	0.75	2	0.31	2	1.35	2
Iodine	1.03	77	1.54	77	1.48	28	1.22	30
Cesium	1.25	9	3.47	9	1.89	9	0.87	9
Barium	1.77	11	1.29	11	1.93	11	1.11	11
Lanthanum	1.23	8	–	–	–	–	–	–
Hafnium	0.63	6	–	–	–	–	–	–
Tantalum	0.97	5	–	–	–	–	–	–
Tungsten	0.92	14	–	–	–	–	–	–
Rhenium	0.76	13	–	–	–	–	–	–
Osmium	0.63	8	–	–	–	–	–	–

Table 8 (continued)

Element	PM6	No.	PM5	No.	PM3	No.	AM1	No.
Iridium	0.96	8	–	–	–	–	–	–
Platinum	1.07	8	–	–	–	–	–	–
Gold	0.78	14	–	–	–	–	–	–
Mercury	0.63	9	0.77	9	0.63	9	0.67	9
Thallium	0.89	3	1.35	3	0.45	3	2.43	3
Lead	0.73	6	0.76	6	0.41	6	0.82	6
Bismuth	0.42	8	3.21	8	1.14	8	3.40	8

Table 9 Average unsigned errors in ionization potential (eV)

Element	PM6	No.	PM5	No.	PM3	No.	AM1	No.
Hydrogen	0.43	226	0.40	226	0.60	226	0.52	217
Lithium	0.89	12	0.88	12	1.29	12	0.59	12
Beryllium	0.52	7	0.29	7	0.93	7	0.45	7
Boron	0.31	11	0.34	11	1.01	11	–	–
Carbon	0.41	230	0.39	230	0.54	230	0.54	227
Nitrogen	0.55	43	0.45	43	0.53	43	0.48	42
Oxygen	0.62	72	0.56	72	0.63	72	0.69	69
Fluorine	0.64	67	0.65	67	0.74	67	0.85	65
Sodium	0.34	5	0.34	5	1.43	5	0.51	4
Magnesium	0.97	4	1.05	4	1.10	4	1.41	4
Aluminum	0.62	3	0.29	3	0.40	3	0.69	3
Silicon	0.43	11	0.81	11	0.70	11	0.68	11
Phosphorus	0.49	13	0.47	13	0.64	13	0.56	13
Sulfur	0.52	46	0.51	46	0.48	46	0.62	46
Chlorine	0.48	62	0.58	62	0.57	60	0.61	57
Potassium	0.23	4	0.50	4	0.54	4	0.34	3
Calcium	0.74	1	1.24	1	0.52	1	0.41	1
Scandium	3.73	1	–	–	–	–	–	–
Titanium	0.09	1	–	–	–	–	–	–
Zinc	0.32	5	0.35	5	0.99	5	0.49	5
Gallium	0.52	3	0.73	3	1.28	3	1.16	3
Germanium	0.70	13	0.49	13	0.93	13	1.05	13
Arsenic	0.69	5	0.31	5	0.62	5	0.79	5
Selenium	0.38	10	0.29	10	0.47	10	1.22	10
Bromine	0.28	33	0.39	33	1.20	33	0.49	32
Rubidium	0.18	3	0.39	3	0.93	3	0.22	3
Strontium	0.63	1	0.38	1	0.14	1	0.26	1
Cadmium	0.33	5	0.46	5	0.39	5	–	–
Indium	0.63	2	0.86	2	2.06	2	0.83	2
Tin	0.70	14	0.48	14	1.22	14	0.44	14
Antimony	0.44	5	0.90	5	1.16	5	0.54	5
Tellurium	0.43	3	0.20	3	0.25	3	0.70	3
Iodine	0.47	29	0.46	29	0.48	29	0.89	29
Cesium	0.58	4	0.71	4	1.37	4	1.11	4
Barium	0.08	1	0.97	1	0.08	1	0.75	1
Mercury	0.51	12	0.43	12	0.74	12	0.49	12
Thallium	0.30	3	0.46	3	0.80	3	0.53	3
Lead	0.56	13	0.47	13	0.93	13	0.65	13
Bismuth	0.98	5	1.28	5	0.72	5	1.66	5

Table 10 Average unsigned errors in ΔH_f for various sets of elements (kcal mol^{-1})

Set of elements	No.	PM6	RM1	PM5	PM3	AM1
H, C, N, O	1157	4.64	4.89	5.60	5.65	9.41
H, C, N, O, F, P, S, Cl, Br, I	1774	5.05	6.57	6.75	8.05	12.57
Whole of main group	3188	6.16		15.27	17.76	22.34
70 elements	4492	8.01				

Table 12 Average unsigned errors in angles for various sets of elements (Degrees)

Set of elements	No.	PM6	RM1	PM5	PM3	AM1
H, C, N, O	100	3.1	3.1	3.3	2.5	2.7
H, C, N, O, F, P, S, Cl, Br, I	244	3.2	4.0	4.3	3.8	3.4
Whole of main group	900	8.0		8.6	8.5	8.8
70 elements	1681	7.9				

unstable in the gas phase, at least at the PM6 level of calculation. In most cases, such moieties had a high formal charge, therefore, without any countercharge, their instability in isolation is understandable. When an intrinsically unstable ion was identified, it was removed from further consideration.

Most transition metal compounds also have extensive UV-visible properties, arising from *d-d* transitions and from charge-transfer excitations, the presence of these absorption bands being indicative of the existence of low-lying electronic excited states. The self-consistent field (SCF) equations frequently did not converge unless special techniques were used. One of these, using the direct inversion of the iterative sub-space [37], or DIIS, would frequently yield an SCF when other methods failed. However, as a result of the way it works, the DIIS converged the wavefunction to the nearest stationary point, not necessarily to the lowest energy point. Because of the potential existence of multiple low-lying excited states, special care had to be taken when the DIIS technique was used. Conversely, the tendency to converge to the nearest stationary point was an advantage when electronic states of transition metal atoms were being optimized. In several instances, the lowest energy wavefunction corresponded to a hybrid of *s*, *p* and *d* atomic orbitals that did not transform as any irreducible representation of the group of the sphere. In those cases, the wavefunction could be induced to converge to the correct spherical harmonic solution by using the DIIS procedure.

Table 11 Average unsigned errors in bond lengths for various sets of elements (Å)

Set of elements	No.	PM6	RM1	PM5	PM3	AM1
H, C, N, O	413	0.025	0.022	0.033	0.021	0.031
H, C, N, O, F, P, S, Cl, Br, I	712	0.031	0.036	0.044	0.037	0.046
Whole of main group	2636	0.085		0.121	0.104	0.131
70 elements	5154	0.091				

Sets of transition metals

For the purpose of discussion, the set of 30 transition metals can be partitioned into eight of the groups of the Periodic Table, with each group containing one or more triads of elements. A detailed discussion of each element is impractical because of the wide range of compounds in transition metal chemistry. The following section, therefore, will be limited to systems where PM6 does not work well, and to systems illustrative of the structural chemistry of specific elements.

Group IIIA: Scandium, Yttrium, Lanthanum, and Lutetium

Possibly because of its scarcity, only a few experimental thermochemical reference data for scandium compounds were available for use in the parameterization. What reference data existed were augmented by the results of DFT calculations and with a large number of atomic energy levels for the neutral and ionized atom. Only the chemistry of Sc^{III} was studied. Most bond lengths involving scandium were reproduced with good accuracy (for example tri (η^5−cyclopentadienyl)-scandium, Fig. 2), the exception being the coordination complex $[Sc(H_2O)_9]^{3+}$ which PM6 predicts to decompose to $[Sc(H_2O)_7]^{3+}$ plus two water molecules.

As with scandium, very few thermochemical reference data were found for yttrium or lanthanum. To compensate for this, extensive use was made of the CSD.

Table 13 Average unsigned errors in dipole moments for various sets of elements (D)

Set of elements	No.	PM6	RM1	PM5	PM3	AM1
H, C, N, O	55	0.38	0.22	0.31	0.26	0.26
H, C, N, O, F, P, S, Cl, Br, I	131	0.37	0.33	0.50	0.36	0.38
Whole of main group	313	0.60		0.86	0.72	0.65
70 elements	569	0.85				

Table 14 Average unsigned errors in I.P.s for various sets of elements (eV)

Set of elements	No.	PM6	RM1	PM5	PM3	AM1
H, C, N, O	99	0.45	0.40	0.41	0.51	0.45
H, C, N, O, F, P, S, Cl, Br, I	229	0.47	0.41	0.44	0.51	0.56
Whole of main group	383	0.50		0.49	0.68	0.63
70 elements	385	0.50				

The chemistry of lutetium is similar to that of lanthanum, with the principal difference being that whereas La^{III} has an empty 4*f* shell, in Lu^{III} that shell is completely filled. Since the 4*f* shell is, at least chemically, virtually inert, lutetium could be regarded as a conventional transition metal, and was therefore included in this work.

Group IVA: Titanium, Zirconium, and Hafnium

In contrast to all the elements of Group IIIA, titanium is plentiful, and an abundance of reference data on Ti^{III} and the more common Ti^{IV} is available. These data include many tetrahedral and octahedral inorganic complexes as well as organotitanium compounds. Most bond lengths are reproduced with good accuracy, the exceptions being the Ti-H bond in TiH_4, where the predicted value, 1.36 Å, is 0.37 Å shorter than the reference, and coordination complexes which involve oxygen forming a purely dative bond to titanium. In this latter case, the Ti-O bond is typically too long by 0.1 to 0.3 Å.

The behavior of zirconium and hafnium is similar to that of titanium.

Group VA: Vanadium, Niobium, and Tantalum

Most of the structural chemistry of vanadium in its five common oxidation states, 0, II, III, IV, and V, are reproduced with good accuracy. The common VO_5 structure which occurs in bis(Acetylacetonato)-oxo-vanadium (iv), where vanadium forms a double bond to one oxygen atom and single bonds to the other four, is reproduced accurately, the V=O distance being 1.58 Å (reference, 1.56), the V-O distance 2.03 Å (1.97), and the O-V=O angle: 104.5 ° (105.9).

Not all systems were reproduced with such accuracy. When there are several ligands around a vanadium atom, the effects of steric crowding are over-emphasized, and PM6 incorrectly predicts that one of the metal-oxygen bonds would break. An example is bis(bis(μ^2-trifluoroacetato-O,O′)-(η^5-cyclopentadienyl)-vanadium), where each vanadium atom extends bonds to four oxygen atoms and one cyclopentadienyl. In this system, PM6 predicts that one of the V-O bonds would break.

In the heavier elements there is an increased tendency to form highly symmetric polynuclear complexes. An example is the tantalum dication, $[Ta_6Cl_{12}]^{2+}$. This is predicted to have an octahedral structure in modest agreement with the DFT result (Fig. 3).

Transition metal complexes usually have one or more unpaired electrons; such systems can only be modeled using an open shell method such as unrestricted Hartree Fock (UHF) or restricted Hartree Fock followed by a configuration interaction (RHF-CI) correction. The UHF method is faster and more reliable, and is the method of choice when only simple properties such as heats of formation or geometries are of interest. For $[M_6X_{12}]^{2+}$, M = Nb or Ta, X = Cl or Br, UHF predicts an almost octahedral complex, a very slight distortion lowering the symmetry to D_{4h}. This distortion is also reflected in the asymmetric charge distribution. When RHF-CI is used, the geometry converges on the exact O_h structure.

Group VIA: Chromium, Molybdenum, and Tungsten

Most Cr–O and Cr–N bonds are reproduced well, as illustrated by $[Cr^{III}(EDTA)]^-$ in Fig. 4. The organometallic bond in chromium hexacarbonyl is 1.90 Å, which is in good agreement with the crystal structure, 1.92 Å, found in FOHCOU01[21].

The octacyano-molybdate(IV) moiety, $[Mo^{IV}(CN)_8]^{4-}$, is a stable eight-coordinate organometallic molybdenum complex ion whose geometry in the crystal is that of a slightly distorted square antiprism. Rather unexpectedly, this structure was reproduced by PM6, the expectation being that in the absence of crystal field forces the structure would have optimized to a geometry which has

Table 15 Average unsigned errors in phosphorus, sulfur, and chlorine

	ΔH_f (kcal mol^{-1})			Bond length (Å)			Dipole (D)			I.P. (eV)			Angles (Degrees)		
	PM6	AM1*	No.	PM6	AM1*	No.	PM6	AM1*	No.	PM6	AM1*	No.	PM6	AM1*	No.
Phosphorus	8.3	19.1	90	0.022	0.051	56	0.57	0.49	10	0.51	0.81	12	2.5	3.3	19
Sulfur	6.5	10.6	199	0.029	0.060	71	0.36	0.64	14	0.52	0.50	45	3.1	4.1	34
Chlorine	6.1	18.2	156	0.025	0.106	69	0.55	0.60	10	0.52	0.62	25	3.4	14.6	4

Table 16 Statistical analysis of errors in predicted ΔH_f for various methods (kcal mol^{-1})

Statistic	PM6	B3LYP*	HF*
Median	3.26	3.75	5.10
AUE	4.44	5.19	7.37
RMS	6.23	7.42	10.68

No. of molecules in set: 1373
* Basis set: 6–31G*

a higher symmetry, i.e., converged to the exact D_{4d} geometry. The predicted Mo-C distance was 2.22 versus 2.16 Å, again in unexpectedly good agreement for an ion with such a large formal charge.

Molybdenum forms the cluster anion $[Mo_6(\eta_3\text{-}Cl_8)Cl_6]^{2-}$ in which the six molybdenum atoms form a regular octahedron. PM6 successfully reproduces this structure, and predicts the following distances: Mo-Mo: 2.30 (2.63), Mo-η_3Cl: 2.75 (2.56), and Mo-Cl: 2.50 (2.43 Å).

The trioxide of molybdenum can form polyoxometalates, a typical example of which is the α-keggin heteropolyoxyanion $[SiO_4@Mo^{VI}{}_{12}O_{36}]^{4-}$. In this structure, shown in Fig. 5, each Mo forms a double bond with one oxygen, single bonds to four other oxygen atoms, and what can only be described as a third of a bond to a sixth oxygen that is part of the SiO_4 unit. Despite the apparently high symmetry, Td, this system has only a center of inversion. This low symmetry is reproduced by PM6.

PM6 predicts the structures of all three hexacarbonyls with good accuracy, but gives qualitatively the wrong structures for the dinuclear decacarbonyls. This failure to qualitatively predict the structure of the polynuclear carbonyls occurred frequently during the survey of the transition metals.

Group VIIA: Manganese, Technetium, and Rhenium

Like many other transition metals, manganese can form sepulchrates, closo polyhedral complexes of general structure 3, 6, 10, 13, 16, 19-hexaaza-bicyclo(6.6.6)icosane. In contrast to the more common open hexadentate chelates of manganese, e.g. $[Mn^{II}(EDTA)]^{2-}$, the metal atom in a sepulchrate is extremely tightly bound, and cannot be removed without destroying the organic framework. A simple sepulchrate is shown in Fig. 6. PM6 predicts the Mn-N distance with good accuracy but gets the twist angle incorrect. A DFT calculation reproduced the twist angle found in the crystal, which suggests that the error in the twist angle cannot be attributed to the neglect of crystal packing forces.

Although there is a large amount of structural information on technetium compounds, there is a distinct shortage of thermochemical data. To make up for this, almost all the reference heats of formation of representative technetium compounds were derived from DFT calculations. Only one heat of formation was used in this derivation, that of the isolated technetium atom, therefore the reference values used almost certainly include a systematic error that may amount to many kilocalories per mole. Consequently, the reference heats of formation and the errors in PM6 predicted heats of formation of technetium compounds should be taken *cum granus salis*. However, this should not be construed as implying that they are meaningless: because reactions are balanced, when heats of reaction are evaluated, any systematic errors in the heats of formation are cancelled out.

One of the more important technetium species is the pertechnetate ion, $[TcO_4]^-$, used in nuclear medicine. In this ion, PM6 predicts the Tc-O distance to be 1.73 Å, in good agreement with the DFT value of 1.76 Å.

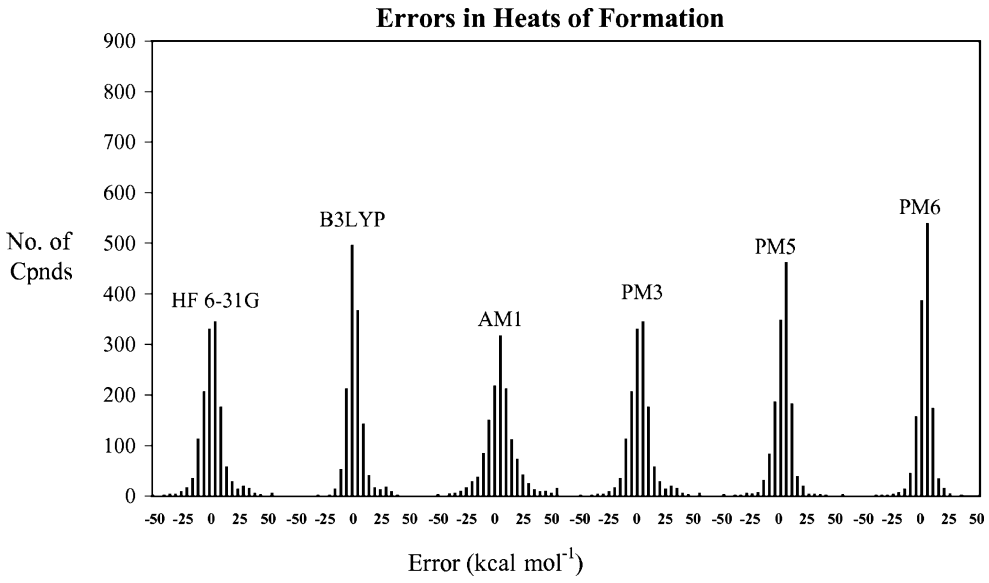

Fig. 1 Histogram of errors in calculated ΔH_f

Table 17 Relative energies of conformers of water dimer

Structure	Ref.	Relative ΔH_f (kcal mol^{-1})			
		PM6	PM5	PM3	AM1
(Non-planar open C_s)*	−5.00	−3.96	−0.24	−2.79	−2.81
1 (Non−planar open C_s)	0.00	0.00	0.00	0.00	0.00
2 (Open C_i)$^+$	0.52	0.83	0.50	0.91	0.64
3 (Planar Open C_s)	0.57	0.66	0.25	0.93	0.46
4 (Cyclic C_i)	0.70	0.29	0.11	2.10	−0.94
5 (Cyclic C_2)	0.95	0.77	0.39	2.63	−0.51
6 (Cyclic C_{2h})	0.99	0.59	0.21	2.71	−0.67
7 (Triply Hydrogen Bonded	1.81	0.93	−1.85	1.16	−0.95
8 (Non-planar Bifurcated	3.57	2.67	−0.83	1.71	1.26
9 (Non-planar Bifurcated	1.79	0.73	−1.95	1.15	−0.87
10 (Planar Bifurcated C_{2v}	2.71	1.42	−1.77	1.28	−0.05

*: Relative to two isolated water molecules
+: Structures 2 – 10 are relative to Structure 1

Group VIIIA: Iron, Cobalt, Nickel, Ruthenium, Rhodium, Palladium, Osmium, Iridium, and Platinum

The geometries of most compounds of this large group were reproduced with modest to good accuracy, including the iron-porphorin complex, Fig. 7, of the type found in heme. The main exception is iron pentacarbonyl, $Fe(CO)_5$, which in its equilibrium geometry is known unambiguously to be of point-group D_{3h}, and which PM6 predicts to be equally unambiguously C_{4v}. When this error was discovered, attempts were made to correct the fault by adding a rule to the training set for iron. This rule stated that "The C_{4v} geometry was 28.7 kcal mol^{-1} higher in energy than the D_{3h} geometry," 28.7 kcal mol^{-1} being the difference between the energies of the two structures calculated using DFT. However, even when a very large weighting factor, 20.0, was used, the C_{4v} structure remained more stable than the D_{3h}, albeit the error in the relative energies was decreased. During this optimization errors in all other iron compounds increased significantly. Rather than accept a general deterioration in the predicted properties of iron compounds, the rule was removed from the training set.

The well-known red complex nickel dimethylglyoxime is normally encountered in the quantitative analysis of inorganic nickel in solution. At the center of the molecule is the planar structure NiN_4 structure, which is frequently found in nickel compounds in biochemical systems. PM6 predicts this with good accuracy (Fig. 8).

One of the first polyhapto organometallic complexes discovered was Zeise's salt. In the anion, $[PtCl_3(\eta^2\text{-}C_2H_4)]^-$, platinum forms a synergic bond with an ethylene molecule. The calculated and X-ray structures of this complex are shown in Fig. 9.

Table 18 Comparison of B3LYP and PM6 hydrogen bond energies (kcal mol^{-1})

Hydrogen-bonded system	Ref	PM6	PM5	PM3	AM1
Ammonia - ammonia	−2.94	−2.34	−0.77	−0.67	−1.41
Water - methanol	−4.90	−5.12	−2.59	−0.20	−4.52
Water - acetone	−5.51	−5.25	−2.43	−2.22	−4.09
Water, dimer, linear (O–H–O = 180°)	−5.00	−3.69	−1.57	−3.49	−3.16
Water, dimer	−5.00	−4.88	−2.43	−1.95	−5.01
Benzene dimer, T-shaped	−2.34	−0.83	−0.22	−0.56	−0.07
Water - acetate anion	−19.22	−18.72	−12.28	−15.77	−15.91
Water - formaldehyde	−5.17	−4.22	−2.17	−2.73	−3.40
Water - ammonia	−6.36	−4.32	−2.75	−1.53	−2.90
Water - formamide	−8.88	−7.60	−4.14	−4.33	−7.54
Formic acid, dimer	−13.90	−10.03	−4.75	−8.65	−6.44
Water - methylammonium cation	−18.76	−14.90	−8.94	−10.48	−14.36
Formamide - formamide	−13.55	−10.83	−4.46	−6.08	−8.14
Acetic acid, dimer	−14.89	−10.33	−4.50	−8.70	−6.44

Table 19 Average errors in pyramidalization of nitrogen (Torsion angle about nitrogen, in degrees)

Statistic	PM6	PM3	AM1	RM1
Average signed error	−1.7	−13.6	0.2	9.7
Average unsigned error	5.0	15.0	3.5	19.1

Group IB: Copper, Silver, and Gold

Copper phthalocyanine is an extremely stable blue dyestuff. As with nickel dimethylglyoxime, the planar CuN_4 moiety at the center of the porphyrin ring is typical of many copper species of importance in biochemistry. PM6 reproduces it with very good accuracy (Fig. 10).

Dimethyl gold cyanide tetramer provides a good example of a square-planar Au^{III} complex. In this system, each gold atom forms covalent single bonds of length 1.99 Å(2.01) to the carbons of the methyl groups, a weaker, longer bond of length 2.12 Å(2.23) to the carbon of the cyanide group, and a still longer bond, 2.27 Å(2.23) to the nitrogen atom.

Gold also forms small planar clusters. PM6 predicts that neutral clusters of up to about nine gold atoms should be planar, an example being the D_{6h} Au_7 cluster, in which the Au-Au distance is predicted to be 2.71 Å(2.01). Clusters of up to 12 gold atoms are also predicted to be stable, provided the cluster has a single negative charge.

Group IIB Zinc, Cadmium, and Mercury

These elements have completely filled *d* shells; therefore the valence shell can be limited to the *s* and *p* orbitals. As such, they behave like main-group elements.

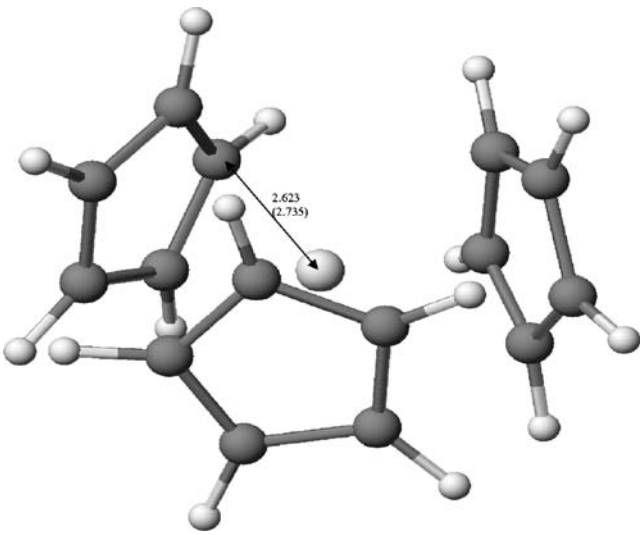

Fig. 2 Tri(η^5−cyclopentadienyl)-scandium Reference value in parenthesis

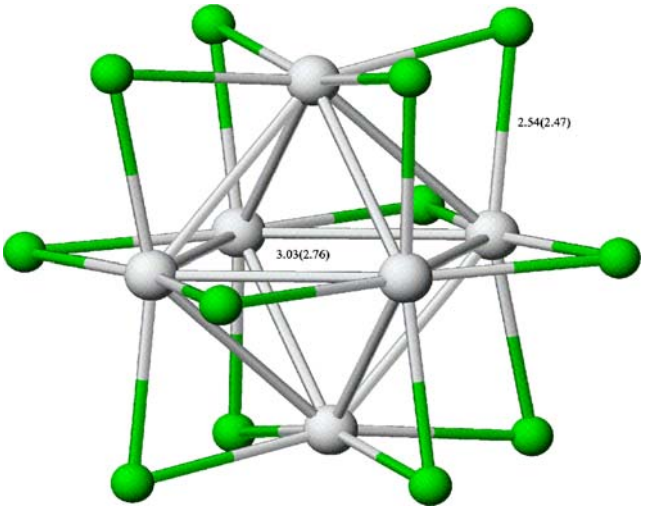

Fig. 3 Calculated structure of the complex ion $[Ta_6Cl_{12}]^{2+}$ Reference value in parenthesis

Discussion

Methodological changes

During the development of PM6, only very minor changes were made to the set of approximations. The main change was in the construction of the training set used for

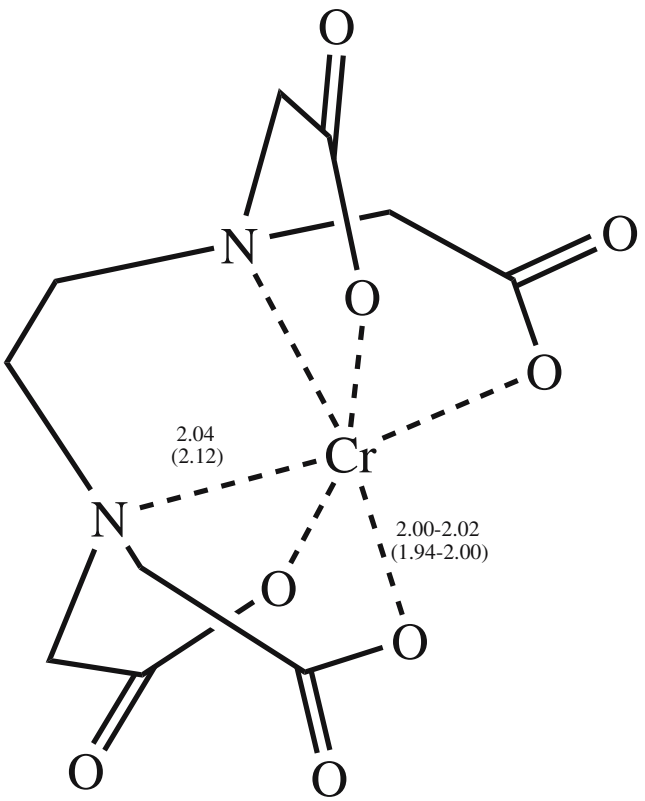

Fig. 4 Chromium Ethylenediaminetetraacetate anion, [Cr(III)(EDTA)]$^-$

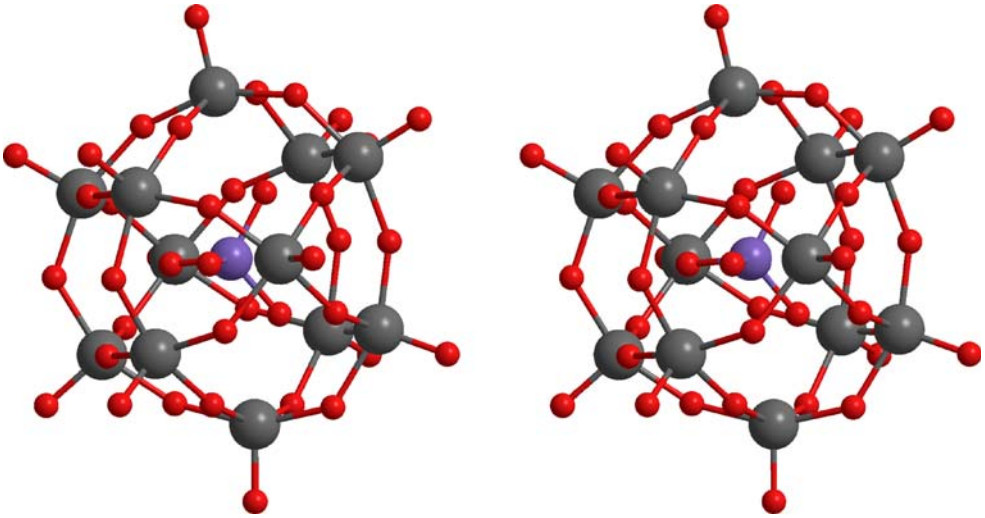

Fig. 5 α-Keggin structure of tetraconta-oxo-silicon-dodeca-molybdenum, $[SiO_4@Mo_{12}O_{36}]^{4-}$ Crossed-eyes stereo; Mo=O: 1.77 Å (1.69), Mo-O: 2.00 (1.85), Si–O: 1.52 (1.64) (Ref. in parentheses)

parameter optimization. One of the most important changes was the use of rules in the training set to define chemical information that was not a function of any single molecule. In earlier methods the training set had included only standard reference data. Of their nature, such data could not allow for chemical facts that were independent of any one moiety. For example, the strength of a hydrogen bond is of great importance in biochemistry, but it could not be expressed in terms of a single species. By use of rules, the value of some chemical quantity could be related to that of another. In the case of hydrogen bonding, the heat of formation of the water dimer was made a function of the heat of formation of two separated water molecules.

Rules were particularly useful when elements of the three transition metal series were being optimized. Many complexes of these elements are highly labile, and, in the early stages of parameter optimization, there was a strong tendency for the optimized geometry of such complexes to be qualitatively incorrect. Faults of this kind could not be corrected by simply increasing the weight assigned to the correct geometry, so rules were developed to indicate that the faulty geometries were indeed incorrect. Specific points on the potential energy surface were selected, and from single-point high level calculations, the relative energy of these points above the minimum was evaluated. The points selected were precisely those qualitatively incorrect geometries resulting from the use of the then-current parameters. The fact that the incorrect geometry was predicted by high level methods to be of higher energy than the correct geometry was then added to the set of rules. A good example of such a rule was the rule concerning $Fe(CO)_5$ mentioned above, in which the only datum that was defined referred to the relative energies of the compound in two different symmetries. No reference was made to the bond lengths, or bond angles. With such a rule in place, the

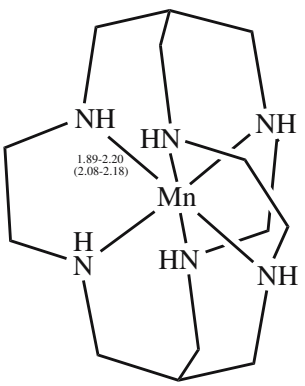

Fig. 6 [Sepulchrate-manganese (III)]$^{3+}$ (3,6,10,13, 16,19-Hexaaza-bicyclo(6.6.6) icosane)-manganese(III) §: CSD entry: HAFBUL

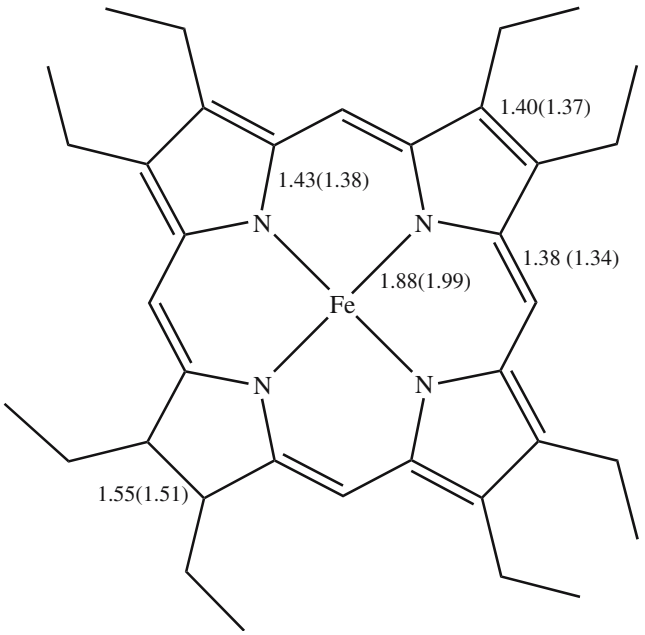

Fig. 7 trans-7,8-Dihydro-2,3,7,8,12,13,17,18-octaethylporphyrinato-iron (II) Reference value (CSD entry BUYKUB) in parenthesis

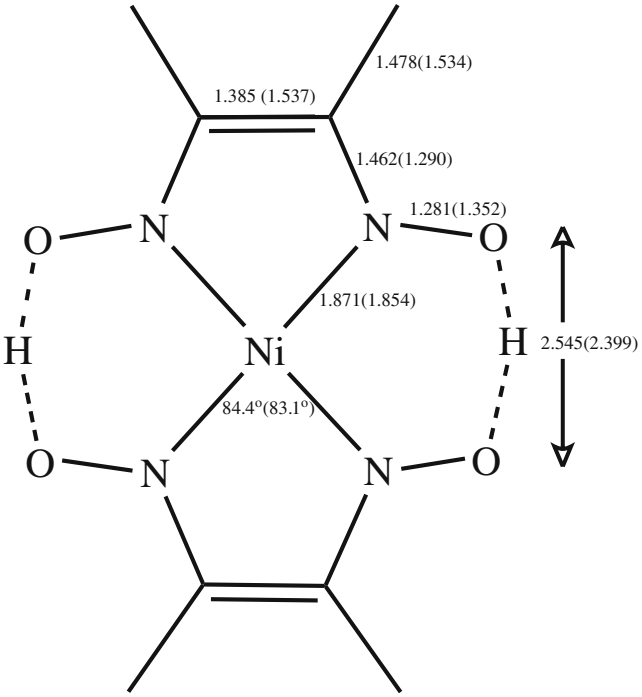

Fig. 8 Nickel Dimethylglyoxime Reference value (CSD entry NIMGLO10) in parenthesis

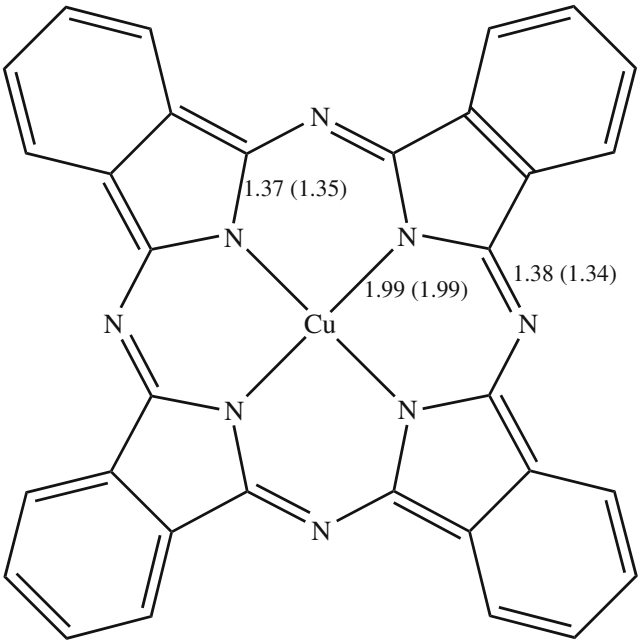

Fig. 10 Copper phthalocyanine Reference value (CSD entry CUPOCY16) in parenthesis

parameters could be re-optimized to minimize the error arising from the rule, with the effect that the energy of the incorrect symmetry increased relative to that of the correct symmetry. In the majority of cases, one rule of this type was sufficient; less frequently, two rules were used, and, in rare cases, even more rules were necessary.

Another change was the use of very large reference data training sets. In earlier parameterizations, the training set used was deliberately made as small as possible. Only when the resulting method was used in a survey of species not used in the training set could the predictive power of the method be determined. The training set used in the development of PM6 was designed to be considerably larger than the survey set. The rationale for this was that, by including in the training set reference data for unconventional species, e.g., non-equilibrium and hypothetical species, a greater region of the error-function surface could be defined. This would in turn, result in a better definition of the values of the parameters. That this is useful can be evidenced by the recent work in parameterizing chlorine at the AM1* level, where the compound 1,1′,2-trichloro-1,2,2′-trifluoroethane, $C_2Cl_3F_3$ has a reported ΔH_f of –173.7 kcal mol^{-1}, but the value predicted using AM1* was –273.9 kcal mol^{-1}. That is, the AM1* value was in error by over 100 kcal mol^{-1}. If this compound had been included in the training set, it is highly likely that the error would have been significantly reduced.

Although over 10,000 reference data were used in the PM6 training set, there are several indications that even this large number is still inadequate for the definition of the values of the parameters, and that an even larger training set would be highly desirable. In light of this, work has begun on identifying species to be added to the training set. During the testing of PM6, several faults were found in the method. Some of these were quickly traced to specific core-core parameters. One of the hydrogen atoms in the complex $[Sc^{III}(H_2O)_7]^{3+}$ was predicted to readily move toward the central atom with the result that a Sc-H bond was formed. Such faults could easily be corrected by the addition to the training set of appropriate reference data from high-level calculations. This was done in several instances, and the specific error was corrected, but this action then also required all the testing to be re-started. Because this was a time-consuming process, when faults were found near the end of the testing phase, the decision was taken that the fault should be noted, as in the Sc-H error mentioned here, and to take no further action at that time.

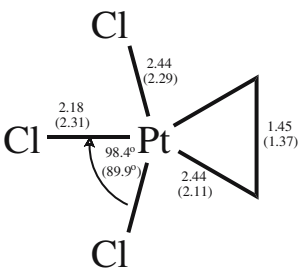

Fig. 9 Zeise's Salt, trichloro-(η^2-ethene)-platinate Reference value (CSD entry XIVSAK) in parenthesis

A different type of error, found only near the end of testing, was the unrealistically large *p* electron population of some transition metals. The values of the parameters that determine the *p* population are defined using two very different groups of reference data: atomic energy levels and conventional properties of polyatomics. If atomic energy levels were excluded from the parameter optimization, then the *p* population would become very small; but if atomic energy levels were excluded, then the resulting method would not be suitable for reproducing such levels. The decision to use all available atomic energy levels in the training set was a value judgement. In the next training set, it is likely that the result of this decision-making process will be different.

Detecting faults in semiempirical methods is difficult, and rather than wait until all errors of this type were found and fixed, a process that could potentially take several more years, the decision was made to freeze the parameters at their current value. Obviously, PM6 still has many errors; some have already been described. Work has already started in an attempt to correct them.

Elimination of computational artifacts

Earlier NDDO methods, particularly PM3 and AM1, produced artifacts in potential energy surfaces as a result of unrealistic terms in the core-core approximation, specifically in the set of Gaussian functions used. In PM6, only one Gaussian-type correction to the core-core potential is allowed, and, consequently, the potential for these artifacts has been reduced. On the other hand, because PM6 uses diatomic parameters, the likelihood of readily-characterized errors involving specific pairs of atoms, e.g. Sc and H, as mentioned earlier, is increased. Errors of this type can be easily eliminated by a re-parameterization of the faulty diatomic.

There are over 450 sets of diatomic interactions parameterized in PM6, covering most of the common types of chemical bonds. But the number of potential bonds is much larger: given 70 elements, there are almost 2500 diatomic sets. If a molecule contains two elements for which the diatomic interaction parameters are missing, then, provided the elements are well separated, say by more than 4 Ångstroms, the absence of the parameters will not be important. If the two elements were near to each other, then the diatomic core-core parameters would be needed. This would involve generating a small training set of reference data that included a few examples of the type of interaction involved, and optimizing the two terms in the diatomic interaction.

This ability to add diatomic parameter sets to PM6 without modifying the underlying parameterization has the advantage that more and more types of interaction can be added without changing the essential nature of the method.

Accuracy

PM6, being the most recent member of the NDDO family of approximate semiempirical methods, is understandably the most accurate. The development of each new method has been guided by the knowledge of the documented faults found in the earlier methods. This is reflected in the steady decrease in AUE of simple organic compounds, from 12.0 kcal mol^{-1} for AM1 to 4.9 kcal mol^{-1} for PM6.

Several low-energy phenomena are predicted more accurately by PM6, with the most important of these being the prediction of the energies and geometries involved in hydrogen bonding. One consequence of this increased accuracy is that the lowest energy conformer of acetylacetone is now correctly predicted to be the ene-ol structure, and not the twisted di-one configuration.

Despite the improvement in hydrogen bonding, a significant error was found in the balance of energies involved in forming zwitterions of hydroxyl and amine groups. This is best illustrated by the dimer of 2-aminophenol, where PM6 predicts that the zwitterion should be 3.6 kcal mol^{-1} more stable than the neutral form, but higher level calculations indicate that the neutral form should be 17.7 kcal mol^{-1} more stable than the zwitterion. In the solid state, CSD entries AMPHOM01 – AMPHOM10 [21], 2-aminophenol exists as the neutral species.

In general, however, average unsigned errors in ΔH_f have steadily decreased as semiempirical methods have evolved. Earlier NDDO methods such as PM3 and AM1 had AUE significantly larger than the 6–31G* Hartree Fock method. With the advent of PM5 and RM1 errors were intermediate between HF and B3LYP. In the current work, AUE in ΔH_f are lower than those of both B3LYP and HF 6–31G*. This increase in accuracy of prediction of ΔH_f relative to higher level methods should not be construed as disparaging those methods: semiempirical methods in general, and PM6 in particular, were parameterized to reproduce ΔH_f. The performance of these methods when applied to non-equilibrium systems, in particular transition states, is likely to be very inferior to that of B3LYP or HF 6–31G*.

As a result of the current work, there is a clear strategy for further improving the accuracy of semiempirical methods. All three potential sources of error need to be addressed. Regarding reference data, considerably more data are needed than were used here. This would likely come from increased use of high-level theoretical methods: methods significantly more accurate than those used here would obviously be needed in any future work. Parameter optimization can be performed with confidence and reliability, particularly when well-behaved systems are used. In all cases examined where problems were encountered in parameter optimization, problems also occurred in the normal SCF calculation in MOPAC2007. This implies

that as faults in the SCF procedure are corrected, faults in parameter optimization would also be removed.

Permanent errors

Notwithstanding the optimism just expressed, not all errors can be eliminated by better data and better optimizations. Despite strenuous efforts, some calculated quantities persistently failed to agree with the reference values. Many potential causes for these failures were investigated. In each case the weight for the offending quantity was increased considerably and the parameter optimization re-run. When that was done, the specific error decreased, but errors elsewhere increased disproportionately. Since the final gradient of the error function was acceptably small, it followed that the parameter optimization was not in error. The reference data were checked to ensure that they were in fact trustworthy. Because two of the three possible origins of error had been eliminated, the inescapable conclusion was that there is a fault in the set of approximations. The most serious of these faults was the qualitatively incorrect prediction of the geometry of the exceedingly simple system, iron pentacarbonyl.

Conclusions

The NDDO method has been modified by the adoption of Voityuk's core-core diatomic interaction parameters. This has resulted in a significant reduction in error for compounds of main-group elements, and, together with Thiel's *d*-orbital approximation, allows extension of the NDDO method to the whole of the transition metal block.

The accuracy of PM6 in predicting heats of formation for compounds of interest in biochemistry is somewhat better than Hartree Fock or B3LYP DFT methods, using the 6-31G(d) basis set. For a representative set of compounds, PM6 gave an average unsigned error of 4.4 kcal mol^{-1}; for the same set HF and B3LYP had AUE of 7.4 and 5.2 kcal mol^{-1}, respectively.

The potential exists for further large increases in accuracy. This would likely result from the increased use of accurate reference data derived from high-level methods, and from the development of better tools for detecting errors at an early stage of method development.

Acknowledgments This work was funded by the National Institute of General Medical Sciences (NIGMS) Grant No. 2 R44 GM067327–02.

References

1. Pople JA, Santry DP, Segal GA (1965) J Chem Phys 43:S129–S135
2. Pople JA, Beveridge DL, Dobosh PA (1967) J Chem Phys 47:2026–2033
3. Dewar MJS, Thiel W (1977) J Am Chem Soc 99:4907–4917
4. Dewar MJS, Thiel W (1977) J Am Chem Soc 99:4899–4907
5. Bingham RC, Dewar MJS, Lo DH (1975) J Am Chem Soc 97:1285–1293
6. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107:3902–3909
7. Stewart JJP (1989) J Comp Chem 10:209–220
8. Stewart JJP (1989) J Comp Chem 10:221–264
9. Stewart JJP (1991) J Comp Chem 12:320–341
10. Anders E, Koch R, Freunscht P (1993) J Comp Chem 14:1301–1312
11. Stewart JJP (2004) J Mol Model 10:155–164
12. Thiel W, Voityuk AA (1992) Theor Chim Acta 81:391–404
13. Thiel W, Voityuk AA (1996) J Phys Chem 100:616–629
14. Voityuk AA, Rösch N (2000) J Phys Chem A 104:4089–4094
15. Winget P, Clark T (2005) J Mol Model 11:439–456
16. Giese TJ, Sherer EC, Cramer CJ, York DM (2005) J Chem Theory and Comp 1:1275–1285
17. Tejero I, González-Lafont A, Lluch JM (2007) J Comp Chem 28:997–1005
18. Arantes GM, Loos M (2006) Phys Chem Chem Phys 8:347–353
19. McNamara JP, Sundararajan M, Hillier IH, Ge J, Campbell A, Morgado C (2006) J Comp Chem 27:1307–1323
20. Afeefy HY, Liebman JF, Stein SE (2003) Neutral Thermochemical Data In: Linstrom PJ, Mallard, WG, Eds. NIST Chemistry WebBook, NIST Standard Reference Number 69 http://webbook nistgov/chemistry National Institute of Standards and Technology, Gaithersburg MD, 20899
21. Allen FH (2007) Acta Cryst B 58:380–388
22. Stewart JJP (2004) J Phys Chem Ref Data 33:713–724
23. Stewart JJP (2004) J Mol Modelling 10:6–12
24. http://webbook.nist.gov/cgi/cbook.cgi?Source=2004STE6-10&Units=SI&Mask=1
25. Becke AD (1993) J Chem Phys 37:5648–5652
26. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789
27. Perdew JP, Chevary JA, Vosko SH, Jackson KA, Pederson MR, Singh DJ, Fiolhais C (1992) Phys Rev B 46:6671–6687
28. Wang B, Ford G (1994) J Comp Chem 15:200–207
29. Moore CE Atomic energy levels, Vol. I (Hydrogen through Vanadium); Natl. Bur. Stand. (U.S.) Cir. No. 467 (U.S. GPO, Washington, D.C.)], 1946
30. Moore CE Atomic Energy Levels, Vol. II (Chromium through Niobium); Circular of the National Bureau of Standards 467, U.S. Government Printing Office: Washington, DC, 1952
31. Moore CE Atomic Energy Levels, Vol. III (Molybdenum through Lanthanum and Hafnium through Actinium); Circular of the National Bureau of Standards 467, U.S. Government Printing Office: Washington, DC, 1958
32. Lennard-Jones JE (1931) Proc Phys Soc 43:461–482
33. Stewart JJP; 7.101W; Stewart Computational Chemistry: Colorado Springs, 2007
34. Dewar MJS, Jie C, Zoebisch EG (1988) Organometallics 7:513–521
35. Rocha GB, Freire RO, Simas AM, Stewart JJP (2006) J Comp Chem 27:1101–1111
36. Tschumper GS, Leininger ML, Hoffman BC, Valeev EF, Schaefer HF, Quack M (2002) J Chem Phys 116:690–701
37. Pulay P (1980) Chem Phys Lett 73:393–398