Representation of van der Waals (vdW) Interactions in Molecular Mechanics Force Fields: Potential Form, Combination Rules, and vdW Parameters

Thomas A. Halgren

Contribution from the Molecular Systems Department, Merck Research Laboratories, Rahway, New Jersey 07065. *Received October 18, 1991*

Abstract: This paper explores the premise that insights gained from studying the well-characterized van der Waals (vdW) interactions of rare-gas atoms can be used advantageously in formulating the representation of vdW nonbonded interactions in molecular mechanics force fields, a subject to which little attention has been given to date. We first show that the commonly used Lennard–Jones and Exp-6 potentials fail to account for the high quality rare-gas data but that a relatively simple distance-buffered potential (Buf-14-7, eq 10) accurately reproduces the reduced rare-gas potentials over the range of interatomic separations of primary interest in molecular mechanics calculations. We also show that the standard arithmetic- and geometric-mean combination rules used in molecular mechanics force fields perform poorly, and we propose alternative "cubic-mean" and "HHG" combination rules for minimum-energy separations R^*_{ij} and well depths ϵ_{ij} (eqs 12, 14) which perform significantly better. We then make further use of the known behavior of the rare gases by developing a formalism for relating ϵ and R^* to experimentally derived data on atomic polarizabilities and on the Slater–Kirkwood "effective number of electrons" for the interacting atoms (eqs 27, 35). This formalism yields the vdW parameters (Table XIII) which we propose to use in the Merck Molecular Force Field (MMFF) being developed in our laboratories. Comparisons to other force fields such as MM2, our laboratory's MM2-based MM2X, AMBER, VFF, CHARMM, and MM3 demonstrate wide variations in vdW parameters from force field to force field but reflect broad agreement with the calculated MMFF values apart from a tendency of the MMFF formalism (i) to yield slightly larger minimum-energy separations and (ii) essentially in agreement with MM2 and MM3 but contrary to AMBER, VFF, and CHARMM to give vdW well depths which do not depend on the chemical environment.

1. Introduction

Molecular mechanics force fields are widely used in computational simulations of organic, bioorganic, and polymeric systems. In pharmaceutical chemistry, such simulations promise to allow quantitative predictions of phenomena such as the differential free energy of binding to a macromolecular receptor of one ligand relative to another. Indeed, encouraging results based on the free-energy perturbation approach have already been reported in systems of chemical interest.[1] Nevertheless, the widely used empirical potentials (force fields) neglect important physical interactions, such as those arising from molecular polarizability, or retain them in overly simplified form, an example being the use of simple atom-centered charges as opposed to more general and more accurate representations based on atom-centered multipoles. This situation is changing. In particular, much has been learned in recent years about how to represent molecular charge distributions,[2] and several efforts are underway to find suitable means for incorporating polarizability into molecular mechanics and dynamics calculations.[3]

The rapid progress being made on these fronts offers the promise that far more accurate methods for evaluating intermolecular interactions can be developed. Reaching this promise, however, may well require that van der Waals (vdW) interactions, the other principal class of nonbonded interactions, also be described comparably well. Nevertheless, to date relatively little attention has been given to their representation in molecular mechanics force fields. This paper will attempt to redress the balance by presenting a framework within which such interactions can be characterized and systematically related to experimentally well-understood interactions involving simpler atomic systems, principally ones involving rare-gas atoms. Our premise is that molecular mechanics potentials which account accurately for the rare-gas data will better be able to describe vdW interactions in polyatomic molecules and to contribute to the definition of empirical force fields which model physical reality closely enough to justify confidence in their use.

The approach taken in this paper will focus on defining on what might be termed "true" vdW parameters—these being those parameters that would be appropriate for use in a molecular mechanics force field in which electrostatic and other significant physical interactions are described accurately. Accordingly, our approach will differ markedly from the standard but notoriously difficult and problematic approach in which vdW parameters are obtained by fitting to data on intermolecular interactions in the hope that errors and omissions made in the description of other physical terms can thereby be counterbalanced. As it happens, we ourselves will first employ the vdW parameters generated by the new approach in a molecular mechanics force field which still describes electrostatic effects relatively crudely. Nevertheless, we wish to emphasize that the vdW parameters developed here are ultimately intended for use in a more complex, physically superior force field, and indeed may perform optimally only in such a context.

We will begin in Section 2 by using high quality data for pairwise interactions of the rare-gas atoms helium, neon, and argon to assess the accuracy of the vdW forms currently employed in molecular mechanics force fields. We will define a new form for the vdW potential which is significantly more accurate but still simple enough for such use. In Section 3 we will then examine the *combination rules* currently used to represent the interaction between atoms of types i and j in terms of the parameters which describe the i,i and j,j like-pair interactions. We will contrast these rules with those needed to accurately account for the rare-gas data, and will present the basis for an alternative formulation which achieves considerably higher accuracy. Section 4 will focus on the "well depth" and "vdW separation" parameters for the rare gases. We will present an algorithm which defines these parameters for an atom of specified type in terms of its assigned atomic polarizability and will show that this algorithm accurately re-

(1) See, for example: (a) Merz, K. M., Jr.; Murcko, M. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 4484–4490. Note, however, the correction in Merz, K. M., Jr.; Murcko, M. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1992**, *114*, 1128. (b) Beveridge, D. L.; DiCupa, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431. (c) Kollman, P. A.; Merz, K. M., Jr. *Acc. Chem. Res.* **1990**, *23*, 246–252. (d) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184–189.

(2) See, for example: (a) Dinur, U. *J. Comput. Chem.* **1991**, *12*, 469–486. (b) Dinur, U. *J. Comput. Chem.* **1991**, *12*, 91–105. (c) Dinur, U. *J. Am. Chem. Soc.* **1990**, *112*, 5569–5671.

(3) See, for example: (a) Corongiu, G.; Migliore, M.; Clementi, E. *J. Chem. Phys.* **1989**, *90*, 4629. (b) Dang, L. X.; Rice, J. E.; Caldwell, J.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 2481–2486, and references therein. (c) Sprik, M. *J. Phys. Chem.* **1991**, *95*, 2283–2291, and references therein. (d) Zhu, S.-B.; Yao, S.; Zhu, J.-B.; Singh, S.; Robinson. G. W. *J. Phys. Chem.* **1991**, *95*, 6211–6217.

0002-7863/92/1514-7827$03.00/0

produces the known vdW parameters for like- and unlike-pair interactions of He, Ne, Ar, Kr, and Xe. In Section 5, we will present and discuss the vdW well depth and separation parameters generated by this algorithm for the atom types defined in the new Merck Molecular Force Field (MMFF)[4] being developed in this laboratory.

2. The Form of the van der Waals Potential

2.1. vdW Potentials Used in Molecular Mechanics Calculations. Among widely used molecular mechanics methods, the MM2[5] and MM3[6] force fields use the Exp-6 form for nonbonded vdW interactions, whereas AMBER[7] and CHARMM[8] use the Lennard-Jones 12-6 form, and VFF[9] can use either the Lennard-Jones 12-6 or 9-6 form. These programs sum vdW interactions over all i,j atom pairs except that each omits 1,2- and 1,3- pairs, and

$$E_{\text{vdW(Exp-6)}} = \sum \epsilon_{ij}\{184\,000 \exp(-12R_{ij}/R^*_{ij}) - 2.25(R^*_{ij}/R_{ij})^6\} \quad (1)$$

$$E_{\text{vdW(LJ-12-6)}} = \sum \epsilon_{ij}\{(R^*_{ij}/R_{ij})^{12} - 2(R^*_{ij}/R_{ij})^6\} \quad (2)$$

$$E_{\text{vdW(LJ-9-6)}} = \sum \epsilon_{ij}\{2(R^*_{ij}/R_{ij})^9 - 3(R^*_{ij}/R_{ij})^6\} \quad (3)$$

AMBER and CHARMM scale the 1,4-interactions. In eqs 2 and 3, $-\epsilon_{ij}$ is the energy for the i,j interaction attained at the minimum-energy separation $R_{ij} = R^*_{ij}$; for the MM3 potential in eq 1 (the similar MM2 potential employs different fixed constants), the interaction energy when $R_{ij} = R^*_{ij}$ is $-1.1195\ \epsilon_{ij}$. VFF and AMBER, as well as the OPLS[10] and GROMOS[11] force fields, implement eq 2 or 3 in the equivalent form

$$E_{\text{vdW(LJ-}n\text{-6)}} = \sum\{A_{ij}/R_{ij}^n - B_{ij}/R_{ij}^6\} \quad (4)$$

where $n = 12$ or 9. As will be discussed in Section 3, each of these force fields uses simple arithmetic- or geometric-mean combination rules to generate the parameters for i,j-pair interactions from the like-pair parameters.

2.2. Potentials Used To Describe Rare-Gas Interactions. In contrast to the narrow range of vdW forms used in molecular mechanics force fields, experimentalists engaged in fitting data on interatomic interactions have employed these and more than 20 additional potential forms.[12,13] The forms which appear to have found greatest use in describing rare-gas interactions in recent years include the HFD (Hartree–Fock + damped dispersion) potential[14] and its variants HFD-B[15] and HFD-C,[16] though other highly accurate forms[17,18] have also been developed.

The HFD potential represents a generalization of the Exp-6 form. It consists of a Born–Mayer repulsive term, sometimes parameterized against the results of ab initio Hartree–Fock calculations, to which is added a damped power series which represents the correlation (dispersion) energy:

$$E_{\text{vdW(HFD)}} = \Delta E_{\text{HF}} + \Delta E_{\text{disp}} \quad (5)$$

$$\Delta E_{\text{HF}} = A \exp(-bR) \quad (6)$$

$$\Delta E_{\text{disp}} = -(C_6/R^6 + C_8/R^8 + C_{10}/R^{10} + \ldots)F(R) \quad (7)$$

$$F(R) = \exp(-(DR^*/R - 1)^2) \qquad R \le DR^* \quad (8)$$

$$= 1 \qquad R \ge DR^*$$

In eqs 7 and 8, the C_n are the dispersion coefficients and D is a damping factor. The damping function $F(R)$, an early form for which was proposed by Brooks,[19] preserves the appropriate long-range behavior of the dispersion energy while preventing the dispersion term from overwhelming the bounded repulsive term at short range. The HFD variants modify the repulsive term by adding a term cR^2 to the exponential factor (in HFD-B) or by multiplying the preexponential factor by R^γ (in HFD-C). In addition to being relatively complex, the HFD forms are not suitable for use in molecular mechanics calculations, however, because the second derivative of the damping factor $F(R)$ is discontinuous at $R = DR^*$. This difficulty is circumvented in related forms which employ continuous and differentiable damping functions[18,20] but at the cost of additional complexity.

While a variety of complex forms are available, no simple forms for vdW interactions which have suitable mathematical properties and describe the rare-gas data accurately have yet been formulated for use in molecular mechanics calculations.[21] The purpose of the remainder of this section is to propose such a form and to demonstrate its ability to account for the well-characterized vdW potentials for rare-gas interactions involving He, Ne, and Ar atoms.

2.3. Buffered 14-7 Potential. The potential proposed here has the following general form for the vdW interaction between atoms i and j:

$$E_{\text{vdW(Buf-}n\text{-}m)_{ij}} = \epsilon_{ij}\left(\frac{1+\delta}{\rho_{ij}+\delta}\right)^{(n-m)}\left(\frac{1+\gamma}{\rho_{ij}^m+\gamma} - 2\right) \quad (9)$$

where δ and γ are *buffering constants* and $\rho_{ij} = R_{ij}/R^*_{ij}$. The specific form we propose has $n = 14$, $m = 7$, $\delta = 0.07$, and $\gamma = 0.12$. We call this the Buffered 14-7 form, or Buf-14-7 for short. Other combinations of n (from 12 to 15) and m (from 6 to 8) were considered but found less satisfactory. Note that the choice $n = 12$, $m = 6$, and $\delta = \gamma = 0$ recovers the LJ-12-6 potential. We will take the constants δ and γ, determined from fits to the rare-gas

(4) The reader should note that the MMFF force field being developed at Merck is not related to the "MMFF option of CHEMLAB-II" cited by Hopfinger and co-workers (cf. Cardozo, M. G.; Kawai, T.; Sugimoto, H.; Yamanishi, Y.; Hopfinger, A. J. *J. Med. Chem.* **1992**, *35*, 590–601).

(5) (a) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *89*, 8127. (b) Bukert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982. (c) Allinger, N. L.; Yuh, Y. *QCPE* **1980**, *12*, 395.

(6) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.

(7) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230–252; Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.

(8) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.

(9) Lifson, S.; Hagler, A. T.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5111–5121, and references therein. Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Proteins* **1988**, *4*, 31–47. This is the force field used in the program DISCOVER distributed by Biosym Technologies, Inc.

(10) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 1489–1496.

(11) Hermans, J.; Berendsen, H. J. C.; van Gunsteren, W. F.; Postma, J. P. M. *Biopolymers* **1984**, *23*, 1513–1518.

(12) Maitland, G. C.; Rigby, M.; Smith, E. B.; Wakeham, W. A. *Intermolecular Forces*; Oxford University Press: New York, 1981.

(13) Kaplan, I. G. *Theory of Molecular Interactions*; Elsevier: New York, 1986.

(14) Hepburn, J.; Scoles, G.; Penco, R. *Chem. Phys. Lett.* **1975**, *19*, 119; Ahlrichs, R.; Penco, R.; Scoles, G. *Chem. Phys.* **1977**, *19*, 119.

(15) Aziz, R. A.; Chen, H. H. *J. Chem. Phys.* **1977**, *67*, 5719.

(16) Aziz, R. A.; Meath, W. J.; Allnatt, A. R. *Chem. Phys.* **1983**, *78*, 295–309.

(17) Pack, R. T.; Valentini, J. J.; Becker, C. H.; Buss, R. J.; Lee, Y. T. *J. Chem. Phys.* **1982**, *77*, 5475–5485.

(18) da Silva, J. D.; Brandao, J.; Varandas, A. J. C. *J. Chem. Soc., Faraday Trans. 2* **1989**, *85*, 1851–1875.

(19) Brooks, F. C. *Phys. Rev.* **1952**, *86*, 92; see also: Ahlrichs, R. *Theor. Chim. Acta* **1976**, *41*, 7–15.

(20) Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* **1984**, *80*, 3726–3741.

(21) A referee has suggested that the Maitland–Smith n(r)-6 potential (cf. ref 12, Appendices 1, 2, and 10) might be sufficiently simple and accurate for such use. However, this potential employs a nonintegral repulsive-term exponent which varies with the interatomic separation. Moreover, in addition to parameters for the well depth and the minimum-energy separation, it employs two ancillary shape parameters whose values also depend on the identities of the interacting atoms. In contrast, the shape parameters γ and δ used in the Buf-14-7 potential are fixed constants. Thus, the n(r)-6 potential is more complex and may well be less tractable mathematically. Very recently, promising results have been shown for the Morse potential (cf. Hart, J. R.; Rappe, A. K. *J. Chem. Phys.* **1992**, *97*, 1109–1115). However, the Morse function employs exponential terms and requires an additional shape parameter whose value depends on the identities of the interacting atoms.

Table I. vdW Parameters for Rare-Gas Interactions

interaction	attractive well depth[a]	minimum-energy distance[b]
He···He[c]	0.02176	2.963
He···Ne[d]	0.04151	3.0355
Ne···Ne[e]	0.08396	3.087
He···Ar[d]	0.05892	3.4777
Ne···Ar[f]	0.13431	3.4889
Ar···Ar[g]	0.28465	3.761

[a] In kcal/mol. [b] In Å. [c] For the HFD-B(HE) potential of ref 25. [d] For the HFD-B potential of ref 27. [e] For the HFD-C2 potential of ref 16. [f] For the HFD-B potential of ref 28. [g] For the HFD-B4 potential of ref 26.

data, as applying to all i,j interactions. Just as for the Exp-6 and Lennard–Jones potentials, then, the Buf-14-7 potential for a given i,j interaction is determined by just *two* interaction-specific parameters, the well depth, ϵ_{ij}, and the minimum-energy distance, R^*_{ij}:

$$E_{vdW(Buf\text{-}14\text{-}7)_{ij}} = \epsilon_{ij}\left(\frac{1.07R^*_{ij}}{R_{ij} + 0.07R^*_{ij}}\right)^7\left(\frac{1.12R^*_{ij}{}^7}{R_{ij}{}^7 + 0.12R^*_{ij}{}^7} - 2\right) \quad (10)$$

Each of these forms also depends on certain *shape parameters*. These parameters include the exponents chosen for the dispersion and repulsion terms; for MM3, they include the dispersion exponent and the two factors used in the repulsion term. Relative to the Lennard–Jones potentials, the Buf-14-7 potential thus employs two additional shape parameters—δ and γ.

For each of the potentials discussed above, a dimensionless *reduced potential* can be defined by dividing either formally (for potentials written in terms of ϵ and R^*) or numerically (for the rare-gas potentials) by ϵ_{ij} (or by 1.1195 ϵ_{ij} for MM3) and by measuring the interatomic separation in multiples of R^*_{ij} (i.e., in terms of ρ_{ij} in eq 9). When buffering constants $\delta = 0.07$ and $\gamma = 0.12$ are used, as in eq 10, the reduced Buf-14-7 potential has a minimum at $\rho = 0.996$ and a value of −1.0006, close to the idealized values of 1 and −1 which hold for the other reduced potentials. In the following subsections, we will examine the degree to which the corresponding reduced forms of the accurate rare-gas potentials conform to one another and to the ones derived from eqs 1–3 and 9 or 10, for a common reduced potential will be appropriate for use in general applications only if actual vdW potentials are conformable.

Before leaving this subsection, it may be helpful to describe some of the qualitative features of the Buf-14-7 potential. The terms involving the buffering constants δ and γ combine to produce the repulsive part of the potential (formal exponent $n = 14$), while the δ-buffered term and the constant term of −2 describe the dispersion interaction (formal exponent $n - m = 7$). The buffered terms keep the potential finite as $R_{ij} \to 0$, avoiding the too-strong divergence found in the unbuffered Lennard–Jones potentials (see below). Similarly motivated buffered terms have previously been employed in a variety of contexts.[22,23] The use of $n - m > 6$ allows the dispersion term to accurately reproduce the power-series expansion of eq 7 for distances of up to a few times R^*_{ij}, while a positive value for δ serves to damp the dispersion term at small R_{ij}, somewhat as does $F(R)$ in the HFD potentials.

If δ in eq 9 is increased, the repulsion energy at small R_{ij} can be greatly diminished without changing the distance R_{ij} at which the potential crosses 0 and without greatly altering either the position or the depth of the energy minimum. This property may allow more strongly buffered versions of the Buf-14-7 potential

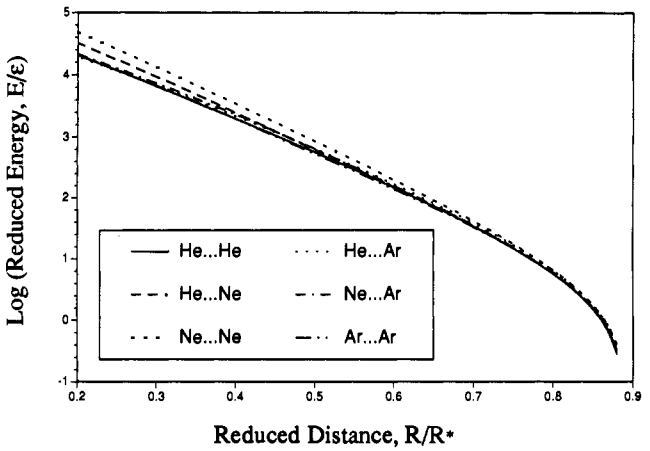

Figure 1. Comparison of reduced potentials $v = E/\epsilon$ for rare-gas interactions at short interatomic separation, where E is the interaction energy and ϵ is the well depth. The dimensionless separation coordinate, designated as ρ in the text, is the interatomic separation R divided by the minimum-energy separation R^*. The comparison shows the degree to which the rare-gas potentials are conformable at short internuclear separation.

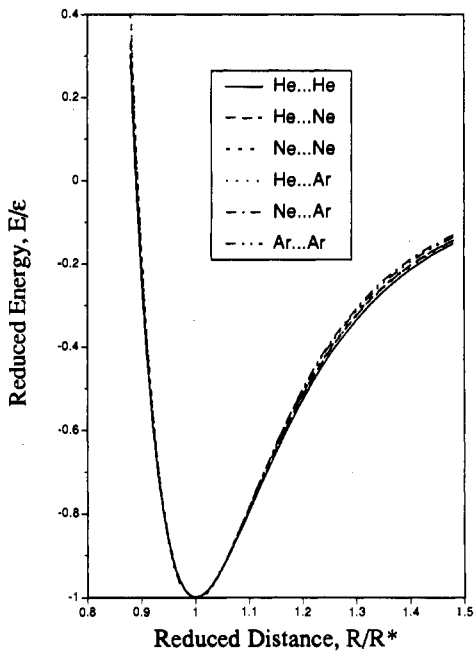

Figure 2. Comparison of reduced potentials for rare-gas interactions near the vdW minimum.

to find advantageous use in molecular mechanics and dynamics calculations, e.g., for optimizing structures having crude initial geometries (as can arise, for example, in X-ray determinations of macromolecules) and for "growing" and "disappearing" structural elements in free-energy perturbation calculations.[24]

2.4. Accurate vdW Potentials for He, Ne, and Ar. To characterize the behavior of accurate vdW potentials, we shall utilize the HFD-B(HE) potential for He···He,[25] the HFD-C2 potential for Ne···Ne,[16] the HFD-B4 potential for Ar···Ar,[26] and the HFD-B potentials for He···Ne, He···Ar,[27] and Ne···Ar.[28] These potentials have recently been derived from simultaneous fits to several kinds of experimental data. In particular, each has been explicitly

(22) (a) Gresh, N.; Claverie, P.; Pullman, A. *Int. J. Quantum Chem.* **1982**, *22*, 199–215. (b) Gresh, N.; Claverie, P.; Pullman, A. *Int. J. Quantum Chem. Symp.* **1979**, *13*, 243. (c) Langlet, J.; Claverie, P.; Caron, F. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 397–429.

(23) (a) Mataga, N.; Nishimoto, K. *Z. Phys. Chem.* **1957**, *13*, 140. (b) Ohno, K. *Theor. Chim. Acta* **1964**, *2*, 219–227. (c) Klopman, J. *J. Am. Chem. Soc.* **1964**, *86*, 4550.

(24) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1989**, *90*, 3300.

(25) Aziz, R. A.; McCourt, F. R. W.; Wong, C. C. K. *Mol. Phys.* **1987**, *61*, 1487–1511.

(26) Aziz, R. A.; Slaman, M. J. *J. Chem. Phys.* **1990**, *92*, 1030–1035.

(27) Keil, M.; Danielson, L. J.; Dunlop, P. J. *J. Chem. Phys.* **1991**, *94*, 296–309.

(28) Barrow, D. A.; Aziz, R. A. *J. Chem. Phys.* **1988**, *89*, 6189–6194.

Table II. Interaction Energies at Half the vdW Minimum-Energy Separation

interaction	MP4/ 6-311G(d,f)[a]	MP4/ 6-311G(3df,3pd)[a]	extrap[b]	exptl[c]
He···He	13.14	12.52	12.14	12.14
	(604)	(575)	(558)	(558)
He···Ne	28.77	26.92	25.57	25.40
	(693)	(649)	(616)	(612)
Ne···Ne	75.82	72.79	71.16	71.04
	(903)	(867)	(848)	(846)
He···Ar	42.67	40.26	38.61	37.52
	(724)	(683)	(655)	(637)
Ne-Ar	104.42	96.89	91.15	84.55
	(778)	(721)	(678)	(630)
Ar···Ar	180.46	163.92	150.33	149.7
	(634)	(576)	(528)	(526)

[a] BSSE-corrected interaction energies in kcal/mol computed at $^1/_2$ the vdW separation listed in Table I, using fourth-order Moeller–Plesset corrections for electron correlation as described in the text. Quantities in parentheses give the ratio, $v(0.5)$, of the interaction energy and the experimentally determined well depth (Table I). For reference, the Hartree–Fock interaction energies for the 6-311G(3df,3pd) basis set are 13.29, 27.96, 73.23, 43.03, 101.56, and 176.74 kcal/mol, respectively. [b] Extrapolated MP4 energy, eq 11. [c] Computed from the experimental vdW potentials cited in Table I.

constructed to fit the highly repulsive region as well as other regions of the potential energy surface.[29] Except for Ar···Ar, each appears to be the one its authors believe to be the most reliable. For Ar···Ar, we have chosen the authors' HFD-B4, rather than HFD-B3, potential because the former was constructed to reproduce very high quality ab initio calculations of McLean et al.[30] which we believe to be definitive.

2.5. Conformability of the Rare-Gas Potentials. To characterize the degree to which the cited pairwise potentials for He, Ne, and Ar are similar in shape, i.e., are conformable, we show in Figures 1 and 2 the *reduced potentials* defined in terms of $v(\rho) = E_{vdw}(R)/\epsilon$ and $\rho = R/R^*$. The experimental values for ϵ and R^* used to generate the rare-gas reduced potentials are listed in Table I. The figures show that all reduced potentials except that for Ne···Ne fall within a relatively narrow range.

To assess whether the repulsive walls are positioned accurately, particularly for the relatively hard Ne···Ne potential, we calculated the interaction energy for each pairwise interaction at $\rho = 0.5$ (i.e., at 50% of the minimum-energy distance R^*) using GAUSSIAN 88.[31] The calculations employed fourth-order Moeller–Plesset corrections for electron correlation based on single, double, triple, and quadruple substitutions from all electrons[32,33] and used 6-311G(d,f) (also known as 6-311G**)[34] and 6-311G(3df,3pd)[35] basis sets. They also incorporated a standard correction for basis-set superposition errors,[36] which gives ΔE_{A-B} as

$$\Delta E_{A-B} = E_{A-B} - E_A - E_B$$

where the calculations on the full system A–B and on the atomic subsystems A and B all use the same A–B dimer basis set.

The results, summarized in Table II, show that both basis sets yield an interaction energy at $\rho = 0.5$ which is systematically too high by comparison to experiment. Remarkably, the simple extrapolation

$$E_{extrap} = 1.02\{2E_{MP4/6\text{-}311G(3df,3pd)} - E_{MP4/6\text{-}311G(d,f)}\} \quad (11)$$

accounts extremely well for the experimental results for four of the six systems. An implication is that the experimental potentials for He···Ne and He···Ar may be slightly too soft at $v(0.5)$.

In agreement with the 1983 HFD-C2 potential,[16] the hardest reduced potential at $\rho = 0.5$ is indeed found to be that for Ne···Ne. To further corroborate the extrapolated energy for this system, we also computed the QCISD[37] and MP4 energies at $\rho = 0.5$ as above but using the Dunning [5s,4p] triple-ζ (TZ) basis set for Ne,[38] augmented by a larger 5d3f set of polarization functions. The additional exponents for the d and f polarization functions

Table III. Accuracy of the Buf-14-7, Exp-6, and Lennard-Jones Potentials[a]

system/region[b]	Buf-14-7[c]	Exp-6[d]	LJ-9-6[e]	LJ-10-6[f]	LJ-12-6[g]
		He···He			
rms (10–20)	0.1969	7.5445	10.7231	28.7787	121.2785
rms (5–10)	0.0647	3.6006	1.8100	7.0635	30.2040
rms (2–5)	0.0416	1.5758	0.2432	1.4906	7.4874
rms (−ϵ–2)	0.0267	0.2531	0.1365	0.0610	0.6059
rms (R^*–3R^*)	0.0002	0.0005	0.0010	0.0007	0.0004
		He···Ne			
rms (10–20)	1.0030	7.4654	2.1893	11.5166	52.4037
rms (5–10)	0.2238	3.4675	0.5572	2.8056	15.1139
rms (2–5)	0.0494	1.5067	0.6655	0.4585	4.1044
rms (−ϵ–2)	0.0255	0.2619	0.1885	0.0551	0.3973
rms (R^*–3R^*)	0.0001	0.0012	0.0021	0.0016	0.0009
		Ne···Ne			
rms (10–20)	3.5954	8.8160	4.7499	0.6373	16.6195
rms (5–10)	1.4015	4.1375	2.7921	1.0249	4.9655
rms (2–5)	0.4808	1.7724	1.4025	0.7412	1.2941
rms (−ϵ–2)	0.0669	0.3090	0.2840	0.1818	0.1060
rms (R^*–3R^*)	0.0002	0.0025	0.0042	0.0033	0.0019
		He···Ar			
rms (10–20)	1.8789	7.9057	1.4664	5.8490	33.5586
rms (5–10)	0.7692	3.7501	1.7492	0.7454	9.0661
rms (2–5)	0.2739	1.6452	1.1072	0.3333	2.2866
rms (−ϵ–2)	0.0420	0.2905	0.2504	0.1402	0.1870
rms (R^*–3R^*)	0.0001	0.0017	0.0029	0.0023	0.0013
		Ne···Ar			
rms (10–20)	0.4722	7.0121	2.7187	2.5524	20.5789
rms (5–10)	0.0447	3.2693	1.9699	0.2606	6.1374
rms (2–5)	0.0528	1.4490	1.1201	0.4160	1.7035
rms (−ϵ–2)	0.0089	0.2743	0.2589	0.1460	0.1645
rms (R^*–3R^*)	0.0009	0.0049	0.0077	0.0062	0.0039
		Ar···Ar			
rms (10–20)	1.6997	5.6654	2.6358	2.0027	16.1761
rms (5–10)	0.6684	2.8163	2.0160	0.4112	4.6380
rms (2–5)	0.2082	1.3379	1.1799	0.5371	1.2667
rms (−ϵ–2)	0.0184	0.2750	0.2854	0.1768	0.1069
rms (R^*–3R^*)	0.0013	0.0098	0.0158	0.0126	0.0077

[a] Based on comparisons at 0.01-Å intervals to the experimentally determined rare-gas potentials cited in Table I. [b] Rms deviations in kcal/mol from the rare-gas potential for experimental-potential interaction energies in the ranges 10–20 kcal/mol, 5–10 kcal/mol, 2–5 kcal/mol, from the energy minimum inward to 2 kcal/mol, and from the energy minimum outward to 3R^*, where R^* is the minimum-energy distance listed in Table I. [c] From eq 9 with $n = 14$, $m = 7$, $\delta = 0.07$, and $\gamma = 0.12$, as shown in eq 10. [d] From eq 1. [e] From eq 3. [f] Defined analogously to the LJ-9-6 and LJ-12-6 potentials of eqs 2 and 3. [g] From eq 2.

(29) For like-pair interactions of Ne, Ar, Kr, and Xe, high-quality multiproperty potentials have also recently been determined from fits which make use of scaled ab initio data (cf. ref 18). For consistency, however, we shall use the HFD-based potentials in this work.

(30) McLean, A. D.; Liu, B.; Barker, J. A. *J. Chem. Phys.* **1988**, *89*, 6339–6347.

(31) Frisch, M. J.; Head-Gordon, M.; Schlegel, H. B.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R.; Kahn, L. R.; Stewart, J. J. P.; Fluder, E. M.; Topiol, S.; Pople, J. A. GAUSSIAN 88, Gaussian, Inc.: Pittsburgh, PA, 1988.

(32) Krishnan, R.; Frisch, M. J.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 4244–4245.

(33) Krishnan, R.; Pople, J. A. *Int. J. Quantum Chem.* **1978**, *14*, 91.

(34) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.

(35) Frisch, M. J.; Pople, J. A. *J. Chem. Phys.* **1984**, *80*, 3265–3269. The f-orbital exponent needed for Ar is not cited in this paper; we have used a value of 1.0, estimated by extrapolation of the values reported for the series Si, P, S, Cl.

(36) Chalasinski, G.; Gutowski, M. *Chem. Rev.* **1988**, *88*, 943–962.

(37) The QCISD (quadratic configuration-interaction with single and double substitutions) method as implemented in GAUSSIAN 88 (ref 31) includes the fourth-order triples contribution to the energy; these contributions are computed after the QCI energy has converged.

(38) Dunning, T. H., Jr. *J. Chem. Phys.* **1971**, *55*, 716–723.

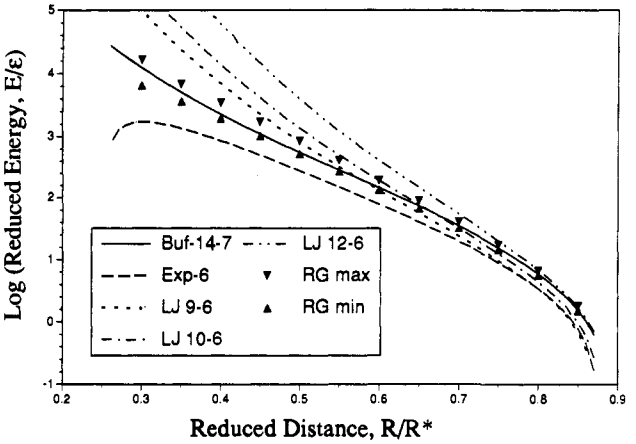

Figure 3. Comparison of reduced potentials at short interatomic separation for vdW potentials used in or intended for use in molecular mechanics force fields. The centers of the filled triangles mark the maximum and minimum values found for the rare-gas potentials.

were generated by multiplying the f exponent and the largest and smallest d exponents of the precursor 3df set by factors of 3 and of $^1/_3$. Reassuringly, the QCISD/TZ(5d3f) value, ΔE = 71.09 kcal/mol, is in full agreement with the value ΔE = 71.04 kcal/mol given by the experimental HFD-C2 potential. The MP4/TZ-(5d3f) value, ΔE = 71.19 kcal/mol, also agrees well.

These ab initio calculations support the conclusion that the reduced potentials displayed in Figure 1 accurately describe the pairwise vdW interactions of He, Ne, and Ar in the vicinity of ρ = 0.5. While conformable at low to moderate energies, the experimental potentials show measurable differences at energies above several hundred times the well depth. Fortunately, however, the behavior of the vdW potential at large repulsion energies is not of great importance in molecular mechanics and dynamics calculations. Hence, the He, Ne, and Ar potentials appear to be sufficiently conformable to justify the expectation that the use of vdW potentials based on a common reduced potential need not lead to unacceptable errors.

2.6. Accuracy of the Buf-14-7, Exp-6, and Lennard-Jones Potentials. Table III summarizes the abilities of the Buf-14-7, Exp-6, and Lennard-Jones 9-6, 10-6, and 12-6 potentials to reproduce the rare-gas potentials in various interaction energy and separation distance regimes. To isolate errors arising from the intrinsic shapes of these potentials from errors arising from the combination rules usually employed in conjunction with them, the experimentally determined ϵ and R^* values were used for both like and unlike pairs in generating the interaction energies on which the table is based. The comparisons thus show how well these simple potentials reproduce the rare-gas potentials when fit to them by fixing the position and depth of the vdW minimum. Figures 3 and 4 similarly compare the corresponding reduced potentials against those for the rare-gas interactions. Here, too, the experimental ϵ_{ij} and R^*_{ij} values have been used to construct the rare-gas reduced potentials. In these figures, the centers of the filled triangles demark the range covered by the experimental reduced potentials in Figures 1 and 2. Clearly the Lennard-Jones potentials are too repulsive at short contact distances, the error being largest for the LJ-12-6 form (eq 2). Conversely, the Exp-6 potential used in MM3[6] (eq 1) appears to be too soft. In contrast, the reduced Buf-14-7 potential reproduces the interaction energies well even up to 20 kcal/mol. In this region, the LJ-9-6 potential affords the second-best choice.

The Buf-14-7 form also performs best in the vicinity of the vdW minimum. As Table III shows, for the $(-\epsilon - 2)$ region (between the vdW minimum of $-\epsilon$ and +2 kcal/mol) the rms deviations for the Buf-14-7 potential are typically 10 times smaller than those for the Exp-6 and Lennard-Jones potentials. Particularly noteworthy is the excellent performance of the Buf-14-7 form at longer distances (between $R = R^*$ and $3R^*$), as is quantified in Table III and is depicted by the behavior near $R = R^*$ shown in Figure 4. Table IV documents this behavior for the particular case of

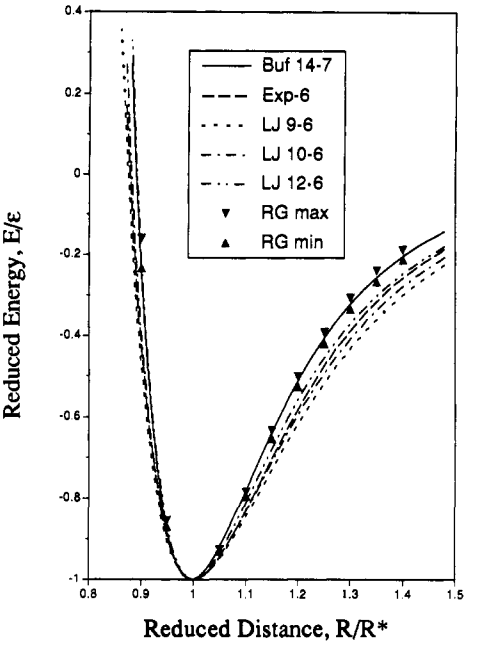

Figure 4. Comparison of reduced potentials near the vdW minimum for vdW potentials used in or intended for use in molecular mechanics force fields. The centers of the filled triangles mark the maximum and minimum values found for the rare-gas potentials.

Table IV. Comparison of Buf-14-7 and Experimentally Determined vdW Potentials for He···Ne (kcal/mol)[a]

		vdW potential	
ρ[b]	R[c]	exptl	Buf-14-7
0.5	1.518	25.4279	23.0598
0.6	1.821	6.2347	6.1237
0.7	2.125	1.4100	1.4682
0.8	2.428	0.2389	0.2468
0.9	2.732	−0.0091	−0.0106
1.0	3.036	−0.0415	−0.0415
1.1	3.339	−0.0329	−0.0324
1.2	3.643	−0.0215	−0.0212
1.4	4.250	−0.0085	−0.0085
1.6	4.857	−0.0037	−0.0036
1.8	5.464	−0.0018	−0.0017
2.2	6.678	−0.0005	−0.0004
2.6	7.892	−0.0002	−0.0001
3.0	9.108	−0.0001	−0.0001

[a] See Table III for citations for the experimental and calculated potentials. [b] Ratio of the He···Ne separation R to the minimum-energy separation R^* = 3.036 Å. [c] In Å.

the He···Ne interaction, which can be taken as a surrogate for H···first-row interactions in organic systems. Evidently, the simple Buf-14-7 form captures the essence of the power series behavior of the damped dispersion term of eq 7. For such longer distances, the LJ-12-6 form provides the second-best, though clearly inferior, choice, while the LJ-9-6 potential is too attractive and represents the poorest choice. Of course, each of the other potentials will eventually fall off properly as R^{-6} but not until well after the true vdW potential has fallen to negligible values.

In summary, only the Buf-14-7 potential performs accurately in all regions of prospective interest in molecular mechanics calculations. Given that this form is not greatly more complicated than are the Lennard-Jones forms, will usually be more efficient than the Exp-6 form to compute, and avoids the divergence of the latter to minus infinity at short contact distances (cf. Figure 1), this is the form we shall use in the Merck Molecular Force Field (MMFF).

3. Combination Rules for vdW Interactions between Unlike Atoms

In a molecular mechanics framework, good results for vdW interactions between unlike atoms will be attainable only if suitable

Table V. Test of Molecular Mechanics Combination Rules for Experimentally Determined Rare-Gas Minimum-Energy Distances R^*

	minimum-energy distance R^* (Å)			
system	exptl	geometric mean[a]	arithmetic mean[b]	cubic mean[c]
He···Ne[d]	3.036	3.024	3.025	3.028
He···Ar[d]	3.478	3.338	3.362	3.455
He···Kr[e]	3.691	3.447	3.487	3.641
He···Xe[e]	3.968	3.595	3.663	3.921
Ne···Ar[d]	3.489	3.407	3.424	3.490
Ne···Kr[f]	3.621	3.519	3.549	3.667
Ne···Xe[f]	3.861	3.670	3.725	3.937
Ar···Kr[g]	3.881	3.884	3.886	3.894
Ar···Xe[h]	4.067	4.051	4.062	4.106
Kr···Xe[h]	4.174	4.183	4.187	4.202
	rms dev	0.165	0.133	0.040

[a] $(R^*_{ii}R^*_{jj})^{1/2}$. [b] $(R^*_{ii} + R^*_{jj})/2$. [c] Equation 12. [d] Cf. Table III. [e] Reference 27. [f] Reference 39. [g] Reference 40. [h] Reference 41.

Table VI. Test of Molecular Mechanics Combination Rules for Experimentally Determined Rare-Gas Well Depths ϵ

	well depth ϵ (kcal/mol)			
system[a]	exptl	geometric mean[b]	harmonic mean[c]	HHG mean[d]
He···Ne	0.0415	0.0427	0.0346	0.0382
He···Ar	0.0589	0.0787	0.0404	0.0534
He···Kr	0.0586	0.0931	0.0413	0.0572
He···Xe	0.0556	0.1105	0.0419	0.0607
Ne···Ar	0.1343	0.1546	0.1297	0.1410
Ne···Kr	0.1422	0.1830	0.1387	0.1578
Ne···Xe	0.1475	0.2171	0.1461	0.1746
Ar···Kr	0.3324	0.3369	0.3322	0.3345
Ar···Xe	0.3748	0.3998	0.3778	0.3884
Kr···Xe	0.4640	0.4732	0.4663	0.4697
	rms dev	0.0350	0.0096	0.0115
	rms % dev	42.05	16.62	8.59

[a] See Table V for references. [b] $(\epsilon_{ii}\epsilon_{jj})^{1/2}$. [c] Equation 13. [d] Equation 14.

combination rules are used to generate the requisite well depth and vdW separation parameters from those which describe the like-pair interactions. This limitation arises because it is not currently feasible to generate independent vdW parameters for all atom-pair interactions of interest.

3.1. Combination Rules Used in Molecular Mechanics Calculations. As indicated in Section 2.1, current molecular mechanics force fields employ either the geometric or the arithmetic mean to express the well depth ϵ_{ij} and the minimum-energy separation R^*_{ij} for the interaction between atoms of types i and j in terms of the parent vdW parameters ϵ_{ii}, ϵ_{jj}, R^*_{ii}, and R^*_{jj}. The accuracy of these combination rules for vdW interactions among the rare-gas atoms He, Ne, Ar, Kr, and Xe is examined in Tables V and VI. In addition to the pairwise potentials for interactions of He, Ne, and Ar previously referenced, these comparisons make use of the HFD-B potentials for He···Kr and He···Xe[27] and for Ne···Kr and Ne···Xe,[39] the HFD-C2 potential for Ar···Kr,[40] the HFD-C potentials for Ar···Xe and Kr···Xe,[41] and the well depths ϵ of 0.3988 and 0.5614 kcal/mol and minimum-energy separations R^* of 4.011 and 4.3627 Å for Kr···Kr and Xe···Xe cited by Barrow and Aziz.[28]

Table V shows that the geometric-mean rule (used by VFF,[9] OPLS,[10] and GROMOS[11] for the A and B coefficients in eq 4 and consequently for both R^* and ϵ) consistently produces too small a minimum-energy distance R^*_{ij} when the like-pair distances R^*_{ii} and R^*_{jj} differ substantially. The error, which can be large, reaches 0.40 Å for the He···Xe interaction. The arithmetic-mean rule, used for R^* by MM2,[5] MM3,[6] AMBER,[7,42] and CHARMM,[8] is subject to the same deficiency to a slightly lesser degree. Evidently, the actual minimum-energy distance falls closer to that for the larger partner than is given by these rules. The final column of Table V shows that this pattern can be reproduced by the "cubic-mean" rule:

$$R^*_{ij} = \frac{R^*_{ii}{}^3 + R^*_{jj}{}^3}{R^*_{ii}{}^2 + R^*_{jj}{}^2} \tag{12}$$

This rule can be viewed as a weighted-average variant of the arithmetic-mean rule, the weight factor for each R^* being the square of its own value. As Table V shows, it reduces the rms error in the minimum-energy distance from ca. 0.15 to 0.04 Å, or by roughly a factor of 4, and thus represents a distinct improvement.

As Table VI shows, the geometric-mean rule, almost always used in molecular mechanics force fields for well depths, leads to even larger departures from the experimental values. This rule consistently overestimates the well depth for unlike-pair interactions, particularly when the two like-pair well depths are markedly dissimilar. Thus, it overestimates the well depth for He···Xe by 100% and can probably be expected to perform equally poorly for comparably dissimilar interactions (such as those between H and Br or I atoms) in organic compounds. In contrast, the Fender–Halsey[43] harmonic-mean rule

$$\epsilon_{ij} = \frac{2\epsilon_{ii}\epsilon_{jj}}{\epsilon_{ii} + \epsilon_{jj}} \tag{13}$$

underestimates the well depths for He interactions but otherwise does well. Also characterized in Table VI is the performance of a new, only slightly more complex rule labeled HHG. In this rule, ϵ_{ij} is given by the harmonic mean of its own harmonic- and geometric-mean values, leading to

$$\epsilon_{ij} = \frac{4\epsilon_{ii}\epsilon_{jj}}{(\epsilon_{ii}{}^{1/2} + \epsilon_{jj}{}^{1/2})^2} \tag{14}$$

This rule gives the smallest rms percent deviation (8.59%) of the three rules examined. The 5-fold improvement over the widely used geometric mean (rms deviation, 42.05%) would seem sufficient to justify its use in molecular mechanics force fields. In Section 4, however, we will present an alternative formulation which achieves still higher accuracy and which, in addition, provides a consistent framework for the assignment of both vdW well depths and minimum-energy separations.

3.2. More Elaborate Combination Rules Used for Rare-Gas and Other Atomic Interactions. In addition to empirical approaches, such as that of Hiza and Duncan,[44] a variety of theoretically-based combination rules have been derived from assumptions about the nature of the vdW interaction and about the form of the vdW potential. In particular, Sikora,[45] Smith,[46] and Kong,[47] among others, have employed atomic-distortion models (related to such quantities as ionization potentials and atomic polarizabilities) for the repulsive part of the potential in conjunction with an assumed overall form (e.g., Lennard-Jones 12-6 or Exp-6) to obtain combination rules in which ϵ_{ij} and σ_{ij} both typically depend on all four like-interaction parameters, ϵ_{ii}, ϵ_{jj}, σ_{ii}, and σ_{jj} (σ measures the interatomic distance at which the vdW potential crosses 0). Diaz Pena et al.[48] have presented and tested generalized expressions for combination rules of this class. These generalized rules identify the attractive dispersion parameter in simple potential forms (such

(39) Barrow, D. A.; Slaman, M. J.; Aziz, R. A. *J. Chem. Phys.* **1989**, *91*, 6348–6358.
(40) Aziz, R. A.; van Dalen, A. *J. Chem. Phys.* **1983**, *78*, 2413–2418.
(41) Aziz, R. A.; van Dalen, A. *J. Chem. Phys.* **1983**, *78*, 2402–2412.
(42) Although the published descriptions (cf. ref 7) suggest to us that AMBER uses the geometric-mean rule on the A and B coefficients in eq 4, Professor Kollman reports that the arithmetic mean is in fact used for R^* (Kollman, P. Personal communication).
(43) Fender, B. E. F.; Halsey, G. D., Jr. *J. Chem. Phys.* **1962**, *36*, 1881.
(44) Hiza, M. J.; Duncan, A. G. *Phys. Fluids* **1969**, *12*, 1531; *AIChE J.* **1970**, *16*, 733; see also Hiza, M. J.; Robinson, R. L., Jr. *J. Chem. Phys.* **1978**, *68*, 4768–4769.
(45) Sikora, P. T. *J. Phys. B* **1970**, *3*, 1475.
(46) Smith, F. T. *Phys. Rev. A* **1972**, *5*, 1708.
(47) Kong, C. L. *J. Chem. Phys.* **1973**, *59*, 968, 1953, 2464.
(48) Diaz Pena, M.; Pando, C.; Renuncio, J. A. R. *J. Chem. Phys.* **1982**, *76*, 325–332, 333–339.

as B_{ij} in eq 4 for the LJ-12-6 potential) with the C_6 dispersion coefficient of eq 7 and employ combination rules which express the C_6 coefficient for the unlike-pair interaction in terms of the C_6 coefficients for the like-pair interactions. The similarly motivated approach recently taken by Tang and Toennies[49] and extended by Bzowski et al.[50] also makes use of a combination rule of this type for C_6.

Diaz Pena et al.[51] have tested a wide variety of combination rules for C_6, C_8, and C_{10}. For C_6, the comparisons show that the geometric mean $(C_{6ii}C_{6jj})^{1/2}$ describes the interactions among rare-gas atoms fairly well but overestimates interactions of the rare gases with alkali atoms by up to 100%. The well-known London approximation[52]

$$C_{6ij} = \frac{3}{2}\frac{\alpha_i\alpha_j I_i I_j}{I_i + I_j} \tag{15}$$

where α represents the static atomic polarizability and I is the ionization potential, showed similar though slightly less pronounced deficiencies. In contrast, the equally well-known Slater–Kirkwood approximation[53]

$$C_{6ij} = \frac{3}{2}\frac{\alpha_i\alpha_j}{(\alpha_i/N_i)^{1/2} + (\alpha_j/N_j)^{1/2}} \tag{16}$$

$$= \frac{2\alpha_i\alpha_j C_{6ii}C_{6jj}}{\alpha_i^2 C_{6jj} + \alpha_j^2 C_{6ii}} \tag{17}$$

gave results for C_{6ij} which were not only far superior but also essentially equivalent to those for the best of the combination rules examined. In eq 16, N represents the effective number of electrons. Equation 17, which has been shown to accurately predict C_6 coefficients by Kramer and Herschbach,[54] is obtained from eq 16 when N_i and N_j are chosen by requiring that eq 16 hold when i = j, giving[55]

$$N_i = 16C_{6ii}^2/9\alpha_i^3 \tag{18}$$

and likewise for N_j. Tang and Toennies[49] and Bzowski et al.[50] also use eq 17 in their combination rules for ϵ_{ij}.

4. Proposed Algorithm and Combination Rules for vdW Parameters

We shall use the cubic-mean rule of eq 12 for R^*. We have already shown (Table V) that this rule performs well; further comparisons will be offered in this section.

For well depths, we shall use the Slater–Kirkwood approximation to formulate an expression in which ϵ_{ij} for both like- and unlike-pair interactions is based on the determination of C_{6ij} coefficients from eqs 16–18. Application of these equations in a molecular framework raises a possible physical difficulty, however, because neither C_6 coefficients nor atomic polarizabilities are rigorously defined for atoms in molecules.[56] Consequently, it is not possible to calculate N_i from known values for C_{6ii} and α_i via eq 18. Still, there is a long history of reasonably successful efforts to account for experimental data in terms of atomic polarizabilities.[57] We shall employ such an approach to obtain the necessary α_i parameters. We shall then present an algorithm for relating the requisite N_i parameters to the atomic number of the atom in question. Given α_i and N_i, we will then compute the C_{6ij} coefficients in a manner consistent with eq 18. Finally, we shall relate the C_{6ij} to the well depths ϵ_{ij} in a way which takes into account the substantial contributions to the dispersion energy at $R = R^*$ made by the higher dispersion coefficients (C_8, C_{10}, ...). This relationship will also make use of an empirical relationship between α and R^*.

Specification of the present approach for determining vdW well depths and separation parameters for like and unlike interactions will thus involve the following steps: (1) define an algorithm for assigning N_i, the Slater–Kirkwood effective number of electrons on an atom of type i, on the basis of that atom's atomic number; (2) define a set of atomic polarizabilities α_i, and show that they and the N_i successfully account for atomic and molecular C_6 dispersion coefficients; (3) define an empirical relationship between atomic polarizability and the vdW minimum-energy separation R^*_{ij}; (4) calibrate the relationship between ϵ, R^*, and C_6 to implicitly include contributions from C_8, C_{10}, C_{12}, ... to the dispersion energy; (5) formulate the complete algorithm for computing ϵ_{ij} from atomic polarizabilities α_i and derived quantities (e.g., the vdW minimum-energy separations R^*_{ii}); and (6) characterize the accuracy of the resultant algorithm for like- and unlike-pair vdW interactions of rare-gas atoms.

In Section 5 we shall then (7) extract atomic polarizabilities α_i for atom types considered in the MMFF force field and (8) obtain the resultant MMFF vdW parameters, and compare them with vdW parameters used in other force fields.

Table VII. Slater–Kirkwood Parameters for Some Simple Atoms and Molecules

	experimental[a]			calculated[b]		
system	α	C_6	N_{eff}	N_{eff}	α	C_6
He	1.385	1.458	1.422	1.42		
Ne	2.669	6.383	3.810	3.81		
Ar	11.08	64.3	5.404	5.40		
Kr	16.79	129.6	6.309	6.30		
Xe	27.16	285.9	7.253	7.25		
H_2	5.428	12.11	1.630	1.60	5.20	11.2
N_2	11.74	73.39	5.918	5.64	12.76	91.2
O_2	10.59	62.10	5.756	6.30	10.85	67.2
NO	11.52	69.78	5.662	5.97	12.32	78.6
N_2O	19.70	184.9	7.950	8.79	19.23	186.0
CO	13.08	81.4	5.264	5.64	11.80	71.4
CO_2	17.51	158.7	8.340	8.78	17.23	157.5
CH_4	17.27	129.6	5.797	5.64	16.78	122.3
NH_3	14.56	89.08	4.570	5.22	14.70	77.8
H_2O	9.642	45.37	4.082	4.75	10.63	47.1
HF	5.601	19.0	3.652	4.28	5.60	18.7
HCl	17.39	130.4	5.748	5.90	17.34	131.6
HBr	23.74	216.6	6.234	6.80	23.7	225.4
			rms % dev	9.6	5.3	9.8

[a] From ref 58: α is the static atomic or molecular polarizability in a_0^3, C_6 is the 6th-power dispersion coefficient in hartrees-a_0^6; N_{eff} is the Slater–Kirkwood effective number of electrons, from eq 18. [b] Computed from the atomic data in Table VIII assuming additivity and, for C_6, using eq 17 for unlike-pair coefficients.

4.1. Slater–Kirkwood Effective Number of Electrons. Table VII presents Slater–Kirkwood parameters for some simple atoms and molecules taken from a recent review by Buckingham et al.[58] The experimental data indicate that α, C_6, and N_{eff} (the S–K effective number of electrons, computed as shown in eq 18) all increase in a fairly regular way with increasing atomic number of the atom or atoms involved. For N_{eff}, however, the value remains far below the number of valence electrons in the atomic or molecular system, in contrast, for example, to values used by CHARMM.[8] Of the quantities listed, it is known that molecular polarizabilities can be represented with reasonable accuracy as

(49) Tang, K. T.; Toennies, J. P. *Z. Phys. D* **1986**, *1*, 91.
(50) Bzowski, J.; Mason, E. A.; Kestin, J. *Int. J. Thermophys.* **1988**, *9*, 131–143.
(51) Diaz Pena, M.; Pando, C.; Renuncio, J. A. R. *J. Chem. Phys.* **1980**, *72*, 5269–5275.
(52) London, F. *Z. Phys. Chem. (Leipzig) B* **1930**, *11*, 222.
(53) Slater, J. C.; Kirkwood, J. G. *Phys. Rev.* **1931**, *37*, 682.
(54) Kramer, H. L.; Herschbach, D. R. *J. Chem. Phys.* **1970**, *53*, 2792–2800.
(55) Wilson, J. N. *J. Chem. Phys.* **1965**, *43*, 2564.
(56) Formally speaking, Bader's approach rigorously defines atomic polarizabilities for atoms in molecules (cf. Laidig, K. E.; Bader, R. W. F. *J. Chem. Phys.* **1990**, *93*, 7213–7224). However, practical objections to the Bader partitioning (cf. Dinur, U. *J. Comput. Chem.* **1991**, *12*, 469–486. Perrin, C. L. *J. Am. Chem. Soc.* **1991**, *113*, 2865–2868) suggest that atomic polarizabilities computed in this way will provide only qualitative guidance.
(57) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533–8542.
(58) Buckingham, A. D.; Fowler, P. W.; Hutson, J. M. *Chem. Rev.* **1988**, *88*, 963–988.

Table VIII. Atomic Data Used in Table VII

atom	N_i[a]	α_i[b]	C_{6ii}[c]
H	0.80	2.60	2.8
C	2.49	6.38	19.1
N	2.82	6.90	22.8
O	3.15	5.42	16.8
F	3.48	3.00	7.3
Cl	5.10	14.8	96.0
Br	6.00	21.1	178.0

[a] Slater–Kirkwood effective numbers of electrons, from eqs 20–24. [b] Atomic polarizability in a_0^3; obtained by fitting to the experimental molecular polarizabilities in Table VII by least-squares, assuming additivity. [c] Atomic C_6 coefficient in hartress-a_0^6, obtained from N_i and α_i as shown in eq 19.

a sum of atomic contributions,[57] and Meath et al.[59] have shown that C_6 coefficients in molecular systems can also be so represented. An additive decomposition of N_{eff} appears feasible as well. Consider, for example, the molecule H_2. Define N_i and α_i for a hydrogen atom in H_2 as being half of the experimental value found for molecular H_2. From eq 18

$$C_{6ii} = {}^3/_4(N_i\alpha_i^3)^{1/2} \tag{19}$$

Therefore the C_6 coefficient for a hydrogen atom in H_2 is one-fourth the C_6 coefficient for molecular H_2. But since the H_2–H_2 interaction involves four such atomic H–H interactions, it follows that the molecular C_6 coefficient can be reconstructed additively as a sum of atomic C_6 coefficients when N_{eff} and α are taken as an additive quantities.

Assuming additivity for N_{eff}, we propose the following algorithm for obtaining N_i, the effective number of electrons on an atom of type i:

$$\text{for H and He} \quad N_i = 0.80 \text{ and } 1.42 \tag{20}$$

$$\text{for C–Ne} \quad N_i = 1.17 + 0.33N_v \tag{21}$$

$$\text{for Si–Ar} \quad N_i = 3.00 + 0.30N_v \tag{22}$$

$$\text{for Ge–Kr} \quad N_i = 3.90 + 0.30N_v \tag{23}$$

$$\text{for Sn–Xe} \quad N_i = 4.85 + 0.30N_v \tag{24}$$

N_v in eqs 21–24 is the number of valence-shell s and p electrons and in each instance ranges from 4 to 8. These equations reproduce the well-established rare-gas value when $N_v = 8$. For the molecular systems listed in Table VII, the calculated values for N_{eff} are given by the sum of the constituent atomic values (listed for convenience in Table VIII). For example, N_{eff} for N_2O is given by 2(2.82) + 3.15 = 8.79.

The calculated values for N_{eff} in Table VII show that eqs 20–24 are moderately successful (rms deviation, 9.6%), though they usually overestimate the experimentally derived values. Better agreement would result from using the following atomic N_i values, obtained from a least-squares fit to the molecular values: H = 0.713; C = 2.747; N = 2.752, O = 2.769, F = 2.939; Cl = 5.025; Br = 5.521. Nevertheless, we will use the values given by eqs 20–24. We need to define a widely applicable scheme despite the fact that data are lacking for many atomic species of interest. Additionally, we feel that atomic N_i values *should* obey systematic relationships akin to those expressed in the above equations.

4.2. Atomic and Molecular C_6 Dispersion Coefficients. We now show that the molecular polarizabilities α and the molecular 6th-power dispersion coefficients C_6 obey additivity reasonably closely. First, we list in Table VII the calculated molecular polarizabilities obtained from the least-squares fitted atomic polarizabilities given in Table VII. Clearly, additivity holds reasonably well (rms deviation, 5.3%), even within the present, restricted scheme in which, for example, the singly-bonded oxygen in H_2O, the oxygen in triplet O_2, and the differently-hybridized, multiply-bonded oxygens in the other systems are all given the same value for the atomic polarizability.

The atomic polarizabilities α_i then allow the atomic C_6 coefficients, C_{6ii}, to be calculated from eq 19, giving the results listed in Table VIII. These atomic C_6 coefficients, in turn, allow the molecular C_6 coefficients to be calculated in the additive framework previously described for H_2. Thus, the molecular C_6 coefficient for H_2 is obtained as four times the atomic C_6 coefficient. Similarly, that for pairwise interaction of two N_2O molecules is obtained as four times the atomic C_6 coefficient for N plus the atomic C_6 coefficient for O plus four times the atomic C_{6ij} coefficient for i = N, j = O. The latter is obtained from eq 16 or from eq 17 using the N_i, α_i, and C_{6ii} values listed in Table VIII. As Table VII shows, the values calculated in this way for the molecular C_6 coefficients also reproduce the experimental values with reasonable accuracy, the overall rms deviation being 9.8%. Clearly, an additive decomposition of dispersion coefficients based on eqs 16–18 is also tenable.

4.3. Atomic Polarizability and vdW Minimum-Energy Separations. We now face the question of how to represent the vdW minimum-energy separations. Ultimately, the related vdW radii might be parameterized against experimental data, such as enthalpies of sublimation for organic crystals, or against the results of ab initio calculations. Suitably high quality calculations of molecular interactions, however, can be prohibitively expensive, and experimental data on enthalpies of sublimation are scarce or lacking entirely for many types of compounds of interest in organic and bioorganic systems. Accordingly, we seek a framework which will yield a useful first approximation for the required vdW separation parameters.

The relationship between polarizability and atomic size has long been recognized. Thus, Pauling speaks of the "polarizability radius" in referring to the cube root of polarizability,[60] and Nagle[61] has recently demonstrated a correlation between the radius defined in this way and the radius in atomic systems at which the maximum density is found for the outermost ground-state orbital. Kirkwood,[62] by way of comparison, relates $\langle r_A^2\rangle$, the expectation value of the square of the distance of an electron from the nucleus, to the atomic polarizability α_A through

$$\alpha_A = 4(\langle r_A^2\rangle/3)^2/a_0 \tag{25}$$

This relationship was cited by Miller and Savchick[63] as suggesting that the vdW radius be taken as proportional to the square root of $\langle r_A^2\rangle$ and hence to the one-fourth power of the atomic polarizability.

To be sure, the relationship between an atom's vdW radius and a measure of its size based on maximum probability or on the expectation value of an electronic radius is not immediately clear either on practical or on theoretical grounds, as the vdW radius would seem to be related to properties of the electron distribution which obtain at considerably longer distances than are characterized by the Nagle or Slater–Kirkwood measures. Nevertheless, a relationship between atomic polarizability and vdW size is conceptually attractive because abundant data are available for the former. Thus, Miller and Savchick[63] show that the radii obtained from $\langle r_A^2\rangle^{0.5}$ when suitably scaled correlate reasonably well with Bondi's crystal-contact radii.[64] Their observation suggests that vdW radii r^* might be correlated with atomic polarizabilities α through a relationship of the form

$$R^* = 2r^* = A\alpha^p + B \tag{26}$$

where A and B are constants of appropriate dimension and the fractional exponent p is of the order of one-third or one-fourth. For the particular case of carbon and neon, for example, reasonable vdW radii are obtained if Nagle's polarizability radius $\alpha_v^{1/3}$ is

(59) Meath, W. J.; Margoliash, D. J.; Jhanwar, B. L.; Koide, A.; Zeiss, G. D. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 101–115.

(60) Pauling, L. *General Chemistry*; W. H. Freeman and Company: San Francisco, 1970; p 395.

(61) Nagle, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 4741–4747.

(62) Kirkwood, J. G. *Phys. Z.* **1932**, *33*, 57. See also: Hirschfelder, J. O.; Curtiss, C. F.; Bird, R. B. *Molecular Theory of Gases and Liquids*; Wiley: New York, 1954; p 946.

(63) Miller, K. J.; Savchick, J. A. *J. Am. Chem. Soc.* **1979**, *101*, 7206–7213.

(64) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

increased by an additive constant of ca. 0.8 Å. In this work, however, we shall use Miller and Savchick's form and thus will take $B = 0$.

For the case $p = {}^1/_4$, this approach leads to

$$R^*_{ii} = A_i\alpha_i^{0.25} \tag{27}$$

for the vdW minimum-energy separations R^*_{ii}, where

$$\text{for H and He} \quad A_i = 4.40 \tag{28}$$

$$\text{for C–Ne} \quad A_i = 3.89 \tag{29}$$

$$\text{for Si–Ar} \quad A_i = 3.32 \tag{30}$$

$$\text{for Ge–Kr} \quad A_i = 3.19 \tag{31}$$

$$\text{for Sn–Xe} \quad A_i = 3.08 \tag{32}$$

In these expressions, the atomic polarizabilities are assumed to be given in units of Å^3. The polarizabilities for the rare gases used in their calibration were taken from Table VII.[65] For comparison, the constant multiplicative factor used by Miller and Savchick[63] and by Miller[57] is 1.05 $(3^{1/2})(a_0{}^{1/4}) = 1.551$ for r^*, or 3.102 for R^*, indicating that the two approaches will give similar radii for higher-row atoms but dissimilar radii for H, He, and first-row atoms. While eqs 28–32 require different multiplicative constants for different rows of the periodic table, it is reassuring that these constants obey a clear systematic trend and appear to be converging toward a well-defined limit.

We shall use eqs 27–32 in the algorithm defined below, but will also consider related expressions obtained for other values of p.

4.4. Higher-Order Dispersion Terms. We are now in a position to define an algorithm for calculating both atomic C_6 coefficients and vdW minimum-energy separations from atomic polarizabilities α_i. In turn, the term $-C_6/R^6$ for a given i,j pair might then be equated to the factor $-2\epsilon R^{*6}/R^6$ which appears in the second term of eq 2 for the Lennard-Jones 12-6 potential or to $-3\epsilon R^{*6}/R^6$ in the LJ-9-6 potential of eq 3 to yield an expression for the well depth ϵ_{ij}. Indeed, Diaz Pena et al.[48] employed a relationship of this sort, and Miller envisioned using his additive atomic polarizabilities in much this way.[57] However, the experimental potentials used to characterize the rare-gas interactions include significant contributions from higher-order dispersion coefficients C_8, C_{10}, C_{12}, ... As we have shown, the Buf-14-7 form accurately reproduces this behavior at distances considerably beyond the minimum-energy separation. Accordingly, we will require that the relationship obtained for ϵ_{ij} reflect the influence of the higher-order dispersion terms.

For the rare gases, comparison of the contribution from the whole dispersion series at $R = R^*$ to that given by the C_6 term alone shows that the higher-order dispersion terms increase the value given by the C_6 term by 44% for He···He, by 43% for Ne···Ne, and by 85% for Ar···Ar.[66] Similarly, when the value of the damping factor $F(R)$ at $R = R^*$ is taken into account, we find that ΔE_{disp} at $R = R^*$ (eq 7) exceeds the value obtained from the undamped C_6 term by 14% for He···He, by 32% for Ne···Ne, and by 85% for Ar···Ar. (Dispersion expansions for small molecules show a similar pattern. Thus, the C_8 and C_{10} terms in the HFD-B spherical potential for H_2[67] increase the dispersion energy at $R = R^*$ relative to that obtained from the undamped C_6 term by ca. 30%.) The resultant correction factors of 1.14 for He···He, 1.32 for Ne···Ne, and 1.85 for Ar···Ar could then be used to scale the well depths obtained from the simplified relationship between ϵ and C_6 suggested in the previous paragraph.

This is essentially the approach we shall take except that the requisite scale factors will be determined empirically, i.e., by determining the constants G_i needed to make the relationship

$$\epsilon_{ii} = {}^1/_2 k G_i{}^2 C_{6ii}/R^*_{ii}{}^6 \tag{33}$$

(65) Division by $0.529168^{-3} = 6.7487$ converts polarizability from units of $a_0{}^3$ to Å^3.

(66) See Table I for citations to the rare-gas potentials used in these comparisons.

(67) Norman, M. J.; Watts, R. O.; Buck, U. *J. Chem. Phys.* **1984**, *81*, 3500–3504.

Table IX. Calibration of Minimum-Energy Distances R^* and Well Depths ϵ for Like-Pair Interactions of Rare-Gas Atoms

	R^* (Å)		ϵ (kcal/mol)	
system[a]	exptl	calcd[b]	exptl	calcd[c]
He···He	2.963	2.961	0.0218	0.0217
Ne···Ne	3.087	3.085	0.0840	0.0840
Ar···Ar	3.761	3.758	0.2846	0.2843
Kr···Kr	4.011	4.006	0.3988	0.3984
Xe···Xe	4.363	4.362	0.5614	0.5614

[a] Experimental values from Table I or ref 28 (Kr···Kr, Xe···Xe). [b] From eqs 27–32 using polarizabilities in Å^3 calculated from values listed in Table VII. [c] From eqs 35–40.

hold for each of the like-pair interactions of the rare gases, where k is an appropriate unit-conversion factor (see below). Based on the just-documented underestimations of the dispersion series when only the C_6 term is used, we expect the values of G_i for He through Ar to range between $(1.14)^{1/2}$ and $(1.85)^{1/2}$, or between 1.07 and 1.36, and hence to increase modestly with atomic number.

4.5. vdW Well Depths. The expression we propose is

$$\epsilon_{ij} = {}^1/_2 k G_i G_j C_{6ij}/R^*_{ij}{}^6 \tag{34}$$

where R^*_{ij} is the cubic mean (eq 12) of the minimum-energy separations R^*_{ii} and R^*_{jj} for the like-pair interactions obtained from eqs 27–32. The C_{6ij} dispersion coefficients in eq 34 are calculated for both like- and unlike-pair interactions from eq 16 in terms of α_i and α_j and of N_i and N_j, the latter being given by eqs 20–24. (Equivalently, eq 16 could be used for C_{6ii} and C_{6jj} and eq 17 then employed to obtain C_{6ij}.) As noted in the previous subsection, G_i and G_j are constants calibrated to reproduce the well depths for the like-pair interactions of rare-gas atoms, and k is a conversion factor which yields ϵ_{ij} in units of kcal/mol when α_i and α_j are in Å^3 and R^*_{ij} is in Å; the required value[68] is $k = 241.55$. Using eq 16, eq 34 can also be written as

$$\epsilon_{ij} = \frac{181.16 G_i G_j \alpha_i \alpha_j}{(\alpha_i/N_i)^{1/2} + (\alpha_j/N_j)^{1/2}} \frac{1}{R^*_{ij}{}^6} \tag{35}$$

For the ideal gases, the following values are found for the scale parameters G_i: 1.209 for He; 1.282 for Ne; 1.345 for Ar; 1.359 for Kr; and 1.404 for Xe. As expected from the arguments given in the previous subsection, these values increase slightly with atomic number. Table IX confirms that they closely reproduce the experimentally determined vdW well depths ϵ_{ii} for the like-pair interactions of the rare-gas atoms and also demonstrates that the minimum-energy separations R^*_{ii} are reproduced by eq 27 when the A_i values in eqs 28–32 are used. This is, of course, as it should be, since the experimental minimum-energy separations R^*_{ii} and well depths ϵ_{ii} have been used to calibrate the A_i and G_i scale parameters. (Note, therefore, that no new disposable parameters have been introduced to describe the vdW potentials for like-pair interactions of rare-gas atoms. Rather, the R^*_{ii} and ϵ_{ii} parameters have simply been recast as the A_i and G_i parameters.) The atomic polarizabilities employed in these calculations were taken from Table VII and were converted[65] to units of Å^3.

Finally, we assume that the value for G_i found for a given rare gas can be employed for other atoms in the same row of the periodic table. Thus:

$$\text{for H and He} \quad G_i = 1.209 \tag{36}$$

$$\text{for C–Ne} \quad G_i = 1.282 \tag{37}$$

$$\text{for Si–Ar} \quad G_i = 1.345 \tag{38}$$

$$\text{for Ge–Kr} \quad G_i = 1.359 \tag{39}$$

$$\text{for Sn–Xe} \quad G_i = 1.404 \tag{40}$$

One curious property of the dimensionality employed for α in eqs 27 and 35 should be noted. When $p = {}^1/_4$, the well depth

(68) This value is obtained in terms of the conversion factors from a_0 to Å (0.529168) and from hartrees to kcal/mol (627.51) as $0.529168^{9/2} \times 627.51 \times 0.529168^{-6} = 241.55$.

Table X. Test of Algorithms and Combination Rules for Minimum-Energy Distances R^* and Well Depths ϵ for Unlike-Pair Interactions of Rare-Gas Atoms

	R^* (Å)		ϵ (kcal/mol)	
system[a]	exptl	calcd[b]	exptl	calcd[c]
He···Ne	3.036	3.027	0.0415	0.0423
He···Ar	3.478	3.453	0.0589	0.0629
He···Kr	3.691	3.637	0.0586	0.0651
He···Xe	3.968	3.920	0.0556	0.0620
Ne···Ar	3.489	3.487	0.1343	0.1292
Ne···Kr	3.621	3.663	0.1422	0.1352
Ne···Xe	3.861	3.935	0.1475	0.1304
Ar···Kr	3.881	3.890	0.3324	0.3307
Ar···Xe	4.067	4.105	0.3748	0.3637
Kr···Xe	4.174	4.200	0.4640	0.4586
rms dev		0.0392		0.0079
rms % dev		1.03		6.99

[a] Experimental values from Table V. [b] From eq 12 using calculated values of R^* for like-pair interactions given in Table IX. [c] From eq 35 using eqs 27–32 and 36–40 or, equivalently, from eq 41.

ϵ_{ii} for a like-pair interaction obtained using eqs 27–32 and 35–40 is

$$\epsilon_{ii} = 90.58 G_i^2 N_i^{1/2} / A_i^6 \qquad (41)$$

and thus is *independent of the value employed for the atomic polarizability* α_i. While a dependence of ϵ_{ii} on α_i does result for any choice for p in eq 27 other than $p = {}^1/_4$, the empirical underpinnings of the present approach suggest that the well depths encountered along any given row of the periodic table may increase uniformly toward the rare-gas value and in any case should not vary widely.

4.6. Tests of vdW Parameters for the Rare Gases. Table X shows that the present formalism describes the unlike-pair interactions of rare-gas atoms quite well. In particular, the vdW minimum-energy separations R^*_{ij} given by the cubic-mean rule (eq 12) in conjunction with R^*_{ii} and R^*_{jj} from eqs 27–32 reproduce the experimental values equally as well as did the same rule when operating on the experimentally determined R^*_{ii} and R^*_{jj} (cf. Table V). Moreover, the well depths ϵ_{ij} obtained from eqs 35–40 reproduce the experimental values even more closely than did the best of the rules previously examined (cf. Table VI). This performance verifies that the combination rule implicit in the Slater–Kirkwood formulation (i.e., eq 17) yields reasonably accurate well depths for unlike-pair interactions when appropriate parameters are employed.

For interactions of Ne with the heavier rare gases, Table XI compares the present Slater–Kirkwood-based approach to the more elaborate combination rules for rare-gas interactions discussed in Section 3.2. These comparisons are based on results presented by Aziz and co-workers.[28,39] The experimental like-pair parameters given in Table IX have been used as input for all the combination rules examined. As is evident, the Sikora approach is relatively poor for minimum-energy separations. Equation 12 is better, but the combination rules of Hiza, Tang and Toennies, and Bzowski are better still. For well depths, the present approach is the least satisfactory of those examined. Still, the errors are not large, particularly when viewed against those made by the commonly used geometric- and arithmetic-mean combination rules (cf. Tables V and VI). In view of the fact that other significant approximations are made in treating nonbonded interactions (see below), the present approach may well be sufficient.

5. vdW Parameters for the Merck Molecular Force Field (MMFF)

5.1. MMFF and MM2X Atom Types. To apply the formalism associated with eq 35 to the Merck Molecular Force Field (MMFF), we first need to define the atomic polarizabilities for the atomic valence state and bonding environments recognized by MMFF, i.e., for the MMFF atom types.

For the most part, the symbolic and numeric MMFF atom types (Table XII) are those used in MM2X (MM2-extended),[69] a force field which has been extensively employed in molecular-modeling applications at Merck. The numeric MM2X and MMFF atom types, which frequently correspond to those defined in MM2[5] and MM3,[6] are used to relate atomic interactions to force-field parameters when OPTIMOL,[69] the host molecular-modeling program for MM2X and MMFF, assembles the energy function for a particular chemical system. The extensive listing of symbolic atom types (ca. 130 entries) reflects the diverse chemical functionality recognized by MMFF, while the smaller set of numeric atom types (ca. 70) arises from instances in which the same parameters currently are assigned in chemical environments which we might at some point wish to treat dissimilarly.

Table XI. Comparison of Slater–Kirkwood and Literature Combination Rules for Unlike-Pair Interactions of Neon and Heavier Rare-Gas Atoms

	Ne···Ar[a]		Ne···Kr[b]		Ne···Xe[b]	
parameter[c]	R^*	ϵ	R^*	ϵ	R^*	ϵ
experiment	3.489	0.1343	3.621	0.1422	3.861	0.1475
Sikora	3.425	0.1335	3.549	0.1416	3.724	0.1411
Hiza	3.491	0.1326	3.656	0.1423	3.892	0.1452
Tang–Toennies	3.477	0.1312	3.645	0.1405	3.890	0.1428
Bzowski	3.466	0.1351	3.634	0.1432	3.871	0.1446
this work[d]	3.487	0.1292	3.663	0.1352	3.935	0.1304

[a] Experimental results and the application of the combination rules in the subsequent four rows are taken from ref 28. [b] Experimental results and the application of the combination rules in the subsequent four rows are taken from ref 39. Bzowski results cited are those in which the repulsive parameter in the Exp-6 potential is 13.0. [c] R^* is the vdW minimum-energy separation in Å, ϵ is the well depth in kcal/mol. [d] From Table X.

5.2. Atomic Polarizabilities for MMFF Atom Types. Table XIII lists the atomic polarizabilities used in MMFF together with the vdW parameters derived from them. To avoid encumbering the discussion, we shall reserve detailed information on the manner in which these polarizabilities were obtained for the Appendix. We note here simply that they have been extracted from data on molecular polarizabilities, where available, in a manner similar to that employed by Miller,[57] or have been estimated by making use of the following trends noted in the polarizabilities obtained from the experimental data: (1) atomic polarizability tends to decrease toward the rare-gas value across a given row of the periodic table; (2) increasing positive charge progressively reduces an atom's polarizability; and (3) multiple bonding usually increases the atomic polarizability.

5.3. Comparison to Miller's Atomic Polarizabilities. For those cases in which the atom types correspond sufficiently closely, Table XIV compares the MMFF atomic polarizabilities with the additive "ahp" values determined by Miller from comparable data on molecular polarizability.[57] As is evident, the two sets of polarizabilities agree closely in most instances. For sp^2-hybridized carbon, however, MMFF distinguishes type C=O from types C=C and CB and assigns a smaller polarizability, whereas Miller uses type CTR in all three environments. Similarly, Miller uses type STE in sulfides, sulfoxides, and sulfones, while MMFF uses types S, SO, and SO2 and assigns polarizabilities which decrease markedly with increasing oxygenation. For trigonal nitrogen, Miller obtains values for types NTR2 and NPI2 which agree reasonably well with the average of the diverse MMFF values for N=C, NPYD, NC=O, NPYL, and NO2. For singly-bonded tetrahedral oxygen, MMFF's OR has nearly the same polarizability as does Miller's OTE. In other environments, however, the MMFF polarizabilities for oxygen are higher. For example, Miller finds OPI2 = 0.27 for the oxygen in furan by assuming additivity,[70]

(69) MM2X, like MMFF, is a force field which differs from MM2 principally in that lone pairs on the heteroatoms are not used and in that electrostatic interactions take place between atom-centered charges, allowing proper treatment of charged systems. MM2X, the resident force field in OPTIMOL, has been parameterized for a wide range of functional groups but shares many parameters with MM2. The molecular modeling program OPTIMOL has been developed at the Merck Research Laboratories by the author and his colleagues.

(70) Miller, K. Personal communication.

Table XII. MMFF and MM2X Symbolic and Numeric Atom Types

atom type		definition [coordination number][a] {formal charge}[b]	atom type		definition [coordination number][a] {formal charge}[b]
symbolic	numeric		symbolic	numeric	
CR	1	alkyl carbon [4]	HNR	23	hydrogen on nitrogen in amines [1]
C=C	2	vinylic carbon [3]	HPYL	23	hydrogen on nitrogen in pyrrole [1]
CSP2	2	generic sp^2 carbon [3]	HNR	23	generic hydrogen on sp^3 nitrogen [1]
C=O	3	generic carbonyl carbon [3]	H3N	23	hydrogen in ammonia [1]
C=N	3	imine-type carbon [3]	HOCO	24	hydroxyl hydrogen in carboxylic acids [1]
CGD	3	guanidine carbon [3]	PO4	25	phosphate group phosphorus [4]
C=OR	3	ketone or aldehyde carbonyl carbon [3]	PO3	25	phosphorus with 3 attached oxygens [4]
C=ON	3	amide carbonyl carbon [3]	PO2	25	phosphorus with 2 attached oxygens [4]
COO	3	carboxylic acid or ester carbonyl carbon [3]	PO	25	phosphine oxide phosphorus [4]
COON	3	carbamate carbonyl carbon [3]	P	26	phosphorus in phosphines [3]
COOO	3	carbonic acid or ester carbonyl carbon [3]	HN=C	27	hydrogen on imine nitrogen [1]
C=OS	3	carbon in thioester with C=O [3]	HN=N	27	hydrogen on azo nitrogen [1]
C=S	3	carbon in thioester with C=S [3]	HNCO	28	hydrogen on amide nitrogen [1]
CSP	4	acetylenic carbon [2]	HNCC	28	hydrogen on enamine nitrogen [1]
=C=	4	allenic carbon [2]	HNCN	28	hydrogen in H—N—C=N moiety [1]
HC	5	hydrogen attached to carbon [1]	HNNN	28	hydrogen in H—N—N=N moiety [1]
HS	5	hydrogen attached to sulfur [1]	HSP2	28	generic hydrogen on sp^2 nitrogen [1]
HSI	5	hydrogen attached to silicon [1]	HOCC	29	enolic or phenolic hydroxyl hydrogen [1]
HP	5	hydrogen attached to phosphorus [1]	HOCN	29	hydroxyl hydrogen in HO—C=N moiety [1]
—O—	6	generic divalent oxygen [2]	CR4E	30	olefinic carbon in 4-membered ring [3]
OR	6	ether oxygen [2]	HOH	31	hydroxyl hydrogen in water [1]
OH2	6	oxygen in water [2]	O2CM	32	oxygen in carboxylate group [1] {$-^1/_2$}
OC=O	6	divalent oxygen, carboxylic acid, or ester [2]	HOS	33	hydrogen on oxygen attached to sulfur [1]
OC=C	6	enolic or phenolic oxygen [2]	NR+	34	quaternary nitrogen [4] {1}
OC=N	6	oxygen in —O—C=N moiety [2]	OM	35	oxide oxygen on sp^3 carbon [1] {−1}
OSO3	6	divalent oxygen in sulfate group [2]	OM2	35	oxide oxygen on sp^2 carbon [1] {−1}
OSO2	6	divalent oxygen in sulfite group [2]	HNR+	36	hydrogen on quaternary nitrogen [1]
OSO	6	divalent oxygen in R(RO)S=O [2]	HNN+	36	hydrogen on imidazolium nitrogen [1]
—OS	6	other divalent oxygen attached to sulfur [2]	HNC+	36	hydrogen on protonated imine nitrogen [1]
OPO3	6	divalent oxygen in phosphate group [2]	HGD+	36	hydrogen on guanidinium nitrogen [1]
OPO2	6	divalent oxygen in phosphite group [2]	CB	37	aromatic carbon, e.g., in benzene [3]
OPO	6	divalent oxygen in R(RO)P=O [2]	NPYD	38	aromatic nitrogen with σ lone pair [2]
—OP	6	other divalent oxygen attached to phosphorus [2]	NPYL	39	aromatic 5-ring nitrogen with π lone pair [2]
O=C	7	generic carbonyl oxygen [1]	NC=C	40	enamine or aniline nitrogen, deloc lp [3]
O=CN	7	carbonyl oxygen in amides [1]	NC=N	40	nitrogen in N—C=N with deloc lp [3]
O=CR	7	carbonyl oxygen in aldehydes and ketones [1]	CO2M	41	carbon in carboxylate anion [3]
O=CO	7	carbonyl oxygen in acids and esters [1]	NSP	42	triply-bonded nitrogen [1]
OS	7	sulfoxide oxygen [1]	NSO2	43	sulfonamide nitrogen [3]
O2S	7	one of 2 terminal Os on sulfur [1] {variable}[c]	STHI	44	aromatic 5-ring sulfur with π lone pair [2]
O3S	7	one of 3 terminal Os on sulfur [1] {variable}[c]	NO2	45	nitrogen in nitro group [3]
O4S	7	one of 4 terminal Os on sulfur [1] {variable}[c]	N=O	46	nitrogen in nitroso group [2]
OP	7	oxygen in phosphine oxide [1]	O=N	47	oxygen in nitroso group [1]
O2P	7	one of 2 terminal Os on P [1] {variable}[c]	ONX	47	oxygen in *N*-oxides [1]
O3P	7	one of 3 terminal Os on P [1] {variable}[c]	O2N	48	oxygen in nitro group [1]
O4P	7	one of 4 terminal Os on P [1] {variable}[c]	O+	49	oxonium oxygen [3] {1}
NR	8	amine nitrogen [3]	HO+	50	hydrogen on oxonium oxygen [1]
N=C	9	imine nitrogen [2]	O=+	51	oxenium oxygen [2] {1}
N=N	9	azo-group nitrogen [2]	HO=+	52	hydrogen on oxenium oxygen [1]
NC=O	10	amide nitrogen [3]	O%+	53	triply-bonded O^+ [1] {1}
NN=N	10	nitrogen in N—N=N with deloc lp [3]	N+=C	54	iminium nitrogen [3] {1}
F	11	fluorine [1]	NCN+	55	either nitrogen in N^+=C—N: [3] {$^1/_2$}
Cl	12	chlorine [1]	NGD+	56	guanidinium ion nitrogen [3] {$^1/_3$}
Br	13	bromine [1]	CIM+	57	aromatic carbon between Ns in imidazolium, [3]
I	14	iodine [1]	CGD+	57	guanidinium carbon [3]
S	15	thiol, sulfide, or disulfide sulfur [2]	CNN+	57	carbon in +N=C—N: resonance structures [3]
S=C	16	sulfur doubly-bonded to carbon [1]	NIM+	58	aromatic nitrogen in imidazolium [3] {$^1/_2$}
SO	17	sulfoxide sulfur [3]	OFUR	59	aromatic 5-ring oxygen with π lone pair [2]
SO2	18	sulfone sulfur [4]	C%-	60	isonitrile {0} or acetylide carbon [1] {−1}
SO2N	18	sulfonamide sulfur [4]	NR%	61	isonitrile nitrogen [2]
SO3	18	sulfonate-group sulfur [4]	NMSO	62	nitrogen in sulfonamide anion [2] {−1}
SO4	18	sulfate-group sulfur [4]	C5A	63	aromatic 5-ring C, α to N:, O:, or S: [3]
SI	19	silicon [4]	C5B	64	aromatic 5-ring C, β to N:, O:, or S: [3]
CR4R	20	aliphatic carbon in 4-membered ring [4]	N5A	65	aromatic 5-ring N, α to N:, O:, or S: [2]
HOR	21	hydroxyl hydrogen in alcohols [1]	N5B	66	aromatic 5-ring N, β to N:, O:, or S: [2]
HOS	21	hydroxyl hydrogen in H–O–S moiety [1]	N2OX	67	sp^2 hybridized *N*-oxide nitrogen [3]
HOP	21	hydroxyl hydrogen in H–O–P moiety [1]	N3OX	68	sp^3 hybridized *N*-oxide nitrogen [4]
HO	21	generic hydroxyl hydrogen [1]	NPOX	69	pyridinium *N*-oxide nitrogen [3]
CR3R	22	aliphatic carbon in 3-membered ring [4]			

[a] Number of attached atoms. [b] Initial full or fractional charge, from which final MMFF or MM2X partial atomic charges are obtained by adding contributions arising from the relative polarity of bonds involving attached atoms. [c] The formal charge is determined by dividing the net ionic charge on the SO_x or PO_x group among the equivalent terminal oxygens.

whereas MMFF, as noted in the Appendix, intentionally chooses a value for OFUR which exceeds the additive one. On the whole, therefore, the agreement is good, but differences arise which can be attributed to a difference in philosophy. Except in one case, Miller assigns the atom type solely on the basis of hybridization, whereas MMFF makes distinctions which also depend on bond

Table XIII. MMFF Atomic Polarizabilities α (Å^3), C_6 Coefficients (kcal/mol-Å^6), Minimum-Energy Separations R^* (Å), and Well Depths ϵ (kcal/mol) (Other Than For Hydrogen Atoms Attached to Heteroatoms)

MMFF atom type						MMFF atom type					
symbolic	numeric	α	C_6	R^*	ϵ	symbolic	numeric	α	C_6	R^*	ϵ
CR	1	1.05	308	3.94	0.068	NPYL	39	0.80	218	3.68	0.072
C=C	2	1.35	448	4.19	0.068	NC=C	40	1.00	304	3.89	0.072
C=O	3	1.10	330	3.98	0.068	CO2M	41	1.30	424	4.15	0.068
CSP	4	1.30	424	4.15	0.068	NSP	42	1.00	304	3.89	0.072
HC	5	0.40	41	3.50	0.016	NSO2	43	1.00	304	3.89	0.072
OR	6	0.65	168	3.49	0.076	STHI	44	2.30	1384	4.09	0.268
O=C	7	0.75	209	3.62	0.076	NO2	45	1.15	375	4.03	0.072
NR	8	1.05	327	3.94	0.072	N=O	46	1.30	451	4.15	0.072
N=C	9	1.40	504	4.23	0.072	O=N	47	0.75	209	3.62	0.076
NC=O	10	1.00	304	3.89	0.072	O2N	48	0.75	209	3.62	0.076
F	11	0.40	86	3.09	0.080	O+	49	0.40	81	3.09	0.076
Cl	12	2.30	1427	4.09	0.276	O=+	51	0.40	81	3.09	0.076
Br	13	3.40	2782	4.33	0.389	O%+	53	0.40	81	3.09	0.076
I	14	5.50	6159	4.72	0.552	N+=C	54	1.00	304	3.89	0.072
S	15	3.15	2219	4.42	0.268	NCN+	55	0.80	218	3.68	0.072
S=C	16	3.90	3056	4.67	0.268	NGD+	56	0.80	218	3.68	0.072
SO	17	2.70	1761	4.26	0.268	CIM+	57	1.00	286	3.89	0.068
SO2N	18	2.10	1208	4.00	0.268	NIM+	58	0.80	218	3.68	0.072
SI	19	4.50	3544	4.84	0.251	OFUR	59	0.50	114	3.27	0.076
CR4R	20	1.05	308	3.94	0.068	C%-	60	1.80	690	4.51	0.068
CR3R	22	1.10	330	3.98	0.068	NR%	61	0.80	218	3.68	0.072
PO4	25	1.60	778	3.73	0.259	NM	62	1.50	558	4.30	0.072
P	26	3.60	2625	4.57	0.259	C5A	63	1.35	448	4.19	0.068
CR4E	30	1.35	448	4.19	0.068	C5B	64	1.35	448	4.19	0.068
O2CM	32	1.00	322	3.89	0.076	N5A	65	1.00	304	3.89	0.072
NR+	34	0.80	218	3.68	0.072	N5B	66	1.00	304	3.89	0.072
OM	35	1.20	423	4.07	0.076	N2OX	67	0.70	178	3.56	0.072
CB	37	1.35	448	4.19	0.068	N3OX	68	0.70	178	3.56	0.072
NPYD	38	1.00	304	3.89	0.072	NPOX	69	0.70	178	3.56	0.072

Table XIV. Comparison of MMFF and Miller Atomic Polarizabilities (Å^3)

MMFF atom type		Miller atom type	atomic polarizability	
symbolic	numeric		MMFF[a]	Miller[b]
HC	5	H	0.40	0.387
CR	1	CTE	1.05	1.061
C=C	2	CTR	1.35	1.352
C=O	3	CTR	1.10	1.352
CB	37	CTR	1.35	1.352
CSP	4	CDI	1.30	1.283
NR	8	NTE	1.05	0.964
N=C	9	NTR2	1.40	1.030
NPYD	38	NTR2	1.00	1.030
NC=O	10	NPI2	1.00	1.090
NPYL	39	NPI2	0.80	1.090
NO2	45	NPI2	1.15	1.090
NSP	42	NDI	1.00	0.956
OR	6	OTE	0.65	0.637
O2N	48	OTE	0.75	0.637
O=C	7	OTR4	0.75	0.569
OFUR	59	OPI2	0.50	0.274
F	11	F	0.40	0.296
PO4	25	PTE	1.60	1.538
S	15	STE	3.15	3.000
SO	17	STE	2.70	3.000
SO2	18	STE	2.10	3.000
S=C	16	STR4	3.90	3.729
STHI	44	SPI2	2.30	2.700
Cl	12	Cl	2.30	2.315
Br	13	Br	3.40	3.013
I	14	I	5.50	5.415

[a] From Table XIII. [b] From ref 57.

multiplicity, ionic charge, and the number and electronegativity of attached neighbors.

5.4. MMFF vdW Parameters. Table XIII also lists the MMFF C_6 coefficients, minimum-energy separations R^*, and well depths ϵ obtained from eq 27, eq 35, and supporting equations in Section 4. As is shown in eq 41, the atomic polarizabilities affect R^* in the present approach (which uses $p = {}^1/_4$ in eq 27) but, for like-pair interactions, do not affect ϵ. Thus, the well depths ϵ_{ii} depend only on scale factors G_i and A_i, which are taken to depend on the row but not the position in the row, and on N_i, the Slater–Kirkwood effective number of electrons. Moreover, N_i in the present approach is determined solely by the atomic number (cf. eqs 20–24), and hence the same well depth is obtained for all atom types defined for a given atomic species. In view of the strong dependence of polarizability on chemical environment, it might be asked whether the N_i should show similar variations. However, even a substantial 40% variation in N_i in eq 41 would produce only a 20% variation in ϵ. The present framework thus suggests that well depths do not depend greatly on chemical environment.

Well Depths. In the first row of the periodic table, the calculated well depths in kcal/mol are 0.068 for carbon, 0.072 for nitrogen, 0.076 for oxygen, and 0.080 for fluorine. As is implied by eq 41, these values increase toward the value of 0.0840 kcal/mol for neon. Similarly, in the second row the well depths are 0.251 for silicon, 0.259 for phosphorus, 0.268 for sulfur, and 0.276 for chlorine, as against 0.2865 for argon. Bromine in the third row and iodine in the fourth are also assigned well depths (0.389 and 0.550) which slightly underlie those for the rare gases (0.399 and 0.561).

Minimum-Energy Separations. In contrast to well depths, the MMFF R^* values depend significantly on the chemical environment. Thus, the value calculated for R^* is just under 4 Å for saturated carbon (CR), but increases to about 4.2 Å for unsaturated carbon in alkenes (C=C) and aromatic compounds (CB). The minimum-energy separations for nitrogen are somewhat smaller except for the multiply bonded N=C, and those for oxygen (OR, O=C) are smaller yet, on the order of 3.5 Å. For the halogens, R^* increases uniformly and reaches ca. 4.7 Å for iodine. As will be shown below, these values agree reasonably well with those employed in existing molecular mechanics force fields.

Aliphatic Hydrogen. For hydrogen the situation is problematic. As shown in Table XIII, the present approach yields $R^* = 3.50$ Å and $\epsilon = 0.016$ kcal/mol for hydrogen bonded to carbon. Comparisons to vdW parameters for other force fields (see below) suggest that the value obtained for R^* may be too large; perhaps the assumption that A_i in eqs 27–32 can be taken as being equal to the rare-gas value is too limiting. Taking A_i for hydrogen as 4.0 in eq 28, rather than 4.4, yields $R^* = 3.18$ Å and $\epsilon = 0.029$

kcal/mol. The value for R^* is close to the MM3 value of 3.24 Å, but the well depth now agrees less well with MM3's value of 0.020 kcal/mol. Given its special role in intermolecular interactions, we will defer the choice of vdW parameters for aliphatic hydrogen until we can assess the performance of the MMFF force field.

Polar Hydrogens. We will also defer the selection of parameters for hydrogens potentially involved in hydrogen-bonding interactions. Because these hydrogens carry an appreciable positive charge, they can be expected to have lower polarizabilities than their aliphatic brethren. Even a 2-fold reduction in polarizability would only decrease R^* by about 15%, however, if the functional dependence in eq 27 is roughly correct. Similarly, a physically plausible reduction of 50% in N_i would diminish the well depth ϵ by only about 30%. These considerations suggest that polar and aliphatic hydrogens should not have greatly dissimilar vdW parameters. Our experience with MM2X, however, indicates that markedly smaller values for ϵ and R^* are needed to reproduce hydrogen-bond energies and geometries. It should be noted, however, that the crude electrostatic models used in current force fields (and in MMFF) obviate the physical significance of the vdW parameters assigned for polar hydrogens. Thus, the quantum-mechanically calculated electrostatic potential near a hydrogen-bond acceptor typically has a minimum in the lone pair region.[71,72] In the simplest model, a proton or polar hydrogen atom having a vanishing vdW radius would be drawn in to this position but no further.[72] In current molecular mechanics representations, in contrast, the calculated electrostatic potential contains no such minimum but rather diverges to negative infinity at the heteroatom nucleus. Accordingly, the vdW parameters assigned for the hydrogen–heteroatom interaction must provide sufficient repulsion to ward off collapse while also helping to compensate for the underestimation of the electrostatic potential in the lone pair region. vdW Parameters which satisfy these requirements cannot be physically based. The MMFF parameters for polar hydrogens will be presented in a subsequent paper, when the energetic and geometric properties of hydrogen-bonding interactions are examined.

Other Values for p in Eq 27. The vdW parameters vary systematically when the assumed dependence of R^* on the atomic polarizability is modified (and the A_i parameters in eqs 28–32 are recalibrated). For example, when p is increased from $^1/_4$ to $^1/_3$ (the exponent used in Nagle's[61] polarizability radius), the minimum-energy separations R^* and well depths ϵ respectively increase and decrease by amounts which depend on how much the atomic polarizability exceeds that of the rare-gas atom in the same row. The well depths are no longer equal for all atom types arising from a given atomic species but do not vary widely. For example, the minimum-energy separations for carbon increase by about 0.3–0.4 Å, and the well depths decrease by about 40%. A similar variation is found for nitrogen and oxygen. Interestingly, in this model the more polarizable atom types for a given atomic species have the *smaller* calculated well depths. Conversely, for $p = {}^1/_5$ the minimum-energy separations decrease (by 0.1–0.2 Å for carbon) and the well depths increase (to 0.090–0.098 kcal/mol for carbon). Again, the variation in well depths for a given atomic species is modest. In this model, however, they increase with increasing polarizability for a given atomic species and need no longer approach the rare-gas value from below.

In summary, while models of the present type provide some flexibility in the determination of vdW parameters, they do not support the assignment of well depths for a given atomic species which vary markedly with chemical environment.

5.5. Comparison to vdW Parameters Used in Other Force Fields. Table XV lists the vdW parameters used in the MM2X force field[69] which was developed at Merck some years ago as an extension of MM2. MM2X uses different vdW forms and separate sets of vdW (and electrostatic) parameters for intra- and intermolecular interactions.[73] For the most part, the intramolecular vdW parameters remain those of MM2. The intermolecular vdW parameters, in contrast, were patterned after those employed in

Table XV. MM2X vdW Parameters[a]

	intramolecular		intermolecular	
atom type	R^*	ϵ	R^*	ϵ
1	4.00	0.040	4.00	0.040
2	3.88	0.044	3.60	0.080
3	3.88	0.044	3.60	0.120
4	3.88	0.044	3.60	0.120
5	3.00	0.047	2.80	0.040
6	3.48	0.050	3.30	0.180
7	3.48	0.066	3.20	0.200
8	3.64	0.055	3.50	0.160
9	3.64	0.055	3.50	0.160
10	3.64	0.055	3.50	0.160
11	3.30	0.078	3.20	0.160
12	4.06	0.240	4.00	0.240
13	4.36	0.320	4.40	0.320
14	4.64	0.424	4.60	0.420
15	4.22	0.202	4.20	0.220
16	4.22	0.202	3.80	0.220
17	4.22	0.202	3.80	0.220
18	4.22	0.202	3.60	0.220
19	4.50	0.140	4.80	0.140
21	2.40	0.036	1.60	0.020
22	3.80	0.044	4.20	0.040
23	2.65	0.034	1.60	0.020
24	1.80	0.015	1.60	0.020
25	4.20	0.200	4.20	0.200
26	3.96	0.034	4.60	0.040
27	3.96	0.034	4.60	0.040
28	1.60	0.020	1.60	0.020
29	1.80	0.015	1.60	0.020
30	3.40	0.060	3.40	0.120
31	1.60	0.020	1.60	0.020
32	3.60	0.066	3.60	0.150
33	1.60	0.020	1.60	0.020
34	3.64	0.055	3.64	0.120
35	3.48	0.066	4.40	0.150
36	1.60	0.015	1.60	0.020
37	3.88	0.044	4.00	0.080
38	3.64	0.055	3.60	0.160
39	3.64	0.055	3.60	0.160
41	3.88	0.044	3.80	0.120
42	3.64	0.055	3.50	0.160
43	3.64	0.055	3.50	0.160
44	4.22	0.202	3.80	0.220
45	3.64	0.055	3.50	0.160
48	3.60	0.066	3.60	0.150
54	3.64	0.055	3.60	0.160
55	3.64	0.055	3.60	0.160
56	3.64	0.055	3.60	0.160
57	3.88	0.044	4.00	0.080
58	3.64	0.055	3.60	0.120
59	3.48	0.050	3.30	0.180
60	3.88	0.044	3.60	0.120
61	3.64	0.055	3.50	0.160
62	3.64	0.055	3.50	0.160

[a] Minimum-energy separations R^* and well depths ϵ in Å and kcal/mol respectively. Intramolecular interactions use the MM2 Exp-6 form so that actual well depths are 1.163 times the listed well depth parameters. Intermolecular interactions use the Lennard–Jones 9-6 form. The numerical atom types correspond to those listed in Table XII. Additional R^*/ϵ values of 2.20/0.200, 3.20/0.200, 2.00/0.200, and 2.00/0.200 are used for Zn^{2+} (type 40) Ca^{2+} (type 63), Cu^{2+} (type 64) and Cu^+ (type 65) in intermolecular interactions.

(71) Luque, F. J.; Illas, F.; Orozco, M. *J. Comput. Chem.* **1990**, *11*, 416–430.

(72) Scrocco, E.; Tomasi, J. *Adv. Quantum Chem.* **1978**, *11*, 115–193; cf. Table IV.

(73) For intramolecular vdW interactions, MM2X uses the MM2 Exp-6 potential; for intermolecular interactions, it uses the Lennard–Jones 9-6 form. Electrostatic interactions are calculated from Coulomb's law using atom-centered charges, obtained for intramolecuar interactions from MM2's bond-dipole parameters in most cases, and for intermolecular interactions obtained from a companion set of bond charge increments chosen to reproduce electrostatic-potential-derived charges.

Table XVI. Some Minimum-Energy Separations R^* and Well Depths ϵ for the AMBER, VFF, CHARMM, and MM3 Force Fields

MMFF atom type	AMBER[a]		VFF[b]		CHARMM[c]		MM3[d]	
	R^*	ϵ	R^*	ϵ	R^*	ϵ	R^*	ϵ
CR	3.60	0.060	4.35	0.039	3.60	0.090	4.08	0.027
C=C			4.06	0.148	3.92	0.040	3.92	0.056
C=O	3.70	0.120	4.06	0.148	3.74	0.141	3.88	0.056
CB	3.70	0.120	4.06	0.148	4.08	0.050	3.92	0.056
CSP					4.20	0.030	3.88	0.056[e]
HC	3.08	0.010	2.75	0.038	2.66	0.042	3.24	0.020
NR	3.70	0.120			3.66	0.090	3.86	0.043
NC=O	3.50	0.160	3.93	0.167	3.66	0.090	3.86	0.043[e]
NPYD	3.50	0.160	3.93	0.167	3.66	0.090	3.86	0.043[e]
NPYL	3.50	0.160	3.93	0.167	3.66	0.090	3.86	0.043[e]
NR+	3.70	0.080	3.93	0.167	3.66	0.090	3.86	0.043[e]
OR	3.30	0.150	3.21	0.228	3.20	0.159	3.64	0.059
O=C	3.20	0.200	3.21	0.228	3.06	0.159	3.64	0.059
O2CM	3.20	0.200	3.21	0.228	3.12	0.152	3.64	0.059[e]
F					3.30	0.078	3.42	0.075[e]
S	4.00	0.200	3.78	0.043	3.78	0.043	4.30	0.202[e]
PO4	4.20	0.200			3.80	0.100	4.40	0.168[e]
CL					4.06	0.240	4.14	0.240[e]
BR					4.36	0.320	4.44	0.320[e]
I					4.30	0.800	4.72	0.424[e]

[a] Reference 7. [b] Reference 9. [c] Parameters for CHARMM atom types CT, CUA1, C, C6R, CUY1, HA, N*, OE, O, OC, XF, S*, P*, XCL, XBR, and XI as accessed by QUANTA 3.2 from MSI. The N* parameters are applied to all listed nitrogen atom types. [d] For CR, C=C, CB and HC, (MM3 types 1, 2, aromatic, and 5) from Allinger, N. L.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8576–8582; for C=O and O=C (MM3 types 3 and 7), from Allinger, N. L.; Chen, Kuohsiang; Pathiaseril, A. *J. Am. Chem. Soc.* **1991**, *113*, 4505–4517; for OR (MM3 type 6), from Allinger, N. L.; Rahman, M.; Lii, J.-H. *J. Am. Chem. Soc.* **1990**, *112*, 8293–8307; for NR (MM3 type 8), from Allinger, N. L.; Schmitz, L. R. *J. Am. Chem. Soc.* **1990**, *112*, 8307–8315. [e] Preliminary MM3 parameters, described respectively by MM3 atom types 4, 9, 9, 9, 48, 47, 11, 15, 60, 12, 13, and 14 (Allinger, N. L. Personal communication).

AMBER. Selected vdW parameters for the AMBER, VFF, CHARMM (Quanta Version 3.2), and MM3 force fields are listed in Table XVI.

Minimum-Energy Separations. Comparison of Tables XIII and XV shows that the MMFF R^* values for carbon are about 0.2 Å larger than the intramolecular MM2X (and MM2) values but that those for oxygen (e.g., for OR and O=C) are comparable. A larger difference arises for nitrogen, where R^* ranges from 3.68 to 4.23 Å in MMFF (for neutral nitrogens) as against 3.62 to 3.64 Å in MM2 and MM2X. The R^* values for carbon in AMBER, VFF, CHARMM, and MM3 range fairly widely and are occasionally as large as those for MMFF but usually are smaller. These force fields also tend to define somewhat smaller oxygens and considerably smaller nitrogens. For the halogens, the CHARMM values behave somewhat irregularly, whereas the MM2/MM2X values increase monotonically and agree fairly closely with the MMFF values. Finally, we note that the oxidation-state-dependent MMFF values for S, SO, and SO2 bracket the constant MM2/MM2X value.

Thus, MMFF usually defines slightly larger vdW minimum-energy separations R^* than those employed in other force fields. The difference is largest for nitrogen, where molecular polarizabilities suggests that nitrogen's atomic polarizability is comparable to that of carbon.

Well Depths. Tables XV and XVI show that the other listed force fields differ more widely on vdW well depths than they do on minimum-energy separations. In particular, MM2 (hence MM2X) and MM3 assign well depths for carbon, nitrogen, and oxygen atoms which are essentially independent of the chemical environment, whereas AMBER and VFF employ significantly larger well depths for atoms involved in multiple bonding. CHARMM well depths also vary strongly, though less regularly, with environmental factors. In MMFF, as previously noted, no dependence on chemical environment arises when $p = {}^1/_4$ is used in eq 27. For carbon, the MMFF value is comparable to the average of the values used for carbon in the other force fields. For nitrogen and oxygen, in contrast, AMBER, VFF, and CHARMM assign well depths which tend to be significantly larger than those of MMFF, whereas MM2 and MM3 assign somewhat smaller values than does MMFF. For the halogens, the MM2/MM3 well depths used in the MM2X intramolecular force field agree well with those obtained in the MMFF formalism, as do the CHARMM well depths except in the case of iodine.

In summary, the MMFF well depths appear reasonable in magnitude but show no dependence on chemical environment, whereas well depths employed in certain other force fields sometimes depend strongly on the environment.

6. Discussion and Conclusions

This paper explores the premise that the well-characterized vdW interactions of rare-gas atoms can be used advantageously in constructing molecular mechanics force fields intended for use on more complex systems. We began by showing that the commonly used Lennard–Jones and Exp-6 potentials account poorly for the high quality rare-gas data. In contrast, we showed that the reduced form of the relatively simple Buf-14-7 distance-buffered potential of eq 10 accurately reproduces the reduced rare-gas potentials over the range of interatomic separations of primary interest in molecular mechanics calculations. We also characterized the performance of the commonly used arithmetic- and geometric-mean combination rules. Here, too, we showed that these approaches account poorly for the behavior found for rare-gas interactions, and we suggested alternative combination rules which perform significantly better. Among these are the cubic-mean rule of eq 12 for minimum-energy separations R^*_{ij} and the HHG rule of eq 14 for well depths ϵ_{ij}.

To make further use of the known behavior of the rare gases, we proposed a formalism for relating vdW well depths and minimum-energy separations to experimentally derived data on atomic polarizabilities. Two key equations were employed. The first, eq 27, correlates the polarizability of an atom in a given chemical environment with the minimum-energy separation R^*. The second, eq 35, expresses the well depth ϵ_{ij} in terms of R^*_{ij}, the atomic polarizabilities α_i and α_j, and the Slater–Kirkwood effective electron numbers N_i and N_j. These equations were calibrated to reproduce the accurately known rare-gas well depths and minimum-energy separations; they yielded the vdW parameters given in Table XIII when we made the fundamental additional assumption that the same calibration constants can be employed for the neighboring atoms in each row of the periodic table. Except for hydrogen atoms, which constitute special cases, we intend to make use of the derived vdW parameters in the Merck Molecular Force Field (MMFF) currently being developed.

Comparisons to vdW parameters employed in such other force fields as MM2, our laboratory's MM2-based MM2X, AMBER, VFF, CHARMM, and MM3 demonstrated that little consensus

on vdW parameters exists. Nevertheless, rough agreement with the calculated MMFF values was found, apart from a tendency of the MMFF formalism (i) to yield slightly larger minimum-energy separations and (ii) essentially in agreement with MM2 and MM3 but in opposition to AMBER, VFF, and CHARMM, to give vdW well depths which do not depend on the chemical environment.

Clear evidence bearing on the utility of the MMFF vdW parameters must wait until the MMFF force field is implemented and tested, as is now being done. As noted in the Introduction, we wish to emphasize that the framework presented in this paper looks beyond the vehicle being provided by MMFF in its current form. In particular, this framework addresses what might be characterized as true vdW parameters, i.e., vdW parameters which would be appropriate if all other physical factors which influence molecular interactions energies were modeled accurately. Neither the initial version of MMFF nor current molecular mechanics force fields provide such a context, however, largely because they model electrostatic interactions crudely. In particular, they almost always neglect induced dipole effects arising from the polarizability of the electron distribution, use only bond dipole or atomic monopole terms to model the molecular electrostatic potential, and ignore geometry-dependent (charge-flux) variations of the electronic charge distribution. Nearly all force fields also use vdW potentials which depend on interatomic distances but not on orientational factors. In view of these deficiencies, any such force field may well ask that its vdW parameters adopt values which help to compensate for these missing or poorly represented physical interactions for the specific experimental or theoretical data used in its calibration. The resultant vdW parameters may then be nonphysical and may even describe data other than that used in their derivation less well than would parameters which describe solely the van der Waals interaction. For these reasons, while the MMFF parameters presented here may prove useful in the context of a relatively simple force field, they may well perform optimally only in the context of a more elaborate, physically superior force field—in particular, one which more accurately represents the molecular charge distribution and properly takes the polarizability of the electronic charge distribution into account.[2,3]

Finally, we note that other possibilities exist within the general framework advanced here. Thus, systematic relations for the quantities expressed through eqs 20–24, 28–32, and 36–40 might be retained without invoking the particular forms used in this work. Also, we have already noted that use of the modified exponent $p = {}^1/_5$ in eq 27 produces smaller values for R^* which better accord with those employed in other force fields and which might prove superior in practice. Alternatively or additionally, the atomic polarizabilities α_i might be allowed to vary systematically from the values derived from molecular polarizabilities, perhaps by fitting them to experimental or theoretical data on molecular interactions.

Much remains to be done before it will be possible to determine the validity and utility of the conceptual framework offered here.

Acknowledgment. I thank Drs. Bruce Bush and Robert Nachbar for carefully reading and making helpful suggestions on this manuscript and for assistance in preparing the figures.

Note Added in Proof. Systematic relationships analogous to, but broader than, eqs 20–24 for N_i, the Slater–Kirkwood effective number of electrons, have recently been described (cf. Cambi, R.; Cappelletti, D.; Liuti, G.; Pirani, F. *J. Chem. Phys.* **1991**, *95*, 1852–1861). That paper also presents an alternative formalism for relating minimum-energy separations R^*, C_6 coefficients, and well depths ϵ to atomic polarizability. We will discuss the relationship between our and their approaches in a forthcoming paper describing the parametrization of MMFF for intermolecular interactions.

Appendix. Assignment of MMFF Atomic Polarizabilities

The atomic polarizabilities for the MMFF atom types are listed in Table XIII. These polarizabilities were extracted from data on molecular polarizabilities in a manner similar to that employed by Miller.[57] In discussing how these values were obtained, our emphasis will partly be on providing a formal documentation and partly on elucidating discernable trends in atomic polarizability. We shall focus on the most general or most commonly occurring chemical environment in those instances in which the symbolic to numeric mapping reflected in Table XII is other than one to one.

Roughly half the atomic polarizabilities listed in Table XIII are based directly on the extensive measurements of molar refractivities of Vogel,[74] who assumed additivity to extract atomic, group, and bond polarizabilities for a wide range of chemical environments. We have not explicitly employed a least-squares approach, but Vogel implicitly has done so; his atomic and group polarizabilities usually represent average values obtained from five or more chemical comparisons. As a result, the approach taken here is comparable to that employed by Miller.[75] Molar refractivity, R_M, is related to molecular polarizability through the Lorenz–Lorentz equation[76]

$$R_M = \left(\frac{\eta^2 - 1}{\eta^2 + 2}\right)\frac{M}{d} = \frac{4\pi}{3}N\alpha$$

where η is the index of refraction, N is Avogadro's number, M is the molecular weight, and d is the density. When R_M is in cm^3/mol, the polarizability α is obtained in units of $Å^3$ as 0.3964 R_M. We have used this relationship to obtain polarizabilities from the atomic and group molar refractivities, R_D, measured by Vogel at the frequency of the sodium D line. Most citations are from Vogel's paper XXIII[74] (V23); many are from his summary Table XXII (V23T). Citations to Miller are to the previously referenced paper.[57]

Hydrocarbons, Halides, and Singly-Bonded Heteroatoms. Polarizabilities listed in Table XIII for the MMFF atom types CR, C=C, HC, CL, BR, I, S, and CR3R are from V23T, usually rounded slightly to avoid implying greater precision than may be warranted. Those of 1.35 and 1.10 for C=C and CR3R add the corrections for C=C double bonds and three-carbon rings to the V23T value for C (in CH_2). For CR4R, saturated carbon in a 4-membered ring, we use the same value as for CR; for CR4E, olefinic carbon in a 4-membered ring, we assume the same value as for C=C. The value 1.35 is also obtained for aromatic carbon, CB, from the V23T value for phenyl by assuming additive contributions from the hydrogen atoms. Thus, it seems appropriate to also use this value for C5A and C5B, aromatic carbons α and β to the unique (π-lone-pair) heteroatom in 5-membered heteroaromatic rings. The value for F in V23T is marked as preliminary; we have elected to use the same value as for HC, noting that the polarizabilities for alkyl fluorides cited by Miller differ little from those for the corresponding alkanes. (Miller's cited polarizabilities for aromatic systems indicate that substitution of F for H significantly reduces the polarizability, implying that F is less polarizable than is HC. However, the molar refractivities for fluorinated benzenes cited by Aroney et al.[77] are consistent with equal polarizabilities.) The listed polarizability for NR reflects an average of Vogel's values for N in secondary and tertiary amines.[78] The value 0.65 for atom type OR lies between that of 0.669 for ethers (V23T) and that of 0.602 for alcohols.[79]

(74) Vogel, A. I. *J. Chem. Soc.* **1948**, 1833–1855. This paper is paper XXIII in the series and is referred to in the text as V23; citations from Table XXII in this paper are referred to as V23T.

(75) Thus, while Miller (ref 57) employed a least-squares approach, not all of the parameters were determined simultaneously. Rather, Miller used a staged approach in which atomic polarizabilities were derived for an initial set of compounds, compounds of additional types were added, and new parameters were then determined by least squares only for the newly encountered atom types (Miller, K. Personal communication).

(76) Schuyer, J.; Blom, L.; van Krevelen, D. W. *Trans. Faraday Soc.* **1953**, *49*, 1391–1401.

(77) Aroney, M. J.; Cleaver, G.; Pierens, R. K.; Le Fevre, R. J. W. *J. Chem. Soc., Perkins Trans. II* **1974**, *2*, 3–5.

(78) Vogel, A. I. *J. Chem. Soc.* **1948**, 1825–1833.

(79) Vogel, A. I. *J. Chem. Soc.* **1948**, 1814–1819; cf. p 1815.

The value for P is from comparisons of nonoxygenated tricoordinate phosphorus in V23. That for SI is estimated by assuming that the polarizability of SI relative to P, S, and Cl (4.5, 3.6, 3.15, 2.70) should be about the same as that of CR relative to NR, OR, and F (1.10, 1.05, 0.65, 0.40). These comparisons indicate that *polarizability tends to decrease toward the rare-gas value* (Ne = 0.3955; Ar = 1.6418) *across a given row of the periodic table.*

Other Oxygens and Oxygenated Heteroatoms. For the carbonyl group, it is difficult to disentangle the contributions made by the carbon and oxygen atoms. We have assigned the value 0.75 to O=C (the carbonyl oxygen) in the expectation that this species should be somewhat more polarizable than is its divalent cousin, OR. The value 1.10 for C=O (carbonyl carbon) then follows from the V23T value for the C=O group in ketones. Similarly, the V23T values for the NO (nitroso) and NO_2 groups lead to the listed NO and NO2 polarizabilities when the same value, 0.75, is assumed for ON and O2N. We also take 0.75 for ONX, oxygen in *N*-oxides. The listed polarizability of 2.7 for SO is the V23 value for the S=O group in sulfoxides minus 0.75 for the terminal oxygen (which uses the same numeric atom type as does O=C); the same procedure would give 1.93 for SO2, but the listed value of 2.1 is based on the polarizability of 7.97 for dimethyl sulfone,[80] less that of 4.47 for ethane (Miller), less 1.50 for the terminal oxygens. The value of 1.6 for PO4 is that for the PO_4 group in V23T, corrected for contributions from three divalent and one terminal oxygen assigned the same polarizabilities as OR and O=C, respectively. We observe that the polarizabilities of 3.15, 2.7, and 2.1 for the series S, SO, SO2 decrease with increasing oxygenation at sulfur, as do the values of 3.6 for P and 1.6 for PO4. The values for C=C (1.35) and C=O (1.10) reflect the same trend. Thus we see that *increasing positive charge progressively reduces an atom's polarizability*, a pattern clearly in accord with expectations based on simple physical arguments. We shall make use of this pattern below in assigning polarizabilities to MMFF atom types for which experimental data are lacking.

Heteroatoms Multiply Bonded to Carbon. The value of 3.9 for S=C derives from V23T's value for the C=S group in alkyl xanthates. That for N=C is obtained from compounds containing N=C double bonds by making use of Vogel's polarizabilities for the NOH group in aldoximes and ketoximes, for the NO group in ketoxime *O*-alkyl ethers, and for the N—N group in aliphatic ketazines, $R_2C{=}N{-}N{=}CR_2$.[81] The molecular polarizability for the latter can be compared directly to that for the corresponding ketone, $R_2C{=}O$, to extract the difference in polarizability of N=C and O=C, given that C=N and C=O use the same numeric atom type. These comparisons favor a somewhat higher polarizability of 1.5–1.6 for N=C; we have adopted a lower value of 1.40 in the expectation that the proper value for C=N should lie between those of C=O and C=C, thus reducing the apparent contribution of N=C to the C=N group. Despite the uncertainty which arises on this point, we see that *multiple bonding usually increases the atomic polarizability*, as is shown in the values N=C = 1.40, NR = 1.05, and S=C = 3.90, S = 3.15, again in accord with simple physical arguments. Note that the previously assigned value of 0.75 for O=C relative to that of 0.65 for OR conforms to this trend, as do the values of 1.35 and 1.05 for C=C and CR, respectively. For triply-bonded nitrogen, however, the V23T value of 2.164 for the CN group in nitriles when paired with CSP = 1.30 yields NSP = 0.86. Given that the CSP value is for alkynes and that carbon triply-bonded to nitrogen probably should be assigned a lower polarizability, we will compensate by taking NSP = 1.00.

Heteroatoms in Delocalized and Aromatic Environments. Vogel does not consider amides. Assuming additivity, the polarizabilities Miller cites for formamide, acetamide, and *N*-methylformamide lead to a value of ca. 1.0 for NC=O. We shall assume the same value for NSO2, the nitrogen in sulfonamides, as well as for the similarly delocalized NC=C enamine nitrogen. These values differ little from those assigned for saturated nitrogen, i.e., NR = 1.05, perhaps reflecting a compensation between the positive-charge effect, which decreases polarizability, and the multiple-bonding effect, which increases it. For heteroaromatic systems, values for the atom types NPYD, NPYL, OFUR, and STHI are problematic. Thus, the average polarizability of 9.25 Å^3 Miller cites for pyridine lies far below the average for benzene of 10.38 Å^3 and implies NPYD 1.35 + 0.40 − 1.13 = 0.62 if additivity is assumed. This value lies disturbingly far below the value N=C = 1.40 given above. Though we would be happier with a still larger value, we will use the compromise value NPYD = 1.0; this larger value assumes that pyridine's partially positively charged carbons should have smaller polarizabilities than the present model assumes and thus that nitrogen makes a larger contribution to the overall polarizability than the comparison to benzene suggests. We assign the same value for N5A and N5B in 5-membered aromatic heterocycles. The molecular polarizability of 7.94 Å^3 Miller cites for pyrrole yields NPYL = 0.54 when adjusted for four additive contributions of type C=C or CB and five of type HC. This value, too, seems too small, though some reduction from the value for, say NC=O, can be understood on the basis of the strong aromatic resonance which places positive charge on nitrogen; we will, somewhat arbitrarily, assign the value NPYL = 0.8. Analogous comparisons for furan and thiophene based on polarizabilities cited by Miller yield uncorrecetd polarizabilities OFUR and STHI of 0.23 and 2.00, respectively; on the basis of a similar rationale, we will use the values OFUR = 0.50 and STHI = 2.30. For the *N*-oxide atom types N2OX, N3OX, and NPOX we assign polarizabilities of 0.7, somewhat reduced from those for the parent amines to reflect the effect of electron withdrawal by the oxide oxygen. We acknowledge the arbitrary nature of many of these assignments, but in seeking to model interatomic interactions as accurately as possible it seems best to use appropriate values for the local atomic polarizabilities even if the resulting molecular polarizabilities are less rigorously additive.

Atoms in Formally-Charged Systems. To our knowledge, no experimental data is available which bears directly on the assignments for cationic species we will make in this paragraph. Rather, we shall base these assignments on atomic polarizabilities obtained for similar chemical environments and on the tendency noted above for increasing positive charge to reduce an atom's polarizability. Evidence exists in support of the related proposition that increasing negative charge enhances polarizability. Thus, the experimentally derived static polarizabilities[82] of 0.759, 2.974, 4.130, and 6.199 Å^3 for the halide anions F^-, Cl^-, Br^-, and I^- consistently exceed the values of 0.4, 2.3, 3.3, and 5.5 Å^3 given in Table XIII for F, CL, BR, and I. The difference, proportionally greatest for the smallest halides, is 90% for F and 30% for Cl. Thus, we expect increased and decreased polarizabilities, respectively, for negatively and for positively charged atoms. For the former, we shall take 1.2 for OM, the oxygen in alkoxide anions, representing an increase of 60% over O=C = 0.75. For carboxylate oxygens, which can be regarded as carrying a formal charge of $-1/2$, we take O2CM = 1.0, reflecting an increase of one-third. We also increase the polarizability for the carboxylate carbon CO2M somewhat, to 1.3. In sulfonamide anions, we take NSOM = 1.5 for the formally negative nitrogen, an increase of 50% over NSO2. In isocyanates, where the triply-bonded carbon and nitrogen carry formal charges of −1 and +1, respectively, we assume C%− = 1.8, NR% = 0.8. Similarly, we take NR+ = 0.8 for positively charged quaternary nitrogen, about 75% of the value NR = 1.05 used in neutral amines. For iminium nitrogen, we take N+=C = 1.0, comparably smaller than the value of N=C = 1.4 for neutral imines. Analogously, we assign a polarizability of 0.8 for nitrogens of types NIM+, NCH+, and NGD+, which carry positive charges of $1/2$, $1/2$, and $1/3$ in resonance theory; we

(80) Aroney, M. J.; Fisher, L. R.; LeFevre, R. J. W. *J. Chem. Soc.* **1963**, 4450–4454. Miller (cf. ref 57) cites polarizabilities for methyl sulfide and methyl sulfoxide which appear to be taken from this work but (apparently erroneously) cites a polarizability of 8.40 Å^3 for dimethyl sulfone rather than the value of 7.97 Å^3 reported in this paper.

(81) Vogel, A. I.; Cresswell, W. T.; Jeffery, G. H.; Leicester, J. *J. Chem. Soc.* **1952**, 514–549.

(82) Tessman, J. R.; Kahn, A. H.; Shockley, W. *Phys. Rev.* **1953**, *92*, 890; from index of refraction measurements for alkali halide crystals.

also assign a somewhat reduced polarizability of 1.0 for CIM+ and the equivalently mapped CGD+, the central carbon atoms in these cationic systems. Finally, for O+, O=+, and O%+, formally charged oxygens of hybridization sp^3, sp^2, and sp, respectively, we take 0.4 for the polarizability. To be sure, these polarizabilities cannot be said to be known accurately, but we include them in Table XIII to suggest how vdW parameters might be expected to behave in such chemical environments.

Continuous Symmetry Measures

Hagit Zabrodsky,[‡] Shmuel Peleg,[‡] and David Avnir*

Contribution from the Departments of Computer Science and Organic Chemistry, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel. *Received November 25, 1991*

Abstract: We advance the notion that for many realistic issues involving symmetry in chemistry, it is more natural to analyze symmetry properties in terms of a continuous scale rather than in terms of "yes or no". Justification of that approach is dealt with in some detail using examples such as: symmetry distortions due to vibrations; changes in the "allowedness" of electronic transitions due to deviations from an ideal symmetry; continuous changes in environmental symmetry with reference to crystal and ligand field effects; non-ideal symmetry in concerted reactions; symmetry issues of polymers and large random objects. A versatile, simple tool is developed as a continuous symmetry measure. Its main property is the ability to quantify the distance of a given (distorted molecular) shape from any chosen element of symmetry. The generality of this symmetry measure allows one to compare the symmetry distance of several objects relative to a single symmetry element and to compare the symmetry distance of a single object relative to various symmetry elements. The continuous symmetry approach is presented in detail for the case of cyclic molecules, first in a practical way and then with a rigorous mathematical analysis. The versatility of the approah is then further demonstrated with alkane conformations, with a vibrating ABA water-like molecule, and with a three-dimensional analysis of the symmetry of a [2 + 2] reaction in which the double bonds are not ideally aligned.

1. Continuous Symmetry Measures. Why Are They Needed?

One of the most deeply-rooted paradigms of scientific thought is that Nature is governed in many of its manifestations by strict symmetry laws. The continuing justification of that paradigm lies within the very achievements in human knowledge it has created over the centuries.[1,2] Yet we argue that the treatment of natural phenomena in terms of "either/or", when it comes to a symmetry characteristic property, may become restrictive to the extent that some of the fine details of phenomenological observations and of their theoretical interpretation may be lost. Atkins writes in his widely-used text on physical chemistry: "Some objects are *more symmetrical* than others",[3] signaling that a scale, quantifying this most basic property, may be in order. The view we wish to defend in this report is that symmetry can be and, in many instances, should be treated as a continuous "gray" property, and not necessarily as a "black or white" property which exists or does not exist. Why is such a continuous symmetry measure important? In short, replacing a "yes or no" information processing filter, which acts as a threshold decision-making barrier which differentiates between two states, with a filter allowing a full range of "maybe's", enriches, in principle, the information content available for analysis.

This report contains four sections. In the next section we develop in some detail the notion of the need for a symmetry scale. It is an important part of the report because the very question at hand is not trivial and is certainly not standard or routine, and some readers may need persuasion that efforts to answer this question are worthwhile and may perhaps lead to a useful framework of discussion of symmetry issues in chemistry.[4-7] Yet, we recall at this point that the door to the questions we pose has been at least partially opened. For instance, Murray-Rust et al. have suggested the use of symmetry coordinates to describe nuclear configurations of MX_4 molecules that can be regarded as distorted versions of the T_d symmetrical reference structure.[8,9] More recently, Mezey and Maurani[10,11] extended the point symmetry concept for quasi-symmetric structures by using fuzzy-set theory (terming it "syntopy" and "symmorphy") and provided a detailed demonstration of its application for the case of the water molecule.[12] Also of relevance are proposals for chirality scales (for some different approaches, see, for example, refs 13 and 14).

In Section 3, we offer a tool for the quantitative assessment of symmetry contents which is, we believe, efficient, easy to implement, and general in the sense that it is applicable to a wide and diverse array of symmetry problems as detailed below. All the required principles and practical aspects of this tool are given in Section 3 using cyclic structures as examples. For the interested reader and for sake of completeness, we provide rigorous mathematical proofs in the Appendix. In Section 4 we demonstrate the implementation of our approach on three additional problems: conformations of open-chain *n*-alkanes, the vibrating water-like molecule, and the symmetry of a [2 + 2] concerted reaction. These three examples are but the tip of the iceberg, the outlines of which

[‡] Department of Computer Science.
* Department of Organic Chemistry.

(1) *Symmetry: Unifying Human Understanding*; Hargittai, I., Ed.; Pergamon Press: New York, 1986. *Symmetry 2: Unifying Human Understanding*; Hargittai, I., Ed.; Pergamon Press: Oxford, 1989.
(2) *Symmetries in Science. III*; Gruber, B., Iachello, I., Eds.; Plenum Press: New York, 1988 and earlier volumes in that series.
(3) Atkins, P. W. *Physical Chemistry*, 3rd ed.; Oxford University Press: Oxford, 1986, p 406.
(4) Hargittai, I.; Hargittai, M. *Symmetry Through the Eyes of a Chemist*; VCH: Weinheim, 1986.
(5) Ezra, G. S. *Symmetry Properties of Molecules*; Springer: Berlin, 1982.
(6) For an excellent collection of classical papers, see: *Symmetry in Chemical Theory*; Fackler, J. P., Jr., Ed.; Dowden, Hutchison & Ross: Stroudsburg, PA, 1973.
(7) Mezey, P. G. *J. Am. Chem. Soc.* **1990**, *112*, 3791.
(8) Murray-Rust, P.; Bürgi, H. B.; Dunitz, J. D. *Acta Cryst.* **1978**, *B34*, 1787.
(9) Luef, W.; Keese, R.; Bürgi, H. B. *Helv. Chim. Acta* **1987**, *70*, 534.
(10) Maruani, J.; Mezey, P. G. *C. R. Hebd. Séances Acad. Sci. Paris*, II, **1987**, *305*, 1051 (Erratum: *Ibid.* **1988**, *306*, 1141). Mezey, P. G.; Maruani, *J. Mol. Phys.* **1990**, *69*, 97.
(11) Mezey, P. G. In *New Theoretical Concepts for Understanding Organic Reactions*; Bertrán, J., Csizmadia, I. G., Eds.; Kluwer: Dordrecht, 1989; pp 55, 77.
(12) The problem of symmetry fuzziness is also described in the Introduction of ref 4 (pp 3–4).
(13) Gilat, G. *J. Phys. A* **1989**, *22*, L545.
(14) Hel-Or, Y.; Peleg, S.; Avnir, D. *Langmuir* **1990**, *6*, 1691.

0002-7863/92/1514-7843$03.00/0