Theor Chem Acc (2005) 113: 133–151
DOI 10.1007/s00214-004-0624-x

FEATURE ARTICLE

Casey P. Kelly · Christopher J. Cramer
Donald G. Truhlar

Accurate partial atomic charges for high-energy molecules using class IV charge models with the MIDI! basis set

Received: 16 July 2004 / Accepted: 4 November 2004 / Published online: 27 January 2005
© Springer-Verlag 2005

Abstract We have recently developed a new class IV charge model for calculating partial atomic charges in molecules. The new model, called charge model 3 (CM3), was parameterized for calculations on molecules containing H, Li, C, N, O, F, Si, S, P, Cl, and Br by Hartree–Fock theory and by hybrid density functional theory (HDFT) based on the modified Perdew–Wang density functional with several basis sets. In the present article, we extend CM3 for calculating partial atomic charges by Hartree–Fock theory with the economical but well balanced MIDI! basis set. Then, using a test set of accurate dipole moments for molecules containing nitramine functional groups (which include many high-energy materials), we demonstrate the utility of several parameters designed to improve the charges in molecules containing both N and O atoms. We also show that one of our most recently developed CM3 models that is designed for use with wave functions calculated at the *m*PW*X*PW91/MIDI! level of theory (where *X* denotes a variable percentage of Hartree–Fock exchange) gives accurate charge distributions in nitramines without additional parameters for N and O. To demonstrate the reliability of partial atomic charges calculated with CM3, we use these atomic charges to calculate polarization free energies for several nitramines, including the commonly used explosives 1,3,5-trinitro-*s*-triazine (RDX) and 2,4,6,8,10,12-hexanitrohexaazaisowurtzitane (HNIW), in nitromethane. These polarization energies are large and negative, indicating that electrostatic interactions between the charge distribution of the molecule and the solvent make a large contribution to the free energy of solvation of nitramines. By extension, the same conclusion should apply to solid-state condensation. Also, in contrast to some other charge models, CM3 yields atomic charges that are relatively insensitive to the presence of buried atoms and small conformational changes in the molecule, as well as to the level of treatment of electron correlation. This type of charge model should be useful in the future development of solvation models and force fields designed to estimate intramolecular interactions of nitramines in the condensed phase.

Keywords Basis set MIDI! · Dipole moments · Materials, high-energy · Nitramines, inversion barrier · Partial charges, class IV

1 Introduction

Unlike many other electronic properties that can be extracted from a quantum mechanical wave function, such as electron density, permanent electrical moments, or the molecular electrostatic potential, the partial atomic charge on an atom in a molecule is not a quantum mechanical observable. As a result, the rules for determining partial atomic charges, which involves assigning quantitative values to the amount of electron density belonging to each atom in a molecule, are ambiguous, and many different methods have been developed and evaluated for accomplishing this [1–79]. We have assigned [73] these various methods to four broad categories, or classes. Class I includes all methods for which partial atomic charges can, in principle, be determined without first calculating a quantum mechanical wave function for the given molecule. Examples include assigning partial atomic charges based on X-ray diffraction data [3], IR stretching frequencies [2,4–8], or, for the case of a diatomic molecule, calculating them directly by dividing the dipole moment by the bond length. Some class I charge models that use only the molecular topology of the given molecule to assign its partial atomic charges have been proposed [11–14], while other class I charge models [15–30] assign partial atomic charges based on the concept of electronegativity equalization [80,81]. Class II includes methods that use a population analysis of a molecular wave function to assign charges, and include Mulliken [31–33], Löwdin [34], and natural [47] population analyses, as well as Bader's "atoms in molecules" (AIM) method [49]. Another class II charge model, called

C.P. Kelly
Department of Chemistry and Supercomputing Institute,
207 Pleasant St. SE, University of Minnesota,
Minneapolis, MN 55455-0431, USA.

C.J. Cramer(✉) · D.G. Truhlar
E-mail: cramer@pollux.chem.umn.edu, cramer@chem.umn.edu
E-mail: truhlar@chem.umn.edu

redistributed Löwdin population analysis [35] (RLPA), is designed to alleviate some of the sensitivity to the inclusion of diffuse functions when calculating partial atomic charges from a Löwdin population analysis. Class III charge models include those methods that assign partial atomic charges by a fit to a wave-function-dependent physical observable. Models that assign partial atomic charges from a least-squares fit to the electrostatic potential calculated at a number of points around the molecule of interest, such as ChElP [60], ChElPG [65], and the Merz–Singh–Kollman scheme [60, 66], are examples of class III charge models. Class IV charge models, which include charge model 1 (CM1) [73] and CM2 [74, 75], differ from class III charge models in that they map partial atomic charges obtained from some other charge-assigning scheme in order to reproduce experimental or converged theoretical charge-dependent observables.

One of the major motivations behind the development of class IV charge models by our group [73–79] was to remedy two major deficiencies of class III models. First, the partial atomic charges obtained from class III charge models are necessarily dependent on the quality of the wave function used to calculate the physical observable against which the partial atomic charges are fit. In many cases, such as for large molecules or for large libraries of molecules, this approach is not practical, as it is often the case that high-level wave functions are required in order to deliver relatively accurate physical observables. Second, for many class III charge models, some partial atomic charges in a given molecule are less well determined than others (i.e. large fluctuations in the values of some charges may have only a small effect on the error function against which the atomic charges are fit). As a result of this second deficiency, class III charge models can in some cases deliver partial atomic charges that are less transferable than those calculated with other models and that also depend too strongly on the conformation of the molecule.

Recently, several molecular mechanics force fields that use partial atomic charges as quantum chemical descriptors have been developed in order to model the solid-state properties of various molecules containing nitramine functionality. For example, Sorescu et al. [82] have used partial atomic charges obtained from the ChElPG fitting procedure to investigate the solid-state properties of several commonly used explosives, including 1,3,5-trinitro-*s*-triazine (RDX), 2,4,6,8,10,12-hexanitrohexaazaisowurtzitane (HNIW) [83], and 1,3,5,7-tetranitro-1,3,5,7-tetraazacyclooctane (HMX) [84] in the solid phase, as well as a variety of other molecules containing nitro-group functionality [85, 86]. Smith et al. [87] have developed a similar type of force field for dimethylnitramine (DMNA) in the solid phase; these workers used partial atomic charges that were obtained from a modified electrostatic fitting procedure subject to the additional constraint that the partial atomic charges of equivalent atoms remain equal during the fit. The charge model used by these workers is similar in spirit to Bayly et al.'s [71] restrained electrostatic potential (RESP) charge fitting scheme in which hyperbolic penalty functions are associated with each atom during the electrostatic fitting procedure.

Recently, we developed a new class IV charge model called charge model 3 (CM3) [77]. This new model has several advantages compared to the models developed earlier within our group [73–76]. First, the functional form is less sensitive to nonphysical fluctuations (when one changes the basis set or level of treatment of electron correlation) in calculated bond orders. Second, the model is based on RLPA charges, which [35] are more suitable for use with diffuse basis functions. This is important because the use of diffuse basis functions is often required for accurate calculations of conformational energies and barrier heights [88]. Third, the CM3 model is based on a larger and more diverse training set than used for previous models. Fourth, it includes an additional term that improves the charges on N and O when these atoms are in the same molecule.

We report in this article a new set of parameters that, when combined with wave functions calculated at the HF/MIDI! [89–91] level of theory, yields accurate partial atomic charges for molecules containing H, Li, C, N, O, F, Si, S, P, Cl, and Br, as judged by comparison of point-charge-derived molecular dipole moments to accurate values. We also show that this model, along with a previously developed CM3 model [76] that is based on wave functions calculated at the *m*PW1PW91/MIDI! level of theory, where *m*PW1PW91 is the hybrid density functional of Adamo and Barone [92], and the "1" denotes 25% Hartree–Fock exchange, can both be used to obtain accurate partial atomic charges for nitramines in the gas phase. Specifically, we show that the partial atomic charges for a test set of molecules containing nitramine functionality are invariant to the level of treatment of electron correlation relative to those atomic charges calculated using several other charge models. In addition, we also show that the partial atomic charges obtained with class IV charge models are less conformation dependent and are far more consistent for atoms sharing similar local molecular environments than those partial atomic charges obtained using the ChElPG fitting procedure (a class III model). The models presented in this work, which are designed to use economical wave functions to give reliable partial atomic charges for molecules, should be useful in future force fields designed to investigate the solid-state properties and solubility of high-energy materials.

2 Background

2.1 Charge model 3 (CM3)

In our previous development of the CM2 and CM3 models [74–79], the charge, q_k, on an atom k was defined to be

$$q_k = q_k^0 + \sum_{k' \neq k} T_{kk'}(B_{kk'}) \tag{1}$$

where the summation goes over all atoms in the molecule, q_k^0 is the partial atomic charge from either a Löwdin population analysis (LPA) [34, 44–46] or a redistributed Löwdin population analysis (RLPA) [35] of the wave function, and

$T_{kk'}(B_{kk'})$ is a quadratic function of the Mayer [93–95] bond order, $B_{kk'}$, between two atoms k and k':

$$T_{kk'}(B_{kk'}) = (D_{Z_k Z_{k'}} + C_{Z_k Z_{k'}} B_{kk'})B_{kk'} \tag{2}$$

where Z_k is the atomic number of atom k, and $C_{ZZ'}$ and $D_{ZZ'}$ are the CM3 parameters. Because the total charge remains constant, the charge transferred from k' to k should be equal but opposite to the charge transferred from k to k'. This conservation of charge is maintained by the following relations:

$$C_{ZZ'} = -C_{Z'Z} \tag{3}$$

$$D_{ZZ'} = -D_{Z'Z} \tag{4}$$

In addition to the mappings of the partial atomic charge given above, we further modify the partial atomic charge on N and O atoms according to

$$\begin{aligned} q_{\mathrm{N}} = q_k^0 + \sum_{\substack{k' \neq \mathrm{N}, \\ k' \neq \mathrm{O}}} B_{\mathrm{N}k'}(D_{\mathrm{N}k'} + C_{\mathrm{N}k'} B_{\mathrm{N}k'}) \\ + \sum_{k'=\mathrm{O}} B_{\mathrm{NO}} \left(D_{\mathrm{NO}} + A_{\mathrm{N}k'} \mathrm{e}^{-\left(B_{\mathrm{NO}}/B_{\mathrm{NO}}^0\right)^2} \right) \end{aligned} \tag{5}$$

where $A_{\mathrm{NO}'}$ and B_{NO}^0 are new parameters to be optimized. To maintain a constant total charge, an analogous definition is employed for the CM3 charge on oxygen (with $D_{\mathrm{NO}} = -D_{\mathrm{ON}}$ and $A_{\mathrm{NO}} = -A_{\mathrm{ON}}$) The final term in Eq. (5), which goes to zero at low bond order, to $B_{\mathrm{NO}}^0(\mathrm{A}_{\mathrm{NO}'} + D_{\mathrm{NO}})$ at some characteristic bond order B_{NO}^0, and to $B_{\mathrm{NO}}^0 D_{\mathrm{NO}}$ for very high bond orders, was introduced in an earlier paper [77] in order to improve the performance of CM3 with semiempirical molecular orbital theory for amides. Here, we found that including this term improved the present model for amides, nitrohydrocarbons, and nitramines.

2.2 Basis sets

All partial atomic charges are based on wave functions calculated with the MIDI! basis set [89–91]. This basis set (also called MIDIX or MIDI!5D) has *d* functions on all atoms (for which it is defined) that are heavier than H with the exception of carbon. It is an especially economical basis set for calculations on large organic systems, but it was designed to give particularly accurate geometries and charge distributions at the Hartree–Fock level. For some calculations in the present work, the more complete MG3S basis set was used. This basis is the “modified G3 semidiffuse” basis set [88], which is obtained from the MG3 basis set [96] (which is identical to the G3LargeMP2 [97] basis set) by deleting diffuse functions on H. We note that for all first-row atoms, the MG3S basis set is identical to the older 6-311 + G(2df, 2p) basis [98,99]. Both MIDI! and MG3S are used with spherical harmonic polarization functions. We also augmented the MG3S basis set described above with additional diffuse functions of *s*, *p*, *d*, and *f* symmetry on O atoms and used it to calculate density dipole moments for a test set of molecules containing nitro groups. These diffuse functions are the same as those used for the aug-cc-pVTZ basis set [100]. We will refer to this augmented basis set as MG3S(+O).

Table 1 Density dipole moments (debyes) and relative gas-phase energies (kcal/mol) for different conformations of nitramide and DMNA calculated at the *m*PW1PW91/MG3S level of theory

Conformer[a]	Dipole (D)[b]	$E_{\mathrm{rel}}^{\mathrm{c}}$	No. Imag[d]
Nitramide I (C_s)	3.94	0.0	0
Nitramide I (C_{2v})	4.31	1.2	1
Nitramide II (C_s)	2.97	11.9	1
Nitramide II (C_{2v})	3.28	21.4	2
DMNA I (C_s)	4.81	0.0	0
DMNA I (C_{2v})	5.04	0.2	1
DMNA II (C_s)	3.43	12.7	1
DMNA II (C_{2v})	3.69	22.9	2

[a] See Fig. 1 and text for descriptions of various conformers

[b] *m*PW1PW91/MG3S density dipole moment in debyes

[c] Relative *m*PW1PW91/MG3S gas-phase energy in kcal/mol

[d] Number of imaginary frequencies

2.3 Theoretical dipole moments

Dipole moments obtained from the full electron density are called density dipole moments to distinguish them from dipole moments calculated from partial atomic charges. All density dipole moments used in the training set described below were obtained as one-electron expectation values of the full electron density from single-point *m*PW1PW91/MG3S calculations.

2.4 CM3 training set

The training set used here contains 397 dipole moments for 389 molecules containing H, C, N, O, F, Si, P, S, Cl, and/or Br, plus the two nonpolar molecules benzene and ethylene, which are used for obtaining the C–H parameter as explained below. This training set is a subset of the previously described CM3 training set [76,77], which contains 398 dipole moments for 390 molecules, along with ethylene and benzene. The experimental dipole moment for dimethylnitramine (DMNA) was removed from the previous training set because the gas-phase geometry of DMNA is a largely unresolved issue. In particular, gas-phase electron diffraction data [101], X-ray diffraction data [102], and some previous ab initio calculations that use 6-31G or smaller basis sets [103–109] all point toward a lowest-energy structure that is planar (C_{2v} point group), whereas more recent ab initio calculations that employ more complete basis sets [87,110,111] (including those carried out as part of this work) suggest that the lowest-energy structure of DMNA has a pyramidal geometry. Furthermore, as the value of the dipole moment varies quite significantly with respect to inversion of the amino group (see Table 1), we feel that fitting a single-geometry dipole moment to experiment is hard to justify for DMNA.

2.5 Geometries of training set molecules

For the CM3 parameterization presented in this article, we use the same geometries for all of the molecules as in our previous CM3 parameterizations [76,77]. Namely, for molecules where an experimental dipole moment is used, except for molecules containing Li and P, *m*PW1PW91/MIDI! geometries are used; for molecules containing Li and P, and for benzene and ethylene, *m*PW1PW91/MG3S geometries are used; and for the eight amides included in the training set, HF/MIDI! geometries are used.

2.6 Software

Geometry optimizations, density dipole moments, and class III partial atomic charges were calculated with the *Gaussian* 03 [112] electronic structure package. The MN-GSM-version 3.1 module [113] was used to calculate class IV partial atomic charges at the HF/MIDI! and *m*PW1PW91/MIDI! levels of theory.

3 Parameter optimization

The first step in the parameterization is to determine D_{HC}. We carry out the determination of this parameter in the same manner as before, i.e., we require that the average charge on H in benzene and ethylene be 0.11, a value that has been justified in a previous paper [74]. Also, as before [76,77], we set $C_{\text{HC}} = 0$.

The remaining CM3 parameters are determined by minimizing the error function, χ, with respect to the CM3 parameters. This is defined to be

$$\chi = \sum_{i=1}^{397} (\mu_i(\text{Calc.}) - \mu_i(\text{Target}))^2 \qquad (6)$$

where the summation goes over all polar molecule data in the training set, and μ_i(Target) is an experimental or theoretical dipole moment. The calculated class IV dipole moment is given by

$$\mu = \sqrt{\left(\sum_k q_k x_k\right)^2 + \left(\sum_k q_k y_k\right)^2 + \left(\sum_k q_k z_k\right)^2} \qquad (7)$$

where q_k is the partial atomic charge, and x_k, y_k, and z_k are the Cartesian coordinates of atom k in a barycentric coordinate system, that is, a coordinate system with the center of mass at the origin. (However, since the present article considers only neutral molecules, the dipole moment is independent on the location of the origin). The optimization of $C_{ZZ'}$ and $D_{ZZ'}$ parameters is done through a nonlinear fit using standard NAG Fortran 90 routines [114].

4 Results and discussion

4.1 CM3 for HF/MIDI!

First, we parameterized CM3 for HF/MIDI! in a fashion analogous to that described in an earlier paper for *m*PW1PW91/MIDI! [76]. Like the previous model, the current model contains parameters for bonds involving H, Li, C, N, O, F, Si, S, P, Cl, and Br atoms. For the current model, we used the CM3 training set described above (397 dipole moments) to optimize the same set of $C_{ZZ'}$ and $D_{ZZ'}$ parameters (with three exceptions) as the ones that were nonzero for *m*PW1PW91/MIDI! wave functions, namely D_{HC}, D_{HN}, D_{HO}, D_{HSi}, D_{HP}, D_{HS}, D_{LiC}, D_{LiN}, D_{LiO}, D_{LiF}, D_{LiS}, D_{LiCl}, D_{CN}, D_{CO}, D_{CF}, D_{CSi}, D_{CP}, D_{CS}, D_{CCl}, D_{CBr}, D_{NO}, D_{NP}, D_{OSi}, D_{OP}, D_{OS}, D_{FSi}, D_{FP}, D_{SiCl}, D_{PS}, D_{PCl}, C_{CO}, C_{OSi}, and C_{PS}. The three exceptions (which are explained in more detail below) are as follows. First, for HF/MIDI!, we set the parameter C_{OP}, which is nonzero for *m*PW1PW91/MIDI!, equal to zero. Second, for HF/MIDI!, the parameter C_{CP}, which was set equal to zero for *m*PW1PW91/MIDI!, was optimized as part of this work. Third, for HF/MIDI!, we treat the N–O charge map using Eq. (5) instead of Eq. (1); for this we optimize a D_{NO} parameter and new A_{NO} and B^0_{NO} parameters, as well as the $D_{\text{N}k}$ parameters.

The application of Eq. (6) was carried out in several stages. First, the parameters for molecules containing at most H, C, N, and O were determined with fixed D_{HC}. Then, with the parameters for H, C, N, and O fixed, the parameters for bonds between C–F, C–Cl, and C–Br were optimized. Then, these parameters were fixed, and the parameters involving sulfur were determined. Then, the parameter optimization for Si was carried out in two steps. First, the D_{HSi} and the D_{CSi} parameters were optimized using the subset of the training set that contains at most C, H, and Si. Second, these parameters were held fixed, and the remaining Si parameters (C_{OSi}, D_{OSi}, D_{FSi}, and D_{SiCl}) were optimized using the rest of the Si training set. Finally, with all of these parameters fixed, the parameters for Li and P were optimized. Contrary to several previous CM3 parameterizations [76,77], we found that for HF/MIDI!, inclusion of a C_{OP} parameter had only a negligible effect on the RMS error of the phosphorus-containing molecules in the CM3 training set, but including a C_{CP} parameter led to a slight decrease in the RMS error for these molecules. As a result, for the present model we include a C_{CP} parameter but not a C_{OP} parameter. Table 2 lists the optimized CM3 parameters for HF/MIDI! obtained according to the procedure described above.

Table 3 gives the root-mean-square (RMS) errors in debyes of the dipole moments calculated using the optimized CM3 parameters for HF/MIDI! over various subsets of the CM3 training set. The model performs quite well for the amides, nitrohydrocarbons, and bifunctional compounds containing H, C, N, and O, giving RMS errors of 0.15, 0.23, and 0.23 debyes for these three subsets of the training set, respectively. The RMS error over the entire CM3 training set is 0.26 debyes, and the mean unsigned error is 0.19 debyes.

4.2 Accurate density dipole moments for nitramines

In order to test different charge models against molecules containing nitramine functional groups, we assembled a test

Table 2 CM3 and CM3.1 parameters for HF/MIDI!

	CM3[a]	CM3.1[b]
	$C_{ZZ'}$	
C–O	0.088	0.090
C–P	−0.034	
O–Si	−0.071	
P–S	−0.120	
	$D_{ZZ'}$	
H–C	0.010	0.010
H–N	0.114	0.121
H–O	0.026	0.025
H–Si	0.073	
H–P	−0.017	
H–S	−0.054	
Li–C	−0.193	
Li–N	0.113	
Li–O	−0.067	
Li–F	0.250	
Li–S	−0.001	
Li–Cl	−0.091	
C–N	0.038	0.037
C–O	−0.168	−0.168
C–F	−0.099	
C–Si	0.121	
C–P	0.035	
C–S	0.008	
C–Cl	0.004	
C–Br	0.064	
N–O	−0.073	−0.081
N–P	−0.105	
O–Si	0.099	
O–P	−0.035	
O–S	0.070	
F–Si	0.025	
F–P	0.009	
Si–Cl	−0.063	
P–S	0.238	
P–Cl	0.084	
	A_{NO}	
N–O	0.524	0.328
	B^0_{NO}	
N–O	0.212	0.392

[a] Presented in Sect. 4.1
[b] Presented in Sect. 4.4

Table 3 Root-mean-square (RMS) errors (in debyes) for CM3 using HF/MIDI! broken down by functional group

Type of compound	No.[a]	RMS error
Inorganic compounds	10	0.25
Alcohols, phenol	13	0.14
Ethers	11	0.18
Aldehydes	5	0.10
Ketones	11	0.29
Carboxylic acids	9	0.29
Esters	6	0.12
Other C, H, O compounds	12	0.26
Aliphatic amines, aniline	13	0.35
Aromatic nitrogen heterocycles	11	0.29
Nitriles	12	0.19
Imines	6	0.53
Other C, H, N compounds	14	0.17
Amides and phenylurea	17	0.15
Nitrohydrocarbons	5	0.23
Bifunctional H, C, N, and O compounds	11	0.23
All H, C, N, and O polar compounds	162	0.25
Fluorine-containing compounds	39	0.20
Chlorine-containing compounds	33	0.15
Bromine-containing compounds	14	0.14
Halogenated bifunctional compounds	23	0.27
Thiols	8	0.16
Sulfides, disulfides	9	0.24
Other sulfur-containing compounds	23	0.49
Phosphorus	10	0.18
Multifunctional phosphorus	13	0.15
Compounds with S and P	7	0.15
C, H, and Si	9	0.12
C, H, O, and Si	9	0.22
C, H, Si, and halogen	18	0.23
Lithium compounds	16	0.55
All polar compounds in CM3 training set	397	0.26[b]

[a] Number of data in the training set for this row
[b] Mean unsigned error = 0.19 debyes

set of 14 high-quality theoretical dipole moments for various conformations of the following molecules: nitramide, DMNA, 1,3,3-trinitroazetidine (TNAZ), RDX, and HNIW. We will refer to this test set of dipole moments as the nitramine test set. The molecular structures of all of the nitramine conformers for which accurate dipole moments were calculated are given in Fig. 1. All of the density dipole moments in the nitramine test set were obtained from wave functions calculated at the *m*PW1PW91/MG3S level of theory, which is the same level of theory that we have used previously [76,77] to obtain accurate density dipole moments for use in our CM3 training set. Table 4 lists these accurate (*m*PW1PW91/MG3S) dipole moments for the nitramine conformers. (Subsequent columns of this table will be discussed in Sects. 4.3 and 4.4) All of the dipole moments given in Table 4 were calculated using geometries optimized at the *m*PW1PW91/MG3S level of theory. It should be pointed out that experimental values for the gas-phase dipole moment *are* available for nitramide [115,116] and DMNA [117], but for the reasons pointed out in Sect. 2.4, theoretical dipole moments were used in their stead. For TNAZ, we included in the test set a theoretical dipole for only one conformation (C_s point group). For the remaining four molecules in the nitramine test set, we included theoretical dipole moments for more than one conformation, which are described below.

For nitramide and DMNA, theoretical dipole moments for four different conformations that involve pyramidilization of and/or twisting of the amino group with respect to the nitro group were included in the nitramine test set. Table 1 gives the calculated dipole moments, absolute gas-phase energies, and number of imaginary frequencies for each conformation of nitramide and DMNA (as obtained at the *m*PW1PW91/MG3S level of theory). The trends obtained here, which are summarized below, are the same as those obtained by Smith and coworkers in an earlier study on nitramide and DMNA in the gas phase [87]. Here, the calculated inversion barrier in the gas phase for the amino group is 1.2 and 0.2 kcal/mol for nitramide and DMNA, respectively. For rotation of the nitro group about the N–N bond axis, the calculated barrier energies are much higher (11.9 and 12.7 kcal/mol for nitramide and DMNA, respectively). The

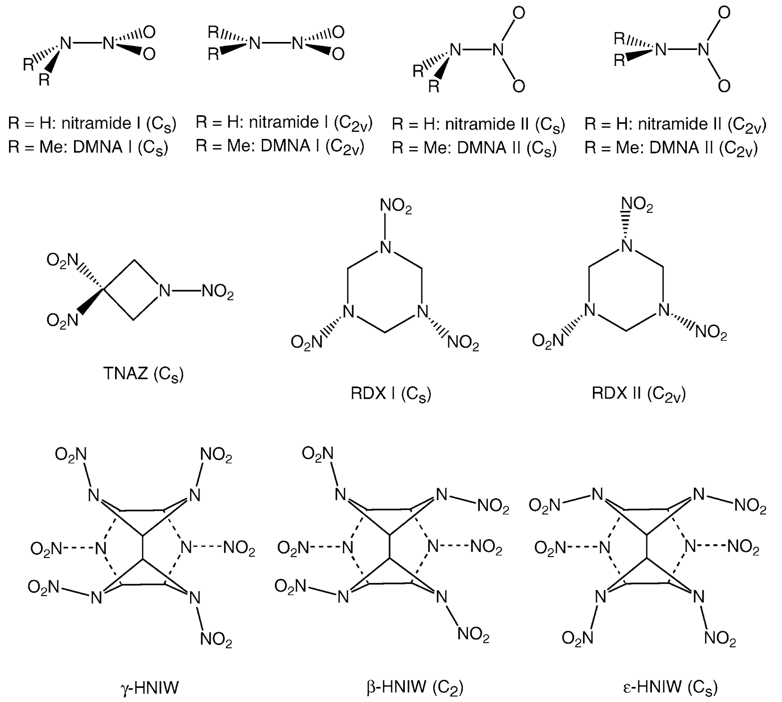

Fig. 1 Molecular structures of conformers in the nitramine test set

Table 4 Values and RMS errors (in debyes) of dipole moments obtained in various ways from HF/MIDI! and *m*PW1PW91/MIDI! wave functions, compared to accurate dipole moments

Molecule[a]	Accurate[b]	HF/MIDI!				*m*PW1PW91/MIDI!		
		Density[c]	Löwdin	CM3	CM3.1	Density[d]	Löwdin	CM3
Nitramide I (C_s)	3.94	3.97	3.71	4.08	3.84	3.59	3.32	3.84
Nitramide I (C_{2v})	4.31	4.31	4.07	4.49	4.22	3.93	3.68	4.19
Nitramide II (C_s)	2.97	3.23	2.94	3.14	2.97	2.71	2.37	2.89
Nitramide II (C_{2v})	3.28	3.61	3.32	3.47	3.28	3.07	2.76	3.27
DMNA I (C_s)	4.81	4.58	4.83	4.95	4.65	4.21	4.36	4.67
DMNA I (C_{2v})	5.04	4.81	5.06	5.18	4.88	4.43	4.59	4.87
DMNA II (C_s)	3.43	3.78	3.75	3.70	3.50	2.99	2.97	3.33
DMNA II (C_{2v})	3.69	4.06	4.16	4.04	3.84	3.38	3.42	3.77
TNAZ (C_s)	0.60	0.41	0.63	0.86	0.67	0.43	0.61	0.66
RDX I (C_s)	5.97	5.97	6.64	6.84	6.29	5.22	5.69	6.20
RDX II (C_{3v})	7.19	7.08	7.97	8.19	7.45	6.22	6.82	7.34
γ-HNIW	1.56	1.50	2.15	2.07	1.83	1.32	1.84	1.80
β-HNIW (C_2)	0.31	0.42	0.86	0.80	0.83	0.42	0.83	0.79
ε-HNIW (C_s)	2.56	2.42	3.16	2.95	2.48	1.95	2.46	2.41
RMS error		0.21	0.42	0.45	0.21[e]	0.49	0.44	0.19[f]

[a] See Fig. 1 and text for descriptions of various conformers

[b] Density dipole moment calculated at the *m*PW1PW91/MG3S level of theory

[c] Density dipole moment calculated at the HF/MIDI!//*m*PW1PW91/MG3S level of theory

[d] Density dipole moment calculated at the *m*PW1PW91/MIDI!//*m*PW1PW91/MG3S level of theory

[e] Mean unsigned error = 0.16 debyes

[f] Mean unsigned error = 0.15 debyes

internal rotational barriers for C_{2v} symmetry (with planar CNCN configuration) are 21.4 and 22.9 kcal/mol for nitramide and DMNA, respectively. These results suggest that the lowest-energy conformations for both nitramide and DMNA in the gas phase are pyramidal (although the energy difference calculated here between the planar and pyramidal conformations for DMNA is relatively small), and correspond to the structures labeled nitramide I (C_s) and DMNA I (C_s) in Fig. 1.

For RDX, theoretical dipole moments for two different conformations that involve the relative position of the nitro groups with respect to the *s*-triazine ring were included in the nitramine test set. The first conformer has two of its three nitro groups in axial positions, and the third nitro group is

Table 5 Experimental gas-phase dipole moments for molecules in the density dipole test set

Molecule	Subclass	Dipole (D)	Ref.
Hydrazine *(gauche)*	Hydrazines	1.75	123
Tetrafluorohydrazine *(gauche)*	Hydrazines	0.257	123
Methyl hydrazine *(inner skew)*	Hydrazines	1.66	124
Methyl hydrazine *(outer skew)*	Hydrazines	1.82	124
Diisopropylidene hydrazine	Hydrazines	1.53	125
1H-pyrazole	N–N azoles	2.20	123
1H-indazole	N–N azoles	1.762	126
1H-1,2,4-triazole	N–N azoles	2.7	123
1H-benzotriazole	N–N azoles	4.34	127
2H-tetrazole	N–N azoles	2.19	128
[1,2,4]-Triazolo-[1,5-a]pyrazine	N–N azoles	4.64	127
Nitrous acid *(cis)*	Nitro compounds	1.423	123
Nitrous acid *(trans)*	Nitro compounds	1.855	123
Nitric acid	Nitro compounds	2.17	123
Nitryl chloride	Nitro compounds	0.53	123
Nitryl fluoride	Nitro compounds	0.466	123
Nitromethane	Nitro compounds	3.46	123
Methyl nitrate	Nitro compounds	3.081	124
Chloronitromethane	Nitro compounds	2.91	129
Trichloronitromethane	Nitro compounds	1.89	129
Trifluoronitromethane	Nitro compounds	1.442	130
Nitroethane	Nitro compounds	3.23	123
Ethyl nitrate *(gauche)*	Nitro compounds	3.23	124
Ethyl nitrate *(trans)*	Nitro compounds	3.39	124
Nitroethylene	Nitro compounds	3.70	124
1-Chloro-1-nitroethane	Nitro compounds	3.33	129
1-Nitropropane	Nitro compounds	3.66	123
2-Nitropropane	Nitro compounds	3.73	123
1-Chloro-1-nitropropane	Nitro compounds	3.52	129
3-Nitropropene *(trans)*	Nitro compounds	4.54	131
Nitrocyclopropane	Nitro compounds	3.95	124
1-Nitrobutane	Nitro compounds	3.61	129
2-Methyl-2-nitropropane	Nitro compounds	3.74	131
Nitrobenzene	Nitro compounds	4.22	123
1-Chloro-2-nitrobenzene	Nitro compounds	4.64	123
1-Chloro-3-nitrobenzene	Nitro compounds	3.73	123
1-Chloro-4-nitrobenzene	Nitro compounds	2.83	123
1-Fluoro-4-nitrobenzene	Nitro compounds	2.87	123
1,3-Dihydroxy-2-nitrobenzene	Nitro compounds	2.331	132

in an equatorial position (C_s point group). The second conformer has all three of its nitro groups in axial positions (C_{3v} point group). The C_s conformation is structurally similar to the α-RDX conformation, which has been characterized by neutron diffraction measurements [118]. The C_{3v} conformation is structurally similar to the geometry proposed for RDX in the gas [119, 120] and liquid [120, 121] phases. At the *m*PW1PW91/MG3S level of theory, the C_s and C_{3v} conformations of RDX are both minima (no imaginary frequencies), with the C_s conformation being lower in energy in the gas phase by 0.4 kcal/mol.

For HNIW, theoretical dipole moments for three different conformations that involve the relative orientation of the nitro groups with respect to the hexaazoisowurtzitane cage were included in the nitramine test set. The geometries of the three conformations used here are structurally similar to the γ-, β-, and ε-HNIW conformations, which have all been resolved by X-ray diffraction [122]. At the *m*PW1PW91/MG3S level of theory, all three of the HNIW conformations are minima. The relative energy ordering in the gas phase is β (0.0 kcal/mol) $<$ γ (+1.3 kcal/mol) $<$ ε (+1.4 kcal/mol).

In order to evaluate the level of theory (*m*PW1PW91/MG3) used above for calculating accurate dipole moments, we assembled a second test set containing 39 compounds for which gas-phase dipole moments have been measured experimentally [123–132]. All of the experimental dipole moments contained in this test set (which we will refer to as the density dipole test set) are for molecules that contain either one or more N–N bonds or one or more nitro groups (five hydrazines, six azoles that have one or more N–N bonds, and 28 nitro compounds). Table 5 gives the experimental values for the gas-phase dipole moments for all 39 compounds in the density dipole test set.

Table 6 shows the RMS errors between experimental dipole moments and density dipole moments calculated at the *m*PW1PW91/MG3S level of theory for the molecules in the density dipole test set. Also shown are RMS errors between density dipole moments calculated at several other

Table 6 Root-mean-square errors (in debyes) between experimental gas-phase dipole moments and density dipole moments calculated at different levels of theory

Method[a]	Hydrazines	N–N azoles	Nitro compounds	All compounds	WRMS error[b]
*m*PW1PW91/MIDI!	0.35	0.36	0.45	0.43	0.39
*m*PW1PW91/6-31G(d)	0.26	0.23	0.24	0.24	0.25
*m*PW1PW91/6-31+G(d)	0.25	0.20	0.42	0.37	0.29
*m*PW1PW91/6-31+G(d,p)	0.21	0.20	0.41	0.37	0.27
*m*PW1PW91/6-311+G(d,p)	0.17	0.20	0.38	0.34	0.25
*m*PW1PW91/6-311++G(d,p)	0.19	0.20	0.38	0.34	0.26
*m*PW1PW91/MG3S[c]	0.11[d]	0.23[d]	0.29	0.26	0.21
*m*PW1PW91/MG3S(+O)[e]	0.11[d]	0.23[d]	0.26	0.24	0.20
*m*PWB95/MG3S(+O)[e]	0.12	0.29	0.26	0.26	0.23
B3LYP/MG3S(+O)[e]	0.14	0.22	0.30	0.27	0.22
B3PW91/MG3S(+O)[e]	0.10	0.23	0.25	0.23	0.19
B97-1/MG3S(+O)[e]	0.15	0.23	0.28	0.26	0.22
B97-2/MG3S(+O)[e]	0.10	0.24	0.24	0.22	0.19
B98/MG3S(+O)[e]	0.18	0.22	0.29	0.27	0.23
*m*PWPW91/MG3S(+O)[e]	0.12	0.28	0.26	0.25	0.22
BLYP/MG3S(+O)[e]	0.10	0.26	0.32	0.29	0.23
HCTH/MG3S(+O)[e]	0.18	0.30	0.22	0.23	0.23

[a] Density dipole moments calculated using wave function and optimized geometry for the level of theory in this row

[b] Weighted root-mean square (WRMS) error; calculated as the average value of the RMS error for each of the three subclasses of molecules in the density dipole test set

[c] This is the method used to calculate accurate density dipole moments for the nitramines

[d] The MG3S and MG3S(+O) basis sets are equivalent for molecules that do not contain O

[e] MG3S(+O) basis set used, where the (+O) denotes that the MG3S basis set has been augmented with diffuse functions for O taken from the aug-cc-PVTZ basis set

levels of theory. For each of the levels of theory tested here, we computed the RMS error over the three subclasses of compounds in the density dipole test set (hydrazines, N–N azoles, and nitro compounds), and the RMS error over all 39 dipole moments in the density dipole test set. Because the majority of the dipole moments in the density dipole test set are for nitro compounds (28 of 39), we also calculated a weighted-root-mean-square (WRMS) error, which we define as the average of the RMS errors for each of the three types of compounds in the density dipole test set. For each method tested, the geometries of the molecules are optimized for the same level of theory used to compute the density dipole moments (e.g. the row labeled *m*PW1PW91/MIDI! in Table 6 gives the RMS and WRMS errors for density dipole moments calculated using *m*PW1PW91/MIDI! wave functions with geometries optimized at the *m*PW1PW91/MIDI! level of theory). Table 6 also gives the RMS and WRMS errors for those dipole moments obtained with the large MG3S(+O) basis set (where the +O denotes aug-cc-PVTZ diffuse functions for O), and it gives results for a variety of pure density functionals, namely *m*PWB95 [92,133] and BLYP [134, 135] and hybrid density functionals, namely B3LYP [136], B3PW91 [137], B97-1 [138], HCTH [136] (this functional is also called HCTH/407), B97-2 [139], and B98 [140]. It is important to point out that none of the five hydrazines or 6 N–N azoles in the density dipole moment test set contain oxygen, so the MG3S and MG3S(+O) basis sets are equivalent for the molecules in these subclasses.

Almost all of the methods tested here give WRMS errors of less than 0.25 debyes. Both the B3PW91 and B97-2 functionals give WRMS errors of 0.19 debyes when used with the MG3S(+O) basis set, the best of any of the methods tested here. For *m*PW1PW91, the WRMS error is 0.20 debyes when it is used with the MG3S(+O) basis set, although this error increases by only 0.01 debyes when the MG3S basis is used instead. Because most of the methods, when used with the large MG3S(+O) basis set, yield similar WRMS errors, because little improvement is gained by the *m*PW1PW91 functional when the MG3S(+O) instead of the MG3S basis is used, and because the molecules in the density dipole test set share structural similarities with the nitramines (N–N bonds and nitro groups), we feel quite confident that the accurate dipole moments reported here for the nitramines, which were calculated at the *m*PW1PW91/MG3S level, are reliable. We will use these dipole moments as the accurate standard in the rest of this article.

4.3 Performance of CM3 for nitramines

Next, we tested the ability of CM3 to reproduce the accurate (*m*PW1PW91/MG3S level of theory) density dipole moments in the nitramine test set. Shown in Table 4 are the RMS errors between the dipole moments calculated using CM3 charges (obtained from both HF/MIDI! and *m*PW1PW91/MIDI! wave functions) and those calculated using Löwdin charges [34,45] (obtained from both HF/MIDI! and *m*PW1PW91/MIDI! wave functions), compared to the accurate dipole moments for all of the conformers in the nitramine test set. Also included in Table 4 are the RMS errors between the accurate dipole moments and the density dipole moments obtained from HF/MIDI! and *m*PW1PW91/MIDI!

wave functions, as well as the RMS errors between the accurate dipole moments and those dipole moments calculated using partial atomic charges from another model (CM3.1) to be presented in Sect. 4.4.

Table 4 shows that the dipole moments calculated using CM3 HF/MIDI! charges yield an RMS error of 0.45 debyes over the nitramine test set, which is slightly higher than the RMS error when Löwdin charges (0.42 debyes) are used, and 0.24 debyes higher than the RMS given by the HF/MIDI! density dipole moments. The charges obtained from the previously developed CM3 *m*PW1PW91/MIDI! model yield the most accurate dipole moments of any of the methods tested here (RMS error = 0.19 debyes), performing much better than the CM3 HF/MIDI! model developed in this paper.

Worth noting here is the quality of the density dipole moments obtained from HF/MIDI! wave functions. The RMS error between the accurate density dipole moments and those obtained with HF/MIDI! wave functions is only 0.21 debyes, whereas this RMS error increases to 0.45 debyes when *m*PW1PW91/MIDI! wave functions are used. The good performance of HF/MIDI! compared to *m*PW1PW91/MIDI! is most likely due to the MIDI! basis set [89–91], which is designed to give particularly accurate charge distributions at the Hartree–Fock level of theory.

4.4 CM3.1 for HF/MIDI!

To see if we could improve the performance of the CM3 HF/MIDI! model for the nitramines, we reoptimized the parameters A_{NO}, B^0_{NO}, C_{CO}, D_{HN}, D_{HO}, D_{CN}, D_{CO}, and D_{NO} (with D_{HC} fixed) against the dipole moments for all molecules in the CM3 training set containing at most H, C, N, and/or O (162 dipole moments), plus the 14 accurate dipole moments in the nitramine test set. We will refer to this training set (which contains 176 dipole moments) as the CM3.1 training set. The parameters obtained using the CM3.1 training set described above are listed in Table 2 alongside the CM3 parameters for HF/MIDI! obtained in Sect. 4.1. Shown in Table 7 are the RMS errors over various subsets of the CM3.1 training set. Using charges obtained with the new CM3.1 parameters (with HF/MIDI! wave functions) to calculate the dipole moments for the nitramines decreases the RMS error by 0.24 debyes compared to when the charges are obtained by using the CM3 parameters obtained earlier (with HF/MIDI! wave functions), and this results in an RMS error of only 0.21 debyes, as shown in Table 4. Comparing the errors in Table 3 against those in Table 7, it can also be seen that the RMS errors given by the CM3 and CM3.1 parameters over the remaining subclasses of compounds in those tables are all within 0.06 debyes of one another. Furthermore, the parameters in Table 2 show that the only significant differences between these two sets of parameters are the values of A_{NO} and B^0_{NO}, which appear in Eq. (5), and that the remaining parameters are all within 0.008 of one another. This result demonstrates the utility of Eq. (5), which was developed to allow for a wider range of functional groups containing N and O atoms to be fit without having a significant impact on the quality of the model for other functional groups. Because reoptimizing the parameters A_{NO} and B^0_{NO} significantly improves the performance of the present model for the nitramines, and because reoptimizing the parameters has little effect on the remainder of the training set, for wave functions obtained at the HF/MIDI! level of theory we will continue the rest of this article using the new CM3.1 parameters for molecules containing C, H, N, and/or O atoms.

Table 7 Root-mean-square errors (in debyes) for CM3.1 using HF/MIDI! broken down by functional group

Type of compound	No.[a]	RMS error
Inorganic compounds	6	0.31
Alcohols, phenol	13	0.14
Ethers	11	0.18
Aldehydes	5	0.13
Ketones	11	0.27
Carboxylic acids	9	0.31
Esters	6	0.13
Other C, H, O compounds	12	0.26
Aliphatic amines, aniline	13	0.36
Aromatic nitrogen heterocycles	11	0.28
Nitriles	12	0.20
Imines	6	0.53
Other C, H, N compounds	14	0.18
Amides and phenylurea	17	0.11
Nitrohydrocarbons	5	0.17
Non-nitramine bifunctionals	11	0.21
Nitramines	14	0.21
All polar compounds in CM3.1 training set	176	0.24[b]

[a] Number of data in the training set for this row
[b] Mean unsigned error = 0.19 debyes

The previously developed CM3 *m*PW1PW91/MIDI! model [76] does not use Eq. (5) for N and O, so we also tried improving this model by using Eq. (5) instead of Eq. (1) to map the atomic charges on N and O atoms and then reoptimizing the same parameters that were reoptimized above for HF/MIDI!. When we did this, though, we found that using the new parameters instead of the original CM3 parameters [76] increased the RMS error for the nitramines slightly (by 0.02 debyes), and that no significant improvement was gained by any of the remaining subclasses in the CM3.1 training set (the RMS errors all changed by less than 0.05 debyes). Therefore, and because we prefer to use the smoothest mapping functions that give good results, we will continue the rest of this article using the original CM3 parameters for *m*PW1PW91/MIDI! (which give an RMS error of 0.19 debyes for the nitramines).

4.5 Partial atomic charges for HNIW conformers

Shown in Table 8 are the values of the partial atomic charges for γ-HNIW obtained using CM3 (parameterized for use with *m*PW1PW91/MIDI! wave functions), CM3.1 (parameterized for use with HF/MIDI! wave functions), and the ChElPG fitting procedure. The ChElPG fitting procedure is a class III charge model that assigns partial atomic charges based on

Table 8 Partial atomic charges for γ-HNIW

Atom	CM3[a]	CM3.1[b]	ChElPG		
			*m*PW1PW91/MIDI!	*m*PW1PW91/MG3S	HF/6-31G(d,p)[c]
H(1)	0.148	0.159	0.111	0.136	0.112
H(2)	0.149	0.161	0.104	0.151	0.118
H(3)	0.143	0.151	0.173	0.202	0.183
H(4)	0.144	0.153	0.149	0.187	0.201
H(5)	0.148	0.160	0.085	0.106	0.081
H(6)	0.147	0.158	0.106	0.119	0.095
Average	0.147	0.157	0.121	0.150	0.132
SD[d]	0.002	0.004	0.030	0.035	0.045
C(1)	0.197	0.221	0.190	0.162	0.257
C(2)	0.209	0.234	0.141	−0.035	0.183
C(3)	0.220	0.249	−0.074	−0.223	0.020
C(4)	0.215	0.243	0.123	0.000	0.089
C(5)	0.206	0.232	0.441	0.441	0.548
C(6)	0.202	0.227	0.270	0.287	0.368
Average	0.208	0.234	0.182	0.105	0.244
SD	0.008	0.009	0.156	0.219	0.176
N(1)	−0.319	−0.313	−0.238	−0.096	−0.197
N(3)	−0.293	−0.286	−0.240	−0.074	−0.358
N(5)	−0.297	−0.295	−0.276	−0.133	−0.327
N(7)	−0.303	−0.295	−0.390	−0.302	−0.392
N(9)	−0.295	−0.286	−0.515	−0.471	−0.578
N(11)	−0.308	−0.301	−0.330	−0.258	−0.235
Average	−0.303	−0.296	−0.331	−0.222	−0.348
SD	0.009	0.009	0.098	0.139	0.123
N(2)	0.414	0.250	0.643	0.638	0.681
N(4)	0.423	0.266	0.692	0.679	0.787
N(6)	0.414	0.258	0.656	0.634	0.755
N(8)	0.417	0.254	0.703	0.704	0.754
N(10)	0.420	0.258	0.760	0.752	0.844
N(12)	0.413	0.251	0.683	0.689	0.697
Average	0.417	0.256	0.690	0.683	0.753
SD	0.004	0.005	0.037	0.040	0.054
O(1)	−0.229	−0.170	−0.316	−0.344	−0.363
O(2)	−0.225	−0.164	−0.320	−0.350	−0.369
O(3)	−0.235	−0.179	−0.336	−0.363	−0.395
O(4)	−0.239	−0.183	−0.334	−0.362	−0.378
O(5)	−0.248	−0.187	−0.309	−0.322	−0.358
O(6)	−0.253	−0.196	−0.351	−0.381	−0.431
O(7)	−0.230	−0.172	−0.331	−0.363	−0.396
O(8)	−0.228	−0.168	−0.328	−0.355	−0.372
O(9)	−0.230	−0.171	−0.335	−0.358	−0.402
O(10)	−0.231	−0.172	−0.344	−0.369	−0.417
O(11)	−0.233	−0.174	−0.330	−0.357	−0.382
O(12)	−0.234	−0.175	−0.334	−0.368	−0.425
Average	−0.235	−0.176	−0.331	−0.358	−0.391
S.D.	0.002	0.003	0.005	0.006	0.019

[a] Obtained using *m*PW1PW91/MIDI! wave function
[b] Obtained using HF/MIDI! wave function
[c] Charges taken from ref. [83]
[d] Standard deviation

a statistical fit to the electrostatic potential calculated at a number of grid points located around the molecule of interest. For all ChElPG calculations in this work, we used the **pop = ChElPG** option in Gaussian 03. Here, we have calculated ChElPG partial atomic charges for γ-HNIW based on *m*PW1PW91/MIDI! and *m*PW1PW91/MG3S wave functions. Also shown in Table 8 are those partial atomic charges reported by Sorescu et al. [83] for γ-HNIW. These workers also used the ChElPG fitting procedure to assign partial atomic charges, for which they used an electrostatic potential calculated at the HF/6-31G(d) level of theory. The partial atomic charges in Table 8 are organized into five different groups, or atom types, which are as follows: H atoms, C atoms, N atoms (amino), N atoms (nitro), and O atoms. The atomic labels in Table 8 correspond to those shown in Fig. 2, and also to those given in ref. [83].

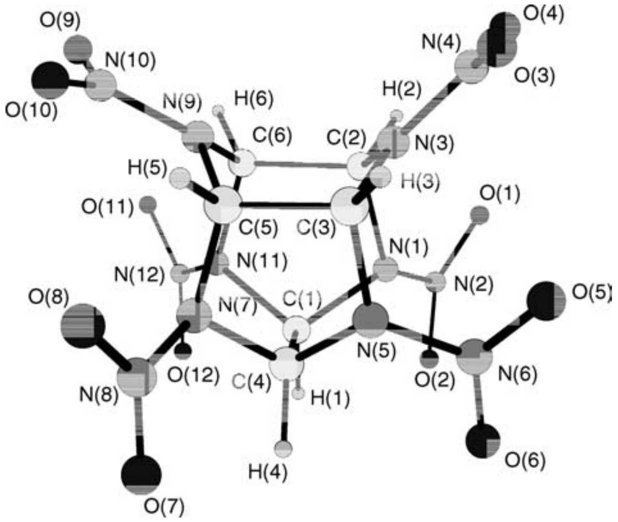

Fig. 2 γ-HNIW; atom labels are consistent with Tables 8 and 9

The γ-HNIW conformer belongs to the low-symmetry C_1 point group, although it is composed of six identically bridged nitramine groups (the C_{2v} symmetry of the hexaazoisowurtzitane cage is broken by the orientation of the nitro groups relative to one another). As a result, many of the atoms in γ-HNIW have almost identical local chemical environments, and so for atoms of a given type, their partial atomic charges should be roughly equal to one another. For example, the partial atomic charges of all of the C atoms in γ-HNIW should be nearly identical due to the similar local chemical environments of all six of these atoms (each of the six C atoms in γ-HNIW is bonded to two N atoms, another C atom, and a H atom). From the data in Table 8, we see that for the CM3 and CM3.1 models, this holds true – the standard deviations of the partial atomic charges within each group of atoms obtained using these two charge models are less than 0.01. In contrast, the standard deviation of the partial atomic charges obtained using the ChElPG fitting procedure on a *m*PW1PW91/MIDI! wave function (ChElPG/ *m*PW1PW91/MIDI! charges; this notation will be used throughout the remainder of this article) for the six C atoms is 0.156 (average = 0.182). When the more complete MG3S basis set is used to fit ChElPG partial atomic charges, the standard deviation from the average for the six C atoms increases to 0.219 (average = 0.105). In this case, the values of the partial atomic charges range from −0.223 for C(3) to 0.441 for C(5). The same types of fluctuations occur for the six amino N atoms, for which ChElPG/ *m*PW1PW91/MIDI! partial atomic charges yield a standard deviation of 0.098 (average = −0.331). Again, this error worsens when the MG3S basis set is used instead of the MIDI! basis set – the standard deviation of the ChElPG/*m*PW1PW91/MG3S partial atomic charges over the six amino N atoms increases to 0.139 (average = −0.222), and the values of these partial atomic charges range from −0.074 for N(1) to −0.471 for N(9). From Table 8, we also see that those partial atomic charges reported by Sorescu et al. show the same trends as the partial atomic charges obtained here with the ChElPG fitting procedure.

The results above provide a dramatic illustration of one of the problems with many class III charge models; namely, that they often deliver unphysical partial atomic charges because of the statistical nature of the fit that they are determined according to. This problem is often quite severe for interior, or "buried", atoms in a molecule [see, for example, refs. 63 and 71], because for these atoms the surface points at which the electrostatic potential is evaluated (which lie outside the van der Waals surface of the molecule) are relatively far away from the interior atoms in a molecule. Indeed, the largest errors for γ-HNIW occur for the six C atoms and the six amino N atoms that are located in the interior region of the molecule. In contrast, the standard deviation of the ChElPG/*m*PW1PW91/MG3S partial atomic charges for the six nitro N atoms decreases to 0.040 (average = 0.683), and for the six H atoms it further decreases to 0.035 (average = 0.150). For the 12 O atoms, which are on the van der Waals surface of γ-HNIW, the standard deviation of the ChElPG/*m*PW1PW91/MG3S partial atomic charges is only 0.006 (average = −0.358).

Although none of the atoms in γ-HNIW are formally symmetrically equivalent to one another, we have argued above that the molecular connectivity and geometry of γ-HNIW yields five distinct atom types, and that the partial atomic charges for different atoms of the same type should be roughly equal to one another. Unlike γ-HNIW, the remaining two HNIW conformers in the nitramine test set do contain atoms that are equivalent to one another by symmetry (the β- and ε-HNIW conformers belong to the C_2 and C_s symmetry point groups, respectively). Shown in Table 9 are the partial atomic charges for each symmetrically unique C and amino N atom in β- and ε-HNIW. In all cases, the partial atomic charge for each pair of symmetrically equivalent atoms was taken as the average partial atomic charge over the pair of atoms, although it is important to point out that the partial atomic charges obtained using CM3, CM3.1, and the ChElPG fitting procedure for each atom within a given pair of symmetrically equivalent atoms differed by no more than 0.004 charge units from one another. From the partial atomic charges shown in Table 9, we see that the ChElPG fitting procedure makes the same types of errors for the interior atoms in these conformers as for the γ-HNIW conformer, although the presence of symmetrically equivalent atoms does alleviate this error somewhat. Even with this improvement though, the fluctuations in the values for the partial atomic charges of the C and amino N atoms are significantly larger for ChElPG than they are for the CM3 and CM3.1 models.

Next, we compare the partial atomic charges of the three HNIW conformers to one another. Again, because these conformers differ only by the orientation of the nitro groups relative to one another, the values for a given partial atomic charge should be roughly equivalent from one conformation to another. (On the other hand, the total dipole moments for these conformers are different, because unlike the value of the partial atomic charge, which should depend only on the

Table 9 Selected partial atomic charges for β-HNIW (C_2) and ε-HNIW (C_s)

Conformer[a]	Atom(s)	CM3[b]	ChElPG[c]
β-HNIW (C_2)	C(1,4)	0.207	0.027
	C(2,5)	0.210	0.156
	C(3,6)	0.216	−0.080
ε-HNIW (C_s)	C(1)	0.231	−0.054
	C(2,6)	0.205	0.071
	C(3,5)	0.224	0.081
	C(4)	0.187	0.161
Average		0.211	0.052
SD[d]		0.014	0.094
β-HNIW (C_2)	N(1,7)	−0.315	−0.100
	N(3,9)	−0.297	−0.232
	N(5,11)	−0.297	−0.207
ε-HNIW (C_s)	N(1,11)	−0.320	−0.102
	N(3,9)	−0.296	−0.224
	N(5,7)	−0.287	−0.203
Average		−0.302	−0.178
SD		0.013	0.061

[a] See Fig. 1 and text for descriptions of various conformers
[b] Partial charges obtained using *m*PW1PW91/MIDI! wave functions
[c] Partial charges obtained using *m*PW1PW91/MG3S wave functions
[d] Standard deviation

Table 10 Density and ChElPG point-charge-derived dipole moments obtained from *m*PW1PW91/MG3S wave functions

Molecule	Density	ChElPG
Nitramide I (C_s)	3.94	3.91
Nitramide I (C_{2v})	4.31	4.27
Nitramide II (C_s)	2.97	2.96
Nitramide II (C_{2v})	3.28	3.30
DMNA I (C_s)	4.81	4.79
DMNA I (C_{2v})	5.04	5.03
DMNA II (C_s)	3.43	3.49
DMNA II (C_{2v})	3.69	3.74
TNAZ (C_s)	0.60	0.53
RDX I (C_s)	5.97	5.98
RDX II (C_{3v})	7.19	7.10
γ-HNIW	1.56	1.55
β-HNIW (C_2)	0.31	0.37
ε-HNIW (C_s)	2.56	2.50

local chemical environment of the atom, the value of the dipole moment depends on both the values and positions of the partial atomic charges in a molecule.) For γ-, β-, and ε-HNIW, the average partial atomic charge on the C atoms obtained using CM3 are 0.208, 0.211, and 0.212. For γ-, β-, and ε-HNIW, the average partial atomic charge on the amino N atoms obtained using CM3 are −0.303, −0.303, and −0.301, respectively. In contrast, those partial atomic charges obtained from the ChElPG fitting procedure are somewhat conformation dependent; for γ-, β-, and ε-HNIW, the average partial atomic charge on C is 0.105, 0.034, and 0.065, respectively, and for the amino N atoms, these values are −0.222, −0.180, and −0.176, respectively.

The results above suggest that the partial atomic charges obtained using CM3 and CM3.1 should be quite useful as quantum chemical descriptors in force fields, because of their invariance to small conformational changes in the molecule. Unlike those partial atomic charges obtained using the ChElPG fitting procedure, CM3 and CM3.1 yield consistent partial atomic charges for atoms sharing similar local chemical environments. This is important because, ideally, one would like the partial atomic charges for two atoms sharing the same or nearly the same local chemical environment to be roughly equal to one another, regardless of nonphysical fluctuations such as the size of the molecule (i.e. whether or not the atom is located in the interior or exterior of the molecule), the level of treatment of electron correlation, or changes in the basis set.

4.6 Dipole moments obtained from the ChElPG fitting procedure

It is useful to reiterate here that the value of the partial atomic charge on an atom in a molecule is not a proper quantum mechanical observable. As a result, one is unable to formally claim that a particular choice of method delivers an “accurate” partial atomic charge. Because of the weaknesses of the ChElPG fitting procedure outlined above, one might assume (incorrectly) that the partial atomic charges obtained from CM3 and CM3.1 should lead to more accurate dipole moments than those obtained from the ChElPG fitting procedure. Indeed, it was shown earlier (Table 4) that CM3 and CM3.1 point-charge-derived dipole moments are significantly more accurate than those dipole moments calculated using partial atomic charges obtained from a Löwdin population analysis. Unfortunately, we cannot gain much insight into the relative quality of those partial atomic charges obtained from the ChElPG fitting procedure, CM3, and CM3.1 based on the accuracy of point-charge-derived dipole moments. This is because for the molecules in the nitramine test set, all of the ChElPG point-charge-derived dipole moments are quite close in value to the corresponding density dipole moment (assuming the density dipole moment is calculated using the same wave function as that used to calculate the electrostatic potential around the molecule); in particular, Table 10 shows that the ChElPG/ *m*PW1PW91/MG3S point-charge-derived moments and the density dipole moments calculated using *m*PW1PW91/MG3S wave functions for each of the molecules in the nitramine test set are all within 0.1 debye of one another. (Although not shown in Table 10, similar results were obtained for ChElPG partial atomic charges obtained from wave functions calculated at several other levels of theory.)

Although the above analysis does not lead to any additional insight into the relative quality of CM3, CM3.1, and the ChElPG fitting procedure because all of these models are able to deliver partial atomic charges that reproduce accurate density dipole moments, the data in Table 10 do illustrate one of the important challenges in the development of reliable class IV charge models; namely, that quite different values for the partial atomic charges in a given polyatomic molecule may lead to identical or very similar values for the molecular dipole moment. It is because of this that for CM3 and CM3.1, as well as for our previous class IV charge models [73–75], we begin the parameterization procedure by first independently determining the bond dipole moment of C–H type bonds (which are present in the majority of the molecules in our training sets), and then continue by optimizing the remaining parameters with the C–H parameters anchored. Furthermore, this is also why we prefer to optimize the remaining parameters in several steps over specific subsets of our data set instead of simultaneously over the entire data set, and why we include only those quadratic $C_{ZZ'}$ parameters (which are more sensitive to nonphysical fluctuations in Mayer bond orders) that are shown to lead to a significant improvement in the accuracy of point-charge-derived dipole moments.

4.7 Polarization contributions to the free energy of solvation

Recently, an SM5 model that uses CM3 charges was introduced [141], and we anticipate that an important application of the CM3 and CM3.1 models will be in estimating intermolecular interactions of nitramines in the liquid and solid phases. A widely used theory for the polarization energy upon solvation is the generalized Born approximation, in which a solvent is represented by a continuum medium characterized by a dielectric constant. The generalized Born polarization energy G_{P} is given as follows [142–145]

$$G_P = -\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_k \sum_{k'} q_k q_{k'} \gamma_{kk'} \tag{8}$$

where ε is the dielectric constant of the solvent, q_k is the partial charge on atom k, and $\gamma_{kk'}$ is a Coulomb integral between atoms k and k'. A useful test for the quality of a charge model is to see whether the atomic charges given by the model lead to reasonable polarization energies when substituted into the above equation. Solvation polarization energy is not a measurable quantity because free energy is not separable into electrostatic and nonelectrostatic contributions. Therefore, we cannot make a direct comparison between a calculated polarization energy and some physical observable. Instead, we will investigate the sensitivity of calculated polarization energies to factors such as level of treatment of electron correlation and choice of method for assigning partial atomic charges. This is important because, ideally, we would like the polarization contribution to the free energy to be relatively insensitive to these variables, at least when electron correlation does not cause a qualitative change in the charge distribution.

Table 11 lists the polarization energies in nitromethane ($\varepsilon = 36.562$ [146]) calculated using gas-phase atomic charges (i.e. these are not self-consistent-reaction-field calculations; there is no solute polarization) obtained from several different charge models. In all cases, the Coulomb integrals are calculated as in SM5.43R [141], which employs the dielectric descreening algorithm of Still et al. [145], but with different parameters. For the nitramines, the RMS difference between polarization energies calculated using CM3 charges (*m*PW1PW91/MIDI! wave functions) and those calculated using CM3.1 charges (HF/MIDI! wave functions) is only 0.26 kcal/mol. Such a good agreement between the polarization energies calculated using these two charge models is quite encouraging, and it further demonstrates the consistency of atomic charges obtained with these models between various levels of theory. Also shown in Table 11 are those polarization energies calculated using Löwdin and ChElPG gas-phase partial atomic charges. For these two models, the difference between the polarization energies obtained using HF/MIDI! and *m*PW1PW91/MIDI! wave-function-based partial atomic charges is much larger (RMS differences of 3.36 and 3.26 kcal/mol for Löwdin and ChElPG charges, respectively). When ChElPG/*m*PW1PW91/MG3S and ChElPG/HF/MG3S partial atomic charges are used to calculate polarization energies, this difference is somewhat smaller (RMS difference = 2.47 kcal/mol), although it is still significantly larger than the difference between those polarization energies calculated with CM3 and CM3.1 gas-phase partial atomic charges.

Next, we compare the polarization energies calculated using CM3 and CM3.1 charges (which were shown above to give roughly equivalent values) to those calculated with ChElPG charges. The polarization energies calculated using CM3/CM3.1 charges agree fairly well with those calculated using ChElPG/*m*PW1PW91/MIDI! charges. The largest disagreement between any of the polarization energies obtained using these three sets of gas-phase charges is for RDX II, for which the use of CM3 and ChElPG/*m*PW1PW91/MIDI! charges leads to a difference of 1.92 kcal/mol (the difference between CM3.1 values and the ChElPG/*m*PW1PW91/MIDI! values is 2.25 kcal/mol). At the HF/MIDI!, HF/MG3S, and *m*PW1PW91/MG3S levels of theory, ChElPG charges yield polarization energies that are significantly more negative than those polarization energies calculated using CM3/CM3.1 charges. The largest differences are for the three HNIW conformers, for which using either ChElPG/HF/MIDI! or ChElPG/*m*PW1PW91/MG3S charges leads to calculated polarization energies that are ~6 kcal/mol more negative than those polarization energies calculated using CM3/CM3.1 charges. At the HF/MG3S level, this difference increases to ~10 kcal/mol.

Based on the above results, it is tempting to suggest that the gas-phase partial atomic charges obtained from the ChElPG fitting procedure (with the exception of those obtained using *m*PW1PW91/MIDI! wave functions) lead to bond dipole moments that are too large, which in turn lead to polarization energies that are too negative. Of course, solvation polarization energy is not a measurable quantity, and although the results of the above calculations indicate that electrostatic interactions between the gas-phase charge distribution of the solute and solvent make a large contribution to the free energy of solvation, mutual polarization of the solute by the solvent (which requires a full SCRF calculation to evaluate), and nonelectrostatic interactions might also be expected to make a contribution to the observable free energy of solvation or materials condensation. We note here that although the treatment above is for liquid nitromethane, similar effects are expected for condensation into solid-state materials [147].

Table 11 Generalized Born polarization energies in nitromethane ($\varepsilon = 36.562$) for nitramines, calculated using partial atomic charges obtained in different ways

Molecule[b]	Polarization energy[a]							
		MIDI! Basis set					MG3S Basis set	
			Löwdin		ChElPG		ChElPG	
	CM3[c]	CM3.1[d]	*m*PW1PW91	HF	*m*PW1PW91	HF	*m*PW1PW91	HF
Nitramide I (C_s)	−8.48	−8.60	−6.00	−7.44	−8.54	−10.40	−9.57	−11.10
Nitramide I (C_{2v})	−9.31	−9.63	−6.87	−8.40	−9.22	−11.04	−10.90	−12.35
Nitramide II (C_s)	−6.65	−6.96	−4.13	−5.95	−8.55	−10.91	−9.67	−11.82
Nitramide II (C_{2v})	−6.61	−6.84	−4.41	−6.18	−8.58	−10.96	−10.95	−11.99
DMNA I (C_s)	−5.92	−5.95	−5.14	−6.26	−5.12	−6.14	−6.65	−7.21
DMNA I (C_{2v})	−6.41	−6.46	−5.63	−6.80	−5.61	−6.66	−7.36	−7.92
DMNA II (C_s)	−3.91	−4.20	−3.08	−4.56	−4.20	−5.68	−5.12	−6.40
DMNA II (C_{2v})	−4.55	−4.76	−3.69	−5.29	−4.16	−5.72	−4.91	−6.17
TNAZ (C_s)	−8.21	−7.61	−7.31	−10.86	−8.29	−11.57	−10.95	−13.38
RDX I (C_s)	−10.22	−10.27	−8.87	−12.15	−9.15	−12.12	−12.34	−14.67
RDX II (C_{3v})	−11.00	−11.33	−9.55	−13.09	−9.08	−11.94	−11.83	−14.08
γ-HNIW	−12.63	−12.43	−11.19	−17.06	−13.43	−19.14	−18.55	−22.91
β-HNIW (C_2)	−13.20	−12.99	−11.74	−17.52	−13.64	−19.25	−18.70	−22.92
ε-HNIW (C_s)	−13.69	−13.65	−12.24	−18.19	−13.92	−19.59	−19.04	−23.31
RMS difference[e]	0.26		3.36		3.26		2.47	

[a] In kcal/mol. All of the polarization energies in this table depend on the values of the Coulomb radii and $d_{kk'}$ parameters in the Coulomb integrals. In order to make a consistent comparison, we used the SM5.43R values [141] for all calculations. Furthermore, we used unrelaxed *m*PW1PW91/MG3S geometries, and we did not allow any of the solute electronic wave functions to relax in solution

[b] See Fig. 1 and text for descriptions of various conformers

[c] polarization free energies calculated with CM3 gas-phase atomic charges obtained from *m*PW1PW91/MIDI! wave functions

[d] polarization free energies calculated with CM3.1 gas-phase atomic charges obtained from HF/MIDI! wave functions

[e] First, second, third, and fourth values listed are the RMS differences (in kcal/mol) between polarization free energies calculated using CM3 (*m*PW1PW91/MIDI! wave functions) and CM3.1 (HF/MIDI! wave functions) charges, Löwdin charges (obtained using *m*PW1PW91/MIDI! and HF/MIDI! wave functions), ChElPG charges (obtained using *m*PW1PW91/MIDI! and HF/MIDI! wave functions), and ChElPG charges (obtained using *m*PW1PW91/MG3S and HF/MG3S wave functions), respectively

4.8 Total solvation free energies

Unlike the solvent polarization energies calculated above, the total free energy of solvation *is* a measurable property, and in this section we will use the SM5.43R solvation model to calculate values for the total free energy of solvation in water based on different sets of solute gas-phase charges. The following solutes will be considered: nitromethane, 1-nitropropane, 2-nitropropane, nitrobenzene, 1,3-dinitrobenzene, 1,3,5-trinitrobenzene, 2,4,6-trinitrotoluene (TNT), 2-amino-4,6-dinitrotoluene, and RDX. For the above solutes, we used geometries optimized at the *m*PW1PW91/MIDI! level of theory, except for RDX, for which we used geometries optimized at the *m*PW1PW91/MG3S level of theory. Also, for RDX we considered both the C_s (RDX I) and C_{3v} (RDX II) conformations explicitly, and used the following equation to calculate the various components to the free energy of solvation:

$$\exp\left[-\frac{\Delta G}{RT}\right] = \sum_{\mathrm{C}} P_{\mathrm{C}} \exp\left[-\frac{\Delta G(\mathrm{C})}{RT}\right] \tag{9}$$

where P_{C} is the equilibrium mole fraction of conformation C in the gas phase. The gas-phase absolute energies required for computing P_{C} in the above equation are for the *m*PW1PW91/MG3S level of theory. At this level of theory, the C_s conformation of RDX is 0.4 kcal/mol lower in free energy than the C_{3v} conformation. The values for the experimental aqueous free energies of solvation for nitromethane, 1-nitropropane, 2-nitropropane, and nitrobenzene were taken from the SM5.43R training set [141]. The values for the remaining five solutes (which are not in the SM5.43R training set) were taken from three different sources [148–150]. For these solutes, the mean value for the free energy of solvation was used in cases where more than one experimental value for a single solute was available. These experimental values, along with those values calculated using CM3 partial atomic charges with the SM5.43R solvent model, are listed in Table 12. Described below are the SM5.43R model for calculating free energies of solvation and the procedure we used for calculating each of the various components to the total free energies of solvation.

For the SM5.43R solvent model, the standard-state free energy of solvation ($\Delta G_{\mathrm{S}}^{\circ}$) is partitioned according to [141]

$$\Delta G_{\mathrm{S}}^{\circ} = G_{\mathrm{P}} + \Delta E + G_{\mathrm{CDS}} + \Delta G_{\mathrm{conc}}^{\circ} \tag{10}$$

where G_{P} is the electronic polarization energy from mutual polarization of the solute and solvent, ΔE is the change in the solute's internal electronic energy when the solute is placed in the solvent, G_{CDS} is a semiempirical term that accounts for all interactions besides bulk electrostatics, and $\Delta G_{\mathrm{conc}}^{\circ}$ accounts for the concentration change between the gas phase and the liquid-phase standard states. Because we use the same concentrations (1 mol/l) in both phases, $\Delta G_{\mathrm{conc}}^{\circ}$ is zero [151]. For the purposes of this work, we will separate the electrostatic polarization energy into two contributions

Table 12 Various components of the aqueous solvation free energy (in kcal/mol) obtained from full SCRF calculations based on CM3 charges, compared to experimental aqueous solvation free energies

Molecule	$G_{\mathrm{P}(q\mathrm{gas})}$	ΔG_{P}	ΔE	$G_{\mathrm{CDS}}^{\mathrm{a}}$	Total	Expt
Nitromethane	−5.57	−2.22	1.20	1.52	−5.07	−3.99
1-Nitropropane	−4.12	−1.44	0.78	1.46	−3.32	−3.34
2-Nitropropane	−3.98	−1.41	0.77	1.79	−2.83	−3.14
Nitrobenzene	−3.90	−1.55	0.88	1.54	−3.03	−4.11
1,3-Dinitrobenzene	−5.55	−2.22	1.25	2.22	−4.30	−5.46
1,3,5-Trinitrobenzene	−6.74	−1.79	0.96	2.92	−4.65	−5.57
TNT	−6.01	−1.44	0.77	2.97	−3.71	−6.62
2-Amino-4,6-dinitrotoluene	−7.96	−3.42	1.95	1.63	−7.80	−9.24
RDX	−10.63	−6.30	3.57	−0.36	−13.72	−12.22
MSE					0.58	
MUE					1.16	

[a] Calculated using parameters developed for use with CM3/*m*PW1PW91/MIDI! partial atomic charges

$$G_{\mathrm{P}} = G_{\mathrm{P}(q\mathrm{gas})} + \Delta G_{\mathrm{P}} \tag{11}$$

where $G_{\mathrm{P}(q\mathrm{gas})}$ is the electronic polarization energy due to the gas-phase charge distribution of the solute and ΔG_{P} is the difference between the electronic polarization energy due to the gas-phase charges of the solute and that due to the relaxed, liquid-phase charges of the solute.

For all of the calculations in this section, the $G_{\mathrm{P}(q\mathrm{gas})}$ term was evaluated in the same way as it was above (i.e. by substituting gas-phase partial atomic charges into Eq. 8) with a value of 78.3 [146] for the dielectric constant of water.

Calculating the ΔG_{P} term requires a full SCRF calculation. The procedure for this has been outlined previously [152, 153], although we note here that it requires evaluating the partial derivatives of the partial atomic charges (which are updated at each step of the SCRF calculation) with respect to the density matrix. The analytic forms of these partial derivatives for CM2 charges (which are the same as for CM3 charges) have been published previously [153] and are used in this work without change. For those charges obtained by the ChElPG fitting procedure, analytical solutions to the above partial derivatives are not available, so we did not perform full SCRF calculations based on these sets of charges. Instead, we calculated ΔG_{P} values based on gas-phase ChElPG charges using the following equation

$$\Delta G_{\mathrm{P}} = G_{\mathrm{P}(q\mathrm{gas})}\,(f - 1) \tag{12}$$

where the prefactor *f* appearing in the above equation depends on the solute and is the ratio of its electronic polarization free energy obtained from a full SCRF calculation based on CM3 charges to that polarization energy, calculated using CM3 gas-phase charges in Eq. (8), i.e.

$$f = \frac{G_{\mathrm{P}}}{G_{\mathrm{P}(q\mathrm{gas,CM3})}} \tag{13}$$

The *f* values for each of the solutes listed above are given in Table 13.

For ΔE, a full SCRF calculation is also required, so for gas-phase ChElPG partial atomic charges, we use ΔG_{P} to calculate ΔE according to

$$\Delta E = -f'\Delta G_{\mathrm{P}} \tag{14}$$

Table 13 Solvation prefactors, based on full SCRF calculations with CM3 charges

Molecule	f	f'
Nitromethane	1.40	0.54
1-Nitropropane	1.35	0.54
2-Nitropropane	1.35	0.55
Nitrobenzene	1.40	0.57
1,3-Dinitrobenzene	1.40	0.56
1,3,5-Trinitrobenzene	1.27	0.54
TNT	1.24	0.53
2-Amino-4,6-dinitrotoluene	1.43	0.57
RDX	1.59	0.57

where the prefactor f' depends on the solute and is calculated using ΔE and ΔG_{P} values obtained from full SCRF calculations based on CM3 charges. The values of these prefactors are also given in Table 13. If one assumes linear response of both the solvent *and* the solute charge distribution, the above prefactor will necessarily equal exactly 1/2. However, for the SM5.43R model (as well as for our previous models based on SCRF theory), we assume only linear response of the solvent, and calculate the polarization energy due to the solute charge distribution explicitly. As a result, the f' values in Table 13 are not required to (and do not) equal exactly 1/2 . These values of f' are also used in the present paper to estimate ΔE for the ChElPG partial atomic charges.

The final contribution to the total aqueous free energy of solvation, G_{CDS}, is calculated according to

$$G_{\mathrm{CDS}} = \sum_{k} \sigma_k A_k \tag{15}$$

where A_k is the solvent-accessible surface area of atom *k*, and σ_k is a functional called an atomic surface tension, which, for water, is itself a function of the local geometry of the solute and a set of parameters called surface tension coefficients. The surface tension coefficients used here have been determined [141] empirically by a linear optimization of "target" G_{CDS} values, which were calculated by subtracting the calculated electrostatic terms from the observable free energies of solvation for a diverse training set of solutes. As a result, the values of the surface tension coefficients themselves

Table 14 Various components of the aqueous solvation free energy (in kcal/mol) calculated using gas-phase ChElPG/ *m*PW1PW91/MIDI! charges, compared to experimental aqueous solvation free energies

Molecule	$G_{\mathrm{P}(q\mathrm{gas})}$	ΔG_{P}	ΔE	$G^{\mathrm{a}}_{\mathrm{CDS}}$	Total	Expt
Nitromethane	−5.30	−2.11	1.14	1.52	−4.75	−3.99
1-Nitropropane	−4.38	−1.53	0.83	1.46	−3.62	−3.34
2-Nitropropane	−3.91	−1.39	0.76	1.79	−2.75	−3.14
Nitrobenzene	−3.66	−1.45	0.83	1.54	−2.75	−4.11
1,3-Dinitrobenzene	−5.68	−2.27	1.28	2.22	−4.45	−5.46
1,3,5-Trinitrobenzene	−7.07	−1.88	1.01	2.92	−5.02	−5.57
TNT	−6.95	−1.67	0.89	2.97	−4.75	−6.62
2-Amino-4,6-dinitrotoluene	−8.87	−3.81	2.17	1.63	−8.88	−9.24
RDX	−9.13	−5.41	3.07	−0.36	−11.83	−12.22
MSE					0.54	
MUE					0.77	

[a] Calculated using parameters developed for use with CM3/*m*PW1PW91/MIDI! partial atomic charges

Table 15 Various components of the aqueous solvation free energy (in kcal/mol) calculated using gas-phase ChElPG/*m*PW1PW91/MG3S charges, compared to experimental aqueous solvation free energies

Molecule	$G_{\mathrm{P}(q\mathrm{gas})}$	ΔG_{P}	ΔE	$G^{\mathrm{a}}_{\mathrm{CDS}}$	Total	Expt
Nitromethane	−6.69	−2.67	1.44	1.52	−6.40	−3.99
1-Nitropropane	−5.87	−2.05	1.11	1.46	−5.35	−3.34
2-Nitropropane	−5.32	−1.88	1.03	1.79	−4.39	−3.14
Nitrobenzene	−4.87	−1.94	1.10	1.54	−4.17	−4.11
1,3-Dinitrobenzene	−8.30	−3.32	1.87	2.22	−7.53	−5.46
1,3,5-Trinitrobenzene	−10.85	−2.88	1.55	2.92	−9.27	−5.57
TNT	−10.60	−2.54	1.36	2.97	−8.81	−6.62
2-Amino-4,6-dinitrotoluene	−10.86	−4.67	2.66	1.63	−11.24	−9.24
RDX	−12.15	−7.20	4.08	−0.36	−15.63	−12.22
MSE					−2.12	
MUE					2.12	

[a] Calculated using parameters developed for use with CM3/*m*PW1PW91/MIDI! partial atomic charges

depend indirectly on the particular set of CM3 parameters that is used in calculating the electrostatic components to the solvation free energy. To further increase the accuracy of our models for predicting solvation free energies, different sets of surface tension coefficients have been developed for use with different sets of CM3 parameters [141]. Hence, the SM5.43R/ *m*PW1PW91/MIDI! model uses those surface tension coefficients that have been determined by an empirical optimization against target G_{CDS} values obtained by subtracting the electrostatic components, calculated at the CM3/ *m*PW1PW91/MIDI! level of theory, from experimental free energies of solvation. Because the G_{CDS} term depends on the particular set of CM3 parameters chosen to calculate the electrostatic components of the total free energy of solvation, and because it is not a measurable property, the choice of what values to use for G_{CDS} in those cases where gas-phase ChElPG charges are used to calculate the electrostatic contributions to the total solvation free energy becomes somewhat arbitrary. Despite this, we feel that some qualitative comparisons can be made between the CM3 and ChElPG fitting procedure by the above analysis, so for consistency (and because in two of the cases the charge models considered here are based on MIDI! wave functions) we used those G_{CDS} values obtained from the SM5.43R/*m*PW1PW91/MIDI! model to calculate all of the total solvation free energies shown in Tables 14 and 15.

From the data in Tables 12 and 14, we see that the solvation free energies, obtained using gas-phase ChElPG/ *m*PW1PW91/MIDI! partial atomic charges and the f and f' values in Table 13, are even more accurate than those solvation free energies obtained from a full SCRF calculation based on CM3 charges. This result is encouraging, because it shows that the above procedure for calculating total solvation free energies based on gas-phase charges obtained from other models can yield accurate solvation free energies. From Tables 12, 14, and 15, we see that those solvation energies obtained from a full SCRF calculation and those obtained from gas-phase ChElPG/*m*PW1PW91/ MIDI! charges are an average of ~0.5 kcal/mol too positive, whereas those solvation free energies obtained from gas-phase ChElPG/mPW1PW91/MG3S are in all cases too negative, by an average of ~2 kcal/mol. These results lend further support to the notion that for the molecules examined as part of this work, the majority of the partial atomic charges obtained using the ChElPG fitting procedure (with the exception of ChElPG/*m*PW1PW91/MIDI!) yield bond dipoles that are too large in magnitude, which in turn lead to polarization energies that are significantly too negative.

The mean signed error in Table 15 is equal to 22% of the mean value of the sum of G_{P} and ΔE. If we apply the same fractional correction (22/122) to the *m*PW1PW91/MG3S values in Table 11, they will apparently be more accurate, and they become less negative by an average of 2.04 kcal/mol. This accounts in most cases for their disagreement from the CM3 and CM3.1 results. We conclude from this that the CM3 and CM3.1 results in Table 11 are more accurate than the *m*PW1PW91/MG3S values.

5 Concluding remarks

In this paper, we have presented a CM3 parameterization for HF/MIDI!. This model is very similar to the one developed previously for *m*PW1PW91/MIDI!, except that it maps charges on N and O using an extended functional form. This functional form allows us to include dipole moments for nitramines (which include many high-energy materials) in our training set without having a significant impact on the quality of the model for molecules containing other functional groups, and we refer to the model extended in this way as CM3.1. The quality of the partial atomic charges obtained from CM3 (which is parameterized for use with *m*PW1PW91/MIDI! wave functions) and CM3.1 (which was parameterized here for use with HF/MIDI! wave functions) compared to those partial atomic charges obtained from a Löwdin population analysis and the ChElPG fitting procedure has also been assessed using a test set of accurate dipole moments for 14 molecules containing nitramine functionality. The above assessment shows that (1) CM3 and CM3.1 point-charge-derived dipole moments are significantly more accurate than point-charge-dipole moments calculated using Löwdin atomic charges and (2) unlike the ChElPG fitting procedure, CM3 and CM3.1 yield partial atomic charges that are relatively invariant to nonphysical fluctuations such as the size of the molecule (i.e. whether or not the atom is located in the interior or exterior of the molecule), small conformational changes in the molecule, or the level of treatment of electron correlation. To demonstrate the applicability of CM3 and CM3.1, we calculated polarization free energies for solvent polarization in liquid nitromethane solvent; these polarization energies are based on gas-phase partial atomic charges obtained from HF/MIDI! and *m*PW1PW91/MIDI! wave functions for 14 molecules containing nitramine functional groups. It was shown that the polarization energies calculated using gas-phase partial atomic charges obtained from CM3 and CM3.1 are much less dependent on the level of treatment of electron correlation than those polarization energies calculated using gas-phase partial atomic charges obtained from either a Löwdin population analysis or the ChElPG fitting procedure. Based on the above results, as well as a comparison of the experimental free energies of solvation in water for a set of nitro-containing compounds (including RDX) to those values calculated with the SM5.43R model, we have suggested that in many cases ChElPG charges lead to calculated polarization free energies that are significantly too negative, and that this is the result of an overestimation of the magnitude of individual bond dipole moments. The partial atomic charges obtained from the models presented in this work should be useful not only for continuum solvation calculations, but also as quantum chemical descriptors in force fields designed to model the solid-state properties of various high-energy materials.

Acknowledgements Special thanks go to Jason Thompson for providing help and insightful comments throughout the course of this work. This work was supported in part by the National Science Foundation and by a Department of Defense (DOD) Multidisciplinary University Research Initiative (MURI) grant managed by the Army Research Office.

References

1. Cioslowski J (1998) In: Schleyer PvR, Allinger NL, Clark T, Gasteiger J, Kollman PA, Schaefer III HF (eds) Encyclopedia of computational chemistry, vol 5. Wiley, New York, p 892
2. King WT, Mast GB, Blanchette PP (1971) J Chem Phys 56:4440
3. Coppens P (1992) Annu Rev Phys Chem 43:663
4. Person WB, Newton JH (1974) J Chem Phys 61:1040
5. King WT, Mast GB (1976) J Phys Chem 80:2521
6. Jolly WL, Perry WB (1973) J Am Chem Soc 95:5442
7. Jolly WL, Perry WB (1974) Inorg Chem 13:2686
8. Wiberg KB, Wendoloski J (1978) J Am Chem Soc 100:723
9. Cioslowski J (1989) J Am Chem Soc 111:8333
10. Cioslowski J, Hay PJ, Ritchie JP (1990) J Phys Chem 94:148
11. Gimarc BM, Ott JJ (1986) J Am Chem Soc 108:4298
12. Gimarc BM, Ott JJ (1986) J Am Chem Soc 108:4303
13. Bonchev D, Kier L (1992) J Math Chem 9:75
14. Galvez J, Garcia R, Salabert MT, Soler R (1994) J Chem Inf Comput Sci 34:520
15. Gasteiger J, Marsili M (1980) Tetrahedron 36:3219
16. Marsili M, Gasteiger J (1980) Croat Chem Acta 53:601
17. Moriter WJ, Van Genechten K, Gasteiger J (1985) J Am Chem Soc 107:829
18. Sanderson RT (1983) J Am Chem Soc 105:2259
19. Abraham RJ, Griffiths L, Loftus P (1982) J Comput Chem 3: 407
20. Abraham RJ, Smith PE (1987) J Comput Chem 9:288
21. Mullay J (1985) J Am Chem Soc 108:1770
22. Mullay J (1988) J Comput Chem 9:764
23. Mullay J (1991) J Comput Chem 12:369
24. No KT, Grant JA, Scherega HA (1990) J Phys Chem 94:4732
25. No KT, Grant JA, Jhon MS, Scherega HA (1990) J Phys Chem 94:4740
26. Rappé AK, Goddard III WA (1991) J Phys Chem 95:3358
27. Oliferenko AA, Palyulin VA, Pisarev SA, Neiman AV, Zefirov NS (2001) J Phys Org Chem 14:355
28. De Proft F, Van Alsenoy C, Peeters A, Langenaeker W, Geerlings P (2002) J Comput Chem 23:1198
29. Bultinck P, Langenaeker W, Lahorte P, De Proft F, Van Alsenoy C, Tollenaere JP (2002) J Phys Chem A 106:7895
30. Bultinck P, Langenaeker W, Carbo-Dorca R, Tollenaere JP (2003) J Chem Inf Comput Sci 43:422
31. Mulliken RS (1935) J Chem Phys 3:564
32. Mulliken RS (1955) J Chem Phys 23:1833
33. Mulliken RS (1962) J Chem Phys 36:3428
34. Löwdin P-O (1950) J Chem Phys 18:365
35. Thompson JD, Winget P, Truhlar DG (2001) Phys Chem Comm 4:72. DOI: 10.1039/b105076c
36. Cusachs LC, Politzer P (1968) Chem Phys Lett 1:529
37. Stout EW, Politzer P (1968) Theor Chim Acta 12:379
38. Politzer P, Harris RR (1970) J Am Chem Soc 92:6451
39. Politzer P (1971) Theor Chim Acta 23:203
40. Politzer P, Mulliken RS (1971) J Chem Phys 55:5135
41. Politzer P, Reggio PH (1972) J Am Chem Soc 94:8308
42. Politzer P, Politzer A (1973) J Am Chem Soc 95:5450
43. Rousseau B, Peeters A, Van Alsenoy C (2001) THEOCHEM 538: 235
44. Golebiewski A, Rzescowska E (1974) Acta Phys Pol 45:563
45. Baker J (1985) Theor Chim Acta 68:221

46. Kar T, Sannigrahi AB, Mukherjee DC (1987) THEOCHEM 153:93
47. Reed AE, Weinstock RB, Weinhold F (1985) J Chem Phys 83: 735
48. Reed AE, Curtiss LA, Weinhold F (1988) Chem Rev 88:899
49. Bader RFW, Larouche A, Gatti C, Carroll MT, MacDougal PJ, Wiberg KB (1987) J Chem Phys 87:1142
50. Coulson CA, Redei LB, Stocker D (1962) Proc R Soc London Ser A 270:357
51. Popelier PLA (2001) Theor Chem Acc 105:393
52. Schwartz ME, Coulson CA, Allen LC (1970) J Am Chem Soc 92: 447
53. Hirshfeld FL (1977) Theor Chim Acta 44:129
54. Smit PH, Derissen JL, Van Duijneveldt (1979) Mol Phys 37: 521
55. Kollman PA (1977) J Am Chem Soc 99:4875
56. Momany FA (1978) J Phys Chem 86:592
57. Cox SR, Williams DE (1981) J Comput Chem 2:304
58. Singh UC, Kollman PA (1984) J Comput Chem 5:129
59. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta SJ, Weiner P (1984) J Am Chem Soc 106:765
60. Weiner SJ, Kollman PA, Nguyen DT, Case DA (1986) J Comput Chem 7:230
61. Chirlian LE, Francl MM (1987) J Comput Chem 8:894
62. Francl MM, Carey C, Chirlian LE, Gange DM (1996) J Comput Chem 17:367
63. Francl MM, Chirlian LE (2000) In: Lipkowitz KB, Boyd DB (eds) Reviews in computational chemistry, vol 14. Wiley, New York, p 1
64. Wiberg KB, Rablen PR (1993) J Comput Chem 14:1504
65. Breneman CM, Wiberg KB (1990) J Comput Chem 11:361
66. Besler BH, Merz KM, Kollman PA (1990) J Comput Chem 11:431
67. Merz KM (1992) J Comput Chem 13:749
68. Westbrook JD, Levy RM, Krogh-Jesperson K (1992) J Comput Chem 13:979
69. Montagani R, Tomasi J (1993) THEOCHEM 279:131
70. Beck B, Clark T, Glen RC (1997) J Comput Chem 18:744
71. Bayly CI, Cieplak P, Cornell W, Kollman PA (1993) J Phys Chem 97:10629
72. Sigfridsson E, Ryde U (1998) J Comput Chem 19:377
73. Storer JW, Giesen DJ, Cramer CJ, Truhlar DG (1995) J Comput-Aid Mol Des 9:87
74. Li J, Zhu T, Cramer CJ, Truhlar DG (1998) J Phys Chem A 102:1820
75. Li J, Williams B, Cramer CJ, Truhlar DG (1999) J Chem Phys 110:724
76. Winget P, Thompson DJ, Xidos JD, Cramer CJ, Truhlar DG (2002) J Phys Chem A 106:10707
77. Thompson DJ, Cramer CJ, Truhlar DG (2003) J Comput Chem 24: 1291
78. Brom JM, Schmitz BJ, Thompson JD, Cramer CJ, Truhlar DG (2003) J Phys Chem A 107:6483
79. Kalinowski JA, Lesyng B, Thompson JD, Cramer CJ, Truhlar DG (2004) J Phys Chem A 108:2545
80. Sanderson RT (1976) Chemical bonds and chemical energy. Academic, New York
81. Politzer P, Weinstein H (1979) J Chem Phys 71:4218
82. Sorescu DC, Rice BM, Thompson DL (1997) J Phys Chem B 101: 798
83. Sorescu DC, Rice BM, Thompson DL (1998) J Phys Chem B 102: 948
84. Sorescu DC, Rice BM, Thompson DL (1998) J Phys Chem B 102: 6692
85. Sorescu DC, Rice BM, Thompson DL (1999) J Phys Chem A 103: 989
86. Sorescu DC, Boatz JA, Thompson DL (2001) J Phys Chem A 105: 5010
87. Smith GD, Bharadwaj RK, Bedrov D, Ayyagari (1999) J Phys Chem B 103:705
88. Lynch BJ, Zhao Y, Truhlar DG (2003) J Phys Chem A 107:1384
89. Easton RE, Giesen DJ, Welch A, Cramer CJ, Truhlar DG (1996) Theor Chim Acta 93:281
90. Li J, Cramer CJ, Truhlar DG (1998) Theor Chim Acta 99:192
91. Thompson JD, Winget P, Truhlar DG (2001) Phys Chem Comm. DOI: 10.1039/b105076c
92. Adamo C, Barone V (1998) J Chem Phys 108:664
93. Mayer I (1983) Chem Phys Lett 97:270
94. Mayer I (1985) Chem Phys Lett 117:396
95. Mayer I (1986) Int J Quantum Chem 29:73
96. Fast PL, Sanchez ML, Truhlar DG (1999) Chem Phys Lett 306: 407
97. Curtiss LA, Redfern PC, Raghavachari K, Rassolov V, Pople JA (1999) J Chem Phys 110:4703
98. Krishnan R, Binkley JS, Seeger R, Pople JA (1980) J Chem Phys 72:650
99. Frisch MJ, Pople JA, Binkley JS (1984) J Chem Phys 80:3265
100. Woon DE, Dunning Jr. TH (1993) J Chem Phys 98:1358
101. Stolevik R, Rademacher P (1969) Acta Chem Scand 23:672
102. Krebs B, Mandt J (1979) Acta Crystallogr B35:402
103. Sumpter BG, Thompson DL (1988) J Chem Phys 88:6889
104. Politzer P, Sukumar N, Jayasuriya K, Ranganathan S (1988) J Am Chem Soc 110:3425
105. Habibollahzadeh D, Murray JS, Redfern PC, Politzer P (1991) J Phys Chem 95:7703
106. Kohno Y, Maekawa K, Tsuchioka T, Hashizume T, Imamura A (1993) Chem Phys Lett 214:603
107. Roszak SJ (1994) Mol Struct 304:269
108. Khaikin LS, Grikina OE, Vilkov LV, Palafox AM, Boggs JE (1993) J Struct Chem 34:2
109. Khaikin LS, Grikina OE, Vilkov LV, Boggs JE (1993) J Struct Chem 34:9
110. Harris JN, Lammerstsma K (1997) J Phys Chem A 101:1370
111. Johnson MA, Truong TN (1999) J Phys Chem A 103:8840
112. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr. JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalamani G, Rega N, Petersson GA, Nagatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C and Pople JA (2003) Gaussian 03, revision A.1. Gaussian, Pittsburgh
113. Xidos JD, Li J, Thompson JD, Hawkins GD, Winget PD, Zhu T, Rinaldi D, Liotard DA, Cramer CJ, Truhlar DG (2003) MN-GSM, version 3.1. University of Minnesota, Minneapolis
114. NAG Fortran 90 Library, 4th edn. (2000) The Numerical Algorithms Group, Inc., Oxford
115. Tyler JK (1963) J Mol Spectrosc 11:39
116. Hunter EC, Partington JR (1933) J Chem Soc 309
117. George MV, Wright G (1958) J Am Chem Soc 80:1200
118. Choi CS, Prince E (1972) Acta Crystallogr B28:2857
119. Shishkov IF, Vilkov LV, Kolontis M, Rozsondai B (1991) Struct Chem 2:57
120. Karpwicz RJ, Brill TB (1984) J Phys Chem 88:348
121. Filhol A, Clement C, Forel M-T, Paviot J, Rey-Lafon M, Richoux G, Trinquecoste C, Cherville J (1971) J Phys Chem 75:2056
122. Nielson AT, Chafin AP, Christian SL, Moore DW, Nadler MP, Nissan RA, Vanderah DJ (1998) Tetrahedron 54:11793
123. Lide DR (ed) (2000) CRC handbook of chemistry and physics, 80th edn. CRC Press, Boca Raton
124. Stark B (1974) In: Hellwege K-H, Hellwege AM (eds) Molecular constants from microwave spectroscopy, Landolt-Börnstein, new series, group II, vol 6. Springer, Berlin Heidelberg, New York, p 261
125. Nelson RD, Lide DR, Maryott AA (1967) Natl stand ref dat ser, United States National Bureau of Standards NSRDS-NBS 10. Washington

126. Velnio B, Cané E, Trombetti A, Corbelli G (1992) J Mol Spectrosc 155:1
127. Velnio B, Cané E, Gagliardi L, Trombetti A, Caminati W (1993) J Mol Spectrosc 161:136
128. Krugh WD, Gold LP (1974) J Mol Spectrosc 49:423
129. McClellan AL (1963) Tables of experimental dipole moments. WH Freeman, San Francisco
130. Stark B (1967) In: Hellwege K-H, Hellwege AM (eds) Molecular constants from microwave spectroscopy, Landolt-Börnstein, new series, group II, vol 4. Springer, Berlin Heidelberg, New York, p 136
131. Stark B (1982) In: Hellwege K-H, Hellwege AM (eds) Molecular constants from microwave spectroscopy, Landolt-Börnstein, new series, group II, vol 14a. Springer, Berlin Heidelberg New York, p 261
132. Caminati W, Velnio B, Danieli R (1993) J Mol Spectrosc 161:208
133. Becke AD (1997) J Chem Phys 107:8554
134. Becke AD (1988) Phys Rev A 38:3098
135. Lee C, Yang W, Parr RG (1987) Phys Rev B 37:785
136. Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) J Phys Chem 98:11623
137. Becke AD (1993) J Chem Phys 98:5648
138. Hamprecht FA, Cohen AJ, Tozer DJ, Handy NC (1998) J Chem Phys 109:6264
139. Wilson PJ, Bradley TJ, Tozer DJ (2001) J Chem Phys 115: 9233
140. Schmider HL, Becke AD (1998) J Chem Phys 108:9264
141. Thompson JD, Cramer CJ, Truhlar DG (2004) J Phys Chem A 108:6532
142. Hoijtink GJ, Boer ED, Meij PHvD, Weijland WP (1956) Recl Trav Chim Pays-Bas 75:487
143. Peradejori F (1963) Cah Phys 17:393
144. Tucker SC, Truhlar DG (1989) Chem Phys Lett 157:164
145. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) J Am Chem Soc 112:6127
146. Winget P, Dolney DM, Giesen DJ, Cramer CJ, Truhlar DG (1999) Minnesota solvent descriptor database. http://comp.chem.umn.edu/solvation/mnsddb.pdf
147. Thompson JD, Cramer CJ, Truhlar DG (2003) J Chem Phys 119:1661
148. Walsh ME, Jenkins TF, Thorne PG (1995) J Energ Mater 13: 357
149. Syracuse research corporation (1994) Physical/chemical property database (PHYSPROP), SRC Environmental Science Center, Syracuse
150. Leo AJ (1994) Masterfile from MedChem Software, BioByte Corp., Claremont
151. Ben-Naim A (1987) Solvation thermodynamics. Plenum, New York
152. Tapia O (1980) In: Daudel R, Pullman A, Salem L, Viellard A (eds) Quantum theory of chemical reactions. Reidel, Dordrecht, p 25
153. Zhu T, Li J, Hawkins GD, Cramer CJ, Truhlar DG (1998) J Chem Phys 109:9117