Ab Initio-Quality Electrostatic Potentials for Proteins: An Application of the ADMA Approach

Thomas E. Exner[†,‡] and Paul G. Mezey*[,†,§]

Mathematical Chemistry Research Unit, Department of Chemistry, University of Saskatchewan, 110 Science Place, Saskatoon, SK, Canada S7N 5C9, and Department of Physical Chemistry, Darmstadt University of Technology, Petersenstrasse 20, 64287 Darmstadt, Germany

Received: June 17, 2002; In Final Form: July 29, 2002

The electrostatic potential around a molecule is often used to describe reactions, binding, and catalysis mechanisms or to serve as a descriptor in structure−activity relationships and molecular similarity studies. Often, very accurate descriptions of this property are needed that traditionally can be obtained, at least for small molecules, by quantum chemical calculations. The aim of this paper is to extend ab initio-quality quantum chemical accuracy to larger molecules such as proteins. The additive fuzzy density fragmentation (AFDF) principle and the adjustable density matrix assembler (ADMA) method are used to divide large molecules into fuzzy fragments, for which quantum chemical calculations can be done directly using smaller, "custom-made" parent molecules including all the local interactions within a preset distance limit. In the next step, the obtained density matrices of electron density fragments are combined to approximate the global density matrix and the electron density of the whole molecules. These ADMA electron densities are then used to calculate ab inito-quality electrostatic potentials of the large molecules. The accuracy of the method is analyzed in detail by two test cases of a penta- and a hexapetide, and the efficiency of the technique is demonstrated by the calculation of the electrostatic potential of the protein crambin.

I. Introduction

The electrostatic potential that is created in the space around the molecule by its nuclei and electrons has been proven to be a useful tool in explaining chemical reactivity and molecular interactive behavior (see, for example, refs 1−5 and references therein). It is through this potential that a molecule is first "seen" or "felt" by another approaching chemical species.[1−3] Thus, it is often used to describe reaction, binding, and catalysis mechanisms[1,4,5] or as a descriptor in many research areas such as molecular structure,[6−8] solvation,[9−21] crystalline state,[22−27] force-field parametrization,[28−37] quantitative structure−activity relationships (QSAR),[38−54] and molecular similarity studies.[55−67] It also has a large impact in rational drug design as a tool for "lead" optimization and pharmacophore searches.[44,45,57,59,68−71] Even though it is a physical property and can be determined experimentally, the electrostatic potential used in these applications is calculated theoretically most of the time. For small molecules, calculations are routinely done by standard quantum chemical methods such as Hartree−Fock and post-Hartree−Fock methods or, increasingly in the electrostatic potential context, using density functional theory.[1−3] But these methods are not feasible for larger molecules such as proteins because of the strong dependence of the computational time and memory on the number of basis functions and therefore the number of atoms in the molecule. Semiempirical methods can increase the range of computations to include larger molecules,[72−75] but even with those methods, the number of treatable atoms per system is limited. Therefore, the development of approximate methods to obtain a reasonable accuracy of the electrostatic potential of these large molecules is extremely important and is a major field of research.

Historically, the most widely used approximations are those based upon multipole expansion.[1−3,76] The quality of these approaches depends strongly on the number of terms (i.e., quadrupole, octapole, etc.) used in the expansion. One prominent example is using a set of point charges, which corresponds to terminating the expansion after the monopole term, which is used in many empirical force fields[34−37,77−79] as well as in molecular modeling studies of biomolecular processes such as protein docking[80,81] or activity relationships.[82,83] Because of the simplicity of the method, the advantage of the use of partial charges is that solvation effects based on the continuum theory can be included by using the Poisson−Boltzmann equation with a reasonable amount of computational effort.[5,84−89] However, it is a disadvantage that the parametrization of the atomic partial charges makes it impossible to include polarization effects due to intra- or intermolecular interactions with other chemical groups because the partial charge obtained by calculations on these molecules is assigned to every atom of the same kind independently of the chemical surroundings.

Another methodology for computing the electrostatic potential involves the representation of a large system as a combination or superposition of contributions from their constituent units or fragments. One possibility is realized in the hybrid quantum mechanical/molecular mechanical (QM/MM) methods.[90−94] Therein, the region of special interest, such as the active site of a protein, is treated quantum mechanically, whereas the other parts playing the role of essentially electrostatic perturbation are represented as point charges. Another possibility is to rely

* Corresponding author. E-mail: mezey@sask.usask.ca. Phone: (306) 966-4661. Fax: (306) 966-04730.
† University of Saskatchewan.
‡ Darmstadt University of Technology.
§ Current address: Szent-Györgyi Albert Professor, Department of Organic Chemistry, Eötvös Loránd University, Pázmány Péter Sétány 1/A, H-117 Budapest, Hungary.

10.1021/jp0263166 CCC: $22.00
Published on Web 11/09/2002

on the fact that the same chemical group has similar properties in different molecular surroundings. In this way, molecular fragments can be defined that are transferable between different molecules and that are calculated using small molecules and are recombined to build the larger molecule. In Náray-Szabó's group, this is achieved using the fragment self-consistent field (FSCF) method, which relies on strictly localized molecular orbitals (SLMO).[94] A number of approaches[48,95−101] uses Bader's theory of atoms in molecules[102] to define additive, approximately transferable fragments with fixed boundaries, hence discontinuities at these boundaries. Although a readjustment of fixed boundaries in order to match them without gaps as well as with equal function values is theoretically possible, this has not been achieved. Breneman[48,95,96] developed the transferable atom equivalent (TAE) method, which defines atom fragments according to Bader's theory with an adjustable parametrization of the molecular electron-density properties such as the electrostatic potential. This method faces the same difficulties of discontinuities as the general model of Bader.

The new approach for the calculation of ab initio-quality electrostatic potentials described in this paper also belongs to the group of methods using fragmentation, exploiting, however, a fuzzy-set approach where no boundary readjustment is required and no discontinuities occur. It is based on the partitioning of the large system into fuzzy subsystems according to the additive fuzzy-density fragmentation (AFDF) principle and the adjustable density-matrix assembler (ADMA) approach.[103−113] The advantage of the method is that the quality of the approximation can be controlled by the user by applying a simple distance criterion so that it can be adjusted to the accuracy needed in the particular application and to the computer power available. The very good qualitative as well quantitative agreement of the method with direct quantum mechanical calculations is demonstrated using smaller test cases.

II. Electrostatic Potential Calculations Using the ADMA Approach

The additive fuzzy-density fragmentation (AFDF) principle and the adjustable density-matrix assembler (ADMA) method[103−113] provide effective computational tools for the calculation and analysis of macromolecular electron densities. Using the conventional Hartree−Fock−Roothaan−Hall formalism, these MO-based, additive fuzzy-density fragmentation methods avoid artificial fragment boundaries and provide local molecular fragments that are fully analogous to complete molecules. The mathematical background and a large number of applications are well documented in the literature,[103,104,111,112] and thus only a short description is given here.

The electron density $\rho(\vec{r})$ of a molecule can be expressed in terms of a basis set of n atomic orbitals $\varphi_i(\vec{r})$ ($i = 1, 2, ..., n$) used for the expansion of the molecular wave function and the density matrix P_{ij} determined for the given nuclear configuration using this basis set:

$$\rho(\vec{r}) = \sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij} \cdot \varphi_i(\vec{r}) \cdot \varphi_j(\vec{r}) \tag{1}$$

Following the AFDF scheme, the first step in the generation of local fuzzy electron-density fragments is to subdivide the molecules into a set of mutually exclusive families of nuclei denoted by f^k with $k = 1, ..., m$. A formal membership function $m^k(i)$, which indicates if a given atomic orbital (AO) basis function $\varphi_i(\vec{r})$ belongs to the set of AOs centered on a nucleus of the family f^k, is defined:

$$m^k(i) = \begin{cases} 1 & \text{if AO } \varphi_i(\vec{r}) \text{ is centered on one nuclei of set } f^k \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Using this membership function, the elements P_{ij}^k of the density matrix of the kth fragment are calculated according to the Mulliken−Mezey scheme:[104,112]

$$P_{ij}^k = 0.5[m^k(i) + m^k(j)] \cdot P_{ij} \tag{3}$$

Since the fragment density matrices are taken from a direct calculation of the large "target" molecule with standard Hartree−Fock or post-Hartree−Fock methods, the strictly additive properties of the AFDF scheme guarantee that the original density matrix and therefore the original electron density can be restored exactly:

$$P_{ij} = \sum_{k=1}^{m} P_{ij}^k \tag{4}$$

A good approximation of the density matrix of the target molecule can be reached if the fragment density matrices are computed indirectly using a set of small "parent" molecules.[104,112] Each of these parent molecules contains one of the nuclear families f^k with the same local nuclear geometry and the same local surroundings as are found in the target molecule. Therefore, the accuracy of this approach depends only on the reproducibility of the local surroundings of each macromolecular fragment density within the parent molecules, which can be improved to any desired accuracy by increasing the size of the parent molecules. In this way, ab initio-quality electron densities can be calculated even for very large molecules where a direct quantum chemical calculation is not practical. This was shown in a number of applications.[103−113] The computation can even be speeded up if a database of fragment density matrices is used according to the ADMA technique.[103,104,110−113] Direct quantum chemical calculations are required only for those molecular fragments not yet available in this database.

On the basis of the earlier approach of quantum crystallography, developed by Karle and Massa,[114−116] involving a transformation of density matrices analogous to the Löwdin-transform−inverse Löwdin-transform method employed in transformations of ADMA-based densities,[113] the application of the ADMA method has also been proposed for the X-ray crystallographic structure−refinement process.[117]

The possibility of a fast generation of ab initio-quality electron densities has motivated the use of these electron densities in molecular similarity studies for the elucidation and prediction of molecular reactivity and biological activity.[118−123] Nowadays, this wide field of study is known as quantitative shape−activity relationships (QShAR).[118,123] As mentioned in the Introduction, another important aspect often used in reactivity studies is the electrostatic potential of a molecule $\phi(\vec{r})$.[1−5] It is defined as the electrostatic force acting on a unit charge caused by the nuclei and the electrons of the considered molecule and can be calculated for one particular point in space $\vec{r}$ according to the following equation:

$$\phi(\vec{r}) = \phi_{\text{nucl}}(\vec{r}) - \phi_{\text{el}}(\vec{r}) = \sum_{A=1}^{N} \frac{Z_A}{|\vec{R}_A - \vec{r}|} - \int \frac{\rho(\vec{r}\,')}{|\vec{r}\,' - \vec{r}|}\, \mathrm{d}\vec{r}\,' \tag{5}$$

with

$\phi(\vec{r})$: electrostatic potential at point $\vec{r}$
$\phi_{nucl}(\vec{r})$: electrostatic potential arising from the nuclei
$\phi_{el}(\vec{r})$: electrostatic potential arising from the electrons
$\vec{R}_A$: position of nucleus A
Z_A: charge of nucleus A
N: number of nuclei
$\rho(r')$: electron density function

Using eq 1 to express the electron density in terms of a specified basis set and the corresponding density matrix, the electrostatic potential arising from the electrons is

$$\phi_{el}(\vec{r}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \int \frac{P_{ij} \cdot \varphi_i(\vec{r}\,') \cdot \varphi_j(\vec{r}\,')}{|\vec{r}\,' - \vec{r}\,|} \, d\vec{r}\,' \qquad (6)$$

In the same manner as for the electron density, an approximation of the electrostatic potential of a macromolecule can be obtained with density matrices of fuzzy molecular fragments calculated from small parent molecules:

$$\phi_{el}(\vec{r}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \int \frac{\sum_{k=1}^{m} P_{ij}^{k} \cdot \varphi_i(\vec{r}\,') \cdot \varphi_j(\vec{r}\,')}{|\vec{r}\,' - \vec{r}\,|} \, d\vec{r}\,' \qquad (7)$$

The fragment density matrix can again be taken from the ADMA database or can be calculated when needed. In addition to the total electrostatic potential, the fragment electrostatic potentials, which have the same strictly additive properties as the electron density fragments, can also be calculated. Therefore, only the K nuclei defining the fragment f^k are used for the calculation of $\phi_{nucl}^{k}(\vec{r})$:

$$\phi^{k}(\vec{r}) = \sum_{A=1}^{K} \frac{Z_A}{|\vec{R}_A - \vec{r}\,|} - \sum_{i=1}^{n}\sum_{j=1}^{n} \int \frac{P_{ij}^{k} \cdot \varphi_i(\vec{r}\,') \cdot \varphi_j(\vec{r}\,')}{|\vec{r}\,' - \vec{r}\,|} \, d\vec{r}\,' \qquad (8)$$

and

$$\phi(\vec{r}) = \sum_{k=1}^{m} \phi^{k}(\vec{r}) \qquad (9)$$

III. Test Calculations on Small Oligopetides

In this section, the accuracy of the above-developed approximation of the electrostatic potential is tested with one hexapetide and one pentapeptide. These test cases are large enough that the dependence of the approximation on the size of the surroundings can be examined but small enough to compare the results to direct quantum chemical calculations.

The atomic positions of both peptides are taken from crystallographic structures stored in the Protein Data Bank,[124] in which the peptides are bound to the active site of a larger protein. The missing coordinates of the hydrogen atoms are added and optimized using the program CHARMM (version 24).[125] For the definition of the molecular fragments, an automated procedure was developed. Each heavy atom with the attached hydrogen and double-bonded oxygen atoms constitutes a fragment. If the heavy atom belongs to an aromatic system, then the whole aromatic system defines the fragment. Also, the $NHC(NH_2)_2$ group of arginine, the CO_2 groups of aspartic and glutamic acid, and the $CONH_2$ groups of asparagine and glutamine are considered to be one fragment. All atoms of the target molecule in a predefined radius around the fragment are considered to be surroundings. This also includes atoms not bonded covalently to the fragment. To obtain parent molecules that satisfy the target−parent compatiblility condition,[104,112] only complete nuclear families defining the target fragments are included in the surroundings. If two nuclear families of the surroundings are separated by only one atom in the target molecule not included in the surroundings, then the nuclear family corresponding to this atom is also included in the surroundings. The remaining missing valences are filled by adding hydrogen atoms in appropriate locations. In Figure 1, two representative fragments plus their surroundings are shown to illustrate once again the fragmentation scheme used. The resulting fragments plus surroundings are saved in pdb format and converted to a z-matrix file, which is suitable for a Hartree−Fock self-consistent-field calculation using the Gaussian 98 program[126] and the 6-31G** basis set.[127−131] The quantum chemical calculations for all fragments are then initiated automatically. In the present version of the software, no fragment from the ADMA database was used because in this study our aim is to investigate how dependent the approximation is on the accuracy and size of the surroundings of the fragments taken into account.

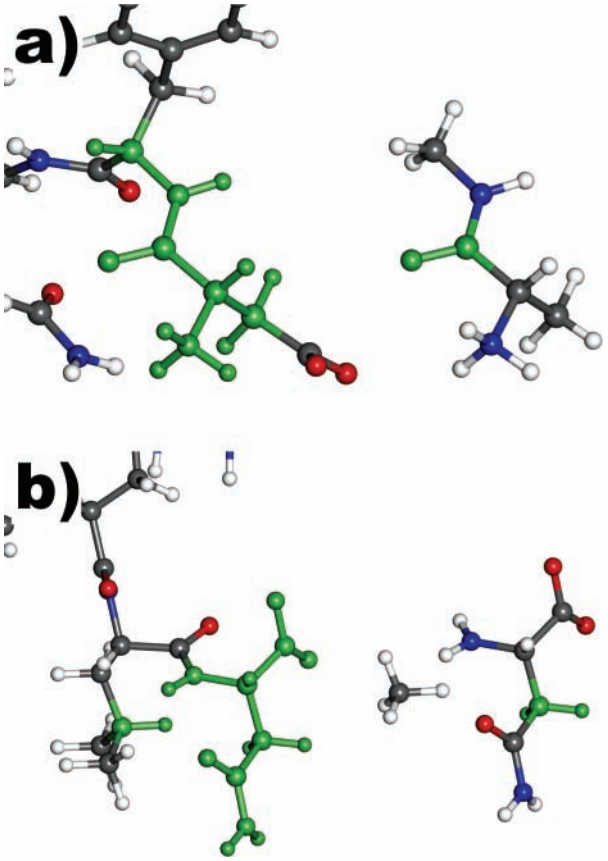

Figure 1. Two representative fragments and their 3.0-Å surrounding. On the left side, one part of the whole molecule including the fragment is shown. The green part of the molecule corresponds to the atoms used for the fragment and the surrounding. On the right side, the resulting fragment is shown. Fragment a is a CO group, and fragment b, a CH_2 group; these are shown in green with the atoms of the surroundings color-coded by atom type. It can be seen that atoms that are less than 3.0 Å apart but are not bonded covalently to the fragment are also included in the surroundings. Also, the hydrogen atoms filling the remaining missing valences can be seen.

The calculation of the electrostatic potential is performed for each fragment separately with the program "cubegen" distributed with the Gaussian 98 program package.[126] This program uses a

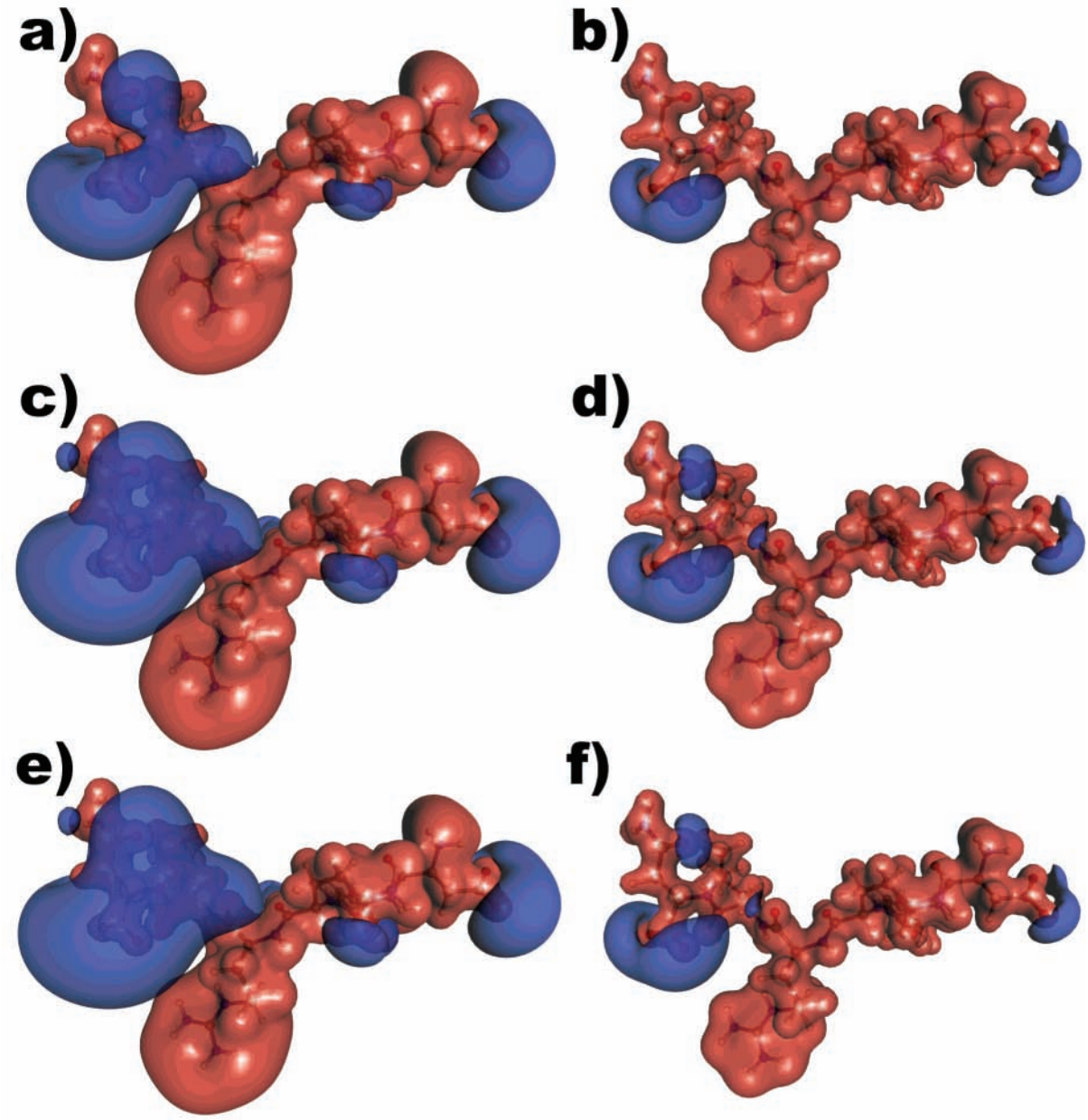

Figure 2. Positive (red) and negative (blue) isosurfaces of the electrostatic potential of the Tgn38 internalization peptide Dyqrln. Isovalues of +0.07 and −0.07 au are shown on the left side, and isovalues of +0.14 and −0.14 au, on the right side. Calculated with (a, b) the Gaussian 98 program; (c, d) the ADMA approach without normalization (surroundings of 3.0 Å); and (e, f) with normalization (surroundings of 3.0 Å).

formatted checkpoint file[126] as input in which the nuclear positions and charges, the parameters of the basis set, and the density matrix elements are specified. Normally, these formatted checkpoint files are generated from files produced in a run of the Gaussian 98 program by a conversion utility. In our case, only the fragment density matrix according to the ADMA approach and the charges of one family of the nuclei should be used in the calculations. Thus, the density matrix in the formatted checkpoint file is substituted by the fragment density matrix, and the charge of every nucleus not included in the fragment is set to zero. The electrostatic potential of the whole molecule is finally obtained by adding the fragment electrostatic potentials for all families of nuclei of the target molecule.

The first molecule used in the studies described here is the Tgn38 internalization peptide Dyqrln. This hexapeptide has the sequence Asp-Tyr-Gln-Arg-Leu-Asn and is composed of 110 atoms. The atomic positions were taken from the pdb entry 1BXX,[132] in which the peptide forms a complex with the internalization signal-binding domain of the Ap2 adaptor. The electrostatic potential calculated with surroundings of 3.0 Å is shown in Figures 2 and 3 in different representations (all figures in this publication are created using the MOLCAD II module[133−135] of the SYBYL molecular modeling package[136]). In Figure 2, isosurfaces are shown that are generated by connecting points with the same value of the electrostatic potential. Figure 2a and b are the results of direct quantum chemical calculations with the Gaussian 98 program, and Figure 2c and d are the results from the ADMA approach described here. It can be seen that, even with the relatively small surroundings employed, a good qualitative and almost quantitative agreement can be achieved. However, the electrostatic potential is somewhat more negative in the approximation than in the direct calculation. This is shown especially by the larger space that is surrounded by the negative (blue) isosurface in Figure 2c and d and can be explained partly by the following reasons: The integration of the electron density over the total

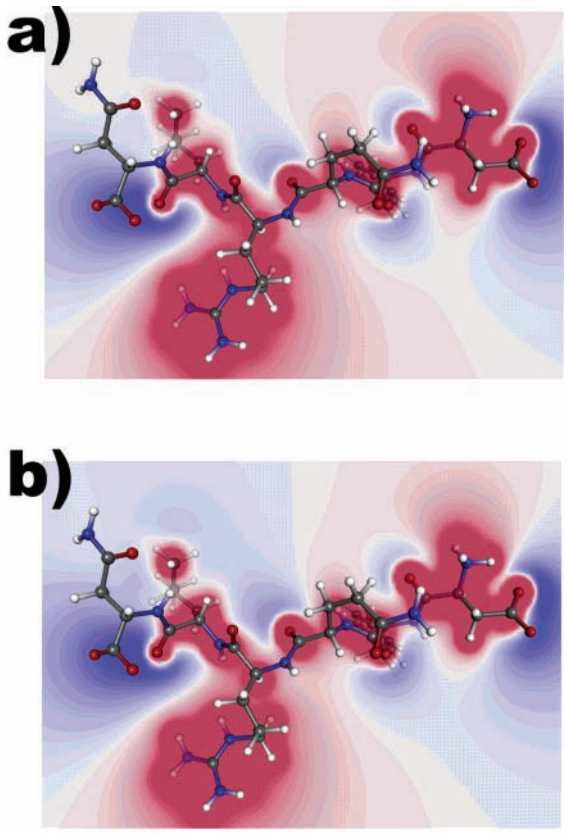

Figure 3. Slicing plane through the electrostatic potential of the Tgn38 internalization peptide Dyqrln. Positive parts are shown in red, and negative parts, in blue. (a) As calculated by the Gaussian 98 program and (b) with the ADMA approach (surroundings of 3.0 Å with normalization).

space should result in the total number of electrons. Because of the combination of fragment density matrices, this condition is not always exactly fulfilled. In our test case, a value of 430.014 electrons results from the integration compared to the 430 electrons contained in the molecule. This excess of 0.014 electrons gives the molecule a partial negative charge, which induces the decrease of the electrostatic potential. To remove this random error, each density matrix element can be multiplied by the quotient of the real number of electrons and the number obtained by the integration of the fragment electron density:

$$P_{ij}^{\text{norm}} = P_{ij} \frac{n_{\text{el}}}{\int \rho(\vec{r})\, \mathrm{d}\vec{r}} \tag{10}$$

The results obtained by this modified electron-density matrix are shown in Figure 2e and f. Even if the electrostatic potential is still too negative on the left side of the molecule, a small refinement of the approximation is achieved by this normalization of the electron-density matrix. This is even more the case when the size of the target molecule increases and thus the number of excess electrons (the difference between the real number of electrons and the number of electrons achieved by the ADMA approach) is also more likely to increase.

In Figure 3, the electrostatic potential is shown along a plane slicing through the molecule with the direct calculation in the upper part and the ADMA approach with normalization in the lower part of the Figure. The good qualitative agreement between the direct calculations and the ADMA approach is demonstrated once again by this representation. Only small differences can be seen in the upper left corner and in the middle part of the molecule, where a negative electrostatic potential is predicted by the ADMA approach, whereas the direct calculation gives an almost neutral potential. The differences between the ADMA approach with and without normalization are too small to be seen in this representation; therefore, the picture of the ADMA approach without normalization is omitted.

The remaining errors in the approximations generated by the ADMA approach are caused by the limited size of the fragments. Thus, the electrons of a fragment are limited to the basis functions located at the position of the family of nuclei of the fragment and are therefore more localized as compared to the case when the whole molecule is considered. This result shows an accumulation of electrons in negatively polarized regions of the molecule and a reduction of the electron density in positively polarized regions, which corresponds to an increase in the absolute value in the negative as well as in the positive region of the electrostatic potential. This localization can be reduced by increasing the size of the surroundings of each fragment. In Figure 4, the electrostatic potentials for surroundings of 4.0 and 5.0 Å are compared to the potentials with surroundings of 3.0 Å and the potential calculated directly with the Gaussian 98 program, which have already been shown in Figure 2. It can be seen that the errors are significantly reduced by the larger surroundings and that almost no differences can be seen between the electrostatic potential with 5.0-Å surroundings and the directly calculated potential. (Figure 4a and b and Figure 4 g and h)

The second oligopeptide is the c-terminal fragment of the chemotaxis receptor in the conformation in the complex with chemotaxis receptor methyltransferase taken from the pdb entry 1BC5.[137] This pentapeptide has the sequence Asn-Trp-Glu-Thr-Phe and is composed of 90 atoms. The electrostatic potential calculated with the Gaussian 98 program as well as with the ADMA approach with surroundings of 3.0, 4.0, and 5.0 Å is shown in Figure 5. This example demonstrates once more that the ADMA approach gives reasonable approximations of the electrostatic potential and that the accuracy can be increased by using larger surroundings for each fragment.

Another method beside the visual inspection of isodensity surfaces shown so far is the numerical comparison of electrostatic potentials and thus, the verification of the results of the various approximations by the calculation of quantum molecular similarity measures (see refs 138−140 and references therein). Carbó-Dorca and co-workers[141−143] proposed a scheme for the generation of quantum similarity measures based on the electron density distribution. One of these similarity measures, the Coulomb quantum similarity measure, is strongly related to the electrostatic potential and is defined as[141−143]

$$z_{AB} = \int\int \rho_A(\vec{r}_1)\, |\vec{r}_1 - \vec{r}_2|\, \rho_B(\vec{r}_2)\, \mathrm{d}\vec{r}_1\, \mathrm{d}\vec{r}_2 \tag{11}$$

The well-known Carbó similarity measure[144−146] is produced by the transformation into a number lying within the interval (0;1], where z_{AA} and z_{BB} are the quantum self-similarity measures of molecule A and B, respectively:[141−143]

$$r_{AB} = \frac{z_{AB}}{\sqrt{z_{AA} z_{BB}}} \tag{12}$$

Because we want to compare the approximated electrostatic potential to the direct calculations, we use in this paper a related similarity measure defined with the electronic part of the electrostatic potential:

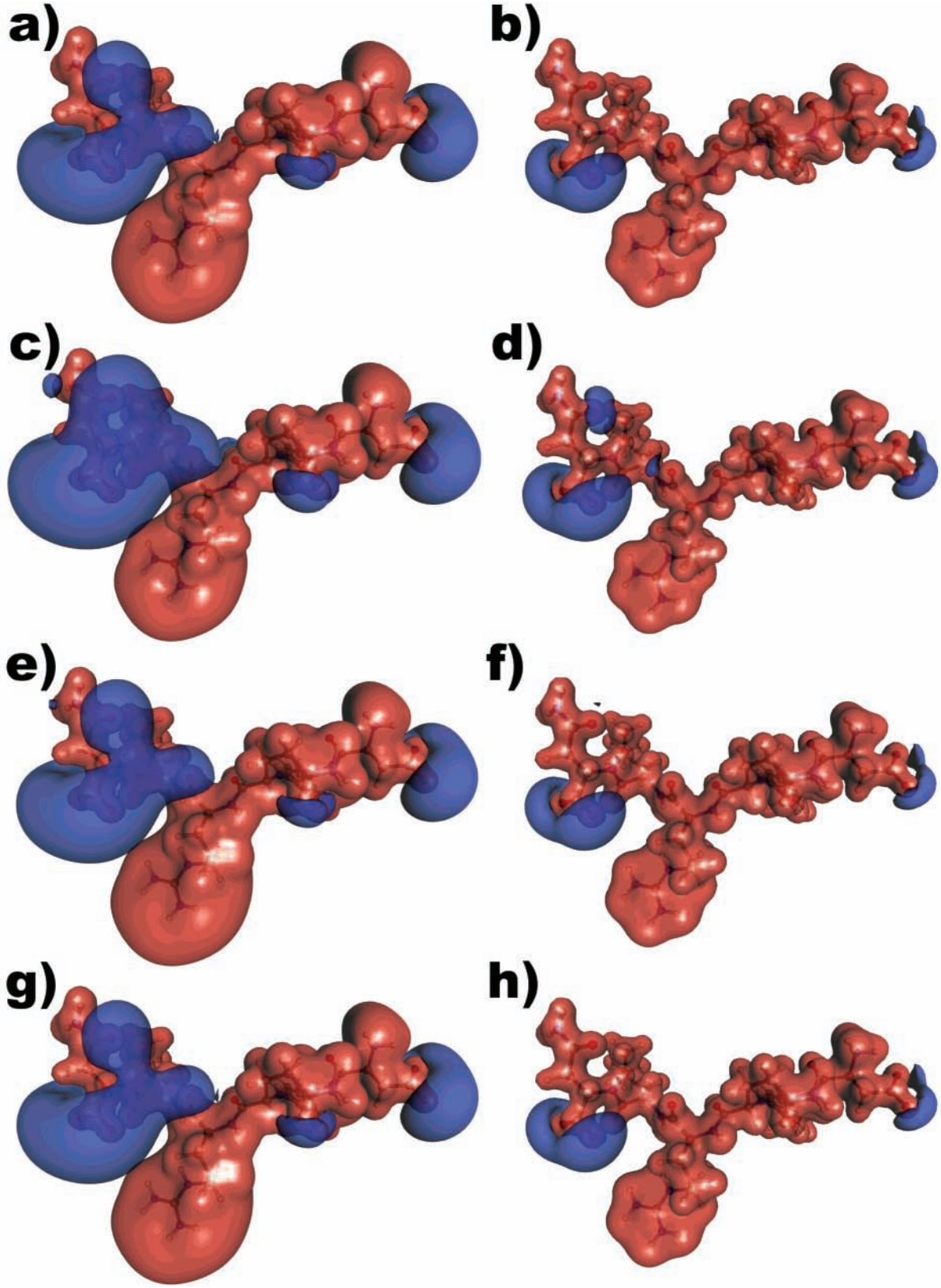

Figure 4. Positive (red) and negative (blue) isosurfaces of the electrostatic potential of the Tgn38 internalization peptide Dyqrln. Isovalues of +0.07 and −0.07 au are shown on the left side, and of +0.14 and −0.14 au, on the right side. Calculated by (a, b) the Gaussian 98 program; (c, d) the ADMA and surroundings of 3.0 Å; (e, f) with surroundings of 4.0 Å; and (g, h) with surroundings of 5.0 Å.

$$r_{AB} = \frac{\int \phi_{\mathrm{el},A}\phi_{\mathrm{el},B}\,\mathrm{d}\tau}{(\int \phi_{\mathrm{el},A}\phi_{\mathrm{el},A}\,\mathrm{d}\tau)^{1/2}(\int \phi_{\mathrm{el},B}\phi_{\mathrm{el},B}\,\mathrm{d}\tau)^{1/2}} \qquad (13)$$

The reason for this definition of a quantum similarity measure is on one hand that the nuclear part of the electrostatic potential is the same for the approximate and the direct calculations; therefore, a similarity of 1 results when these parts are compared. On the other hand, the electrostatic potential is a balance of large positive and large negative contributions of the nuclei and

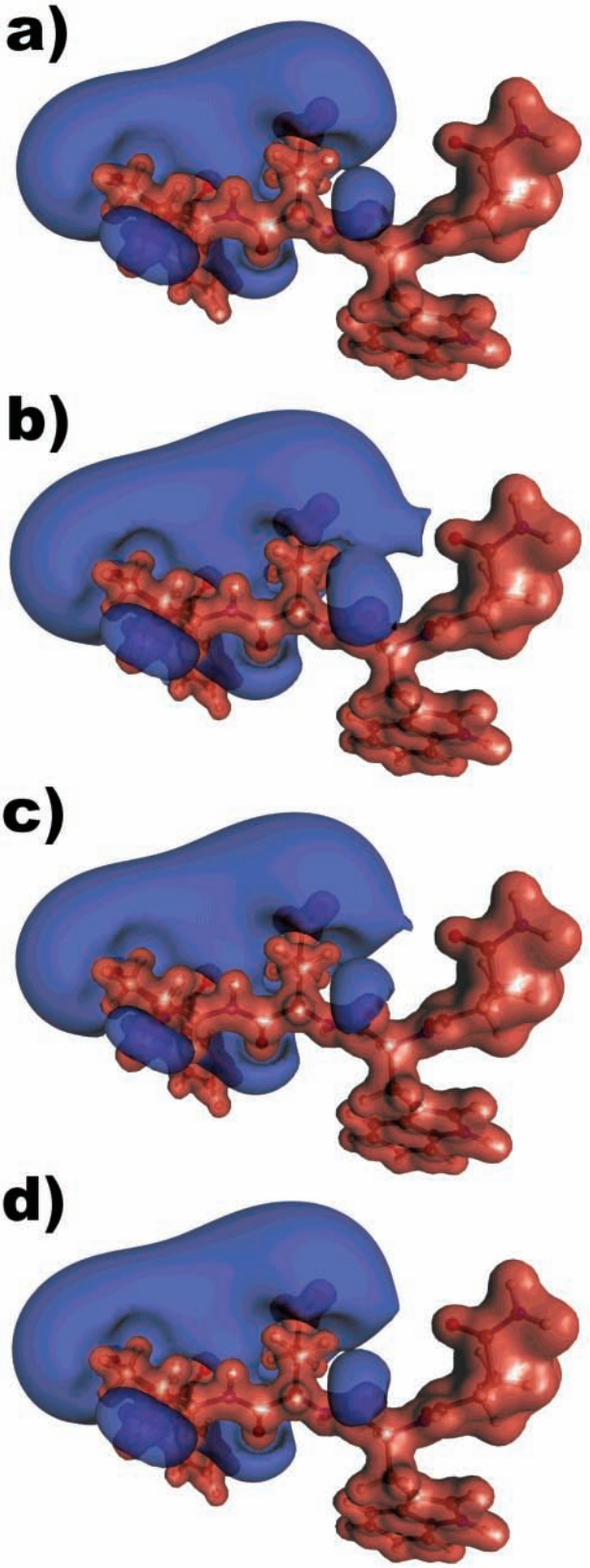

Figure 5. Positive (red) and negative (blue) isosurface of the electrostatic potential of the c-terminal fragment of the chemotaxis receptor. Isovalues of +0.07 and −0.07 au are shown. Calculated by (a) the Gaussian 98 program, (b) with the ADMA and surroundings of 3.0 Å, (c) with surroundings of 4.0 Å, and (d) with surroundings of 5.0 Å

the electrons, respectively. A similarity measure for the total electrostatic potential would therefore make the differences between the approximation and the direct calculations more evident. But because of the nondefinite positive nature of the electrostatic potential, such a measure is not possible using the Carbó similarity index.

TABLE 1: Quantum Molecular Similarity Measure of the Approximative Electrostatic Potential for Two Oligopepties Compared to Direct Quantum Chemical Calculations[a]

surroundings (Å)	A[b]	B[c]
3.0	0.999997	0.999998
4.0	0.999999	0.999999
5.0	1.000000	1.000000

[a] Values corresponding to the Carbó similarity index as defined in eq 13. [b] Tgn38 internalization peptide Dyqrln. [c] c-Terminal fragment of chemotaxis receptor.

The values of the similarity measure for both oligopeptides with different surroundings are summarized in Table 1. It can be seen that values near 1 result even for the smallest surroundings taken, which demonstrates once again the usefulness of the proposed ADMA approximation. It can also be seen that the values are getting even better with the increasing size of the surroundings. Consequently, the quality of the approximation can be controlled by choosing an appropriate size of the surroundings so that it can be adjusted to the accuracy needed in the particular application and to the computer power available.

IV. Electrostatic Potential of Crambin

To demonstrate that the new approach is able to calculate the electrostatic potential even for large molecules, whose sizes have precluded traditional quantum chemical calculations so far, the protein crambin (pdb entry 1CNR[147]) is chosen as an example. The results using surroundings of 4.0 Å are shown in Figure 6. In the Figure, an isosurface of the electron density is also shown. This isosurface roughly describes the shape of the molecule and hides the electropositive core of the molecules. Thus, only those parts of the electrostatic potential can be seen that correspond to highly polarized regions and therefore are of interest.

With 330 heavy atoms, crambin is a rather small protein, but it is already close to the limit of standard quantum chemical calculations. In this paper, quantum chemical calculations are carried out for every fragment of the target molecule. With the use of the ADMA database, only fragments not included in the database would be calculated, and the stored electron-density matrix would be used otherwise. For the effective use of the database, not all possible conformations of a fragment can be stored in the database. This problem can be circumvented by approximating the electron-density matrix using similar conformations stored in the database using different transformation methods described elsewhere.[148,149] The database approach would result in a large decrease in the computational time to obtain the electron density matrices. But more than half the time is spent for the calculation of the electrostatic potential. This is due to the large grid size that is used to cover the 3D space around the molecule and the large number of basis functions for this molecule. On one hand, using a larger spacing between the grid points and an interpolation algorithm to calculate the values between the points can circumvent this. A smaller grid size can also be achieved if one is interested in only a small part of the protein, such as the active site, and the electrostatic potential is calculated only for this part. On the other hand, if a grid point and a nucleus are far apart, then the basis functions corresponding to this nucleus have almost no influence on the

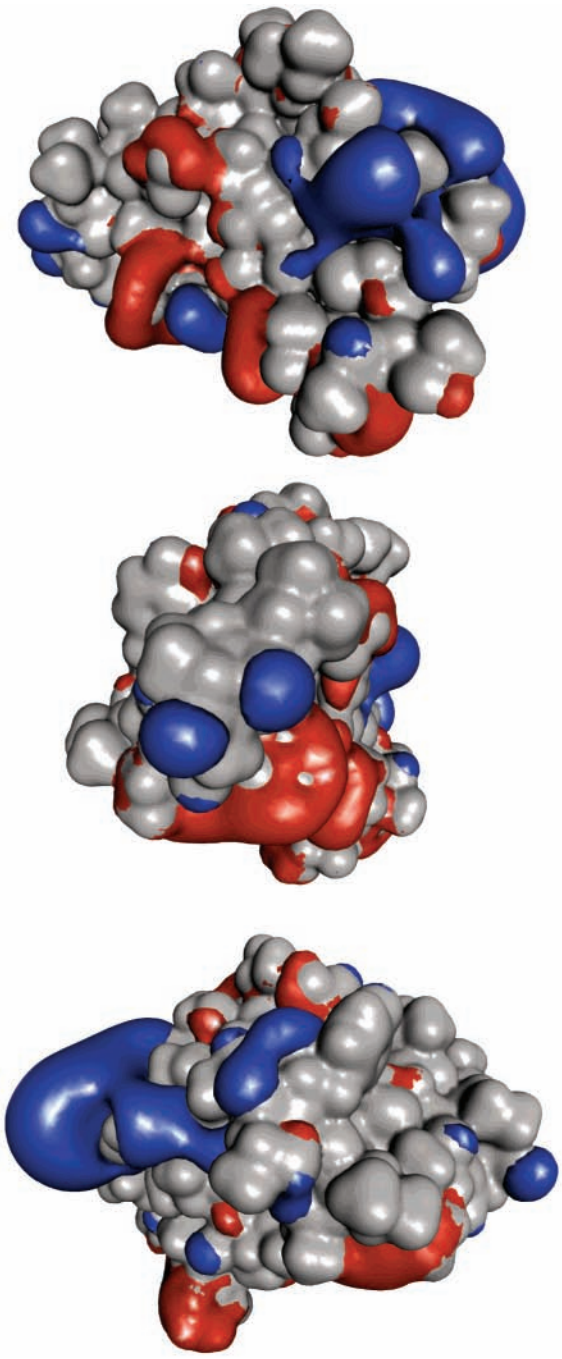

Figure 6. Electrostatic potential of crambin in three different orientations. An isosurface of the electron density (gray, isovalue 0.002 au) and a positive (red, isovalue +0.05 au) and a negative (blue, isovalue −0.05 au) isosurface of the electrostatic potential are shown.

electrostatic potential at the location of the grid point. Therefore, only basis functions located on nuclei near a grid point may be considered in the calculation of the electrostatic potential for this grid point. This would result in a much smaller number of basis functions for each grid point and a noticeable speedup in the calculation.

V. Conclusion

In this paper, a method for the calculation of ab initio-quality electrostatic potentials is introduced. It is based on the additive fuzzy-density fragmentation (AFDF) and the adjustable density-matrix assembler (ADMA) approaches in which large systems are partitioned into fuzzy subsystems. Quantum chemical calculations are performed on these small subsystems with tailor-made surroundings, and the fragments are recombined to obtain the approximate electrostatic potential of the whole system. The calculations of these approximations require a fraction of the computational time that would be needed for conventional ab initio techniques, and because of the linear scaling properties of the ADMA approach, they are feasible even for very large systems. An additional large decrease in computational time can be obtained if an ADMA database of precalculated fragments is used in the calculations.

In the first part of this paper, the very good qualitative as well as quantitative agreement of the method with direct quantum mechanical calculations is demonstrated with smaller test cases. This is done by visual inspection of isosurfaces and slicing planes as well as numerically by quantum molecular similarity measures using the approximative electrostatic potential of two oligopeptides. In these two examples using the 6-31G** basis set,[127−131] a surrounding of 4.0 to 5.0 Å was sufficient for obtaining these good results. This number should probably been increased if basis sets with diffuse functions, such as the 6-31+G[150] or aug-ccpVXZ basis sets,[151,152] are used because of the increasing influence of basis functions centered on atoms further away on the electron density of the fragment. In the second part, it is demonstrated that the new approach is able to calculate the electrostatic potential even for large molecules whose sizes have precluded ab initio-quality quantum chemical calculations so far. This is an important extension to the existing ADMA method because it is now possible to calculate not only the electron density but also the electrostatic potential on the basis of this approach. The ADMA electron densities are used to describe molecular shape in similarity studies such as topological shape analysis and quantitative shape−activity relationships (QShAR). These studies can now be extended to electrostatic properties so that another important factor of molecular recognition in addition to shape can be investigated. The treatment of other properties (e.g., dipole moment, intermolecular forces) is presently under investigation in our laboratory.

The other advantage of the method is that the quality of the approximation can be controlled by the user so that it can be adjusted to the accuracy needed in the particular application and to the computer power available. In this sense, very accurate calculations can be performed in regions of the molecule that are of special interest, such as an active site in a protein. The regions that are further away and thus have a smaller influence on the properties of the active site can be calculated less accurately with smaller surroundings. This kind of approach can be compared to hybrid quantum mechanical/molecular mechanical (QM/MM) methods in the sense that in these methods a part of the molecule is also calculated very accurately where the influences of the other regions are regarded only as approximative. But in the ADMA approach, there are no difficulties in combining the two parts because both are treated quantum chemically. In the QM/MM methods, the region between the two subsystems requires special treatment, which represents a major problem connected with the QM/MM approach.[94]

Acknowledgment. We sincerely thank Professor J. Brickmann for supplying the MOLCAD II module and the Natural Sciences and Engineering Research Council of Canada for financial support of this work. T.E.E. gratefully acknowledges the Alexander von Humboldt Foundation for a Feodor Lynen research fellowship.

References and Notes

(1) Politzer, P.; Murray, J. S. Molecular Electrostatic Potentials and Chemical Reactivity. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1991; Vol. 2, pp 273−312.

(2) Murray, J. S.; Politzer, P. Electrostatic Potentials: Chemical Applications. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; Wiley & Sons: Chichester, U.K., 1998; pp 912−920.

(3) Murray, J. S.; Politzer, P. The Molecular Electrostatic Potential: A Tool for Understanding and Predicting Molecular Interactions. In *Molecular Orbital Calculations for Biological Systems*; Sapse, A.-M., Ed.; Oxford University Press: New York, 1998; pp 49−84.

(4) Náray-Szabó, G. Electrostatic Catalysis. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; Wiley & Sons: Chichester, U.K., 1998; pp 905−912.

(5) Madura, J. D.; Davis, M. E.; Gilson, M. K.; Wade, R. C.; Luty, B. A.; McCammon, J. A. Biological Applications of Electrostatic Calculations and Brownian Dynamics Simulations. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1994; Vol. 5, pp 229−267.

(6) Politzer, P. *J. Am. Chem. Soc.* **1980**, *102*, 3027.

(7) Ritchie, J. P. *J. Am. Chem. Soc.* **1985**, *107*, 1829.

(8) Politzer, P.; Murray, J. S. In *Molecular Electrostatic Potentials: Concepts and Applications*; Murray, J. S., Sen, K., Eds.; Elsevier: Amsterdam, 1996; pp 649−660.

(9) Oliveira Neto, M. *J. Comput.Chem.* **1986**, *7*, 617.

(10) Gilson, M. K.; Honig, B. *Proteins* **1988**, *4*, 7.

(11) Murray, J. S.; Politzer, P. *J. Org. Chem.* **1991**, *56*, 6715.

(12) Tomasi, J.; Mennucci, B.; Cammi, R. In *Molecular Electrostatic Potentials: Concepts and Applications*; Murray, J. S., Sen, K., Eds.; Elsevier: Amsterdam, 1996; pp 1−85.

(13) Luque, F. J.; Barril, X.; Orozco, M. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 139.

(14) Conçalves, P. F. B.; Strassen, H. *J. Comput. Chem.* **2002**, *23*, 706.

(15) Matrin, M. E.; Sánchez, M. L.; Olivares del Valle, F. J.; Aguilar, M. A. *J. Chem. Phys.* **2002**, *116*, 1613.

(16) Allen, R.; Hansen, J.-P.; Melchionna, S. *Phys. Chem. Chem. Phys.* **2001**, *3*, 4177.

(17) Zou, J.; Yu, Q.; Shang, Z. *J. Chem. Soc., Perkin Trans.* **2001**, *2*, 1439.

(18) Tobias, D. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 253.

(19) DelValle, F. J. O.; Aguilar, M. A.; Mendoza, M. L. S. *Recent Res. Dev. Phys. Chem.* **1998**, *2*, 575.

(20) Murray, J. S.; Abu-Awward, F.; Politzer, P. *J. Phys. Chem. A* **1999**, *103*, 1853.

(21) Hummer, G.; Pratt, L. R.; Garcia, A. E.; Berne, B. J.; Rick, S. W. *J. Phys. Chem. B* **1997**, *101*, 3017.

(22) Spackman, M. A.; Stewart, R. F. In *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Politzer, P., Truhlar, D. G., Eds.; Plenum Publishing: New York, 1981; pp 407−415.

(23) Weber, H. P.; Craven, B. M. *Acta Crystallogr., Sect B* **1990**, *46*, 532.

(24) White, J. C.; Hess, A. C. *J. Phys. Chem.* **1993**, *97*, 6398.

(25) Yatsenko, A. V.; Paseshnichenko, K. A. *Chem. Phys.* **2000**, *262*, 293.

(26) Gavezzotti, A. *J. Phys. Chem. B* **2002**, *106*, 4145.

(27) Stone, A. J.; Tsuzuki, S. *J. Phys. Chem. B* **1997**, *101*, 10178.

(28) Momany, F. A. *J. Phys. Chem.* **1978**, *82*, 592.

(29) William, D. E. *J. Comput. Chem.* **1988**, *9*, 745.

(30) Ferenzcy, G. G.; Reynolds, C. A.; Richards, W. G. *J. Comput. Chem.* **1990**, *11*, 159.

(31) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.

(32) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.

(33) Bonaccorsi, R.; Scrocco, E.; Petrongolo, C.; Tomasi, J. *Theor. Chim. Acta* **1984**, *20*, 331.

(34) Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236.

(35) Dykstra, C. E. *Chem. Rev.* **1993**, *93*, 3.

(36) Wang, W.; Donini, O.; Reges, C. M.; Kollman, P. A. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211.

(37) Jalaie, M.; Lipkowitz, K. B. In *Reviews in Computional Chemistry;* Lipkowitz, K., Boyd, D., Eds.; Wiley-VCH: New York, 2000; Vol. 14, published force field parameters of molecular mechanics, molecular dynamics, and Monte Carlo simulations.

(38) Náray-Szabó, G.; Surjan, P. R. In *Theoretical Chemistry of Biological Systems*; Náray-Szabó, G., Ed.; Elsevier: Amsterdam, 1986; pp 1−30.

(39) Loew, G. H.; Berkowitz, D. S. *J. Med. Chem.* **1975**, *18*, 656.

(40) Hayes, D. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1976**, *98*, 335.

(41) Hayes, D. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1976**, *98*, 7861.

(42) Weinstein, H.; Osman, R.; Topiol, S.; Green, J. P. *Ann. N.Y. Acad. Sci.* **1981**, *367*, 434.

(43) Petrongolo, C.; Tomasi, J. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1975**, *2*, 181.

(44) Orozco, M.; Canela, E. I.; Franco, R. *Mol. Pharmacol.* **1989**, *35*, 257.

(45) Orozco, M.; Canela, E. I.; Franco, R. *Eur. J. Biochem.* **1990**, *188*, 155.

(46) Warshel, A.; Åqvist, J. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267.

(47) Náray-Szabó, G.; Nagy, P. *Int. J. Quantum Chem.* **1989**, *35*, 215.

(48) Breneman, C. M.; Martinov, M. In *Molecular Electrostatic Potentials: Concepts and Applications*; Murray, J. S., Sen, K., Eds.; Elsevier: Amsterdam, 1996; pp 143−180.

(49) Winkler, D. A. *Briefings in Bioinformatics* **2002**, *3*, 73.

(50) Gupta, S. P. *Prog. Drug Res.* **2001**, *56*, 121.

(51) Katritzky, A. R.; Petruhkin, R.; Tatham, D.; Basak, S.; Benfenati, E.; Karelson, M.; Maran, U. *J. Chem. Inf. Comput. Sci.* **2002**, *41*, 679.

(52) Warne, M. A.; Nicholson, J. K. *Prog. Environ. Sci.* **1999**, *1*, 327.

(53) Martin, Y. C. *Perspect. Drug Discovery Des.* **1998**, *12*−*14*, 3.

(54) Kubinyi, H. *Drug Discovery Today* **2002**, *2*, 457.

(55) Besalú, E.; Carbó, R.; Mestres, J.; Solà, M. Foundations and Recent Developments on Molecular Quantum Similarity. In *Topics in Current Chemistry: Molecular Similarity I*; Sen, K., Ed.; Springer-Verlag: Berlin, 1995; Vol. 173, pp 31−62.

(56) Burt, C.; Huxley, R.; Richards, W. G. *J. Comput. Chem.* **1990**, *11*, 1139.

(57) Rogríguez, J.; Manault, F.; Sanz, F. *J. Comput. Chem.* **1993**, *14*, 922.

(58) Sanz, F.; López, E.; Rogríguez, J.; Manault, F. *Quant. Struct.-Act. Relat.* **1994**, *13*, 281.

(59) Sanz, F.; Manault, F.; Rogríguez, J.; Lozoya, E.; López, E. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 337.

(60) Charlton, M. K.; Thomson, C. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 3533.

(61) Blomberg, N.; Gabdoulline, R. R.; Nilges, M.; Wade, R. C. *Proteins* **1999**, *37*, 379.

(62) Thorner, D. A.; Wild, D. J.; Willett, P.; Wright, P. M. *Perspect. Drug Discovery Des.* **1998**, *9*−*11*, 301.

(63) Thorner, D. A.; Wild, D. J.; Willett, P.; Wright, P. M. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 900.

(64) Richards, W. G.; Robinson, D. D. *IMA Vol. Math. Its Appl.* **1999**, *108*, 39.

(65) Hodgkin, E.; Richards, W. G. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1987**, *14*, 105.

(66) deCaceres, M.; Villa, J.; Lozano, J. J.; Sanz, F. *Bioinformatics* **2000**, *16*, 568.

(67) Sanz, F.; Manaut, F.; Rodriguez, J.; Lozoya, E.; Lopez-de-Brinas, E. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 337.

(68) Mason, J. S.; Good, A. C.; Martin, E. J. *Curr. Pharm. Des.* **2001**, *7*, 567.

(69) Tropsha, A.; Zheng, W. *Comput. Biochem. Biophys.* **2001**, 351.

(70) Klebe, G. *J. Mol. Med.* **2000**, *78*, 269.

(71) Apostolakis, J.; Caflisch, A. *Comb. Chem. High Throughput Screening* **1999**, *2*, 91.

(72) Zerner, M. C. Semiempirical Molecular Orbital Methods. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1991; Vol. 2, pp 313−365.

(73) Stewart, J. J. P. Semiempirical Molecular Orbital Methods. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1990; Vol. 1, pp 45−81.

(74) Khandogin, J.; Hu, A.; York, D. M. *J. Comput. Chem.* **2000**, *21*, 1562.

(75) Bliznyuk, A. A.; Rendell, A. P.; Allen, T. W.; Chung, S.-H. *J. Phys. Chem. B* **2001**, *105*, 12674.

(76) Williams, D. E. Net Atomic Charge and Multipole Models for the Ab Initio Molecular Electric Potential. In *Reviews in Computational Chemistry*; Lipkowitz, K., Boyd, D., Eds.; Wiley-VCH: New York, 1991; Vol. 2, pp 219−271.

(77) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Cadwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(78) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1664.

(79) Mackerell, A. d.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, D.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(80) Muegge, I.; Rarey, M. Small Molecule Docking and Scoring. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 2001; Vol. 17, pp 1−60.

(81) Ehrlich, L. P.; Wade, R. C. Protein−Protein Docking. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 2001; Vol. 17, pp 61−97.
(82) Oprea, T. I.; Waller, C. L. Theoretical and Practical Aspects of Three-Dimensional Quantitative Structure−Activity Relationships. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 1997; Vol. 11, pp 127−182.
(83) Greco, G.; Novellino, E.; Martin, Y. C. Approaches to Three-Dimensional Quantitative Structure−Activity Relationships. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 1997; Vol. 11, pp 183−240.
(84) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157*, 671.
(85) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K. A.; Honig, B. H. *Proteins* **1986**, *1*, 47.
(86) Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J. Comput. Chem.* **1987**, *9*, 327.
(87) Nicholls, A.; Honig, B. H. *J. Comput. Chem.* **1991**, *12*, 435.
(88) Honig, B. H.; Nicholls, A. *Science (Washington, D.C.)* **1995**, *268*, 1144.
(89) Bruccoleri, R. E.; Novotny, J.; Davis, M. E.; Sharp, K. A. *J. Comput. Chem.* **1997**, *18*, 268.
(90) Gao, J. Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1995; Vol. 7, pp 119−185.
(91) Mlinsek, G.; Novic, M.; Hodoscek, M.; Solmajer, T. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1286.
(92) Burton, N. A.; Harrison, M. J.; Hart, J. C.; Hillier, I. H.; Sheppard, D. W. *Faraday Discuss.* **1998**, *110*, 463.
(93) Hernández, B.; Luque, F. J.; Orozco, M. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 329.
(94) Náray-Szabó, G. *Comput. Chem. (Oxford)* **2000**, *24*, 287.
(95) Breneman, C. M.; Thompson, T. R.; Rhem, M.; Dung, M. *Comput. Chem. (Oxford)* **1995**, *19*, 161.
(96) Breneman, C. M.; Rhem, M. *J. Comput. Chem.* **1997**, *18*, 182.
(97) Matta, C. F. *J. Phys. Chem. A* **2001**, *105*, 11088.
(98) Biegler-König, F. W.; Nguyen-Dang, T. T.; Tal, Y.; Bader, R. F. W.; Duke, A. J. *J. Phys. B: At. Mol. Opt. Phys.* **1981**, *14*, 2739.
(99) Biegler-König, F. W.; Bader, R. F. W.; Tang, T. H. *J. Comput. Chem.* **1982**, *13*, 317.
(100) Chang, C.; Bader, R. F. W. *J. Phys. Chem.* **1992**, *96*, 1654.
(101) Bader, R. F. W.; Martin, F. J. *Can. J. Chem.* **1998**, *76*, 284.
(102) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, U.K., 1990.
(103) Mezey, P. G. *Adv. Quantum Chem.* **1996**, *27*, 163.
(104) Mezey, P. G. *Int. Rev. Phys. Chem.* **1997**, *16*, 361.
(105) Walker, P. D.; Mezey, P. G. *J. Am. Chem. Soc.* **1993**, *115*, 12423.
(106) Walker, P. D.; Mezey, P. G. *Can. J. Chem.* **2001**, *72*, 2531.
(107) Walker, P. D.; Mezey, P. G. *J. Am. Chem. Soc.* **1994**, *116*, 12022.
(108) Walker, P. D.; Mezey, P. G. *J. Math. Chem.* **1995**, *17*, 203.
(109) Walker, P. D.; Mezey, P. G. *J. Comput.Chem.* **1995**, *16*, 1238.
(110) Mezey, P. G. *Pharm. News* **1997**, *4*, 29.
(111) Mezey, P. G.; Walker, P. D. *Drug Discovery Today* **1997**, *2*, 6.
(112) Mezey, P. G. *J. Math. Chem.* **1995**, *18*, 141.
(113) Mezey, P. G. *Int. J. Quantum Chem.* **1997**, *63*, 39.
(114) Karle, J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10099.
(115) Massa, L.; Huang, L.; Karle, J. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1995**, *29*, 371.
(116) Huang, L.; Massa, L.; Karle, J. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1996**, *30*, 1691.
(117) Mezey, P. G. *Adv. Mol. Struct. Res.* **1998**, *4*, 115.
(118) Mezey, P. G. *Shape in Chemistry: An Introduction to Molecular Shape and Topology*; VCH Publishers: New York, 1993.
(119) Mezey, P. G. *Adv. Mol. Similarity* **1998**, *2*, 79.
(120) Zimpel, Z.; Mezey, P. G. *Int. J. Quantum Chem.* **1996**, *59*, 379.
(121) Mezey, P. G. Local Shape Analysis of Macromolecular Electron Densities. In *Computational Chemistry: Reviews and Current Trends*; Leszczynski, J., Ed.; World Scientific: Singapore, 1996; Vol. 1, pp 109−137.
(122) Walker, P. D.; Arteca, G. A.; Mezey, P. G. *J. Comput. Chem.* **1993**, *14*, 1172.
(123) Mezey, P. G. Molecular Surfaces. In *Reviews in Computional Chemistry*; Lipkowitz, K., Boyd, D., Eds.; VCH Publisher: Weinheim, Germany, 1990; Vol. 1, pp 265−294.
(124) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
(125) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(126) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; 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.; 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. *Gaussian 98*, revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.
(127) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.
(128) Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209.
(129) Gordon, M. S. *Phys. Lett.* **1980**, *76*, 163.
(130) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
(131) Binning, R. C., Jr.; Curtiss, L. A. *J. Comput. Chem.* **1990**, *11*, 1206.
(132) Owen, D. J.; Evans, P. R. *Science (Washington, D.C.)* **1998**, *282*, 1327.
(133) Waldherr-Teschner, M.; Goetze, T.; Heiden, W.; Knoblauch, M.; Vollhardt, H.; Brickmann, J. *MOLCAD*—Computer Aided Visualization and Manipulation of Models in Molecular Science; Second Eurographics Workshop on Visualization in Scientific Computing, Delft, Netherlands, 1991.
(134) Brickmann, J.; Keil, M.; Exner, T. E.; Marhöfer, R.; Moeckel, G. Molecular Models: Visualization. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. C., Clark, T., Gasteiger, J., Kollmann, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; Wiley & Sons: Chichester, U.K., 1998; pp 1679−1693.
(135) Brickmann, J.; Keil, M.; Exner, T. E.; Marhöfer, R. *J. Mol. Model.* **2000**, *6*, 328.
(136) *SYBYL*, version 6.7.1. Tripos Inc.: St. Louis, MO.
(137) Djordjevic, S.; Stock, A. M. *Nat. Struct. Biol.* **1998**, *5*, 446.
(138) *Fundamentals of Molecular Similarity*; Carbo-Dorca, R., Girones, X., Mezey, P. G., Eds.; Kluwer/Plenum: New York, 2001.
(139) *Topics in Current Chemistry: Molecular Similarity I;* Sen, K., Ed.; Springer-Verlag: Berlin, 1995; Vol. 173.
(140) *Topics in Current Chemistry: Molecular Similarity II;* Sen, K., Ed.; Springer-Verlag: Berlin, 1995; Vol 174.
(141) Robert, D.; Gironés, X.; Carbó-Dorca, R. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 839.
(142) Robert, D.; Amat, L. I.; Carbó-Dorca, R. *Int. J. Quantum Chem.* **2000**, *80*, 265.
(143) Gironés, X.; Amat, L. I.; Robert, D.; Carbó-Dorca, R. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 477.
(144) Carbó, R.; Leyda, L.; Arnau, M. *Int. J. Quantum Chem.* **1980**, *17*, 1185.
(145) Carbó, R.; Calabuig, B. *Int. J. Quantum Chem.* **1992**, *42*, 1681.
(146) Carbó, R.; Besalú, E.; Amat, L. I.; Fradera, X. *J. Math. Chem.* **1996**, *19*, 47.
(147) Yamano, A.; Teeter, M. M. *J. Biol. Chem.* **1994**, *269*, 13956.
(148) Mezey, P. G. *J. Comput. Chem.* **1998**, *19*, 1337.
(149) Exner, T. E.; Mezey, P. G. *J. Phys. Chem. A* **2002**, *106*, 5504.
(150) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.
(151) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.
(152) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.