Similarity and Chirality: Quantum Chemical Study of Dissimilarity of Enantiomers

Greet Boon,[†] Christian Van Alsenoy,[‡] Frank De Proft,[†] Patrick Bultinck,[§] and Paul Geerlings*[,†]

Free University of Brussels, Faculteit Wetenschappen, Eenheid Algemene Chemie, Pleinlaan 2, B-1050 Brussels, Belgium, University of Antwerp, Department of Chemistry, Universiteitsplein 1, B-2610 Antwerp, Belgium, and Ghent University, Department of Inorganic and Physical Chemistry, Krijgslaan 281 (S-3), B-9000 Gent, Belgium

Received: July 4, 2003; In Final Form: October 6, 2003

In the analysis of molecular similarity and chirality of enantiomers performed in this paper, we introduce a new local similarity index based on the Hirshfeld partitioning. In the framework of conceptual density functional theory and considering the enantiomers of the halomethane CHFClBr and of the amino acids alanine and leucine, this index is used to investigate the dissimilarity of chiral molecules. Furthermore, we illustrate Mezey's holographic electron density theorem.

1. Introduction

Until the middle of the last century, chemistry consisted mainly in comparing and classifying molecular properties and chemical reactions and rationalizing them using empirical concepts. Since the development of quantum chemistry, experimental observations and raised theories can be rationalized from first principles. In pharmacology for example, one uses this approach to describe resemblances in the physiological activity of molecules using molecular similarity.[1]

A lot of pharmacologically important molecules are chiral structures,[2] leading upon interaction with a chiral partner (e.g., a receptor) to diastereoisomeric transition states, complexes, and reaction products with different energies and properties, which has enormous consequences for the (difference in) activity of chiral pharmaca. Via the use of computer-aided molecular design, Richards et al.[3] demonstrated the existence of a quantitative structure activity relation between the potency ratio of two enantiomers, called the eudismic ratio (ER), and the "chiral coefficient" of the enantiomer pair. The ER is defined as the ratio of the potencies of the more potent enantiomer (eutomer) and the less potent one (distomer). The chiral coefficient is a quantitative index of dissimilarity between enantiomers defined as "1 − molecular similarity". Such a correlation permits the prediction, within a homologous series, of the ER of new pairs of enantiomers, which can be of great use for medicinal chemists. Here, the introduction of the concept of local chirality, instead of considering global chirality, can be of great importance as well.

Often, one considers chirality as a discrete, black−white property; a molecule is either chiral or not chiral. However, Avnir and co-workers[4] extended the treatment of symmetry as a continuous molecular structural property to chirality considering chirality as a more continuous concept, saying that one molecule can be more or less chiral compared to another one. Petitjean[5] evaluated the continuous measure of quantitative chirality as well. At this point, one can also ask if there is a connection between the difference in behavior of enantiomers and their degree of chirality.[3] In this work, we indeed assume that the degree of chirality is linked to the (dis)similarity of two enantiomers reducing the problem to how to quantify molecular similarity.

In recent years several similarity indices were proposed, among which were the indices defined by Carbó, Hodgkin, and Richards based on the electron density.[6,7] These indices describe the global similarity of the total systems. It can however be important to have an idea about the local similarity of certain regions of the systems under consideration but, until now, only a few local similarity indices were proposed. With an expression analogous to the Carbó index, Lee and co-workers, for example, defined such a local similarity index in terms of fragments of electron densities of the molecules.[8] Also Mezey et al.[9] studied similarity locally, comparing the values of ab initio quantum similarity measures (QSM), calculated from the electron densities of arbitrary molecular fragments, with their self-similarity measures (vide infra). Furthermore, Mezey used in ref 9 the MQSM as a measure of molecular chirality as well, going along with an idea of Carbó reported in ref 10.

The scope of this study is, next to using global indices, to propose and use a new local similarity index in the analysis of molecular similarity and, as a consequence, in the analysis of molecular chirality. To the best of our knowledge, only Mezey[9] published a study in this framework, where, however, the comparison of global and local indices is not performed on a uniform ab initio or DFT type level.

Furthermore, in this work, the proposed local index is used to investigate the relationship between the dissimilarity and the optical activity of chiral molecules and for numerical tests on Mezey's holographic electron density theorem.[11]

To develop a methodology to quantify the dissimilarity of enantiomers, we studied the prototype of all chiral molecules, CHFClBr and the amino acids alanine and leucine, all of them containing only one chiral center, which reflects the situation of many active compounds in pharmacology. We considered these systems for different reasons. As the textbook example

* Corresponding author. E-mail: pgeerlin@vub.ac.be.
[†] Free University of Brussels.
[‡] University of Antwerp.
[§] Ghent University.

10.1021/jp035930u CCC: $25.00
Published on Web 11/15/2003

chiral molecule,[12] the halomethane CHFClBr is in a first phase an ideal test molecule to study the dissimilarity and local chirality of its enantiomers. Leucine and alanine, among the simplest "biomolecules", are then considered in order to investigate the relationship between their degree of chirality, and thus their dissimilarity, and their optical activity.

This work is performed within the framework of conceptual density functional theory (DFT).[13−16] In a first step, we concentrate here on the use of the electron density based similarity indices, whereas future studies may be devoted to evaluate (dis)similarity, and thus chirality, of enantiomers using earlier-introduced DFT-based reactivity-related similarity indices (e.g., using the local softness)[17,18] as well.

2. Theory and Computational Details

2.1. Similarity Indices. As mentioned in the Introduction, Carbó was the first to define a similarity index R_{AB}[6] between two molecules A and B with electron densities $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$, based on the idea of minimizing the following expression

$$\epsilon_{AB} = \int |\rho_A(\mathbf{r}) - \rho_B(\mathbf{r})|^2 \, d\mathbf{r} = \int \rho_A{}^2(\mathbf{r}) \, d\mathbf{r} + \int \rho_B{}^2(\mathbf{r}) \, d\mathbf{r} - 2\int \rho_A(\mathbf{r})\rho_B(\mathbf{r}) \, d\mathbf{r} = Z_{AA} + Z_{BB} - 2Z_{AB} \quad (1)$$

involving the overlap integral Z_{AB} between the electron densities of molecule A and B.

Z_{AA} and Z_{BB} are called the self-similarities of molecules A and B.[19]

Furthermore

$$\epsilon_{AB} = D_{AB}{}^2 \quad (2)$$

with D_{AB} as the Euclidean distance[6] between the densities.

The resulting expression for R_{AB}, the Carbó index, is written as

$$R_{AB} = \frac{\int \rho_A(\mathbf{r})\rho_B(\mathbf{r}) \, d\mathbf{r}}{((\int \rho_A{}^2(\mathbf{r}) \, d\mathbf{r})(\int \rho_B{}^2(\mathbf{r}) \, d\mathbf{r}))^{1/2}} = \frac{Z_{AB}}{(Z_{AA}Z_{BB})^{1/2}} \quad (3)$$

On the basis of the fact that R_{AB} was shown to measure only the "shape similarity", Hodgkin and Richards proposed the index H_{AB}[7]

$$H_{AB} = \frac{2\int \rho_A(\mathbf{r})\rho_B(\mathbf{r}) \, d\mathbf{r}}{\int \rho_A{}^2(\mathbf{r}) \, d\mathbf{r} + \int \rho_B{}^2(\mathbf{r}) \, d\mathbf{r}} = \frac{2Z_{AB}}{Z_{AA} + Z_{BB}} \quad (4)$$

This index now describes the similarity of shape and extent of the electron distributions.

2.1.1. Global Similarity Indices for Enantiomers. For the *R* and *S* enantiomers of a chiral molecule, one can obviously write

$$Z_{RR} = Z_{SS} \quad (5)$$

which, used in eqs 3 and 4, yields for two enantiomers

$$R_{RS} = H_{RS} \quad (6)$$

Up until now, an important drawback of these Carbó and Hodgkin−Richards similarity indices was, together with the dependence of their value on the molecular superposition,[20,21] the time-consuming three-dimensional integration. Mestres et al. developed the program MESSEM (Mesures de Semblanca)[22,23] in order to analytically evaluate the integrals needed. In this work however, we used a highly efficient implementation of the integrals, both numerically (program STOCK)[24] and analytically (program BRABO), which allows calculating similarity integrals at a negligible computational cost. Both programs are part of the BRABO program package developed by Van Alsenoy et al.[25,26]

2.1.2. Local Similarity Indices for Enantiomers. Following the extension of the theorems of Hohenberg and Kohn by Riess and Münch,[27] Mezey proposed the holographic electron density theorem,[11] stating that the ground-state electron density $\rho_\Omega(\mathbf{r})$ of a finite but otherwise arbitrary subdomain Ω uniquely determines all ground-state properties (among others the electron density) in Ω, in any other subdomain Ω', and in the total domain of the boundaryless system. This inherent property of molecular electron densities, called the holographic property, provides a strong basis for local quantum chemical similarity analyses. According to our work, this implies that in evaluating molecular similarity we could focus in principle on a certain molecular region in the chiral system (for example the atomic region around the asymmetric carbon atom). Notice however that, as a consequence of Mezey's theorem, also regions around other, nonchiral atoms contain all the information about the system, thus also about its chirality.

In this work we put this into practice by restricting ourselves to certain atomic regions using the Hirshfeld partitioning discussed in the following part.

Hirshfeld partitioning. The partitioning of the electron density developed by Hirshfeld,[28] and recently used by the present authors in the study of atomic charges, dipole moments, and Fukui functions,[29−31] partitions the total electron density $\rho(\mathbf{r})$ of a molecule in atomic contributions $\rho_A(\mathbf{r})$.

These atomic contributions are proportional to the weight $w_A(\mathbf{r})$ of the electron density of the isolated molecule in the so-called promolecular density;[28] this weight is defined as the ratio of the electron density of the isolated atom and the density built up from the superposition of the isolated densities of all the atoms located on the same positions as those in the molecule ("the promolecular density").

Indeed, the contribution of atom A, $\rho_A(\mathbf{r})$, to the electron density $\rho(\mathbf{r})$ is given as

$$\rho_A(\mathbf{r}) = w_A(\mathbf{r})\rho(\mathbf{r}) \quad (7)$$

with

$$w_A(\mathbf{r}) = \frac{\rho_A^0(\mathbf{r})}{\sum_X \rho_X^0(\mathbf{r})} \quad (8)$$

where $\rho_A^0(\mathbf{r})$ is the electron density of the isolated atom A and where $\Sigma_X \rho_X^0(\mathbf{r})$ is the density built from the superposition of the densities of the isolated atoms placed on the same positions as in the molecule ("the promolecular density"). It is clear from eq 8 that the weight coefficients *w* are always positive.

In this study, we propose to convert the global Carbó index into a local index using this Hirshfeld partitioning.

Analogous to the expressions (eqs 7 and 8) and considering the *R* and *S* enantiomers of a chiral molecule, the contribution of, for example, the asymmetric carbon atom to the total electron density $\rho_R(\mathbf{r})$ of the *R* enantiomer can be written as

$$\rho_{C,R}(\mathbf{r}) = w_{C,R}(\mathbf{r})\rho_R(\mathbf{r}) \quad (9)$$

with

$$w_{C,R}(\mathbf{r}) = \frac{\rho^0_{C,R}(\mathbf{r})}{\sum_X \rho^0_{X,R}(\mathbf{r})} \tag{10}$$

and the contribution of the asymmetric carbon atom to the total electron density $\rho_S(\mathbf{r})$ of the *S* enantiomer as

$$\rho_{C,S}(\mathbf{r}) = w_{C,S}(\mathbf{r})\rho_S(\mathbf{r}) \tag{11}$$

with

$$w_{C,S}(\mathbf{r}) = \frac{\rho^0_{C,S}(\mathbf{r})}{\sum_Y \rho^0_{Y,S}(\mathbf{r})} \tag{12}$$

In the numerical calculation of the overlap integral $\int \rho_R(\mathbf{r})\rho_S(\mathbf{r})\,d\mathbf{r}$, the integrand $\rho_R(\mathbf{r})\rho_S(\mathbf{r})$ is evaluated on a combined grid of the aligned enantiomers. The contribution of one particular kind of atom of the grid, e.g., the asymmetric carbon, to the total integral thus needs the consideration of two contributions, the carbon atom in the *R* enantiomer and the one in the *S* enantiomer.

As such, we propose to write the contribution of, for example, the asymmetric carbon atom to the product $\rho_R(\mathbf{r})\rho_S(\mathbf{r})$ as

$$\rho_{C,R+S}(\mathbf{r}) = w_{C,R+S}(\mathbf{r})\rho_R(\mathbf{r})\rho_S(\mathbf{r}) \tag{13}$$

where

$$w_{C,R+S}(\mathbf{r}) = \frac{\rho^0_{C,R}(\mathbf{r}) + \rho^0_{C,S}(\mathbf{r})}{\sum_X \rho^0_{X,R}(\mathbf{r}) + \sum_Y \rho^0_{Y,S}(\mathbf{r})} \tag{14}$$

where $\Sigma_X\, \rho^0_{X,R}(\mathbf{r}) + \Sigma_Y\, \rho^0_{Y,S}(\mathbf{r})$ is the total promolecular superimposed density of the two enantiomers with their asymmetric carbon atoms superimposed.

The numerator Z_{RS} of the Carbó index then becomes

$$Z_{RS}^{\text{local,C}} = \int w_{C,R+S}(\mathbf{r})\rho_R(\mathbf{r})\rho_S(\mathbf{r})\,d\mathbf{r} = \int \left(\frac{\rho^0_{C,R}(\mathbf{r}) + \rho^0_{C,S}(\mathbf{r})}{\sum_X \rho^0_{X,R}(\mathbf{r}) + \sum_Y \rho^0_{Y,S}(\mathbf{r})} \right) \rho_R(\mathbf{r})\rho_S(\mathbf{r})\,d\mathbf{r} \tag{15}$$

where the global index is thus partitioned into atomic contributions.

For this work, the local index (eq 15) based on the Hirshfeld partitioning is, in view of studying local similarity and chirality, implemented numerically in the program STOCK[24] mentioned earlier.

2.1.3. Relative Orientation of the Enantiomers. An important drawback of the Carbó and Hodgkin−Richards index is their dependence on the relative orientation of the molecules under consideration. Several methods have already been proposed to establish a criterion on how molecules might be superposed, such as aligning the molecules according to common physicochemical features (for example, matching different three-dimensional molecular fields such as steric, electrostatic, or hydrophobic fields)[32] or alignment of the molecules based on topological and geometrical features only; a procedure called topo-geometrical superposition algorithm (TGSA)[33] based on comparisons of atom types and interatomic distances. Other methods opt to align the molecules in such a way that the resulting molecular similarity measure is maximized.[34] In this framework, a new algorithm, the quantum similarity superposition algorithm (QSSA), was recently designed by Bultinck et al.,[20,21] expressing the relative position of two molecules in terms of mutual translation in three Cartesian directions and three Euler angles. The quantum similarity overlap, considering the electron densities of the molecules within the promolecular atomic shell approximation (PASA)[35] and considering the atomic electron densities within the atomic shell approximation (ASA),[36,37] is then used to optimize the mutual positions and similarity of the molecules, using a Lamarckian genetic algorithm.

In this work however, we decided to superimpose the asymmetric carbon atom and two of its directly bounded substituents of both enantiomers under consideration, leading to six different orientations. This methodology enables us, as opposed to the usage of, for example, QSSA, to evaluate, next to global similarity, local similarity using our proposed local similarity index (eq 15). Depending on which superposition method we choose, and thus according to what property we are precisely interested in, global and local similarity can be studied complementary to one another.

In this way, the choice of CHFClBr as a pilot system has the additional advantage that intuitively we expected that the optimal relative orientation of the two enantiomers using QSSA would coincide to a large extent with one of our six orientations where the asymmetric carbon atom and two of its substituents are superimposed (see section 3.3.).

3. Results and Discussion

All charge densities used in this work were calculated using the GAUSSIAN 98[38] program at the B3PW91/6-311G* level,[39−41] which can be expected to be adequate for the purpose of our study.[42]

3.1. Global and Local Similarity Indices for Enantiomers. The results for the calculation of the global similarity index (eq 3), and thus also (eq 4), and the proposed local similarity index based on the Hirshfeld partitioning (eq 15) are shown in Table 1 and Table 2.

Together with the similarity index, the numerator Z_{RS} of the Carbó index and the Euclidean distance D_{RS} are given in these (and following) tables, both however having the disadvantage of not being normalized. Therefore, the use of these latter two quantities enables us to evaluate the generated similarity values according to different orientations of one enantiomer pair, and not between different pairs of enantiomers, only. The Euclidean distance D_{RS}[6] is calculated via

$$D_{RS} = \left(\int (\rho_R(\mathbf{r}) - \rho_S(\mathbf{r}))^2\,d\mathbf{r}\right)^{1/2} = (Z_{RR} + Z_{SS} - 2Z_{RS})^{1/2} \tag{16}$$

Note that this quantity can still be termed a Euclidean distance as the triangular condition[20,21] $|D_{RA} - D_{AS}| < D_{RA} < D_{RA} + D_{AS}$ is still fulfilled.

All results concern both analytical and numerical evaluations for the global index, deviating from each other, as shown in **Table 1**, by a factor of maximum 10^{-3}, which indicates that the respective numeric and analytic expressions were correctly implemented, and numerical values for the local index.

From these results it was seen that the core electrons of both the *R* and *S* asymmetric carbon atom contributed in a similar dominant way to the similarity index (cf. Table 2 the value for

TABLE 1: CHFClBr, Alanine, and Leucine: Global Similarity Index (Both Analytical and Numerical) Using Global Densities

superimposed atoms	Carbó index	Euclidean D_{RS}	Z_{RS}
CHFClBr Analytical			
ClCBr	0.990	15.393	11617.0
FCBr	0.915	44.624	10739.8
HCBr	0.906	46.870	10637.1
FCCl	0.098	145.516	1148.1
HCCl	0.089	146.216	1046.0
HCF	0.054	148.974	638.9
CHFClBr Numerical			
ClCBr	0.990	15.393	11615.5
FCBr	0.915	44.624	10737.9
HCBr	0.906	46.870	10635.2
FCCl	0.098	145.507	1148.5
HCCl	0.089	146.207	1046.4
HCF	0.055	148.956	640.2
Alanine Analytical			
COOH C NH_2	0.396	19.342	122.66
H C NH_2	0.389	19.461	120.35
CH_3 C NH_2	0.375	19.682	115.98
H C CH_3	0.357	19.955	110.61
CH_3 C COOH	0.317	20.575	98.06
H C COOH	0.310	20.671	96.06
Alanine Numerical			
COOH C NH_2	0.396	19.342	122.68
H C NH_2	0.387	19.469	119.73
CH_3 C NH_2	0.375	19.682	116.00
H C CH_3	0.361	19.935	112.05
CH_3 C COOH	0.317	20.574	98.07
H C COOH	0.310	20.680	95.89
Leucine Analytical			
H C NH_2	0.326	23.345	132.043
COOH C NH_2	0.311	23.613	125.772
CH_2C... C NH_2	0.288	23.999	116.560
CH_2C... C COOH	0.278	24.174	112.363
H C CH_2C...	0.269	24.317	108.888
H C COOH	0.229	24.983	92.466
Leucine Numerical			
H C NH_2	0.327	23.341	132.076
COOH C NH_2	0.311	23.608	125.857
CH_2C... C NH_2	0.288	23.999	116.581
CH_2C... C COOH	0.277	24.182	111.992
H C CH_2C...	0.265	24.345	106.599
H C COOH	0.228	24.986	92.452

Z_{RS} for the contribution of the asymmetric carbon atom to the total integral for CHFClBr, alanine, and leucine is always typically around 31). Also Mezey et al.[9] find these dominant contributions in their local similarity analysis within the PASA model. The approximate analytical calculation of the density overlap integral $Z_{RS} = \int \rho_{R,C}(\mathbf{r})\rho_{S,C}(\mathbf{r})\, d\mathbf{r}$ for the core electrons of the carbon atom is given in Appendix.

In our study however, we tried to avoid this effect by using in the expressions of the similarity indices the product of the density differences $\Delta\rho$**,** well-known to represent bonding characteristics in molecules, instead of the global densities $\rho(\mathbf{r})$ of the two molecules under consideration.

The density difference $\Delta\rho_R$ of the R enantiomer is defined as

$$\Delta\rho_R = \rho_R - \rho_R^0 \tag{17}$$

with ρ_R^0 the promolecular density of the R enantiomer, yielding the following expression for the numerator of the Carbó index

$$Z_{RS} = \int \Delta\rho_R \Delta\rho_S \, d\mathbf{r} = \int (\rho_R - \rho_R^0)(\rho_S - \rho_S^0)\, d\mathbf{r} \tag{18}$$

Here, we did not opt to calculate this overlap integral using valence orbital densities instead of density differences because the selection of the valence orbitals might not always be as straightforward as expected.

TABLE 2: Local Counterparts of the Carbó Index (Eq 2) and Euclidean Distance (Eq 18) and Contribution of a Given Atom Type to ZRS Based on the Hirshfeld Partitioning (Eq 17) Using Global Densities for CHFClBr, Alanine, and Leucine

superimposed atoms	Carbó index	Euclidean D_{RS}	Z_{RS}
A. CHFClBr Asymmetric Carbon Atom			
HCF	0.999	0.272	31.396
FCBr	0.998	0.346	31.381
FCCl	0.998	0.392	31.373
HCBr	0.996	0.479	31.356
ClCBr	0.996	0.484	31.350
HCCl	0.995	0.548	31.338
B. CHFClBr Hydrogen Atom			
HCF	0.991	0.045	0.1162
HCBr	0.980	0.069	0.1155
HCCl	0.971	0.082	0.1150
FCBr	0.289	0.740	0.1116
ClCBr	0.223	0.872	0.1345
FCCl	0.214	0.968	0.1037
C. CHFClBr Bromine Atom			
FCBr	1.00000	0.087	10586.0300
HCBr	1.00000	0.095	10586.0295
ClCBr	1.00000	0.098	10586.0293
HCF	0.03022	142.790	317.6481
HCCl	0.00135	145.374	14.3352
FCCl	0.00006	145.504	0.6006
D. Alanine Asymmetric Carbon Atom			
H C NH_2	1.000	0.174	31.395
H C CH_3	0.999	0.211	31.391
H C COOH	0.999	0.239	31.388
COOH C NH_2	0.999	0.266	31.385
CH_3 C NH_2	0.999	0.267	31.387
CH_3 C COOH	0.998	0.346	31.372
E. Alanine Hydrogen Atom			
H C NH_2	0.994	0.037	0.1143
H C CH_3	0.992	0.042	0.1142
H C COOH	0.991	0.045	0.1141
COOH C NH_2	0.433	0.561	0.1201
CH_3 C NH_2	0.432	0.564	0.1208
CH_3 C COOH	0.353	0.680	0.1260
F. Leucine Asymmetric Carbon Atom			
H C NH_2	1.000	0.154	31.401
H C CH_2C...	0.999	0.219	31.395
H C COOH	0.999	0.251	31.392
COOH C NH_2	0.999	0.267	31.390
CH_2C... C NH_2	0.999	0.271	31.389
CH_2C... C COOH	0.998	0.346	31.377
G. Leucine Hydrogen Atom			
H C NH_2	0.990	0.049	0.1145
H C CH_2C...	0.988	0.054	0.1144
H C COOH	0.986	0.057	0.1143
COOH C NH_2	0.438	0.556	0.4375
CH_2C... C NH_2	0.432	0.565	0.1217
CH_2C... C COOH	0.354	0.678	0.1264

Thus, in this work we can discern four types of similarity indices summarized in the following chart, where reference is made to the corresponding tables in the paper

CHART 1: Different Levels Used in this Study To Evaluate Dissimilarity Between Enantiomers

	global density $\rho(\mathbf{r})$	density difference $\Delta\rho(\mathbf{r})$
global similarity index	Table 1	Table 3
local similarity index	Table 2	Table 4

The results for the calculations of the global and local indices using the density difference are given in Table 3 and Table 4, respectively. Note the possibility of generating a negative value

TABLE 3: Analytical Global Similarity Index Using Density Differences for CHFClBr, Alanine, and Leucine

superimposed atoms	Carbó index	Euclidean D_{RS}	Z_{RS}
	CHFClBr		
FCBr	0.941	0.202	0.324
HCBr	0.910	0.249	0.314
ClCBr	0.874	0.295	0.301
HCF	0.146	0.767	0.050
FCCl	0.125	0.776	0.043
HCCl	0.091	0.791	0.031
	Alanine		
CH_3 C NH_2	0.543	0.355	0.075
COOH C NH_2	0.516	0.371	0.073
H C COOH	0.467	0.389	0.066
H C CH_3	0.441	0.399	0.063
CH_3 C COOH	0.426	0.404	0.061
H C NH_2	0.387	0.418	0.055
	Leucine		
COOH C NH_2	0.398	0.522	0.090
CH_2C... C COOH	0.365	0.536	0.083
CH_2C... C NH_2	0.312	0.558	0.071
H C COOH	0.266	0.576	0.060
H C NH_2	0.249	0.583	0.056
H C CH_2C...	0.243	0.585	0.055

for the Carbó index (and thus for Z_{RS}) as this is, for example, the case for the orientation in part B of Table 4 where the asymmetric carbon atoms and the substituents chlorine and bromine are superimposed. In fact, calculating similarity is measuring the distance between density functions, in this case, the distance between density differences. This results in searching for a least-squares fit which can also be performed using functions which can be smaller or equal to zero, as, for example, the density difference $\Delta\rho$.

$$D_{RS} = \left(\int (\Delta\rho_R - \Delta\rho_S)^2 \, d\mathbf{r}\right)^{1/2} = \left(\int \Delta\rho_R{}^2 \, d\mathbf{r} + \int \Delta\rho_S{}^2 \, d\mathbf{r} - 2\int \Delta\rho_R \Delta\rho_S \, d\mathbf{r}\right)^{1/2} \quad (19)$$

For two enantiomers, one obviously has $\int \Delta\rho_R{}^2 \, d\mathbf{r} = \int \Delta\rho_S{}^2 \, d\mathbf{r}$ and the expression for D_{RS} becomes

$$D_{RS} = \left(2\int \Delta\rho_R{}^2 \, d\mathbf{r} - 2\int \Delta\rho_R \Delta\rho_S \, d\mathbf{r}\right)^{1/2} \quad (20)$$

Minimizing D_{RS} corresponds with searching maximum similarity values. Note that, in the case of using density differences, the similarity index still has a maximum value of 1 but can be smaller than zero as densities are positive definite while density differences are not.

The results given in Table 2 and Table 4 are a numerical illustration of Mezey's holographic electron density theorem; when evaluating regions around nonasymmetric atoms using the Hirshfeld technique as described above, for example, the region around the hydrogen atom, we clearly notice that these nonasymmetric atomic regions as well contain information about the chirality of the systems.

Taking a look at part A of Table 2, we nearly do not see any difference between the generated values for the Carbó index while the results in part A of Table 4, using the density differences, show a range of values for this parameter. This points out that using density differences in the expressions of the indices, instead of using global densities, gives different and complementary information about the similarity of the systems because here the identical dominant contribution of the core electrons is eliminated.

Looking at parts B, E, and G of Table 2 and also looking at parts B, E, and G of Table 4 shows that highest values of similarity are generated for the orientations where the hydrogen atoms are superimposed while the three other orientations give rather low similarity values. Analogous part C Tables 2 and 4 for CHFClBr are giving the highest similarity values for the orientations where the bromine atoms are superimposed.

From the results in Table 1 for CHFClBr, the highest similarity is obtained for the orientation where the asymmetric carbon and the substituents chlorine and bromine are superimposed, while this orientation in part A of Table 4 generates the lowest value of similarity. The opposite is true for orientation where the asymmetric carbon and the substituents hydrogen and fluorine are superimposed.

It seems that in the global similarity index, the core size of the substituents is dominant while the local indices with density differences clearly contain different and complementary information.

TABLE 4: Local Counterparts of the Carbó Index (Eq 2) and Euclidean Distance (Eq 18) and Contribution of a Given Atom Type to ZRS Based on the Hirshfeld Partitioning (Eq 17) Using Density Differences for CHFClBr, Alanine, and Leucine

superimposed atoms	Carbó index	Euclidean D_{RS}	Z_{RS}
	A. CHFClBr Asymmetric Carbon Atom		
HCF	0.995	0.012	0.0157
HCBr	0.951	0.039	0.0150
HCCl	0.940	0.044	0.0149
FCBr	0.623	0.109	0.0098
FCCl	0.620	0.111	0.0100
ClCBr	0.446	0.127	0.0065
	B. CHFClBr Hydrogen Atom		
HCF	1.000	0.002	0.00619
HCBr	0.996	0.007	0.00617
HCCl	0.995	0.008	0.00616
FCBr	0.375	0.065	0.00126
FCCl	0.359	0.071	0.00142
ClCBr	−0.191	0.084	−0.00057
	C. CHFClBr Bromine Atom		
HCBr	1.000	0.009	0.2919
FCBr	1.000	0.014	0.2918
ClCBr	1.000	0.016	0.2916
HCF	0.021	0.668	0.0048
HCCl	0.008	0.755	0.0024
FCCl	0.003	0.764	0.0009
	D. Alanine Asymmetric Carbon Atom		
H C NH_2	0.956	0.039	0.0166
CH_3 C NH_2	0.954	0.041	0.0170
H C COOH	0.953	0.041	0.0165
COOH C NH_2	0.937	0.047	0.0163
H C CH_3	0.926	0.051	0.0159
CH_3 C COOH	0.865	0.068	0.0148
	E. Alanine Hydrogen Atom		
H C NH_2	0.989	0.011	0.00521
H C COOH	0.988	0.011	0.00520
H C CH_3	0.984	0.013	0.00518
CH_3 C NH_2	0.793	0.044	0.00365
COOH C NH_2	0.570	0.061	0.00248
CH_3C C OOH	0.287	0.072	0.00104
	F. Leucine Asymmetric Carbon Atom		
H C NH_2	0.957	0.039	0.016748
H C COOH	0.942	0.045	0.016413
COOH C NH_2	0.938	0.047	0.016410
CH_2C... C NH_2	0.935	0.048	0.016383
H C CH_2C...	0.925	0.051	0.015983
CH_2C... C COOH	0.863	0.069	0.014818
	G. Leucine Hydrogen Atom		
H C NH_2	0.984	0.013	0.005242
H C COOH	0.983	0.013	0.005236
H C CH_2C...	0.982	0.014	0.005222
CH_2C... C NH_2	0.762	0.047	0.003599
COOH C NH_2	0.592	0.059	0.002541
CH_2C... C COOH	0.287	0.072	0.001038

TABLE 5: CHFClBr, Alanine, and Leucine

molecule	superimposed atoms	Carbó index *a*	*b*	*c*
CHFClBr	Cl C Br	0.990	0.993	0.997
Alanine	COOH C NH_2	0.396	0.743	0.745
Leucine	H C NH_2	0.326	0.563	0.566

[a] Relative orientation of the enantiomers with the asymmetric carbon atoms and two of its directly bounded substituents superimposed. Global similarity index using global densities. [b] Relative orientation of the enantiomers maximized using QSSA. Global similarity index using ASA densities. [c] Relative orientation of the enantiomers maximized using QSSA. Global similarity index using global densities.

3.2. Relationship between the Dissimilarity and the Optical Activity of L- and D-Amino Acids. Supposing that the optical activity, as quantified in a standardized way by the specific rotation $[\alpha]_D$, is a good experimentally measurable quantity for the degree of chirality of a molecule and also assuming that the dissimilarity between enantiomers describes their chirality, one can expect a link between the dissimilarity and the optical activity of enantiomers, as was previously shown by Mezey et al.[9]

In a first stage of this study, we chose to consider two aliphatic amino acids, alanine and leucine, with the lowest and highest (absolute value of) $[\alpha]_D$, respectively, and experimentally measured under the same circumstances. $[\alpha]_D$ for alanine is 2.7 and $[\alpha]_D$ for leucine is -10.8, both measured in water as solvent and at the sodium D line.[43]

Here, we could not take into account the study of the chiral molecule CHFClBr, as its $[\alpha]_D$ is measured under different circumstances than those measured for the amino acids, the maximum specific rotation (i.e., the rotation for the pure enantiomer) of CHFClBr (neat, $\rho = 1.91$ kg dm^{-3}) measured at the sodium D line is ± 1.6,[44] and as a consequence, comparing the generated similarities would not be useful at all.

Taking a look at Tables 1 and 3, we see that for all orientations the values for the global similarities for alanine are higher than those for leucine. This is also seen, except, with a small difference of a factor 10^{-3}, for the orientation where the asymmetric carbon, the amine function, and the hydrogen are superimposed, comparing the values given in parts D and E of Table 4 (alanine) with those given in parts F and G of Table 4 (leucine). This could be expected as alanine has a smaller $[\alpha]_D$ than leucine and, as a consequence, the enantiomers of alanine are globally more similar, thus less dissimilar and the similarity index is higher, than the enantiomers of leucine.

The range of differences, both for the halomethane and the amino acids, between the similarity values in Table 2 is not very large in comparison with the values given in Table 4, again pointing out the importance of using density differences, instead of global densities, giving additional information.

As mentioned earlier, global and local similarity indices both contain different information.

3.3. Global Similarity Indices Using QSSA. So far, we calculated in this work global (and local) similarity indices for the enantiomers of CHFClBr, alanine, and leucine, where we have chosen to superimpose the asymmetric carbon atoms and two of its directly bounded substituents generating six possible orientations.

Additionally we considered the calculation of global indices where the QSSA method is used in order to align the molecules according to their maximum similarity value, calculated within the ASA model.[20,21]

Table 5 represents, for the three molecules under consideration, DFT-based global similarity indices (eq 3) calculated using global densities using the earlier mentioned program BRABO. Furthermore, only the results for the orientation yielding the maximum similarity values are given.

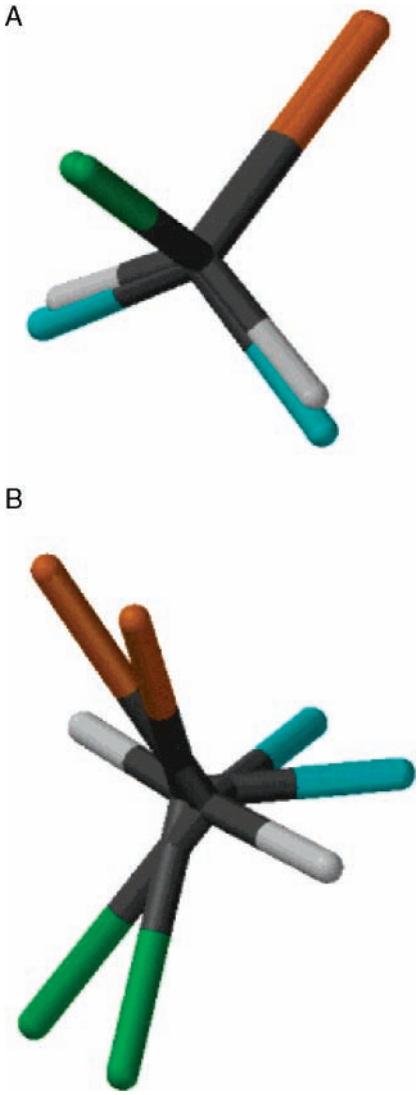

Figure 1. (a) Relative orientation of the *R* and *S* enantiomers of CHFClBr. The asymmetric carbon atoms, chlorine, and bromine are superimposed. (b) Relative orientation of the *R* and *S* enantiomers of CHFClBr maximized using QSSA.

In Table 5, the relative orientation of the molecules is obtained, using the QSSA method where the similarity value given is the maximized one using the ASA model.

These optimal orientations generated within the QSSA model are then used to calculate DFT-based global similarity indices (eq 3) using ab initio global densities and using the program BRABO. The results are given in Table 5.

From these results, it can be seen that the goal of considering the molecule CHFClBr, as mentioned earlier, is indeed justified. First of all, being one of the simplest chiral molecules, CHFClBr could be considered as an ideal test system to analyze both global and local similarity. Furthermore, looking at the results for CHFClBr given in Table 5 and visualizing the generated relative orientations of the enantiomers in parts a and b of Figure 1, it is clearly shown that these orientations coincide to a very large extent, as we intuitively expected.

On the other hand, visualization (parts a and b of Figure 2) of, for example, the orientations for alanine used in Table 5, shows that for the optimal relative orientation obtained using

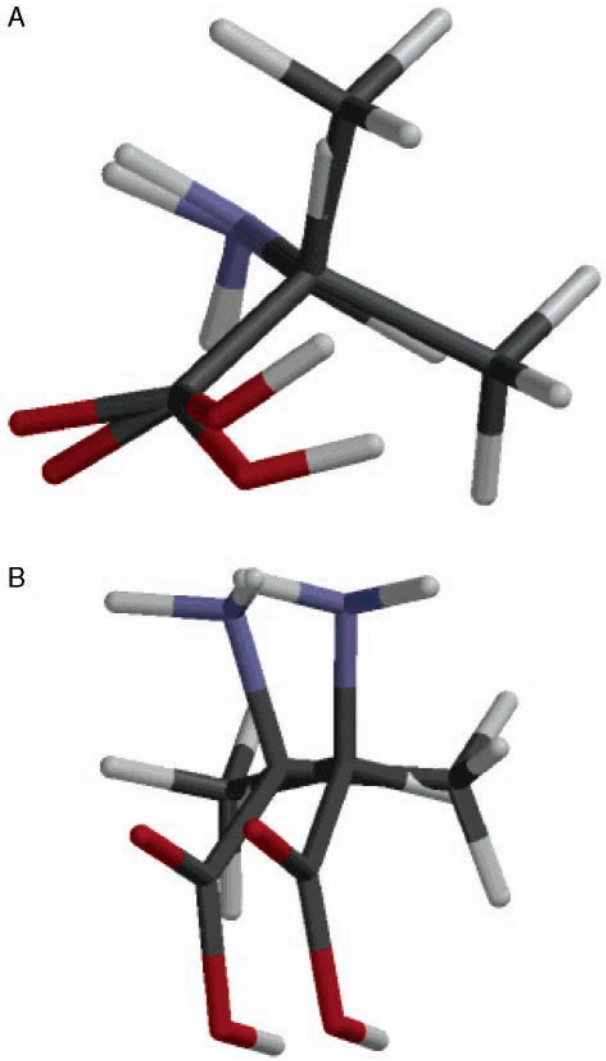

Figure 2. (a) Relative orientation of the *R* and *S* enantiomers of alanine. The asymmetric carbon atoms and the substituents COOH and NH_2 are superimposed. (b) Relative orientation of the R and S enantiomers of alanine maximized using QSSA.

QSSA, the asymmetric carbon atoms of the enantiomers do not coincide, preventing to analyze local similarity using this approach.

Furthermore, for the three similarity analyses given in Table 5, the trends are the same; the values for the similarities for alanine are higher than those for leucine. This, as already explained in section 3.2., could be expected as alanine has a smaller absolute value for $[\alpha]_D$ than leucine and, as a consequence, the enantiomers of alanine are more similar than the enantiomers of leucine.

4. Conclusions

The results for the prototype chiral molecule containing only one chiral center, CHFClBr, show that the proposed methodology is adequate to estimate the dissimilarity between enantiomers.

The study on the amino acids alanine and leucine shows that global dissimilarity between pairs of enantiomers is related to their optical activity (global chirality). The optical activity can be considered as a global property of the whole molecule where the positioning of the atom groups around the asymmetric carbon atom plays an important role. Global (dis)similarity can be studied using global electron densities and, giving complementary results, density differences.

On the other hand, dealing with the local approach of studying similarity, we proposed a new local similarity index (eq 15) based on the Hirshfeld partitioning and we illustrated the holographic electron density theorem. Here, it is important that, even considering nonasymmetric atomic regions, local dissimilarity, and thus "local" chirality, could be demonstrated. Comparison of the results generated for alanine and leucine using global densities in the expression of the local similarity index shows only very small variations while using density differences a better distinction between them can be made, although the variations remain small. These small variations could even be expected realizing that the difference between the electron density around, for example, the asymmetric carbon atom of the enantiomers is not very large.

In fact we can say that, on a micro level, local similarity and local chirality reflect all the aspects of global similarity and global chirality. This opens up new ways for the theoretical study of the different properties of chiral systems of, among others, pharmaceutical importance. The generalization of the ansatz presented in this paper to the case of more than one asymmetric centers is in progress using an atom-by-atom procedure (cf. Cioslowski's work).[45,46]

Appendix

Approximate analytical calculation of the density overlap integral $Z_{RS} = \int \rho_{R,C}(\mathbf{r})\rho_{S,C}(\mathbf{r})\, d\mathbf{r}$ for the core electrons of the carbon atom. The electron density of a core electron pair of a molecule can be written as

$$\rho_{\text{core}} = 2|\psi|^2 \tag{A.1}$$

where ψ is a molecular core orbital.

Simplifying the molecular orbital ψ to an atomic orbital χ_{ns}, we have

$$\psi \approx \chi_{\text{ns}} \tag{A.2}$$

We will represent χ_{ns} by the following normalized Slater-type atomic orbital

$$\chi_{\text{ns}} = \frac{(2\zeta)^{n+(1/2)}}{((2n)!)^{1/2}} r^{n-1} e^{-\zeta r} \frac{1}{(4\pi)^{1/2}} \tag{A.3}$$

For the 1s orbital of carbon and for an orbital exponent $\zeta_{1s} = 5.70$, derived using the Slater rules,[47] we have

$$\chi_{1s} = \frac{(2\zeta)^{(3/2)}}{((2)!)^{1/2}} \mathrm{e}^{-\zeta r} \frac{1}{(4\pi)^{1/2}} \tag{A.4}$$

and the overlap integral Z_{RS} becomes

$$\begin{aligned} \int \rho_{R,C}(\mathbf{r})\rho_{S,C}(\mathbf{r})\, d\mathbf{r} &= 4\int |\chi_{1s}|^2 |\chi_{1s}|^2\, d\mathbf{r} \\ &= 4\int \frac{(2\zeta)^6}{2^2} \mathrm{e}^{-4\zeta r} \frac{1}{(4\pi)^2} d\mathbf{r} \\ &= 29.5 \text{ for } \zeta = 5.70 \end{aligned} \tag{A.5}$$

This calculated value of 29.5 indeed covers the main dominant part of the generated values of 31 for Z_{RS} in Table 2, justifying the use of density differences.

Acknowledgment. P.G. wishes to thank the Free University of Brussels for a generous computer grant and the Fund for Scientific Research—Flanders for continuous support. P.G., G.B.,

and F.D.P. thank Dr. W. Langenaeker and Prof. J. P. Tollenaere (Janssen Research Foundation) for many stimulating discussions in an initial phase of the work and Drs. J. Baert for his contribution to the initial calculations. P.G. thanks Prof. R. Kieboom from the Mathematics Department for an in-depth discussion on the existence of the Z_{AA}, ... integrals. C.V.A. acknowledges support by the University of Antwerp under Grant GOA-BOF-UA nr 23. PB wishes to thank the *Fund for Scientific Research—Flanders* (Belgium) for their grants to the Ghent Quantum Chemistry Group and acknowledges the *European Community—Access to Research Infrastructure action* of the *Improving Human Potential Program*, allowing the use of the CEPBA infrastructure at the PolyTechnical University of Catalonia (Spain).

References and Notes

(1) Rouvray, D. H. In *Molecular Similarity I (Topics in Current Chemistry)*; Sen, K. D., Ed.; Springer-Verlag: Berlin Heidelberg, 1995; Vol 137, p 2.

(2) Eliel, E. L.; Wilen, S. H.; Mander, L. N. *Stereochemistry of Organic Compounds*; Wiley: New York, 1994.

(3) Seri-Levy, A.; Richards, W. G. *Tetrahedron: Asymmetry* **1993**, *4*, 1917.

(4) Zabrodsky, H.; Avnir, D. *J. Am. Chem. Soc.* **1995**, *117*, 462.

(5) Petitjean, M. C. R. *Acad. Sci. Paris, Chimie/Chemistry* **2001**, *4*, 331.

(6) Carbó, R.; Arnau, M.; Leyda, L. *Int. J. Quantum Chem.* **1980**, *17*, 1185.

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

(8) Lee, C.; Smithline, S. *J. Phys. Chem.* **1994**, *98*, 1135.

(9) Mezey, P. G.; Ponec, R.; Amat, L.; Carbó-Dorca, R. *Enantiomer* **1999**, *4*, 371.

(10) Carbó, R.; Calabuig, B. *Int. J. Quantum Chem.* **1992**, *42*, 1681.

(11) Mezey, P. G. *Mol. Phys.* **1999**, *96*, 169.

(12) Morrison, R. T.; Boyd, R. N. *Organic Chemistry*, 6th ed.; Prentice-Hall: Englewood Cliffs, New Jersey, 1992.

(13) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press, New York, Clarendon Press, Oxford, 1989.

(14) Parr, R. G.; Yang, W. *Annu. Rev. Phys. Chem.* **1995**, *46*, 701.

(15) (a) Geerlings, P.; De Proft, F.; Langenaeker, W. *Adv. Q. Chem.* **1999**, *33*, 303. (b) Geerlings, P.; De Proft, F.; Langenaeker, W. *Chem. Rev.* **2003**, *103*, 1793. (c) Geerlings, P.; De Proft, F. *Int. J. Mol. Sci.* **2002**, *3*, 276.

(16) Chermette, H. *J. Comput. Chem.* **1999**, *20*, 129.

(17) Boon, G.; De Proft, F.; Langenaeker, W.; Geerlings, P. *Chem. Phys. Lett.* **1998**, *295*, 122.

(18) Boon, G.; Langenaeker, W.; De Proft, F.; De Winter, H.; Tollenaere, J. P.; Geerlings, P. *J. Phys. Chem. A* **2001**, *105*, 8805.

(19) The existence of Z_{AA} has, to the best of our knowledge, never been addressed in the large series of papers on molecular similarity. Its existence is in our view assured by the existence of $\int \rho(\mathbf{r})\, d\mathbf{r}$. The integrand of the latter integral is continuous and finite $\forall(\mathbf{r})$, and shows an exponential decay at infinity due to the fact that in a first approximation and at large r values the molecular density can be written as a sum of atomic densities, each of which decays as $\exp(-(8I)^{1/2})$ for large r (I is the first ionization energy; see: Morrell, M. M.; Parr, R. G.; Levy, M. *J. Chem. Phys.* **1975**, *62*, 549). The integral $\int \rho^2(\mathbf{r})\, d\mathbf{r}$ has an integrand which decays faster to 0 than ρ for large r, $\rho(\mathbf{r})$ is evidently also continuous and finite $\forall(\mathbf{r})$. Consequently, $\int \rho^2(\mathbf{r})\, d\mathbf{r}$ exists. Using the Schwarz inequality (Courant, R.; John, F. Introduction to Calculus and Analysis; Springer Verlag: New York, 1984; Vol. I, p 197), the existence of Z_{AB} then follows from the existence of Z_{AA} and Z_{BB}.

(20) Bultinck, P.; Kuppens, T.; Gironés, X.; Carbó-Dorca, R. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1143.

(21) Bultinck, P.; Carbó-Dorca, R.; Van Alsenoy, C. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1208.

(22) Mestres, J.; Solà, M.; Besalú, E.; Duran, M.; Carbó, R. *MESSEM*; Girona, CAT, 1993.

(23) Besalú, E.; Carbó, R.; Duran, M.; Mestres, J.; Solà, M. *Modern Techniques in Computational Chemistry: METTEC-95*; ESCOM: Leiden, 1995.

(24) Rousseau, B.; Peeters, A.; Van Alsenoy, C. *Chem. Phys. Lett.* **2000**, *324*, 189.

(25) Van Alsenoy, C.; Peeters, A. *THEOCHEM* **1993**, *105*, 19.

(26) Van Alsenoy, C. *J. Comput. Chem.* **1988**, *9*, 620.

(27) (a) Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864. (b) Riess, J.; Münch, W. *Theor. Chim. Acta* **1981,** *58*, 295.

(28) Hirshfeld, F. L. *Theor. Chim. Acta* **1977**, *44*, 129.

(29) De Proft, F.; Van Alsenoy, C.; Peeters, A.; Langenaeker, W.; Geerlings, P. *J. Comput. Chem.* **2002**, *23*, 1198.

(30) De Proft, F.; Vivas-Reyes, R.; Peeters, A.; Van Alsenoy, C.; Geerlings, P. *J. Comput. Chem.* **2003**, *24*, 463.

(31) Bultinck, P.; Langenaeker, W.; Lahorte, P.; De Proft, F.; Geerlings, P.; Van Alsenoy, C.; Tollenaere, J. P. *J. Phys. Chem. A* **2002**, *106*, 7895.

(32) Mestres, J.; Roher, D. C.; Maggiora, J. M. *J. Comput. Chem.* **1997**, *18*, 934.

(33) Gironés, X.; Robert, D.; Carbó-Dorca, R. *J. Comput. Chem.* **2001**, *22*, 255.

(34) Constans, P.; Amat, L.; Carbó-Dorca, R. *J. Comput. Chem.* **1997**, *18*, 826.

(35) Amat, L.; Carbó-Dorca, R. *Int. J. Quantum Chem.* **2002**, *87*, 59.

(36) Constans, P.; Carbó-Dorca, R. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 1046.

(37) Amat, L.; Carbó-Dorca, R. *J. Comput. Chem.* **1997**, *18*, 2023.

(38) 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.6; Gaussian, Inc.: Pittsburgh, PA, 1998.

(39) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(40) Perdew, J. P.; Wang, Y. *Phys Rev. B* **1992**, *45*, 609.

(41) Stevens, P. J.; Delvin, F. J.; Chablaoski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(42) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*, 2nd ed.; Wiley-VCH: Weinheim, 2000.

(43) *Handbook of Chemistry and Physics*, 63rd ed.; Weast, R. C. Ed.; CRC Press: 1982−1983.

(44) Wilen, S. H.; Bunding, K. H.; Kasheres, C. M.; Wieder, M. J. *J. Am. Chem. Soc.* **1985**, *107*, 6997.

(45) Cioslowski, J.; Stefanov, B. B.; Constans, P. *J. Comput. Chem.* **1996**, *17*, 1352.

(46) Stefanov, B.; Cioslowski, J. *Adv. Mol. Similarity* **1996**, *1*, 43.

(47) Slater, J. C. *Phys. Rev.* **1930**, *36*, 57.