

# Calculation of Electrostatic Interaction Energies in Molecular Dimers from Atomic Multipole Moments Obtained by Different Methods of Electron Density Partitioning

ANATOLIY VOLKOV, PHILIP COPPENS

Department of Chemistry, State University of New York at Buffalo,  
Buffalo, New York, 14260-3000

Received 1 September 2003; Accepted 23 November 2003

**Abstract:** Accurate and fast evaluation of electrostatic interactions in molecular systems is still one of the most challenging tasks in the rapidly advancing field of macromolecular chemistry, including molecular recognition, protein modeling and drug design. One of the most convenient and accurate approaches is based on a Buckingham-type approximation that uses the multipole moment expansion of molecular/atomic charge distributions. In the mid-1980s it was shown that the pseudoatom model commonly used in experimental X-ray charge density studies can be easily combined with the Buckingham-type approach for calculation of electrostatic interactions, plus atom–atom potentials for evaluation of the total interaction energies in molecular systems. While many such studies have been reported, little attention has been paid to the accuracy of evaluation of the purely electrostatic interactions as errors may be absorbed in the semiempirical atom–atom potentials that have to be used to account for exchange repulsion and dispersion forces. This study is aimed at the evaluation of the accuracy of the calculation of electrostatic interaction energies with the Buckingham approach. To eliminate experimental uncertainties, the atomic moments are based on theoretical single-molecule electron densities calculated at various levels of theory. The electrostatic interaction energies for a total of 11 dimers of  $\alpha$ -glycine, *N*-acetylglycine and L-(+)-lactic acid structures calculated according to Buckingham with pseudoatom, stockholder and atoms-in-molecules moments are compared with those evaluated with the Morokuma–Ziegler energy decomposition scheme. For  $\alpha$ -glycine a comparison with direct “pixel-by-pixel” integration method, recently developed Gavezzotti, is also made. It is found that the theoretical pseudoatom moments combined with the Buckingham model do predict the correct *relative* electrostatic interactions energies, although the absolute interaction energies are underestimated in some cases. The good agreement between electrostatic interaction energies computed with Morokuma–Ziegler partitioning, Gavezzotti’s method, and the Buckingham approach with atoms-in-molecules moments demonstrates that reliable and accurate evaluation of electrostatic interactions in molecular systems of considerable complexity is now feasible.

© 2004 Wiley Periodicals, Inc. J Comput Chem 25: 921–934, 2004

**Key words:** electron density; electrostatic interaction energy; density partitioning; atomic moments; molecular moments; pseudoatom model

## Introduction

According to the Morokuma–Ziegler energy partitioning scheme<sup>1–3</sup> the total intermolecular monomer-monomer interaction energy  $E_{\text{int}}$  is written as

$$E_{\text{int}} = E_{\text{es}} + E_{\text{Pauli}} + E_{\text{oi}} \quad (1)$$

where  $E_{\text{es}}$  is the electrostatic interaction energy,  $E_{\text{Pauli}}$  is the Pauli, or exchange repulsion, which accounts for steric repulsion, and  $E_{\text{oi}}$

(orbital interaction energy) includes charge transfer and polarization effects that occur upon the relaxation of the interacting system to its final state. This energy decomposition scheme is similar to the Kitaura–Morokuma analysis,<sup>4,5</sup> which explicitly partitions the

**Correspondence to:** Dr. A. Volkov; e-mail: volkov@chem.buffalo.edu

This article includes Supplementary Material available from the authors upon request or via the Internet at <http://www.interscience.wiley.com/jpages/0192-8651/suppmat>.

interaction energy into electrostatic ( $E_{\text{es}}$ ), exchange, polarization, and charge transfer contributions.

In most of the Morokuma-type energy decomposition approaches the electrostatic interaction energy  $E_{\text{es}}$  represents a classic Coulombic interaction between two unperturbed (unmodified) monomer charge distributions  $\rho^A$  and  $\rho^B$ , which is defined as<sup>6</sup>

$$E_{\text{es}} = - \int \int \frac{\rho^A(r_A)\rho^B(r_B)}{|r_A - r_B|} dr_A dr_B \quad (2)$$

Although this is an exact expression, it is not readily suitable for practical applications because of the complicated 6D integration involved and, if evaluated numerically, requires a precise knowledge of the charge distributions and the ability to calculate them at any point in space.

Among other methods explored are the generalized Gaussian quadrature integration in a curvilinear coordinate system<sup>7</sup> and the direct numerical integration with the “pixel-by-pixel” approach,<sup>8</sup> both of which require a significant amount of computational time. A much simpler form can be obtained when expanding  $|r_A - r_B|^{-1}$  as a Taylor series. In the resulting expression<sup>9,10</sup>

$$\begin{aligned} E_{\text{es}} = & Tq^A q^B + T_{\alpha}(q^A \mu_{\alpha}^B - q^B \mu_{\alpha}^A) \\ & + T_{\alpha\beta}(\frac{1}{3} q^A \Theta_{\alpha\beta}^B + \frac{1}{3} q^B \Theta_{\alpha\beta}^A - \mu_{\alpha}^A \mu_{\alpha}^B) \\ & + T_{\alpha\beta\gamma}(\frac{1}{15} q^A \Omega_{\alpha\beta\gamma}^B - \frac{1}{15} q^B \Omega_{\alpha\beta\gamma}^A - \frac{1}{3} \mu_{\alpha}^A \Theta_{\beta\gamma}^B + \frac{1}{3} \mu_{\alpha}^B \Theta_{\beta\gamma}^A) + \dots \quad (3) \end{aligned}$$

detailed knowledge of the two charge distributions  $\rho^A$  and  $\rho^B$  is not a necessity and instead only the permanent multipole moments  $q$ ,  $\mu$ ,  $\Theta$ ,  $\Omega$  ... (monopole, dipole, quadrupole, octopole) of the unperturbed molecular charge distributions are required. Parameters  $T_{\alpha,\beta,\gamma} \dots$  are the symmetrical interaction tensors of the form<sup>9</sup>

$$T_{\alpha\beta\gamma} \dots = \nabla_{\alpha} \nabla_{\beta} \nabla_{\gamma} \dots R^{-1} \quad (4)$$

where  $R$  is the vector from the origin of  $A$  to  $B$  and  $\alpha, \beta, \gamma$  represent  $x$ ,  $y$ , and  $z$ , using the Einstein convention for summation over the indices. This approximation approaches the exact expression (2) when the distance between the two molecular charge distributions  $A$  and  $B$  is much larger than the size of the distributions and when a sufficient number of high-order terms is included in the expansion. In practice, it is more convenient to replace the molecular moments by a summation of contributions from the individual atoms constituting the molecules. Then, expression (3) can be rewritten in terms of pairwise interactions that run over all atoms  $i = 1 \dots N(A)$  and  $j = 1 \dots N(B)$  constituting molecules  $A$  and  $B$ , respectively:<sup>11</sup>

$$\begin{aligned} E_{\text{es}} = & \sum_i^{N(A)} \sum_j^{N(B)} Tq_i q_j + T_{\alpha}(q_i \mu_{\alpha,j} - q_j \mu_{\alpha,i}) + T_{\alpha\beta}(\frac{1}{3} q_i \Theta_{\alpha\beta,j} \\ & + \frac{1}{3} q_j \Theta_{\alpha\beta,i} - \mu_{\alpha,j} \mu_{\alpha,i}) + T_{\alpha\beta\gamma}(\frac{1}{15} q_i \Omega_{\alpha\beta\gamma,j} - \frac{1}{15} q_j \Omega_{\alpha\beta\gamma,i} \\ & - \frac{1}{3} \mu_{\alpha,i} \Theta_{\beta\gamma,j} + \frac{1}{3} \mu_{\alpha,j} \Theta_{\beta\gamma,i}) + \dots \quad (5) \end{aligned}$$

Unfortunately, the atomic moments in expression (5) cannot be determined uniquely as neither the wave function nor the electron density can be unambiguously partitioned into atomic fragments. In the approximations used, the partitioning is done on either the orbitals that construct the wave function or on the density as calculated directly from the wave function or determined experimentally. Partitioning methods based on the electron density are more suitable for the purpose of calculating the electrostatic interactions because they directly lead to the definition of the moments of atomic charge distributions from the molecular electron density.

Two types of electron density partitioning methods, referred to as fuzzy and discrete boundary partitionings, can be distinguished.<sup>11</sup> In the first case the charge density at each point is assigned to overlapping functions centered at different locations, while in the second all the density at a given point is assigned to a specific center. Two examples of fuzzy boundary methods are the stockholder<sup>12</sup> and the pseudoatom partitioning.<sup>13,14</sup> In the current study the latter is represented by the Hansen–Coppens expansion.<sup>11,15</sup>

Within the atoms-in-molecules (AIM) theory developed by Bader<sup>16–18</sup> the atomic basin of a nucleus  $\Omega$  is defined as a discrete region in space bounded by interatomic surfaces (IASs) that satisfy the boundary condition of zero flux:<sup>18</sup>

$$\nabla\rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \quad \forall \mathbf{r} \in \text{surface} \quad (6)$$

where  $\nabla\rho(\mathbf{r})$  is the gradient vector field of the electron density and  $\mathbf{n}(\mathbf{r})$  is the vector normal to the surface at  $\mathbf{r}$ .

The atomic moments are then evaluated as an integral of the form

$$A(\Omega) = \int_{\Omega} \hat{A}\rho(\mathbf{r}) d\tau \quad (7)$$

where  $\hat{A}$  is the moment operator and  $A(\Omega)$  is the atomic moment integrated over the basin  $\Omega$ . The application of AIM moments to the calculation of intermolecular electrostatic interaction energies  $E_{\text{es}}$  has been extensively studied by Popelier et al. based on supermolecular calculations on a number of organic molecules of different sizes and complexity,<sup>19–21</sup> including DNA base pairs. The convergence of the calculated energies as a function of the cutoff of the order of the moments was shown to be somewhat slower for the AIM (unless the AIM moments were distributed over extra nonnuclear sites via a shifting procedure) than for Stone’s DMA moments,<sup>22</sup> which are widely used in force field calculations, but the resulting energies from the two methods agreed within 1.3 kJ/mol when higher-order ( $l \geq 4$ ) terms were included in the expansion. The electrostatic interaction energies calculated with the Buckingham-type expansion were also compared with “exact” values obtained by a 6D double-basin integration of (2) for each atom–atom pair in the supermolecule. In this definition the exact values also include the penetration energy that occurs upon overlapping of two charge distributions at close distances. Thus, the electrostatic energies calculated from the Buckingham-type expansion can be considered only as an approximation to the exact energies calculated by double-basin integration. Nevertheless, a good agreement between AIM-based and

exact values of  $E_{\text{es}}$  has been reported with the differences never exceeding 6 kJ/mol with an average discrepancy of 1–2 kJ/mol.

Hirshfeld's fuzzy-boundary stockholder partitioning<sup>12</sup> can equally well be applied to theoretical and experimental electron densities. According to this method, the actual molecular density  $\rho(\mathbf{r})$  at each point is divided among the atoms of the molecule in proportion to their respective contributions to the promolecule density at that point:

$$\rho_i^{\text{stock}}(\mathbf{r}) = w_i(\mathbf{r})\rho(\mathbf{r}) = \frac{\rho_i^{\text{free atom}}(\mathbf{r})}{\rho_i^{\text{promol}}(\mathbf{r})}\rho(\mathbf{r}) \quad (8)$$

where the promolecule density is defined as  $\rho^{\text{promol}} = \sum_i \rho_i^{\text{free atom}}(\mathbf{r})$  and  $\rho_i^{\text{free atom}}(\mathbf{r})$  are the densities of the free spherically averaged atoms. The application of stockholder moments in the calculation of intermolecular interactions in hydrogen-bonded small organic systems has been discussed in detail by Spackman.<sup>6</sup> According to the formalism used in his studies, the total intermolecular interaction (binding) energy is partitioned as follows:

$$E_{\text{int}} = E_{\text{es}} + (E_{\text{rep}} + E_{\text{disp}}) + E_{\text{pen}} \quad (9)$$

where the electrostatic interaction energy  $E_{\text{es}}$  is calculated according to the Buckingham-type expansion (5), the repulsion ( $E_{\text{rep}}$ ) and dispersion ( $E_{\text{disp}}$ ) forces are approximated using *exp*-6 potentials derived earlier,<sup>35</sup> and  $E_{\text{pen}}$ , the penetration electrostatic contribution, is defined<sup>6</sup> as the interaction of the spherical charge distribution of one molecule with the deformation charge density of the second.<sup>23</sup> The *total* intermolecular bonding energies obtained using stockholder atoms were compared with those calculated from Stone's DMA expansion<sup>22,24</sup> and the generalized scattering factor approach (GSF).<sup>25</sup> The results show good agreement in total binding energies between the Buckingham-type approximation with stockholder, DMA, and GSF moments and those obtained directly from supermolecular calculations, the differences being only about 2 kJ/mol. The inclusion of octapolar and hexadecapolar moments was shown to be an important factor in reaching convergence in the energy. However, no attempts were made to analyze the accuracy of the evaluation of the purely electrostatic component of the total binding energies.

The multipolar model of Hansen and Coppens<sup>11,15</sup> is commonly used in experimental X-ray determinations of charge densities in molecular crystals. The static charge density in the crystal is described by the superposition of aspherical nucleus-centered *pseudoatoms*  $\rho_i^{\text{pseudo}}(\mathbf{r})$

$$\rho_i^{\text{pseudo}}(\mathbf{r}) = \rho_i^{\text{core}}(\mathbf{r}) + \rho_i^{\text{val}}(\mathbf{r}) + \rho_i^{\text{def}}(\mathbf{r}) \quad (10)$$

where the terms  $\rho^{\text{core}}(\mathbf{r})$  and  $\rho^{\text{val}}(\mathbf{r})$  represent the spherically averaged core and valence densities, respectively, and  $\rho^{\text{def}}(\mathbf{r})$  describes the aspherical deformation density. Detailed descriptions of the parameters included in Hansen–Coppens pseudoatom expansion are given in refs. 11 and 15.

The remarkable transferability of the pseudoatom parameters determined from experimental X-ray data has been long noted. In the mid-1990s Lecomtes group reported a databank with the most significant pseudoatom parameters for light atoms, based on ex-

perimental charge density studies of a series of amino acids and oligopeptides.<sup>26</sup> Application of the databank parameters resulted in more accurate thermal and positional parameters than in spherical atom refinements of crystal structures. Also, the electrostatic potential for several molecules, calculated based on the databank parameters only, was shown to agree well with results based on the actual pseudoatom parameters from aspherical refinement of experimental X-ray structure factors.

Koritsanszky et al.<sup>27</sup> pointed out that for such a databank pseudoatom parameters based on the refinement of theoretical single-molecule, rather than experimental structure factors, provide an alternative. With theoretical data it is possible to obtain pseudoatom parameters free of experimental errors and not affected by possible anisotropic nuclear motions. It also allows a great variety of atom types and systems to be studied. Examination of a series of polypeptides showed that pseudoatom parameters are highly transferable and rather invariable on rotation around single bonds in the peptide framework.

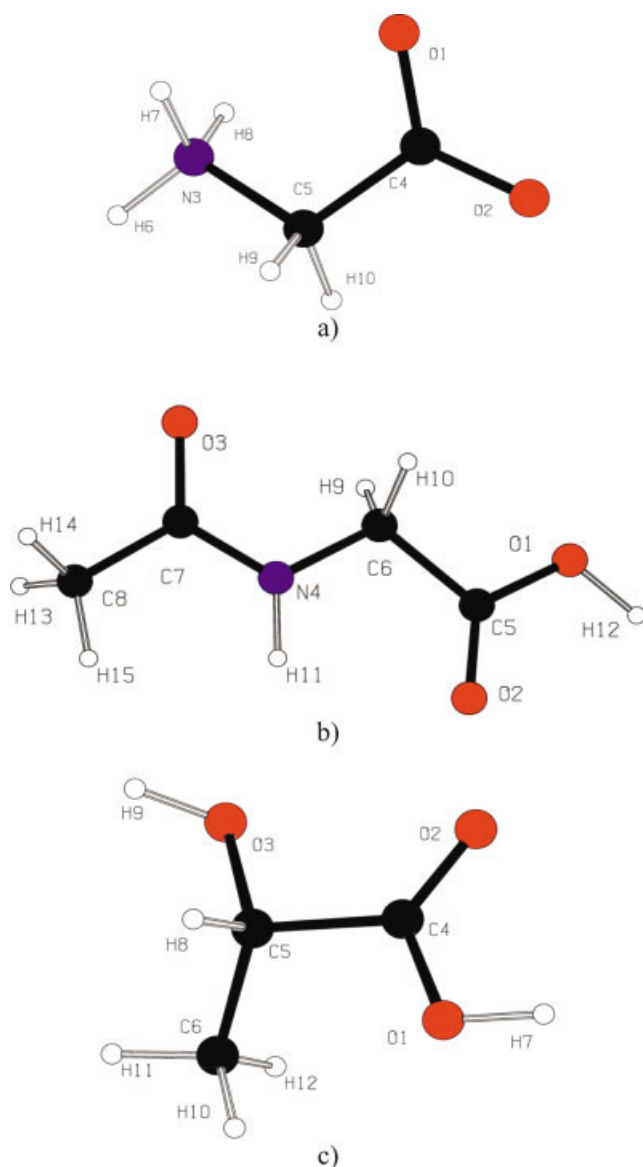
Calculation of electrostatic interactions based on both theoretical and experimental pseudoatoms have been reported by several authors.<sup>23,28–34</sup> However, most of these studies dealt with either the total interaction energies in dimers or with the molecular binding and/or lattice energies in crystals. As in the stockholder moment study by Spackman, the electrostatic interaction was not examined separately but rather as a part of the total interaction energy. With different types of atom–atom potentials<sup>35–37</sup> used to approximate repulsion and dispersion forces the consistency in the calculation of electrostatic energies could not be tested.

As transferability has already been established, the current study is a logical next step in the development of the pseudoatom databank. It deals with the practical application of the databank pseudoatom parameters in the calculation of the monomer molecular moments and the electrostatic contribution to the interaction energies in dimers. The Morokuma–Ziegler energy partitioning scheme included in the density functional theory (DFT) package ADF<sup>38–40</sup> allows the evaluation of the purely electrostatic interaction energy, which should be directly comparable with electrostatic energies calculated via a Buckingham-type expansion. Thus, the ambiguity of using the atom–atom potentials (required for calculation of the total intermolecular binding energy, which is usually compared with supermolecular calculations) is eliminated in this study. To estimate the effect of the partitioning method on the calculation of the total molecular moments and electrostatic interaction energies, the results calculated with databank and model pseudoatoms are compared to those obtained with stockholder and AIM moments derived from the same starting wave functions. This approach allows separation of the effects of the approximations used in calculation of the electrostatic interaction energies from those introduced by the partitioning model itself.

## Computational Details

### Crystal Data

In this study the unoptimized molecular and dimer geometries directly extracted from the crystal structures of  $\alpha$ -glycine (Gly), *N*-acetylglycine (AcG), and L-(+)-lactic acid (Lac) were used.



**Figure 1.** Molecular geometries of (a) Gly, (b) AcG, and (c) Lac. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

Structural parameters for Gly [Fig. 1(a)] ( $C_2H_5NO_2$ , space group  $P2_1/n$ ) were taken from a recent accurate X-ray charge density study at 23 K by Destro et al.<sup>41</sup> The positions of the hydrogen atoms determined in this study agree within 2.3 e.s.d.s with the 15-K neutron results. The crystal structure of Gly shows six distinct dimers (Fig. 2). Four of the six dimers involve N-H...O interactions of  $NH_3^+$  and  $COO^-$  groups. Two are head-to-tail dimers with  $R_{H...O} = 1.75$  (Gly1) and 1.82 Å (Gly2), two are symmetrical dimers with two contacts N-H...O related by a center of symmetry with  $R_{H...O} = 2.39$  and 2.04 Å for Gly3 and Gly4, respectively, while the remaining two are characterized by C-H...O interactions between  $CH_2$  and  $COO^-$  groups, with  $R_{H...O} = 2.41$  and 2.34 Å for dimers Gly5 and Gly6, respectively.

AcG [Fig. 1(b)] ( $C_4H_7NO_3$ , space group  $P2_1/c$ ) crystal data were taken from a room-temperature neutron study.<sup>42</sup> Two distinct dimers in the AcG crystal are illustrated in Figure 3. The first (AcG1) is formed by N-H...O contacts between NH and COOH groups ( $R_{H...O} = 2.07$  Å) and C-H...O ( $R_{H...O} = 2.51$  Å) contacts between  $CH_3$  and COOH groups with the same oxygen O(2) acting as an acceptor. Because of the center of symmetry there are two equivalent sets of contacts in dimer AcG1. In the second dimer (AcG2) there is a short O-H...O interaction ( $R_{H...O} = 1.53$  Å and  $R_{O...O} = 2.56$  Å).

Structural parameters for Lac [Fig. 1(c)] ( $C_3H_6O_3$ , space group  $P2_12_1$ ) were taken from a 100-K low-temperature X-ray study.<sup>43</sup> To compensate for the shortening of X-ray determined bonds, the H atoms were placed at 1.059 Å from C along C—H, 1.015 Å from O in the carbonyl group, and 0.967 Å from O in the hydroxyl group along O—H bonds, which correspond to the averaged C—H and O—H distances in C—C—H<sub>3</sub>, O = C—O—H and C—O—H groups, respectively.<sup>44</sup> Three distinct dimers in the crystal structure of Lac are shown in Figure 4. The first dimer (Lac1) is formed by the O-H...O interaction between hydroxyl and COOH groups with  $R_{O...O} = 2.63$  Å. The second dimer (Lac2) is an example of a bifurcated H bond with the hydroxyl group being a donor and both acceptor oxygen atoms belonging to a COOH group with  $R_{OH...OH} = 2.34$  Å and  $R_{OH...O} = 1.88$  Å. The third dimer is characterized by a CH...O interaction between a secondary CH group and a carboxyl oxygen atom of a COOH group ( $R_{CH...O} = 2.52$  Å).

More detailed information on bond distances, angles, and symmetry operations required to generate symmetry-equivalent molecules in all dimers is given in Table 1.

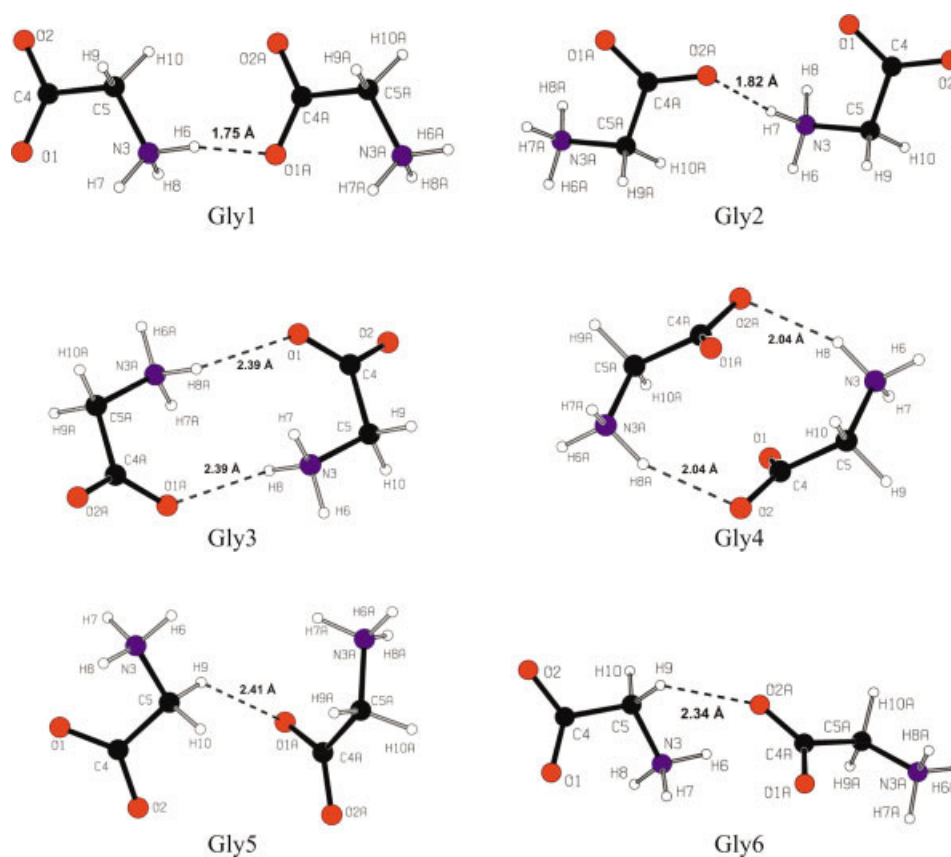
### Theoretical Calculations

Single-molecule calculations were performed with the GAUSSIAN98<sup>45</sup> (G98) and ADF2003.01<sup>38–40</sup> programs at the DFT level. A fundamental difference between the two programs is their use of different primitive functions in the atomic basis set expansions. GAUSSIAN98 uses Gaussian-type functions  $\exp(-\alpha r^2)$  while the Slater-type functions  $\exp(-\alpha r)$  are used in ADF. GAUSSIAN98 calculations employed the split-valence double-exponential 6-31G\*\*<sup>46</sup> and triple-exponential 6-311++G\*\*<sup>47</sup> basis sets with polarization functions, the latter also containing diffuse functions.<sup>48</sup> In ADF the standard double- (DZP) and triple-zeta exponential (TZP) basis sets with polarization functions were used. DFT calculations in GAUSSIAN98 were performed with Becke's three-parameter hybrid method,<sup>49</sup> combined with the nonlocal correlation functional of Lee, Yang and Parr<sup>50</sup> (B3LYP keyword in GAUSSIAN98). In ADF the Hartree–Fock exchange is not implemented, which excludes use of hybrid-type functionals such as B3LYP. Therefore, Becke's 1988 exchange<sup>51</sup> combined with Lee, Yang, and Parr's correlation functional<sup>50</sup> (BLYP) was used in all ADF calculations.

Basis set superposition errors (BSSEs) were accounted for with the counterpoise correction method.<sup>52</sup>

### Multipole Refinements of Theoretical Structure Factors

Complex static valence-only structure factors in range  $0 < \sin\theta/\lambda < 1.1 \text{ \AA}^{-1}$  were obtained by Fourier transform of the molecular charge densities for reciprocal lattice points correspond-



**Figure 2.** Dimers in the crystal structure of Gly. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

ing to a pseudocubic cell with 30-Å edges. In the refinement of the static theoretical data, performed with the XD package,<sup>53</sup> temperature factors and atomic positions were not refined, thus eliminating an important source of correlation between parameters. Both  $\kappa$  and  $\kappa'$  parameters<sup>11</sup> were refined independently for each atom, while chemically similar hydrogen atoms were allowed to share the same  $\kappa$  and  $\kappa'$  parameters. The multipole expansion was truncated at the hexadecapolar level ( $l_{\max} = 4$ ) for the nonhydrogen atoms and at the quadrupole level ( $l_{\max} = 2$ ) for hydrogens, for which only bond-directed multipolar functions<sup>11</sup> with  $l, m = 1, 0$  and  $2, 0$  were refined. To reduce the number of refined parameters, local symmetry constraints were applied to some atoms. A molecular electroneutrality constraint was applied in all refinements.

#### Multipolar DataBank

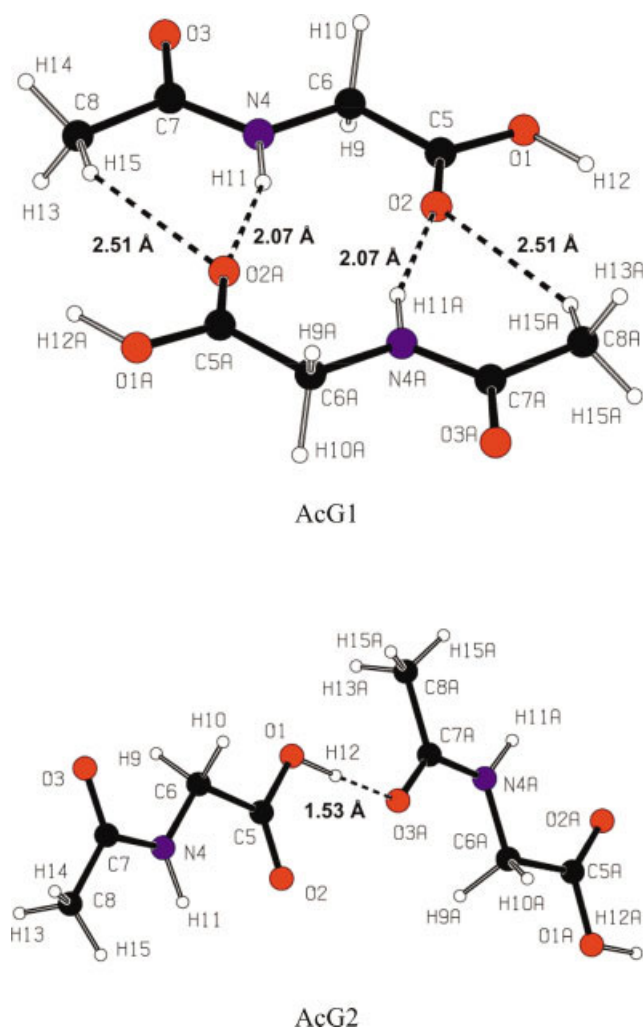
The strategy in building the multipolar databank has been discussed elsewhere.<sup>27</sup> However, in contrast to the previous study, in which molecular geometries were optimized at the molecular mechanics level, all single-molecule and dimer calculations used unoptimized crystal molecular geometries from accurate X-ray or neutron data extracted from the Cambridge Structural Database.<sup>54</sup> Every entry in the databank contains a full set of pseudoatom parameters sufficient to represent an atom within the Coppens–Hansen pseudoatom formalism, averaged over a number of calcu-

lations on different molecules. Only parameters statistically different from zero are included for each atom.

#### Calculation of Atomic Moments

For both GAUSSIAN98 and ADF wave functions calculation of Cartesian unbridged atomic moments<sup>11,55</sup> based on the stockholder partitioning method was performed using a FORTRAN77 program specifically written for this study. The free atomic wave functions of H, C, N, and O atoms for the most stable electronic states were calculated separately for each functional and basis set. The integration of the stockholder atoms was implemented using an adaptive subroutine DCUHRE<sup>56</sup> (module 698 in the public domain TOMS package)<sup>57</sup> for the numerical integration of a vector of integrals over a given hyperrectangular region. The integration for each atom was done in a cubic volume of dimensions  $6 \times 6 \times 6$  Å with the nucleus located in the center of the cube, giving a numerical integration accuracy of about  $10^{-4}$  a.u. The integration was done separately for each atomic moment while forcing the adaptive numerical integration to approach the desired accuracy.

For GAUSSIAN98 wave functions atomic moments based on the AIM partition were calculated in the program PROAIMV,<sup>58</sup> locally modified to integrate Cartesian unbridged (instead of traceless) moments, and expanded to include  $l = 3$  and 4 moments. Determination and integration of AIM atomic basins from ADF



**Figure 3.** Dimers in the crystal structure of AcG. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

wave functions were performed with a newly developed program TOPADF, which is based on the programs TOPOND98<sup>59</sup> and TOPXD.<sup>60</sup> In TOPADF the first and second derivatives of the density are evaluated analytically. As in case of PROAIMV, the Cartesian unabridged moments with  $l \leq 4$  were integrated.

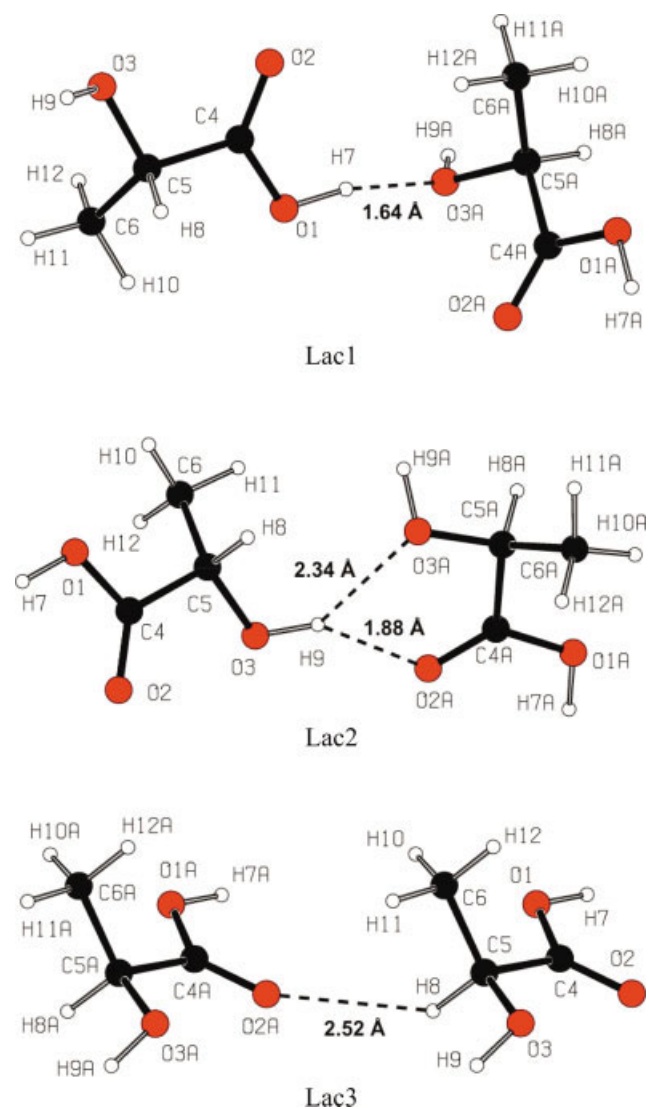
The computation of unabridged stockholder moments from the Hansen–Coppens pseudoatom model was performed with a new version of the XDPROP program,<sup>53</sup> which uses the subroutine DCUHRE for numerical integration, following the same procedure as described above.

The integrated Cartesian unabridged moments were subsequently converted to traceless and, finally to spherical tensor form.<sup>10</sup>

#### Calculation of Electrostatic Interaction Energies in Dimers Using Atomic Moments

The calculation of the electrostatic interaction energies between monomers in dimers was performed using the program MIN16.<sup>61</sup>

The program uses the Buckingham-type expansion in terms of Cartesian traceless atomic moments, although it requires spherical moments in its input. The program can calculate all interaction tensors up to order  $L = l_1 + l_2 = 8$  with moments up to the hexadecapolar level ( $l = 4$ ) (interactions with  $L = 5 + 3, 6 + 2, 7 + 1$ , and  $8 + 0$  are not supported). Results obtained with MIN16 are in an excellent agreement (within 1 kJ/mol) with those from ORIENT3.2,<sup>62</sup> which uses spherical moments but is limited to interactions with  $L = 5$ , suggesting that interactions of order  $L = 4 + 2, 3 + 3, 4 + 3$ , and  $4 + 4$  either do not have a significant contribution to the electrostatic interaction energy for the selected systems or cancel each other. Nevertheless, these interactions were included in all calculations.



**Figure 4.** Dimers in the crystal structure of Lac. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

**Table 1.** Geometric Parameters of Dimers in Crystal Structures of Gly, AcG, and Lac.

Dimer	D	H	A	Symmetry code	D—H (Å)	H ··· A (Å)	D ··· A (Å)	D—H ··· A (°)
Gly1	N(3)	H(6)	O(1)	$x, y, -1 + z$	1.04	1.75	2.77	167
Gly2	N(3)	H(7)	O(2)	$1 + x, y, z$	1.02	1.82	2.84	170
Gly3 <sup>a</sup>	N(3)	H(8)	O(1)	$1 - x, -y, 2 - z$	1.02	2.39	2.94	113
Gly4 <sup>a</sup>	N(3)	H(8)	O(2)	$-x, -y, 2 - z$	1.02	2.04	3.01	156
Gly5	C(5)	H(9)	O(1)	$-1/2 + x, 1/2 - y, -1/2 + z$	1.08	2.41	3.30	138
Gly6	C(5)	H(9)	O(2)	$1/2 + x, 1/2 - y, -1/2 + z$	1.08	2.34	3.22	138
AcG1	N(4)	H(11)	O(2)	$-x, 1 - y, 1 - z$	1.01	2.07	3.04	162
AcG2	C(8)	H(15)	O(2)	$-1 + x, 1/2 - y, 1/2 + z$	1.03	2.51	3.37	141
Lac1	O(1)	H(7)	O(3)	$1/2 - x, -y, -1/2 + z$	1.01	1.64	2.63	167
Lac2	O(3)	H(9)	O(2)	$1/2 + x, 1/2 - y, 1 - z$	0.97	1.88	2.71	143
Lac3	O(3)	H(9)	O(3)	$1 + x, y, z$	0.97	2.34	3.15	141
Lac3	C(5)	H(8)	O(2)	$1 + x, y, z$	1.10	2.52	3.37	133

<sup>a</sup>Symmetrical dimers, i.e., there are two identical H bonds related by the center of symmetry.

## Results

The flowchart of the procedure followed is depicted in Figure 5. Results of each of the steps are discussed below. The most important conclusions are summarized at the end of the article.

### Total Molecular Moments

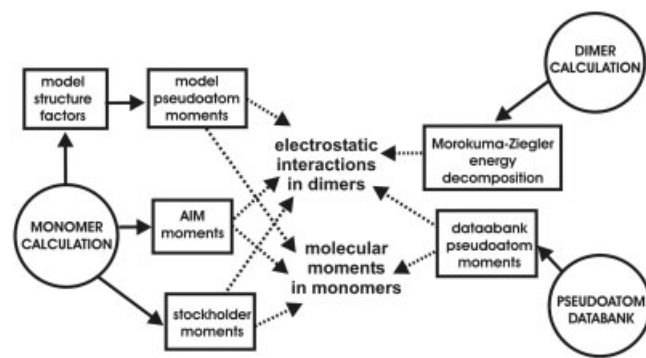
For a given charge density distribution of an isolated molecule the molecular moments should be independent of the type of partitioning used to separate individual atomic contributions (Tables S1–S3). Because it is not possible to obtain the unabridged moments of order 2 and higher directly from pseudoatom parameters, we decided to apply the stockholder partitioning for charge density distributions calculated with the pseudoatom model. Below, moments referred to as obtained from the pseudoatom model are those calculated via stockholder partitioning of the pseudoatom electron densities. To our knowledge this is the first time unabridged

moments of order 2 and higher based on the pseudoatom model are compared with those calculated directly from the wave function.

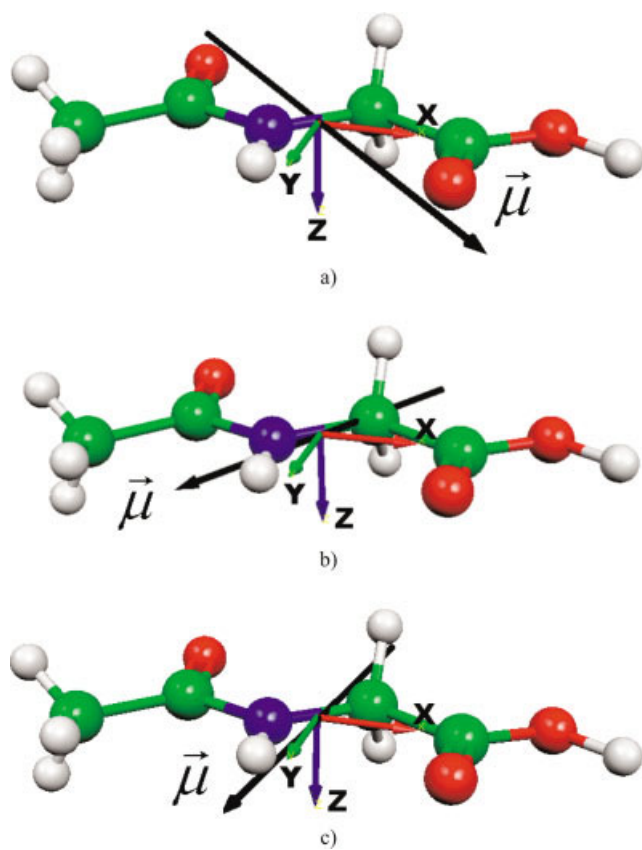
Inspection of Tables S1–S3 shows that summations of individual atomic stockholder, pseudoatom and AIM moments based on original wave function densities yield essentially the same total molecular moments as calculated analytically from corresponding wave functions except for Lac and AcG, for which exceptions occur both in the case of pseudoatom moments from aspherical atom refinements of theoretical structure factors and for the databank.

For Gly all molecular moments based on the pseudoatom model agree well with those calculated by other methods. For the dipole moment (including all three of its components) the agreement is almost perfect. Even for moments of orders 2–4 the observed differences are small and the sign of the components is usually the same. The most striking feature of Table S1 is the fact the molecular moments calculated from the databank pseudoatoms are of the same order of accuracy as the model pseudoatoms.

The differences between pseudoatom and other partitioning methods for Lac and AcG are more pronounced. It is disturbing that the X component of the AcG molecular dipole moment calculated from model pseudoatoms, although small, has the opposite sign than that from the other methods. For the other two components (Y and Z), the signs are correctly retained but the values are slightly underestimated. As a result, the magnitude of the small molecular dipole moment vector in AcG from model pseudoatoms (1.8 D) is underestimated compared to other methods (2.8–3.0 D) and is almost perpendicular to that from the original wave function (Fig. 6). All three components of the dipole moment vector calculated from databank pseudoatoms are closer to the exact values than to those from model pseudoatoms, although the X component still has the incorrect sign. As a result, the direction of the dipole moment vector calculated from the databank lies somewhere between the exact and the model directions. In Lac



**Figure 5.** Flowchart of the study.



**Figure 6.** Molecular dipole moment  $\vec{\mu}$  in AcG calculated from (a) the wave function, (b) pseudoatoms after the refinement of model structure factors, and (c) databank. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

with both model and databank pseudoatoms the total molecular moments, 4.2 and 4.0 Debye, respectively, are slightly overestimated compared to the other methods (3.0–3.4 Debye) because of a small overestimation of each of the X, Y, and Z components. The

second molecular moments of AcG and Lac calculated from both databank and model pseudoatoms are in good agreement with the exact values, while significant differences are observed for some of the components of the third and fourth moments. Again, the databank performs at the same level of accuracy as the model pseudoatoms and the discrepancies with the databank are observed for exactly the same components of the higher moments as with the model pseudoatoms. The origin of the discrepancies seems to be related to the procedure used in the aspherical atom refinement. It is clear that the fitting of the pseudoatom parameters to the structure factors (Fourier transform of the density) in reciprocal space does not always provide as accurate a description of the molecular density in real space (in terms of molecular moments at least) as do the other partitioning methods, which operate in real space. The fact that the model structure factors are free of errors and thermal motion effects does not seem to improve the performance. However, it may be assumed that the differences in third and fourth moments will not have any significant contributions to the electrostatic energies calculated with the Buckingham expression because these contributions are usually small, so the accuracy in determination of first and second moments is much more important.

The bias in the density introduced by the pseudoatom model upon refinement of structure factors has been discussed before.<sup>60,63–67</sup> The bias manifests itself in the topological properties (such as net charges and volumes and properties of  $\rho$  at critical points) of the density represented by the pseudoatom model when compared to the model density directly calculated from the wave function. Its origin has been attributed to the limited flexibility of the single-exponential radial functions used to describe the deformation part of the density in the pseudoatom model.

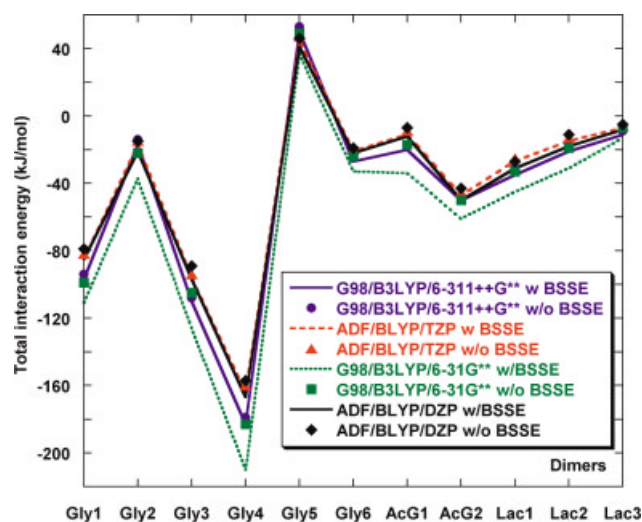
#### Total Interaction Energies

Table 2 and Figure 7 summarize the total dimer interaction energies  $E_{\text{int}}$  for all calculations. In general all calculations employing DZP basis sets uncorrected for BSSE overestimate  $E_{\text{int}}$ , especially with the Gaussian basis set 6-31G\*\*, for which the average BSSE is 13 kJ/mol and the largest BSSE of 27 kJ/mol is observed for

**Table 2.** Total Interaction Energies (kJ/mol) of Dimers in Test Crystal Structures Corrected for BSSE (Numbers in Parentheses).

Dimers	G98/B3LYP/ 6-311++G**	ADF/BLYP/TZP	G98/B3LYP/ 6-31G**	ADF/BLYP/ DZP
Gly1	−94 (−13)	−82 (−3)	−99 (−12)	−79 (−6)
Gly2	−14 (−3)	−15 (−2)	−22 (−15)	−15 (−6)
Gly3	−108 (−2)	−94 (−2)	−105 (−21)	−89 (−8)
Gly4	−179 (−4)	−160 (−4)	−183 (−27)	−157 (−10)
Gly5	+53 (−3)	+47 (−2)	+49 (−12)	+46 (−5)
Gly6	−25 (−2)	−20 (−1)	−24 (−9)	−19 (−2)
AcG1	−17 (−3)	−9 (−2)	−17 (−17)	−7 (−5)
AcG2	−49 (−1)	−45 (−2)	−50 (−10)	−43 (−7)
Lac1	−31 (−4)	−25 (−2)	−33 (−12)	−27 (−4)
Lac2	−18 (−3)	−13 (−2)	−19 (−12)	−11 (−7)
Lac3	−9 (−2)	−6 (−1)	−7 (−6)	−5 (−3)





**Figure 7.** Total interaction energies from dimer calculations. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

dimer Gly4. On the other hand, for the Slater-type DZP basis set the BSSEs are much smaller—about 6 kJ/mol on average and at most 10 kJ/mol also for the dimer Gly4. These results are in agreement with those published before<sup>68</sup> in which the double-zeta Slater basis set with polarization functions was shown to produce a smaller BSSE compared to Gaussian 6-31G\*\* basis sets for a number of H-bonded neutral dimers of small molecules. For the triple-zeta basis sets the BSSEs are almost negligible (2–4 kJ/mol) and the difference in BSSEs between Slater and Gaussian functions practically vanishes.

Differences in total interaction energies between GAUSSIAN98 and ADF calculations appear more significant than the BSSEs. These differences arise from the different type and number of primitive functions and from the different DFT functionals used in the GAUSSIAN98 and ADF calculations. Indeed, when comparing the triple-zeta G98/B3LYP/6-311++G\*\* and ADF/BLYP/TZP calculations the former seems to produce larger absolute values for  $E_{\text{int}}$  in all dimers (except Gly2), which defines stronger attractive and stronger repulsive interactions. For Gly2 both calculations predict essentially the same interaction energy of 14–15 kJ/mol. The differences seem to increase with the absolute values of  $E_{\text{int}}$ , reaching a maximum of 19 kJ/mol for Gly4, with the largest  $E_{\text{int}}$  –179 kJ/mol for G98 and –160 kJ/mol for the ADF results. However, the mean difference between triple-zeta G98 and ADF calculations for all 11 dimers is only about 7 kJ/mol.

Because it was not possible to obtain the electrostatic interaction energies of monomers in dimers, defined by expression (2), directly from the G98 calculations, we will concentrate on our comparison of  $E_{\text{int}}$  with  $E_{\text{es}}$  in the ADF calculations. In general, the electrostatic energy makes a significant contribution and when attractive is always larger in magnitude than the total interaction energy (Table 3). The other two contributions to the total interaction energy (based on Morokuma–Ziegler energy partitioning), exchange repulsion and orbital interaction, are also large but virtually cancel each other for several dimers (Gly4, Gly5, Gly6,

and Lac3), which brings  $E_{\text{es}}$  within 10 kJ/mol of the  $E_{\text{int}}$  in these cases. Dimer Gly5 is the only one in this study for which the electrostatic interaction shows a significant repulsive character. Strong exchange repulsion forces (55–133 kJ/mol) are observed in all other dimers, yet the sum of  $E_{\text{es}}$  and  $E_{\text{oi}}$  outweighs the repulsion and the total interaction character in these dimers is attractive. These results are in qualitative agreement with the previously performed studies of the energy decomposition in DNA bases<sup>69</sup> using Morokuma–Ziegler energy decomposition and with a number of small molecule organic dimers with the Kitaura–Morokuma approach.<sup>70</sup> In the first study the reported absolute values of  $E_{\text{es}}$ ,  $E_{\text{Pauli}}$ , and  $E_{\text{oi}}$  were smaller and the  $E_{\text{es}}$  values were consistently closer to  $E_{\text{int}}$  (within 6–18 kJ/mol) compared to the energies in this study. In the second study, the absolute values of the energies were much smaller than those calculated in this study, yet  $E_{\text{es}}$  always accounted for at least 50% of the total interaction energy.

Thus, our results confirm that  $E_{\text{es}}$  provides a good estimate of the bonding strength in the selected molecular dimers, although the use of  $E_{\text{es}}$  alone would overestimate the attractive component of the total interaction energy.

### Electrostatic Interaction Energies

The relative order of the electrostatic interaction energies (attractive or repulsive) for all dimers determined with the Morokuma–Ziegler method is independent of the basis sets in the cases examined (Table 4). The exact values calculated with ADF at the BLYP level with the TZP and DZP basis sets differ from each other only by about 4 kJ/mol on average and at most 9 kJ/mol for dimer AcG2, the TZP values being always lower. Such an agreement suggests that the electrostatic interactions in selected systems based on the unperturbed densities of the isolated molecules are not much dependent of the basis set chosen, at least when based on the exact expression. In the following we will use the results from the TZP calculations for the exact reference values of  $E_{\text{es}}$  as they should be more accurate than those from the DZP basis.

In general, the Buckingham-type calculations of  $E_{\text{es}}$  predict the correct relative electrostatic interaction energies in the majority of cases with the exception of the pair of dimers Gly2 and Gly6, for

**Table 3.** Contributions to the Total Dimer Interaction Energies (kJ/mol) According to the Morokuma–Ziegler Decomposition for ADF/BLYP/TZP Calculations.

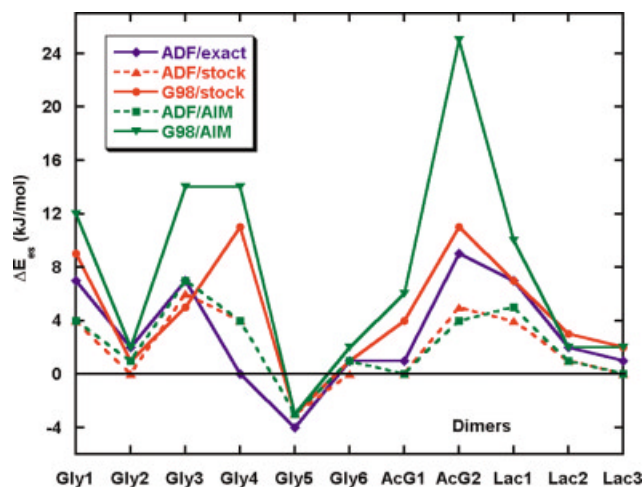
Dimer	$E_{\text{es}}$	$E_{\text{Pauli}}$	$E_{\text{oi}}$	$E_{\text{Pauli}} + E_{\text{oi}}$	$E_{\text{int}}$
Gly1	–115	+89	–56	+33	–82
Gly2	–37	+69	–48	+21	–15
Gly3	–109	+54	–40	+14	–94
Gly4	–165	+84	–79	+5	–160
Gly5	+39	+22	–14	+8	+47
Gly6	–27	+18	–12	+6	–20
AcG1	–48	+60	–21	+39	–9
AcG2	–99	+133	–79	+54	–45
Lac1	–70	+95	–50	+45	–25
Lac2	–44	+55	–24	+31	–13
Lac3	–13	+11	–3	+8	–6

**Table 4.** Electrostatic Interaction Energies Between Two Monomers in Each of the Dimers (kJ/mol).

Dimer	Buckingham-type approach											
	Morokuma–Ziegler		Stockholder moments				AIM moments				Pseudoatom moments	
	ADF-1	ADF-2	ADF-1	ADF-2	G98-1	G98-2	ADF-1	ADF-2	G98-1	G98-2	G98-2	Databank
Gly1	-115	-108	-86	-82	-101	-92	-92	-88	-106	-94	-83	-86
Gly2	-37	-35	-15	-15	-15	-14	-16	-15	-15	-13	-10	-4
Gly3	-109	-102	-74	-68	-87	-82	-95	-88	-110	-96	-78	-83
Gly4	-166	-165	-134	-130	-154	-143	-142	-138	-163	-149	-134	-133
Gly5	+43	+35	+47	+44	+55	+52	+48	+45	+55	+52	+53	+53
Gly6	-26	-26	-20	-20	-23	-22	-24	-23	-26	-24	-20	-19
AcG1	-48	-47	-27	-27	-31	-27	-31	-31	-35	-29	-27	-25
AcG2	-99	-90	-68	-63	-74	-63	-92	-88	-94	-69	-30	-31
Lac1	-70	-63	-46	-42	-52	-45	-53	-48	-56	-46	-36	-28
Lac2	-44	-42	-27	-26	-30	-27	-25	-24	-28	-26	-21	-21
Lac3	-13	-14	-9	-9	-11	-9	-9	-9	-11	-9	-16	-15

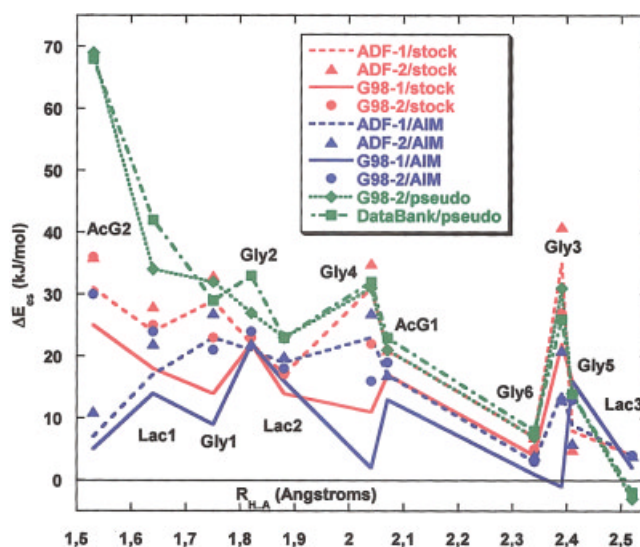
ADF-1, ADF/BLYP/TZP; ADF-2, ADF/BLYP/DZP; G98-1, GAUSSIAN98/B3LYP/6-311++G\*\*; G98-2, GAUSSIAN98/B3LYP/6-31G\*\*.

which the relative order of electrostatic energies is always reversed compared to the exact values. For example, the symmetrical dimer Gly4 is correctly predicted to be the most stable in terms of  $E_{es}$ , by all methods, with a range of  $-130$  to  $-166$  kJ/mol. The only dimer with a repulsive electrostatic interaction, Gly5, is also identified by all methods ( $E_{es} = 35$ – $55$  kJ/mol). In general, when attractive, the electrostatic interaction energies calculated with the Buckingham-type expansion are somewhat underestimated compared with the exact values. The magnitude of underestimation varies and seems to be dependent on several factors, including the geometry of the dimers and, more importantly, the partitioning method itself.

**Figure 8.** Differences in electrostatic interactions energies (kJ/mol) between DZP and TZP calculations for all test dimers.

#### Basis Set Effects

Inspection of Table 4 shows the basis set dependence of the  $E_{es}$  calculated with the Buckingham-type expansion (Fig. 8). Indeed, for both stockholder and AIM moments the double-zeta basis sets always underestimate  $E_{es}$  compared to the triple-zeta sets, exactly as for exact  $E_{es}$  energies, although the discrepancies between triple-zeta and double-zeta values can be larger in the Bucking-

**Figure 9.** Difference between the exact electrostatic interaction energy and that calculated from the Buckingham-type expansion ( $\Delta E_{es}$ ) for different atomic partitions as a function of the closest intermolecular contact. Calculations: ADF-1, ADF/BLYP/TZP; ADF-2, ADF/BLYP/DZP; G98-1, GAUSSIAN98/B3LYP/6-311++G\*\*; G98-2, GAUSSIAN98/B3LYP/6-31G\*\*.

**Table 5.** Total Charge- $l$  Contributions to the  $E_{es}$  Calculated From the Buckingham-Type Expansion From AIM Atomic Moments Based on GAUSSIAN98 B3LYP/6-311++G\*\* (First Row) and B3LYP/6-31G\*\* (Second Row) Densities for Selected Dimers.

Dimers	Charge-charge	Charge-dipole	Charge-quadrupole	Charge-octapole	Charge-hexadecapole
Gly1	-144	<b>62</b>	-4	-12	-4
	-144	<b>69</b>	-4	-9	-2
Gly3	-133	<b>17</b>	7	0	-2
	-133	<b>31</b>	5	0	-1
Gly4	-230	<b>92</b>	-23	-1	-4
	-233	<b>104</b>	-22	-1	-4
AcG2	<b>-97</b>	<b>36</b>	4	-18	<b>-11</b>
	<b>-103</b>	<b>51</b>	5	-15	<b>-2</b>
Lac1	-80	<b>49</b>	<b>-15</b>	-4	-2
	-83	<b>58</b>	<b>-10</b>	-2	-2
Gly2	14	-19	-4	-7	-1
	10	-19	-1	-4	-1

The largest discrepancies appear in bold.

ham-type analysis. Thus, for negative  $E_{es}$  the triple-zeta energies are closer to exact while for the dimer Gly5 in which the  $E_{es}$  is positive the double-zeta calculations agree better with the exact  $E_{es}$ . In general, for ADF moments the differences between triple-zeta and double-zeta  $E_{es}$  are relatively small—on average about 2 kJ/mol for stockholder atoms and 3 kJ/mol for AIM partitioning. The largest discrepancy is only 7 kJ/mol for the dimer Gly3 from ADF/AIM calculation. These differences are comparable to those obtained for the exact  $E_{es}$  calculated in ADF with TZP and DZP basis sets. For GAUSSIAN98 moments from stockholder partitioning the differences between  $E_{es}$  from triple-zeta and double-zeta basis sets are also not large: at most 11 kJ/mol for dimers Gly4 and AcG2 and on average about 5 kJ/mol. The most significant basis set dependence of  $E_{es}$  calculated with the Buckingham-type expansion is observed with GAUSSIAN98 AIM moments. The largest discrepancy between triple-zeta and double-zeta is as large as 25 kJ/mol for dimer AcG2 and larger than or equal to 10 kJ/mol for dimers Gly1, Gly3, Gly4, and Lac1 (10–14 kJ/mol). Because the discrepancies for other dimers are small, the average difference in  $E_{es}$  between triple-zeta and double-zeta basis sets for GAUSSIAN98 AIM moments is only 8 kJ/mol. Detailed analysis of the contributions to  $E_{es}$  (Table 5) shows that these discrepancies mainly result from different charge-dipole contributions for Gly1, Gly3, Gly4, and Lac1 dimers and also a different charge-hexadecapole term for the AcG2 dimer. These contributions almost completely account for observed differences in the  $E_{es}$ . It should be noted that these five dimers are either the ones with the shortest H...A distances (AcG2, Lac1, Gly1) or symmetrical dimers (Gly3 and Gly4). This indicates a pronounced effect of the dimer geometry that will be discussed below. In general, our results seem to contradict claims of the relative basis set independence of the AIM partitioning method.<sup>71</sup> Our results, however, show that stockholder moments are less basis set dependent than the AIM moments as clearly indicated by the calculated intermolecular electrostatic energies.

#### Effect of Partitioning

The choice of density partitioning method from which the atomic moments are calculated is clearly the most important factor affecting the calculation of  $E_{es}$  from the Buckingham-type expansion. Despite the fact that in most of the cases the atomic moments correctly add up to the total molecular moment, as discussed before, the individual atomic moments of an atom may differ greatly among the partitioning methods,<sup>72</sup> as illustrated in Table 6 for the net atomic charges. On the other hand, the pseudoatom and stockholder charges, both based on fuzzy boundary electron density partitioning, are similar but differ considerably from the discrete boundary AIM charges, especially for the carbon and oxygen atoms of  $\text{COO}^-$  and the nitrogen atom of  $\text{NH}_3^+$  groups.

For the same molecular density, the  $E_{es}$  calculated with AIM moments are always (in four of four calculations, i.e., ADF/BLYP/TZP, ADF/BLYP/DZP, G98/B3LYP/6-311++G\*\*, and G98/B3LYP/6-31G\*\*) closer to the exact values than those calculated with

**Table 6.** Net Atomic Charges ( $e$ ) From AIM, Stockholder, and Pseudoatom Partitioning of GAUSSIAN98/B3LYP/6-31G\*\* Molecular Density of Gly.

Atom	$q_{\text{AIM}}$	$q_{\text{stockholder}}$	$q_{\text{pseudoatom}}$
O1	-1.26	-0.41	-0.31
O2	-1.21	-0.40	-0.30
N3	-1.09	+0.06	-0.12
C4	+1.81	+0.11	-0.05
C5	+0.29	+0.01	+0.30
H6	+0.43	+0.17	+0.19
H7	+0.49	+0.17	+0.21
H8	+0.48	+0.18	+0.21
H9	+0.02	+0.05	-0.06
H10	+0.03	+0.05	-0.07

Table 7. Electrostatic Interaction Energies (kJ/mol) in Dimers of Gly From Different Methods.

Dimer	Morukuma–Ziegler ADF/BLYP/TZP	Gavezzotti <sup>a</sup> G98/MP2/6- 31G** total density	Buckingham-type expression			
			Stockholder G98/ B3LYP/6- 311++G**	AIM G98/ B3LYP/6- 311++G**	Pseudoatoms experiment <sup>b</sup>	Pseudoatoms G98/B3LYP/ 6-31G**
Gly1	-115	-122	-101	-106	-139	-83
Gly2	-37	-23	-15	-15	+2	-10
Gly3	-109	-108	-87	-110	-149	-78
Gly4	-166	-152	-154	-163	-231	-134
Gly5	+43	+57	+55	+55	+94	+53
Gly6	-26	-22	-23	-26	-38	-20

<sup>a</sup>Ref. 8a.<sup>b</sup>Ref. 41.

stockholder moments. The average differences between the exact  $E_{es}$  and values calculated with AIM moments are 9, 14, 16, and 17 kJ/mol for G98/6-311++G\*\*, ADF/TZP, ADF/DZP, and G98/6-31G\*\* densities, respectively, while for the stockholder moments the discrepancies are 15, 21, 23, and 20 kJ/mol for the same densities. For GAUSSIAN98 B3LYP/6-31G\*\* calculation, the only one for which the pseudoatom moments have been also obtained, the accuracy of calculated  $E_{es}$  decreases in the sequence AIM > stockholder > pseudoatoms. The largest discrepancies between exact and calculated  $E_{es}$  for each partitioning type are 30 kJ/mol in AcG2, 41 kJ/mol in Gly3, and 69 kJ/mol in AcG2 for AIM, stockholder, and pseudoatom moments, respectively. This suggests that with the Buckingham-type expansion more accurate intermolecular  $E_{es}$  are obtained when using atomic moments based on AIM partitioning, especially those extracted from a high-quality calculation such as B3LYP/6-311++G\*\*. Stockholder and, especially, pseudoatom moments give much less accurate values for the intermolecular  $E_{es}$ , although the sequence of the values is in general reproduced.

As the atomic moments are dependent on the partitioning method (see, e.g., Table 6) it is important to analyze the individual multipole contributions to the electrostatic interaction energy  $E_{es}$  (Tables S4–S14). In all cases the most important contributions are from charge–charge and charge–dipole interactions. For all dimers with an attractive  $E_{es}$  the contribution of charge–charge terms is also attractive, although the absolute values are much larger for AIM moments because of the larger net atomic charges. The large attractive charge–charge term with AIM moments is opposed by strongly repulsive charge–dipole term, while for stockholder and pseudoatoms this term is attractive and for some dimers (AcG1, AcG2, Lac2, and Lac2) even dominates the charge–charge term. The other important, although much less significant, terms for stockholder and pseudoatoms are dipole–dipole (always attractive), charge–quadrupole (mostly attractive), and dipole–quadrupole (mostly attractive). For AIM moments the convergence is clearly much slower than for fuzzy boundary moments with even hexadecapole–hexadecapole contributions being significant for some of the dimers (AcG1 and AcG2). In the electrostatically repulsive dimer Gly5 the charge–charge contribution is strongly repulsive for all methods. Despite significant differences in individual contributions, the total electrostatic interaction energies

calculated with the Buckingham-type expression agree well among different partitioning methods for most of the dimers. The largest disagreement is observed in dimers with short intermolecular contacts, indicating that the geometry may affect the adequacy of a particular approach.

#### Geometry Effects

In Figure 9 the differences between exact  $E_{es}$  and values calculated with the Buckingham-type approach are plotted versus the shortest intermolecular contact for each of the dimers. Inspection of the figure shows that, as for the basis set differences, the largest discrepancies tend to occur at the shortest distances (1.5–1.8 Å) as well as for symmetrical dimers Gly3 (2.04 Å) and Gly4 (2.39 Å) for all but AIM moments. With pseudoatoms (both model and databank) the discrepancies with exact energies in these cases can be as large as 68–69 kJ/mol. For longer interactions the performance of pseudoatoms becomes closer to both stockholder and AIM moments.

#### Pseudoatom DataBank

One of the main goals of this study is to examine the performance of the theoretical pseudoatom databank when applied to calculation of electrostatic interaction energies. Inspection of the last two columns in Table 4 and Figure 9 clearly shows the results obtained from databank pseudoatoms to be comparable in accuracy with the energies calculated from model pseudoatoms. The average difference is only 3 kJ/mol and the largest discrepancy is only 6 kJ/mol for dimer Gly2. In some cases the  $E_{es}$  calculated from databank pseudoatoms are even closer with the exact energies than those from model pseudoatoms, which again demonstrates the remarkable transferability of pseudoatom parameters.

#### Comparison with Gavezzotti's Method and with Experimental Values for Gly

Table 7 summarizes the electrostatic interaction energies in dimers of Gly for selected calculations performed in this study, those published in the literature based on a fully theoretical approach

pioneered by Gavezzotti [semiclassic density sum (SCDS)],<sup>8</sup> and those from experimental pseudoatom moments obtained from an accurate X-ray charge density study using the Buckingham-type approximation.<sup>41</sup> It is fascinating to find a good agreement for electrostatic interaction energies obtained with the theoretical Morokuma–Ziegler, Gavezzotti, and Buckingham (based on theoretical AIM B3LYP/6-311++G\*\* moments) approaches despite the significant differences in Hamiltonians and basis set expansions used for monomer calculations. The agreement for the Buckingham-type approach using stockholder moments is slightly worse than for AIM moments. While the increased interaction energies obtained with experimental pseudoatoms may be attributed to an additional polarization of the molecular charge distribution in the crystal because of intermolecular interactions,<sup>8</sup> this explanation cannot be used to explain the discrepancies between theory and results obtained with theoretical (both model and databank) pseudoatoms that have been obtained based on unperturbed ground-state molecular charge distributions.

## Discussion and Conclusions

In this study we investigated the calculation of electrostatic and total intermolecular interaction energies in a total of 11 dimers in 3 different molecular systems. The electrostatic interaction energies were calculated exactly according to expression (2), as implemented in ADF within the Morokuma–Ziegler energy decomposition approach, and with the Buckingham-type approximation (5) using atomic moments derived from three different partitioning methods: stockholder, AIM, and pseudoatoms.

Comparison of the total and electrostatic interactions energies suggests that  $E_{es}$  alone can be used to estimate the relative strength of bonding in dimers despite the fact that it overestimates the strength of attractive and underestimates the strength of repulsive interactions.

Calculations based on the Buckingham-type approximation in general underestimate the attractive and slightly overestimate repulsive electrostatic interactions (although for the latter case there is only one example in our study). However, the discrepancies in  $E_{es}$  produced by the AIM moments are much smaller than those with other partitioning schemes, especially for dimers AcG2 and Lac1, which have the shortest intermolecular contacts (1.53 and 1.64 Å, respectively) among all the systems examined in this study, and the symmetrical dimers Gly3 and Gly4 with two identical, but relatively long, intermolecular contacts at 2.04 and 2.39 Å. This is attributed to the fundamental differences in the definition of the atoms (and thus the atomic moments) in different partitioning methods—discrete for AIM and overlapping for stockholder and pseudoatoms, the latter being the most delocalized. This can account for large discrepancies in  $E_{es}$  at shortest distances and symmetrical dimers. Nevertheless, the pseudoatoms derived from theoretical densities are successful in predicting the *relative* strength of the electrostatic interactions in the test dimers.

On the other hand, the pseudoatom partitioning is rather attractive as the pseudoatoms not only can be used to reconstruct the molecular/crystal electron densities via simple analytic expressions, which allows electrostatic and topological properties be evaluated and studied, but also are highly transferable. The latter is clearly demonstrated by the fact that the molecular moments in monomers and electrostatic

interaction energies of dimers calculated with databank pseudoatoms are in good agreement with those calculated from model pseudoatoms. The problem of the pseudoatoms not correctly determining the total molecular moments in some cases can be attributed to the ambiguity in their determination via a fitting procedure in the discrete reciprocal space (i.e., aspherical refinement of structure factors). We find that even a very good fit to the structure factors (3% on the valence-only structure factors and about 0.6% on the structure factors calculated from the total density) does not always result in the correct multipole moments and can introduce a bias in the total molecular density distribution. This bias is even observed with theoretical model data in which the nuclei positions are fixed, no thermal smearing of the charge density is present, and the structure factors are calculated with 100% accuracy and up to any  $\sin\theta/\lambda$  cutoff. Determination of the pseudoatom parameters by fitting to the model theoretical density in direct space may be preferable and is to be explored. Such an approach would involve a minimization of an integral of the form  $\int(\rho_{wfn} - \rho_{mult})^2 dv$  with optimization schemes such as the Newton–Raphson, steepest descent, or conjugated gradient method.<sup>73</sup> Another possible way to raise the accuracy in the determination of  $E_{es}$  with pseudoatoms to that with the stockholder moments is to first partition the molecular density into stockholder atoms and then fit the pseudoatom parameters to the numerically evaluated stockholder atoms. With such a procedure the multipole atomic and molecular moments will be the same, within the accuracy of the fit, as in stockholder partitioning.<sup>74</sup>

The stockholder atoms can also be used to reconstruct the total molecular density, but because of their numerical nature cannot be used for such purpose as easily as pseudoatoms. It is highly unlikely that the AIM atoms can be used to reconstruct the total density because these are discrete quantities and may leave empty spaces (gaps) in the synthesized density. In addition, because of their complicated numerical nature they cannot be readily represented analytically.

An important result of this study is the good agreement between electrostatic interaction energies in Gly dimers obtained with the Morokuma–Ziegler energy decomposition method and with Gavezzotti's pixel-by-pixel direct numerical integration despite significant differences in molecular monomer wave functions used. The Gavezzotti approach should also be applicable to the electron densities calculated from the databank pseudoatoms and thus may make it possible to obtain readily available estimates of molecular interaction energies.

## Acknowledgments

The authors thank Dr. Irina Novozhilova for help with ADF calculations and Xue Li for the calculation of the pseudoatom databank parameters. Molecular graphics were created with the programs PLATON<sup>75</sup> and MOLEKEL.<sup>76,77</sup> Support from the National Institutes of Health (GM56829) and the National Science Foundation (CHE0236317) is gratefully acknowledged.

## References

1. Bickelhaupt, F. M.; Baerends, E. J. *Rev Comput Chem* 2000, 15, 1.
2. Ziegler, T.; Rauk, A. *Theor Chim Acta* 1977, 46, 1.

3. Ziegler, T.; Rauk, A. *Inorg Chem* 1979, 18, 1755.
4. Morokuma, K. *J Chem Phys* 1971, 55, 1236.
5. Kitaura, K.; Morokuma, K. *Int J Quantum Chem* 1976, 10, 325.
6. Spackman, M. A. *J Chem Phys* 1986, 85, 6587.
7. Cioslowski, J.; Liu, G. *Chem Phys Lett* 1997, 277, 299.
8. (a) Gavezzotti, A. *J Phys Chem* 2002, 106, 4145; (b) Gavezzotti, A. *J Phys Chem* 2003, 107, 2344.
9. Buckingham, A. D. *Adv Chem Phys* 1967, 12, 107.
10. Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: New York, 1997.
11. Coppens, P. *X-Ray Charge Densities and Chemical Bonding*; Oxford University Press: New York, 1997.
12. Hirshfeld, F. L. *Theor Chim Acta* 1977, 44, 129.
13. Hirshfeld, F. L. *Acta Crystallogr B* 1971, 27, 769.
14. Stewart, R. F. *Acta Crystallogr A* 1976, 32, 565.
15. Hansen N. K.; Coppens, P. *Acta Crystallogr A* 1978, 34, 909.
16. Bader, R. F. W. *J Am Chem Soc* 1979, 101, 1389.
17. Bader, R. F. W.; Nguyen, T. T.; Tal, Y. *Rep Progr Phys* 1981, 44, 893.
18. Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Clarendon Press: Oxford, UK, 1990.
19. Popelier, P. L. A.; Joubert, L.; Kosov, D. S. *J Phys Chem* 2001, 105, 8254.
20. Joubert, L.; Popelier, P. L. A. *Mol Phys* 2002, 100, 3357.
21. Joubert, L.; Popelier, P. L. A. *Phys Chem Chem Phys* 2002, 4, 4353.
22. Stone, A. J.; Alderton, M. *Mol Phys* 1985, 56, 1047.
23. Suponitsky, K. Y.; Tsirelson, V. G.; Feil, D. *Acta Crystallogr A* 1999, 55, 821.
24. Stone, A. J. *Chem Phys Lett* 1981, 83, 233.
25. Stewart, R. F.; Bentely, J.; Goodman, B. *J Chem Phys* 1975, 63, 3786.
26. Pichon-Pesme, V.; Lecomte, C.; Lachekar, H. J. *Phys Chem* 1995, 99, 6242.
27. Koritsanszky, T.; Volkov, A.; Coppens, P. *Acta Crystallogr A* 2002, 58, 464.
28. Moss, G.; Feil, D. *Acta Crystallogr A* 1981, 37, 414.
29. Spackman, M. A.; Weber, H. P.; Craven, B. M. *J Am Chem Soc* 1988, 110, 775.
30. Abramov, Y. A.; Volkov, A.; Coppens, P. *J Mol Struct* 2000, 529, 27.
31. Abramov, Y. A.; Volkov, A.; Wu, G.; Coppens, P. *Acta Crystallogr A* 2000, 56, 585.
32. Abramov, Y. A.; Volkov, A.; Wu, G.; Coppens, P. *J Phys Chem B* 2000, 104, 2183.
33. Li, X.; Wu, G.; Abramov, Y. A.; Volkov, A.; Coppens, P. *PNAS* 2002, 99, 12132.
34. de Vries, R. Y.; Feil, D.; Tsirelson, V. G. *Acta Crystallogr B* 2000, 56, 118.
35. Spackman, M. A. *J Chem Phys* 1986, 85, 6579.
36. Cox, S. R.; Hsu, L. Y.; Williams, D. E. *Acta Crystallogr A* 1981, 37, 293.
37. Williams, D. E.; Cox, S. R. *Acta Crystallogr B* 1984, 40, 404.
38. te Velde, G.; Bickelhaupt, F. M.; van Gisbergen, S. J. A.; Fonseca Guerra, C.; Baerends, E. J.; Snijders, J.G.; Ziegler, T. *J Comput Chem* 2001, 22, 931.
39. Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. *Theor Chem Acc* 1998, 99, 391.
40. ADF2003.01, SCM; Theoretical Chemistry, Vrije Universiteit: Amsterdam, (<http://www.scm.com>).
41. Destro, R.; Roversi, P.; Barzaghi, M.; Marsh, R. E. *J Phys Chem A* 2000, 104, 1047.
42. Mackay M. F. *Cryst Struct Commun* 1975, 4, 225.
43. Schouten, A.; Kanters, J. A.; van Krieken, J. *J Mol Struct* 1994, 323, 165.
44. *International Tables for X-Ray Crystallography*, vol. C; Kluwer Academic: Dordrecht, The Netherlands, 1992.
45. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Strattmann, R. E. Jr.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. GAUSSIAN98, revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.
46. Hariharan P.C.; Pople, J.A. *Theor Chim Acta* 1973, 28, 213.
47. Krishnan, R.; Binkley, J. S.; Seeger R.; Pople, J. A. *J Chem Phys* 1980, 72, 650.
48. Clark, T.; Chandrasekhar, J.; Schleyer, P. v. R. *J Comput Chem* 1983, 4, 294.
49. Becke, A.D. *J Chem Phys* 1993, 98, 5648.
50. Lee, C.; Yang, W.; Parr, R. G. *Phys Rev B* 1988, 37, 785.
51. Becke, A. D. *Phys Rev A* 1988, 38, 3098.
52. Boys, S. F.; Bernardi, F. *Mol Phys* 1970, 19, 553.
53. Koritsanszky, T.; Howard, S. T.; Richter, T.; Macchi, P.; Volkov, A.; Gatti, C.; Mallinson, P. R.; Farrugia, L. J.; Su, Z.; Hansen, N. K. XD—A computer program package for multipole refinement and Topological Analysis of charge densities from diffraction data, 2003.
54. Allen F. H. *Acta Crystallogr B* 2002, 58, 380.
55. Laidig, K. E. *J Phys Chem* 1993, 97, 12760.
56. Bernsten, J.; Espelid, T. O.; Gentz, A. *ACM Trans Math Software* 1991, 17, 452.
57. ACM. *ACM Trans Math Software* 1986, 12, 171.
58. Biegler-König, F. W., Bader, R. F.; Tang, T.- H. *J Comput Chem* 1982, 13, 317.
59. Gatti, C. *TOPOND98 Users' Manual*; CNR-CSRSC: Milano, Italy, 1999.
60. Volkov, A.; Gatti, C.; Abramov, Y.; Coppens, P. *Acta Crystallogr A* 2000, 56, 252.
61. Kisiel, Z. PROSPE—Programs for ROTational SPEctroscopy (<http://info.ifpan.edu.pl/~kisiel/prospe.htm>).
62. Stone, A. J.; Dullweber, A.; Hodges, M. P.; Popelier, P. L. A.; Wales, D. J. *Orient: a program for studying interactions between molecules*, version 3.2; Cambridge University Press: Cambridge, UK, 1996.
63. Swaminathan, S.; Craven, B. M.; Spackman, M. A.; Stewart, R. F. *Acta Crystallogr B* 1984, 40, 398.
64. Spackman, M. A.; Byrom, P. G. *Acta Crystallogr B* 1996, 52, 1023.
65. Spackman, M. A.; Byrom, P. G.; Alfredson, M.; Hermansson, K. *Acta Crystallogr A* 1999, 55, 30.
66. Bianchi, R.; Gatti, C.; Adovasio, V.; Nardelli, M. *Acta Crystallogr B* 1996, 52, 471.
67. Volkov, A.; Abramov, Y. A.; Coppens, P.; Gatti, C. *Acta Crystallogr A* 2000, 57, 272.
68. Alagona, G.; Ghio, C. *J Mol Struct Theochem* 1995, 330, 77.
69. Fonseca Guerra, C.; Bickelhaupt, F. M.; Baerends, E. J. *Cryst Growth Design* 2002, 2, 239.
70. Curutchet, C.; Bofill, J. M.; Hernandez, B.; Orozco, M.; Luque, F. J. *J Comput Chem* 2003, 24, 1263.
71. Dobado, J. A.; Martinez-Garcia, H.; Molina, J.; Sundberg, M. R. *Quantum Sys Chem Phys* 2000, 2, 337.
72. Maslen, E. N.; Spackman, M. A. *Aust J Phys* 1985, 38, 273.
73. Mathews J. H. *Numerical Methods for Mathematics, Science & Engineering*; Prentice-Hall: Englewood Cliffs, NJ, 1992.
74. Koritsanszky, T.; Volkov, A. *Chem Phys Lett* 2004, 385, 431.
75. Spek, A. L. *J Appl Crystallogr* 2003, 36, 7.
76. Flükiger, P.; Lüthi, H. P.; Portmann, S.; Weber, J. MOLEKEL 4.3; Swiss Center for Scientific Computing: Manno, Switzerland, 2000–2002.
77. Portmann S.; Lüthi, H. P. *CHIMIA* 2000, 54, 766.