Theoretical Reconstruction of the Electron Density of Large Molecules from Fragments Determined as Proper Open Quantum Systems: The Properties of the Oripavine PEO, Enkephalins, and Morphine

Chérif F. Matta[†]

Chemistry Department, McMaster University, Hamilton, Ontario, Canada L8S 4M1

Received: June 25, 2001; In Final Form: September 11, 2001

A new method of computing the properties of large molecules, not amenable to direct computation, is presented. The large molecule is reconstructed from judiciously chosen fragment molecules which are small enough for direct computation but large enough to retain good approximations to the electron density of the respective moieties in the large molecule. The atoms in the fragment molecules are determined as proper open systems according to the quantum theory of atoms-in-molecules (QT-AIM) developed by Bader. The properties of the atoms in the fragment molecules representing atoms in the large molecule are summed to obtain the properties of the large molecule. The computational effort is, thus, reduced from the *n*th power of a large number (the size of the basis set describing the large molecule) to the sum of the *n*th power of small numbers (the sizes of the basis sets describing each fragment molecule). The method is tested on the complicated oripavine molecule 7-α-[1-(*R*)-hydroxy-1-methyl-3-phenylpropyl]-6,14-*endo*-etheno-tetrahydrooripavine (PEO) and used to study some opioids. The properties of PEO computed from fragments are found to be remarkably close to those computed for the intact molecule. The success of this approach is the result of the transferability as well as the additivity of atomic properties. Three stable conformations of PEO are found to correspond closely to three conformations of enkephalins (Enk), each of the three PEO/Enk conformer pairs having key functional groups in very similar spatial disposition. These pairs can be candidate biologically active conformations at the three opioid receptor subtypes. The relationship between the concepts of transferability and similarity is discussed, and a new real-space "reactive similarity" index based on the Laplacian of the density is proposed. The extension of Carbó's similarity measures to pharmacophoric fragments of molecules is discussed. The results open the door for accurate theoretical studies of a multitude of large complicated biological molecules.

Introduction

Obtaining the properties of large molecules from high levels of theory has been elusive to this day. Even with the rapid advances in computer architectures and the anticipated meteoric growth in computational power, a direct calculation of the properties of a middle-sized protein from a good wave function is hopeless, at least within the foreseeable future, unless new methodologies are developed. The need to gain very detailed physical descriptions of large complicated biological molecules with the goal of understanding and modulating their biological functions is more pressing than ever. This paper describes an attempt to extend the computational arm to such molecules. The properties of a large inaccessible molecule are obtained from smaller accessible fragments utilizing the partitioning of the electron density provided by the quantum theory of atoms-in-molecules (QT-AIM) developed by Bader.[1] The method is subjected to a stringent test on a very complicated opioid which is large enough to be fragmented yet small enough to be computable intact, and hence providing a comparison control.

The Quantum Theory of Atoms-in-Molecules (QT-AIM) and the Conditions for the Theoretical Reconstruction of a Molecule from Fragments: Transferability and Additivity. Computational resources needed to obtain accurate electron densities and the derived properties scale to the *n*th power of the contracted basis set ($n \geq 3$). The method described in this paper reduces the computational burden from the *n*th power of a large number to the sum of the *n*th power of smaller numbers. This is accomplished by breaking the molecule into fragments, each of which is supplanted with hydrogen atoms as required to satisfy the valence of the atoms participating in the severed bonds yielding a set of closed-shell "fragment molecules". Separate self-consistent-field (SCF) calculations are performed on each fragment molecule, and the integrated properties averages of the atoms in the fragments are computed. The properties averages of those atoms in the fragments which correspond to atoms in the large molecule are summed to yield approximations to the molecular properties of the large molecule. The accuracy with which the properties of the large molecule are recovered from such a sum is determined by the simultaneous *additivity* of these properties as well as their *transferability* from the electronic environment in the intact molecule to the corresponding environment in the fragments.

One can anticipate the transferability of the properties of the atoms, which are sufficiently removed from the severed edges of a fragment based on simple chemical intuition. Indeed, a central chemical concept such as a "functional group" is an empirical expression of transferability: That groupings of atoms retain characteristic properties which are slightly perturbed by their neighbors in different molecules. QT-AIM[1] has been shown to recover the experimental transferability of group contributions to the heats of formation,[2] magnetic susceptibility,[3−5] electric polarizabilities,[6,7] partial molar volumes at infinite dilution,[8] and

[†] Fax: (905) 522-2509. E-mail: mattacf@mcmaster.ca.

10.1021/jp012397e CCC: $20.00
Published on Web 11/17/2001

to the heat of transfer from the gas to the solution phase.[8] On the other hand, since the density uniquely determines the external potential up to an additive constant, as required by the Hohenberg and Kohn theorem (HK),[9,10] *perfect transferability* is an unattainable limit since the external potential of the intact molecule is different than that of any fragment molecule. The same HK theorem, however, imposes no restrictions as to how close this limit can be approached. The high degree of transferability of atoms two bonds (or more) away from functional groups undergoing such severe alteration as charge separation, tautomerization, and conformational changes provides an example of the dependence of transferability on the immediate electronic environment and its relatively short-range nature.[11] Another example is provided by the preservation of charge neutrality of amino acid residues determined as proper open systems upon twisting of the peptide backbone.[12] This charge neutrality is a prerequisite for the transferability of amino acid residues from a tripeptide mold to an extended polypeptide chain, since in building a protein from such density residues one should not accumulate charges (except for ionized side-chains).[12]

The other necessary condition for a successful fragments approach is additivity, i.e., the properties of the parts (atoms) must sum up to yield the properties of the whole (molecule):

$$\langle \hat{A} \rangle = \sum_{\Omega} A(\Omega) \tag{1}$$

where $\langle \hat{A} \rangle$ is the molecular expectation value of the operator $\hat{A}$, $A(\Omega)$ is the average value of the property A over an atom Ω, and where the sum runs over all the atoms in the molecule. To compute $A(\Omega)$, the molecular space is partitioned into separate atomic basins bounded by surfaces of local zero-flux in the gradient vector field of the electron density $\rho(\mathbf{r})$:[1,13]

$$\nabla\rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0, \quad \text{for all } \mathbf{r} \text{ on the surface} \tag{2}$$

where $\mathbf{n}(\mathbf{r})$ is a vector normal to the surface. Each property density is integrated over the atomic basins to yield the respective atomic properties. Additivity is a direct physical consequence of this exhaustive partitioning of real space into nonoverlapping atoms leaving no gaps or dead volumes unaccounted.

Proper Open Quantum Systems. Equation 2 is the necessary boundary condition for the application of Schwinger's principle of stationary action to an open system.[14] Only subsystems bounded by zero-flux surfaces[13] are valid "proper" self-contained quantum entities satisfying all the theorems of quantum mechanics.[15] This definition maximizes the transferability of $\rho_{\Omega}(\mathbf{r})$ of a particular subsystem Ω (e.g., functional group) embedded in different closed total systems (molecules).

Every property A is represented by a "dressed" real space property density $\rho_{\mathbf{A}}(\mathbf{r})$ which assigns to every point in real space ($\mathbf{r}$) a value of the property describing the average interaction of a single electron at $\mathbf{r}$ with all the remaining particles in the molecule. The contribution of atom Ω to the molecular property A is given by the integration of $\rho_{\mathbf{A}}(\mathbf{r})$ over the volume of its basin:

$$A(\Omega) = \int_{\Omega} \rho_{\mathbf{A}}(\mathbf{r}) d\mathbf{r} \tag{3}$$

Thus, defining real space boundaries of an atom in a molecule entails the definition of all of its properties, e.g., atomic charge, dipole (and higher multipoles), energy, volume, etc. The sum of these properties over all the atoms in the molecule yield the corresponding experimentally measurable molecular properties. Atomic properties are easily computed[16−18] from theoretical[1] or experimental[19−21] electrons densities.

Theoretical Synthesis of Large Molecules from Fragments

Some Other Related Approaches. It is beyond the scope of this paper to present a comprehensive review of the fragment methodologies: these are briefly reviewed in the literature cited in this section. We will, however, discuss a few that are related to the current method.

Massa, Huang, and Karle[22,23] express the density of a large molecule as a sum of densities of fragments. The atoms of interest constitute the "kernel" of the fragment taken together with the contributions of orbitals centered on atoms of neighboring kernels. The overlaps of basis functions centered on the kernel and it neighbor are equally divided between the two. The orbitals of each fragment are used to construct a density matrix, and the sum of the fragment density matrixes yield the density matrix for the whole molecule which is constrained to be idempotent and normalized. The fragment matrixes can be obtained from either theoretical calculations or from experimental X-ray structure factors. In the latter case, additional constraints are imposed by requiring that the predicted density matrix reproduce the measured structure factors.

Lecomte's group[24−26] is building a data bank of transferable multipolar electron density parameters obtained from a fitting of the X-ray structure factors. This is brought about using the MOLLY program[27] which expands the density in terms of a set of overlapping atom-centered basis functions. They have shown that similar atoms in different environments, such as the atoms of the four different peptide groups of Leu-enkephalin, have statistically identical experimental multipolar density parameters. In other words, within this modeling of the density, similar atoms are transferable between different environments. Their strategy is essentially to transfer the multipolar charge-density parameters obtained from ultrahigh resolution X-ray diffraction experiments ($d < 0.5$ Å) on amino acids and small peptides to similar atoms in a protein, significantly improving the refinement statistics and its electron density description.[28] The method was shown to dramatically enhance the resolution of two proteins: crambin[29] and scorpion toxin.[30] The same group used the multipolar densities so obtained to compute the molecular electrostatic potential of an octapeptide, a potential that was shown to be significantly different from one obtained from a density generated by spherical atoms refinement.[28]

Chang and Bader[31] have shown that the density of a polypeptide can be reconstructed by linking of amino acid residues (fragments) matched at their amidic zero-flux surfaces that bound and define each amino acid residue. In this approach, there is no question of overlapping basis functions from neighboring fragments. Chang and Bader[31] have demonstrated that the properties of di- and tripeptides can be closely approximated by linking such amino acid density fragments. In an extension to this method, Martín and Bader[32,33] obtained a complete library of accurate densities of tripeptides of the type Gly|Aa|Gly, where Aa stands for any of the 20 genetically encoded amino acid residues and the vertical bars to the zero-flux amidic surfaces bounding the residue within the tripeptide mold. They showed that the shape of these amidic zero-flux surfaces and the density they bound are insensitive to a change in the nature of the side chain,[34] i.e., the main chain backbone atoms are essentially transferable and so are the side chain atoms. Exploiting this transferability and the almost perfect

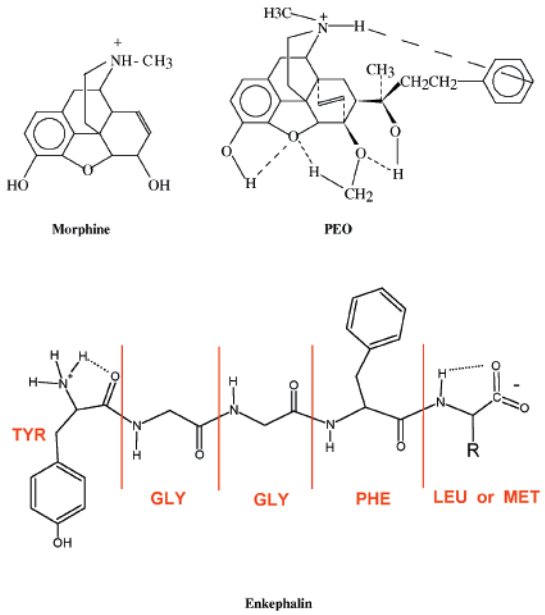
matching of the amino acid residues' zero-flux surfaces obtained in the Gly|Aa|Gly molds, a software was developed to link these amino acids residues densities (stored in the library) sequentially in an arbitrary order and thus obtain essentially an excellent approximation to the accurate density of a peptide of arbitrary length in seconds.[32] It has recently been shown how the transferability of these amidic surfaces determines the transferability of the properties of the subsystem bounded by these surface.[34] The visual simultaneous transferability of the zero-flux surfaces and the accompanying density is the result of the simultaneous stationarity of the zero-flux surface, as well as the wave function Ψ and energy of the total system to variations $\delta\Psi$.[34] Thus, the distribution of charge of an open system, and hence all of its properties, change only in response to changes in its bounding surfaces. In the approach of Bader, Martín, and Chang, no attempt is made to correct for the small error resulting from the slight mismatch of the amidic surfaces. This error has been found to be insignificant in the case of amino acid residues linked to form a polypeptide.[31,32]

In a related method, Breneman's group introduced the "transferable atom equivalents" (TAE),[35,36] which are atomic density fragments bounded by zero-flux surfaces but which have the added advantage of being flexible to allow for correction of the surfaces mismatch. In Breneman's approach, the RECON program is used to build densities sequentially by joining TAEs stored in a library at their approximately matching surfaces, translating the nuclei to the positions initially specified in an input **z**-matrix. At each addition of a new TAE to the previous one, the two nearly matching zero-flux surfaces are rotated about the inter-nuclear axis to optimize the matching of their respective zero-flux surfaces. Next, the bond length is adjusted by balancing the surface electronic "pressure" on each side of the surface, a pressure readily computed from the quantum stress tensor. In the process, the two zero-flux surfaces are fused into a single new surface which is constrained to satisfy the zero-flux condition (eq 2). This procedure of "melding" the two surfaces of a new TAE to the previous is repeated until the total molecular density is reconstructed. In building the TAE library, information about the derivatives of each atomic properties with respect to small radial variations in each unique position of the zero-flux surface are also stored. This information is used to correct for the changes in the atomic properties induced by the melding procedure. The TAE approach has been shown to be of great value in quantitative-structure−activity relationships studies.[37]

There is no guarantee that the densities obtained from any of the approaches described above (with the exception of the approach of Massa and Karle) will yield a density which is *N*-representable, i.e., derivable from a properly antisymmetrized wave function. Hernańdez-Trujillo and Bader succeeded in subjecting nearly idempotent fragment-reconstructed densities to an iterative algorithm constrained to yield *N*-representable idempotent density matrices.[38] The latter approach not only provides a physically sound method of melding, but results in densities derivable from properly antisymmetric wave functions as well.

Other excellent strategies for obtaining densities of large molecules have been successfully developed, these include Yang's divide-and-conquer method,[39−42] and the linear scaling approach.[43,44] The reader is referred to the literature on these important approaches, since in this paper we limit ourselves to a method based on the disjoint partitioning of the real space into nonoverlapping atoms which subscribes to a fundamentally different philosophy.

SCHEME 1



Selection of the Test Molecule for the Fragments Method: 7-α-[1-(*R*)-Hydroxy-1-methyl-3-phenylpropyl]-6,-14-*endo*-ethenotetrahydrooripavine (PEO). The success of a related fragment approach in polypeptides[31−33] motivated the extension to molecules with complicated electron densities and zero-flux surfaces which do not consist of simple repetitive units. As an extreme test case, 7-α-[1-(*R*)-hydroxy-1-methyl-3-phenylpropyl]-6,14-*endo*-ethenotetrahydrooripavine (PEO), an oripavine (demethylated thebaine) opioid, was selected for a number of theoretical and biological reasons.

The molecule has a strong theoretical appeal, since it contains aromatic rings, saturated rings, heterocyclic (furanoid) rings, five- and six-membered rings, fused and nonfused rings, cages (regions of spaces completely surrounded by rings), intramolecular hydrogen bonds, aliphatic regions, phenolic and methoxy oxygen atoms, and a quaternized nitrogen. This variety of chemical features represents a stringent benchmark test. Another theoretical advantage is the size of PEO: The molecule is small enough to be amenable intact to an SCF calculation (to compare with the reconstructed molecule), yet it is large enough to be broken into fragments representing different parts of the molecule.

Biological interest in PEO stems from its unsurpassed ability to bind to the three main subtypes of the opioid receptor (μ, κ, and δ). This molecule, a few thousand times more potent than morphine, is believed to be an ideal substrate to the opioids receptor.[45−48] PEO was selected rather than its methylated derivative at the phenolic hydroxyl group, 7-α-[1-(*R*)-hydroxy-1-methyl-3-phenylpropyl]-6,14-*endo*-ethenotetrahydrothebaine (PET), since blockade of the phenolic group drastically reduces the binding of several related opioids to the μ receptor.[49] Methylated opioids must undergo in vivo de-methylation to be converted to the biologically active form.[49] Calculations were performed on the quaternized PEO (with a protonated nitrogen atom), and hence bearing a net +1 au charge, since this is the prevalent state of ionization of opioids at physiological pH (7.2−7.6) and temperature (37 °C).[50]

Leu- and Met-enkephalin in their zwitter ionic forms and morphine in both its quaternized as well as neutral forms were also studied. (See Scheme 1.) The numbering scheme for PEO is shown in Scheme 2.

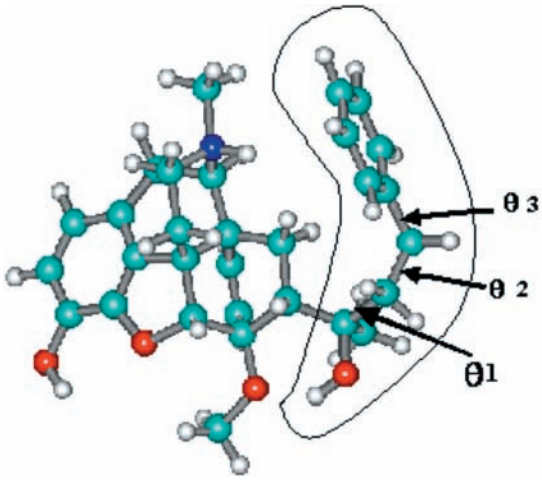
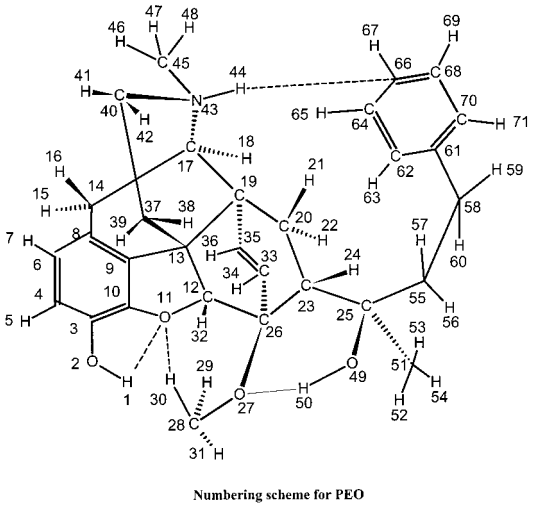


Figure 1. PEO in the scorpion (**A**) conformation, the global minimum. The flexible side chain is encircled and the three torsional degrees of freedom labeled $\theta 1$, $\theta 2$, and $\theta 3$.

SCHEME 2



Computational Strategy and Methods

A search of the literature and of the crystallographic databases revealed that neither the experimental nor a high quality optimized geometry of PEO or PET are available. In view of the size of PEO (71 atoms), a fine grid high-level of theory scan of the potential energy surface is computationally prohibitive, even if restricted to the flexible side chain torsion angles. Alternatively, an initial geometry guess for a local full energy optimization was obtained as follows.

The X-ray crystallographic geometry of a molecule from the oripavine family closely related to PEO, etorphine (in which the phenyl ring of the side chain is replaced by a methyl group),[51] was used as an initial guess to the geometry of the rigid moiety of PEO. It has been proposed that the flexible side chain of potent opioid receptors agonists such as PEO must adopt a "scorpion" conformation, since this is the only common conformation accessible to several agonists.[46,48] See Figure 1. The scorpion conformation is one in which the groups in the flexible side-chain of PEO adopt a favorable staggered conformation. The scorpion conformation was used in the present work to provide the initial geometry guess for the side-chain. The scorpion conformation ensures geometrical conditions favoring the hydrogen bonding of the −O(49)H(50) of the side-chain with the methoxy oxygen O(27). All bonding interactions—symbolized in the structural diagrams in this paper by the usual "stick" linking two atoms—were characterized by the simultaneous presence of a bond path[52] and an associated interatomic zero-flux surface. The O(49)−H(50)···O(27) hydrogen bond was unequivocally characterized at this geometry by a bond path. The presence of this hydrogen bond is consistent with the interpretation of the NMR spectrum of etorphine solution.[53] The scorpion conformation also allows another hydrogen bonding interaction between the electrophilic proton H(44) and carbon atom C(66) of the electron rich phenyl ring. This initial geometry was subjected to a full energy optimization to obtain a minimum (**A**) at the restricted Hartree−Fock (rHF) level using a 3-21G basis set. The fully optimized geometry **A** retains the essential features of the input geometry, albeit with altered values of the geometrical parameters. This minimum (**A**) was then used to probe the potential energy surface (PES) defined by the three torsion angles $\theta 1$ = C(23)−C(25)−C(55)−C(58), $\theta 2$ = C(25)−C(55)−C(58)−C(61), and $\theta 3$ = C(55)−C(58)−C(61)−C(62). See Figure 1. The torsion angle around the C(23)−C(25) bond was frozen to preserve the O(49)−H(50)−O(27) hydrogen bonding, since it occurs in solution.[53] The PES was sampled at 216 points by a rigid scan in 60° steps around $\theta 1$, $\theta 2$, and $\theta 3$, at the rHF level using a 3-21G basis set implemented in Gaussian94.[54] The potential energy scan under these conditions revealed that the scorpion geometry (**A**) is the global minimum. Two other low-energy valleys were also found. The respective conformational minima of these two additional valleys, in increasing order of energy, are labeled **B** and **C**. Conformer **A** was selected as a test case for the fragment method.

The fully optimized geometry of PEO (**A**) was then used for a single point SCF calculation using a polarized basis set to obtain the electron density. The single-point step was performed at the rHF level using a 6-31G* basis set.[54] The resulting density was analyzed using the program EXTREME[16,17,55] to obtain the molecular graph, and the satisfaction of the Poincaré-Hopf relationship[1] was used to ensure that no critical point has been overlooked. Some of the atoms of the intact PEO molecule were subjected to numerical integration using PROAIM[16,17,55] to obtain their respective atomic properties which serve as comparison benchmark for the corresponding integrated atomic properties obtained from the fragments. EXTREME and PROAIM are parts of the AIMPAC[55] suite of programs implementing the theory of atoms-in-molecules.

The fully optimized geometry **A** was used to specify the geometries of the fragments. The PEO molecule was conceptually broken into four fragments representing different parts of the molecule, the fragments having some atoms in common at the boundaries. Atoms at the boundaries of the severed edges are not used in the reconstruction and are present only to provide an electronic environment similar to the one in the intact molecule. Hydrogen atoms were added as necessary (at idealized positions) to the atoms participating in the bonding that was severed in the fragmentation to obtain closed-shell "fragment molecules". ("Fragment molecule(s)" and "fragment(s)", except when clearly distinguished otherwise by the context, are synonymous in this paper.) See Figure 2. Separate single-point SCF calculations were performed at the rHF level using a 6-31G* basis set on each fragment frozen at the geometry of the corresponding moiety in the intact optimized PEO (**A**). The resulting densities were subjected to atomic integrations using AIMPAC[55] to obtain the atomic properties of the atoms in the fragment molecules. The atomic properties of the atoms in the fragments corresponding to those in PEO were summed (eq 1) to obtain the reconstructed properties of the PEO molecule. Some atomic properties of PEO obtained from fragments in

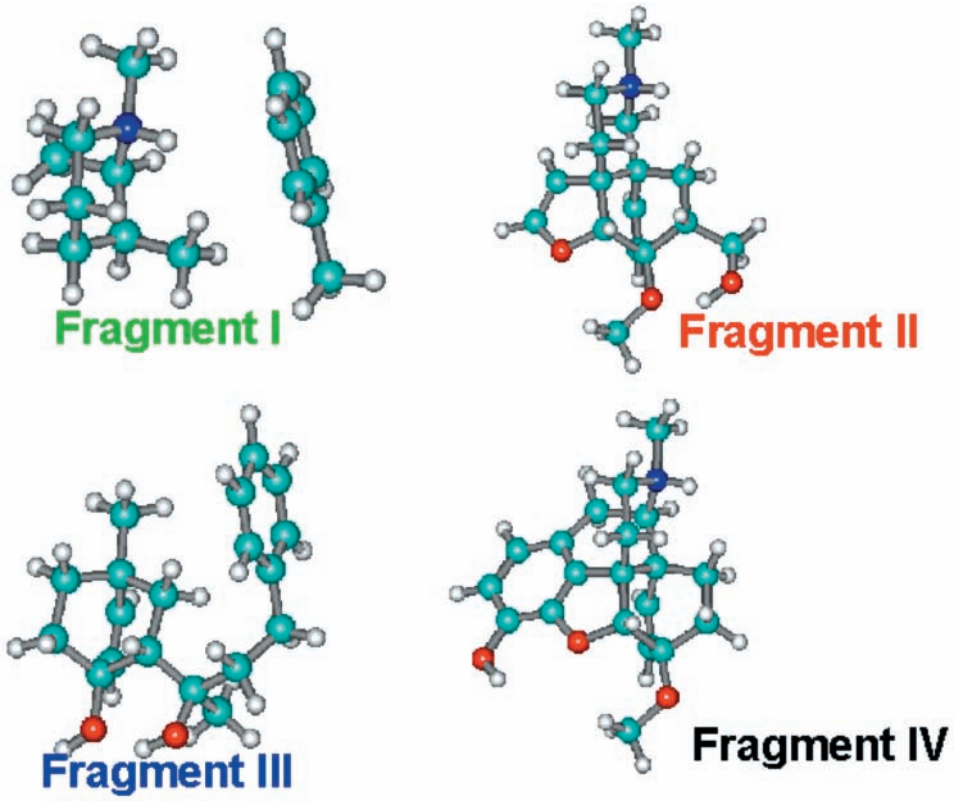


Figure 2. Fragment molecule **I** (green): Mimics the electron density around the quaternized nitrogen atom and its bonded hydrogen atom. It is a dimer of two separate molecules, one being a toluene molecule; all the atoms of the two separate molecules are at the exact positions of the corresponding atoms in PEO. This dimeric "fragment molecule" is designed to reproduce the electronic environment around the N−H group since it is hydrogen-bonded to C(60) of the phenyl ring. The fragment also mimics the density of the upper half of the phenyl ring. Fragment molecule **II** (red): Mimics the density around the hydrogen bond between the side-chain −O(49)H(50) group and the methoxy oxygen O(27). Fragment molecule **III** (blue): Mimics the density of the side chain and the lower half of the phenyl ring. Fragment molecule **IV** (black): Mimics the density of the phenolic ring, the ether oxygen O(11), and the rest of the atoms in this side of the molecule. The color-code of the fragments labels in this figure identifies the fragment utilized to obtain each of the charges displayed (and the rest of the properties which are not displayed) in Figure 4a.

addition to a few benchmark atomic properties obtained from atomic integration of intact PEO are given in Table 2. More properties can be obtained from the Supporting Information.

Different fragments have certain atoms (on the severed edges) representing the same atoms in the intact PEO molecule. These minimize the need for correction of zero-flux surfaces mismatches by providing a "buffer" region on the edges of each fragment. The reconstruction of the density/properties of the large molecule (PEO) is achieved using atoms carved-out from deep inside the fragments—away from the severed edges—to avoid edge effects. This paper demonstrates that this approximation is very good when the fragments are sufficiently large to represent the electronic environment of the atoms used in the reconstruction.

Similar reconstructions of the charge distributions of Leu- and Met-enkephalins are also achieved in the present work. The atomic charges for the amino acid residues Gly and Phe were obtained from the calculations of Martín and Bader[32,33] at the rHF/6-311G**//6-31+G* level. The charges for the terminal amino acids side chains (Tyr, Leu, and Met) were obtained from calculations on the free, neutral, nonzwitter-ionic molecules,[8] and those for the ionized carboxylic and amino groups from calculations on $H_2N-CH2-COO^-$ and $H_3N^+-CH2-COOH$, respectively (all at the same level of theory used by Martín and Bader).

Atomic properties of morphine in both ionization states (quaternized-cationic, and free base) were calculated directly (not from fragments) at the rHF level using a 6-31G* basis set at a geometry fully optimized using a 3-21G basis set.

The theoretical method used in the fragment approach should be size-consistent.[56] Even though it is known that closed-shell Hartree−Fock calculations are size-consistent,[57] the numerical error arising from the SCF convergence criteria was tested. The total SCF energy (at the rHF/6-31G* level) of fragment molecules **I** and **II** separated by 1000 Å was found to be 0.2 kcal/mol less stable than the sum of the total SCF energies of the isolated monomers. Size-consistency has to be ensured in fragment calculations, especially if truncated configuration interaction (CI) methods are used in the future.[57,58]

Results and Discussion

The Biologically Active Conformations of PEO and Enkephalin. Three biologically active conformations are proposed. As stated in the computational section, three valleys (**A**, **B**, and **C**) were found on a rough approximation to the potential energy surface (216 single points at the rHF/3-21G level). As suggested in the literature,[59] an important criterion for candidacy as a biologically active conformation of enkephalin is the coincidence of the geometrical relationships of certain functional groups with the those between corresponding groups in the PEO or in the PET molecules. Following these lines, the X-ray crystallographic geometry of Leu-enkephalin was obtained from the Cambridge Crystallographic Data Centre.[60] The geometry was modified using a molecular modeler[61] to maximize the superimposition on PEO (**A**, **B**, and **C** geometries) while preserving reasonable peptide bond geometries. The result is a striking geometrical resemblance of the three respective conformations of PEO and enkephalin, a resemblance that can be

TABLE 1: Key Geometrical Parameters Defining Three Nearly Coincident Conformations of PEO and Enkephalins[a]

		A	B	C
PEO				
	$\theta 1$	−84	−84	−144
	$\theta 2$	93	121	93
	$\theta 3$	44	−16	44
	τ	77	64	89
	d	8.01	10.19	9.66
Enkephalins				
Tyr [1]	ϕ	73	73	73
	ψ	−97	−97	−97
	ω	−175	−175	−175
	$\chi 1$	178	178	178
	$\chi 2$	46	46	46
	$\chi 3$	177	177	177
Gly [2]	ϕ	−151	−151	−151
	ψ	27	27	27
	ω	−164	−164	−164
Gly [3]	ϕ	−140	−140	−140
	ψ	−21	−21	−84
	ω	−170	−170	−170
Phe [4]	ϕ	−82	−82	−82
	ψ	5	5	5
	ω	179	179	179
	$\chi 1$	−40	−100	93
	$\chi 2$	−1	−1	11
	$\chi 3$	179	179	179
Met or Leu [5]	ϕ	−73	−73	−73
	ψ	−39	−39	−39
	τ	62	108	104
	d	7.66	10.02	10.24

[a] Conformations **A**, **B**, and **C** are defined in the text, they correspond to those in Figure 3. Side-chain angles for Met or Leu are omitted since their conformation was not studied in details. The angles $\theta 1-\theta 3$ are defined in the text and in Figure 1; peptide torsion angles are named in accordance to IUPAC−IUPAB conventions. τ is the dihedral angle between the best regression planes cutting through the two aromatic rings and d the distance between their centroids. Regression planes were determined using a least-squares approach[88] implemented in a spreadsheet. Values in degrees and angstroms.

refined and pursued in the future. A search for similarity of each of these geometrical pairs with known specific opioid receptor subtypes agonists may help in assigning one geometrical pair as a specific biologically active conformation to each of the three receptors subtypes (μ, κ, and δ). The three candidate biologically active PEO/enkephalin conformation pairs of **A**, **B**, and **C** are summarized in Table 1 and Figure 3.

The scorpion conformation **A** of PEO allows the hydrogen bonding between the proton bonded to the quaternary nitrogen [H(44)] and the *para*-carbon atom of the phenyl on the side-chain. Since the atomic properties of the quaternized nitrogen atom and its attached hydrogen atom in enkephalin are very similar to those of the N(43)−H(44) group in PEO, a similar hydrogen bonding is anticipated in enkephalin when the Tyr-NH_3^+ group adopts a similar geometrical arrangement with respect to the Phe aromatic ring. In PEO, $d_{C(66)-H(44)} = 2.93$ Å, $d_{H(44)-N(43)} = 1.02$ Å, and $\angle_{N(34)-H(44)-C(66)} = 137.9°$. In enkephalin, when the Tyr-NH_3^+ group is rotated so that $\angle_{N(Tyr)-H(Tyr)-C(Phe)} = 137.9°$ (**A**), the two other parameters also show a remarkable similarity to the corresponding values in PEO, these are $d_{para-C(Phe)-H(Tyr)} = 3.15$ Å and $d_{H(44)-N(43)} = 1.02$ Å.

The close proximity of the Tyr and Phe aromatic ring of Leu-enkephalin in a tightly folded structure has been observed experimentally where the centroids of the two rings are separated by only 5.0 Å in a nearly orthogonal orientation.[62] Furthermore, the significance of hydrogen bonding between aromatic rings of Phe, Tyr, and Trp residues and peptide backbone −NH groups in stabilizing the peptide secondary structures has been recently emphasized.[63] In their paper, Tóth et al. analyzed 560 different proteins crystallographic structures statistically, and they reported a clustering of the aromatic side-chain angle $\chi 1$ around the range −50° to −100° for the hydrogen-bonds of the type Ar(i)-HN(i-3), an angle which in enkephalin **A** has the value of −40°. They also showed a significantly high propensity for the residue i-2 to be Gly, which is of course the case in both types of enkephalin. These considerations add weight to the proposed tightly folded biological conformation **A** of enkephalins, a conformation which could be used as a reasonable initial guess for a full geometry optimization in the future.

Reconstructed Properties of PEO and Enkephalins, and the Properties of Intact Morphine. The atomic charges of PEO and Leu- and Met-enkephalin from fragments and those of intact morphine in both ionization states are shown in Figure 4. Table 2 compares some of the molecular properties of PEO obtained from fragment using eq 1 with those obtained directly from the SCF calculation on the intact molecule. It is remarkable that when the atomic properties obtained from fragments are summed, they yield the molecular properties of this complicated molecule to such small errors.

Table 3 reveals how similar the atomic electron populations (and atomic charges) of PEO reconstructed from fragments are to the corresponding values in the intact molecule. The same transferability trends are also exhibited by the rest of the properties that were studied (Supporting Information). Moreover, equivalent atoms in reconstructed PEO, intact morphine, and reconstructed enkephalin possess very similar properties. The comparison of atomic properties of these molecules can be as detailed as one wishes. These properties include, for example, atomic (or group) contributions to static or field-induced electrostatic mutipoles.[64] Expansion of the molecular electrostatic potential using QT-AIM atomic multipoles has been shown by Popelier[65] to closely reproduce the exact ab initio electrostatic potential for a number of molecules. As important as is the ability of QT-AIM multipoles to reproduce an electrostatic potential field is their basis set insensitivity, a property not shared by moments expanded in terms of a set of basis functions such as the distributed multipole analysis (DMA) of Stone.[66,67] Since QT-AIM multipoles work so well, and since the individual components of the QT-AIM multipoles are so transferable from fragments to the intact molecule (Supporting Information) one can expand molecular properties such as

TABLE 2: Comparison of Some Molecular Properties of Intact PEO versus Those Obtained from Fragments[a]

PEO molecular property	obtained from calculations on the intact molecule (1)	obtained from calculations on fragments (2)	difference (2)−(1)	% error
total number of electrons	254	254.0199	+0.0199	+0.01
total charge	+1	+0.9801	−0.0199	−1.99
total energy	−1510.703254	−1510.680842	+14.1kcal/mol	+0.001
total volume	565.57 Å^3	566.29 Å^3	+0.73 Å^3	+0.13

[a] Values in au unless indicated otherwise.

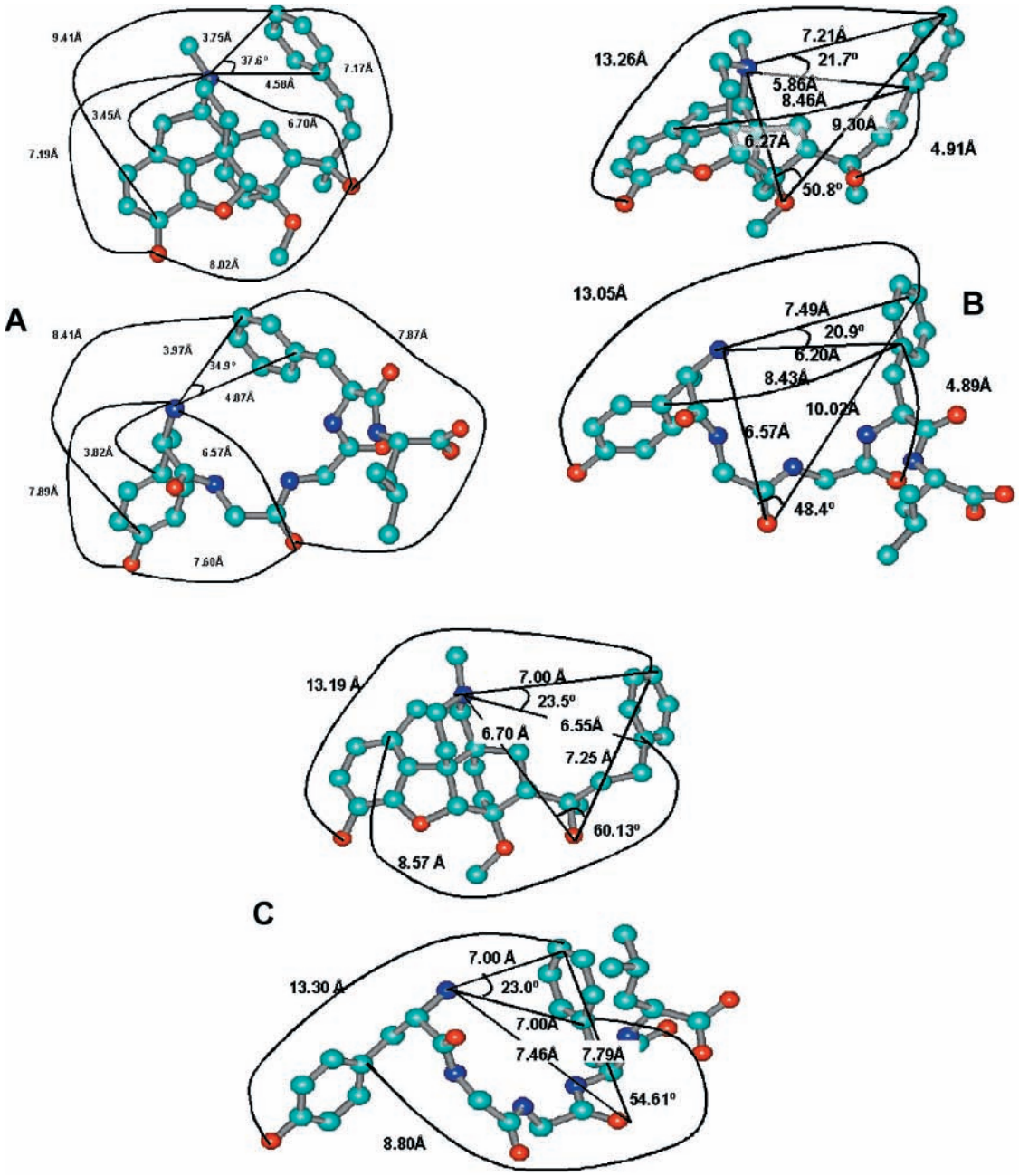


Figure 3. Three coincident conformations (**A**, **B**, and **C**) of PEO (top) and Leu-enkephalin (bottom). Some striking similarities in interatomic distances and angles are indicated for each PEO/enkephalin conformation pair. These are candidates for three biologically-active conformations at the opioid receptors sites. Distances in angstroms and angles in degrees.

the molecular electrostatic potential (MEP), or the electrostatic field, using atomic multipoles computed from the fragment molecules. On these bases, an excellent representation of the MEP of a large molecule is anticipated from QT-AIM atomic multipoles obtained from fragments, but a computer code is needed.

The total volume of the intact PEO molecule was calculated in GAUSSIAN94[54] by calling the self-consistent reaction field SCRF routine, which starts by calculating the solute cavity defined by the 0.001 au isodensity envelope. The agreement between the integrated volume up to the 0.001 au envelope assembled from the different fragments and that calculated for the intact molecule is a direct manifestation of the transferability of the atomic volumes from the fragments to the intact molecule environment and reflect the exhaustive nature of the zero-flux partitioning leaving no voids unaccounted for nor regions where the densities of different open systems overlap. The molecular volume calculated from intersecting spherical (CPK) atoms having standard van der Waal's radii at the geometry **A** is 475.6 Å, i.e., is in error by −16%. It is time to replace models of molecular surfaces and volumes based on intersecting spheres by molecular isodensity envelopes such as the 0.001 au which generally encloses more than 99% of the molecular electronic population and which corresponds to the van der Waal's surface in the gas phase. (In the condensed phase, the 0.002 au represents the molecular size more faithfully.)[1]

The total energy is recovered to within 14kcal/mol, a remarkable accuracy in view of the size of the molecule (error of the order of 1:10[5]). This very small error in the total energy can be due to a combined effect of a cumulative atomic integration error, imperfect transferability, and a type of basis set superposition error.

Atomic integration errors are minimized by imposing an upper limit on acceptable values of the integrated Laplacian density

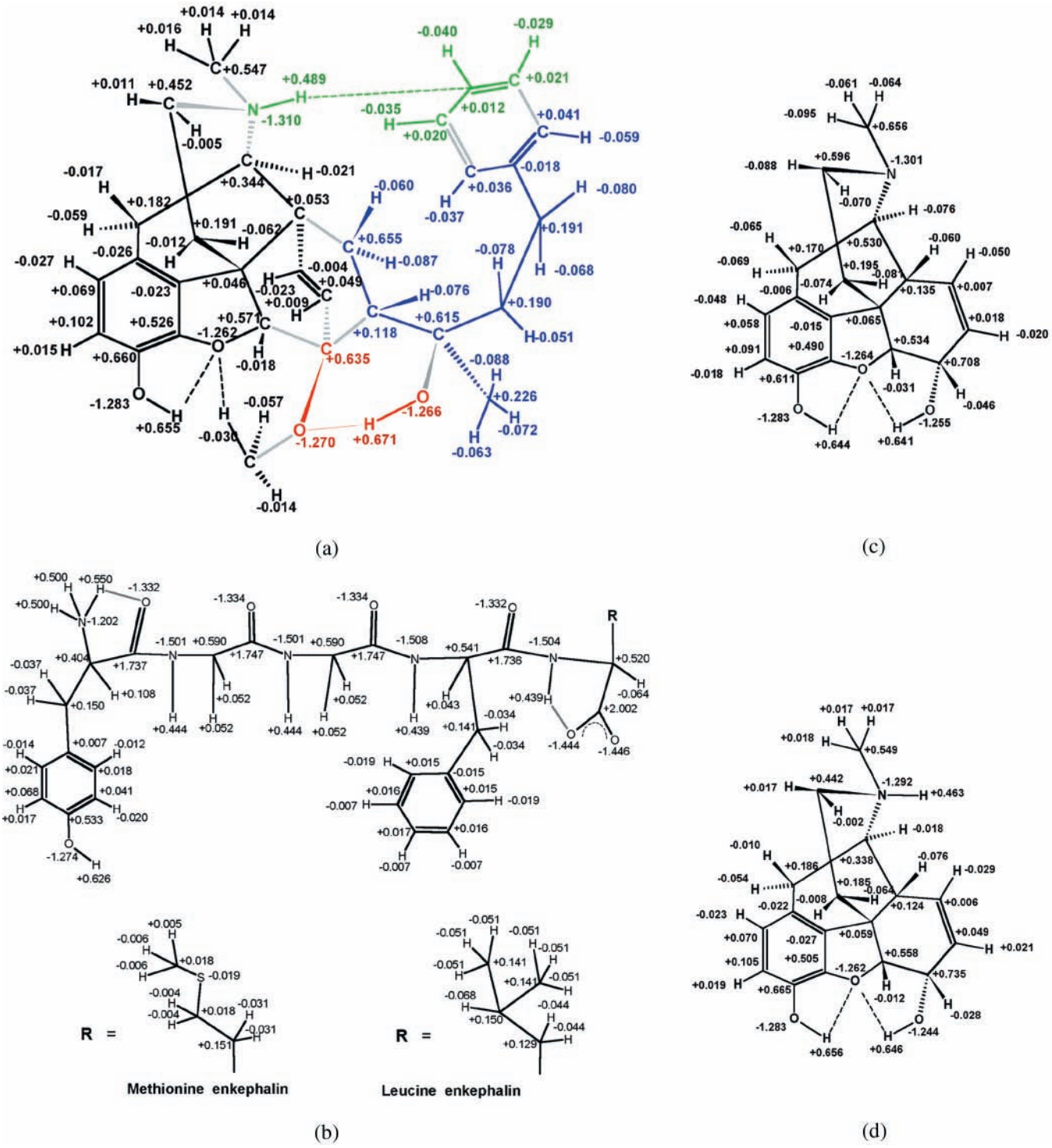


Figure 4. Atomic charges of, PEO, enkephalins, and neutral and quaternized morphine. (a) Atomic charges of an N-quaternized PEO molecule obtained from the four fragments defined in Figure 2. Colors code for the identity of the fragment used to generate the respective atomic charges (and the other atomic properties). Fragments color code: green (fragment **I**), red (fragment **II**), blue (fragment **III**), and black (fragment **IV**). (**b**) Atomic charges of zwitterionic enkephalin molecules (total charge = zero) obtained from fragments (discussed in the text). The sum of the atomic charges are +0.051 and +0.043 for Leu- and Met-enkephalin, respectively. The deviations of the atomic sum from zero are measures of the total integration and imperfect transferability errors. (c) Atomic charges of morphine free base (total charge = 0) obtained from an self-consistent-field (SCF) calculation on the intact molecule. The sum of charges is 0.008. (d) Atomic charges of an N-quaternized morphine molecule (total charge = +1) obtained from a self-consistent-field (SCF) calculation on the intact molecule. The sum of charges is 0.995. Note the striking similarity of the charges of equivalent atoms in these four molecules. All charges in atomic units (au).

over each atom. The Laplacian of the density $L(\mathbf{r})$ vanishes identically when integrated over the volume of a proper open system bounded by zero-flux surfaces, i.e., $L(\Omega) = 0$, and any (positive or negative) deviations are due to numerical integration errors. The average value of $|L(\Omega)|$ over the 71 atoms used in the reconstruction of PEO from fragments is 0.00052 au with $|L(\Omega)| \leq 0.00361$ au for any one atom. (See Table 3.) Obviously, a trade off has to be reached between the desired accuracy of the integration and computational costs. With the rapid advances in computers and the development of faster integration algorithms,[18,65,68,69] the accuracy and speed with which atomic integrations are performed will dramatically improve.

TABLE 3: Some Atomic Properties of PEO from Fragments[a]

	electron population (Ω)							$E(\Omega)$	vol 1(Ω)	$L(\Omega)$
atom no.	intact PEO	fragment **I**	fragment **II**	fragment **III**	fragment **IV**	frag. used	rec PEO	rec PEO	rec PEO	rec PEO
H(1)	0.3466				0.3451	**IV**	0.3451	−0.31762	17.17	0.00001
O(2)	9.2823				9.2825	**IV**	9.2825	−75.37406	117.47	−0.00021
C(3)					5.3396	**IV**	5.3396	−37.44918	57.39	0.00013
C(4)					5.8983	**IV**	5.8983	−37.79589	78.58	−0.00082
H(5)					0.9854	**IV**	0.9854	−0.62911	45.97	0.00007
C(6)					5.9312	**IV**	5.9312	−37.79806	78.59	0.00024
H(7)					1.0268	**IV**	1.0268	−0.64656	48.49	0.00008
C(8)	6.0245				6.0258	**IV**	6.0258	−37.88816	66.64	0.00013
C(9)			6.0062		6.0230	**IV**	6.0230	−38.04155	58.32	0.00015
C(10)			5.3911		5.4739	**IV**	5.4739	−37.60968	58.13	0.00044
O(11)	9.2642		9.2528		9.2618	**IV**	9.2618	−75.36266	97.71	0.00013
C(12)			5.4221	5.8120	5.4291	**IV**	5.4291	−37.50417	39.74	−0.00181
C(13)	5.9526	5.8250	5.9284	5.8299	5.9545	**IV**	5.9545	−37.88836	39.99	0.00028
C(14)		5.7789			5.8185	**IV**	5.8185	−37.71727	52.62	0.00192
H(15)		1.0083			1.0170	**IV**	1.0170	−0.64547	46.41	0.00012
H(16)		1.0564			1.0587	**IV**	1.0587	−0.66785	45.16	0.00012
C(17)		5.6552	5.5849	5.7775	5.6558	**IV**	5.6558	−37.60724	43.52	0.00219
H(18)		1.0251	1.0111	1.0848	1.0205	**IV**	1.0205	−0.65661	45.66	0.00015
C(19)		5.8601	5.9433	5.8963	5.9470	**IV**	5.9470	−37.84567	40.41	0.00007
C(20)		5.7698	5.8324	5.8143	5.8300	**III**	5.8143	−37.74292	48.42	0.00228
H(21)		1.0889	1.1293	1.0598	1.1317	**III**	1.0598	−0.67193	44.51	0.00006
H(22)		1.0526	1.0563	1.0865	1.0521	**III**	1.0865	−0.68778	43.08	0.00000
C(23)			5.8829	5.8824	5.8203	**III**	5.8824	−37.77245	43.59	0.00361
H(24)			1.0619	1.0756	1.0774	**III**	1.0756	−0.68487	44.91	0.00024
C(25)	5.3719		5.2469	5.3855		**III**	5.3855	−37.37188	35.85	0.00032
C(26)			5.3652	5.4220	5.3328	**II**	5.3652	−37.42885	35.12	−0.00031
O(27)	9.2738		9.2700	9.3012	9.2473	**II**	9.2700	−75.39442	89.63	−0.00144
C(28)			5.2700		5.2520	**IV**	5.2520	−37.35735	53.86	0.00084
H(29)			1.0453		1.0566	**IV**	1.0566	−0.66662	48.72	0.00011
H(30)			1.0095		1.0303	**IV**	1.0303	−0.66181	45.34	−0.00023
H(31)			1.0117		1.0143	**IV**	1.0143	−0.64756	47.42	0.00006
H(32)			1.0129	1.0789	1.0181	**IV**	1.0181	−0.66064	44.03	0.00014
C(33)			5.9526	5.8061	5.9515	**IV**	5.9515	−37.88178	73.02	0.00031
H(34)			0.9941	1.0484	0.9915	**IV**	0.9915	−0.63191	45.66	0.00012
C(35)	5.9966		5.9976	5.9936	6.0042	**IV**	6.0042	−37.89329	71.43	0.00057
H(36)			1.0212	1.0456	1.0228	**IV**	1.0228	−0.64453	47.38	0.00011
C(37)		5.8043	5.8118		5.8089	**IV**	5.8089	−37.75720	50.87	0.00247
H(38)		1.0566	1.0618		1.0622	**IV**	1.0622	−0.66632	46.64	0.00023
H(39)		1.0239	1.0094		1.0123	**IV**	1.0123	−0.64011	46.85	0.00010
C(40)	5.5435	5.5404	5.5446		5.5479	**IV**	5.5479	−37.57651	49.82	0.00152
H(41)		1.0109	1.0041		0.9889	**IV**	0.9889	−0.64156	40.94	0.00002
H(42)		1.0077	1.0018		1.0054	**IV**	1.0054	−0.64366	46.00	0.00009
N(43)	8.3120	8.3099	8.2986		8.3026	**I**	8.3099	−55.01652	63.29	0.00018
H(44)	0.5090	0.5112	0.5253		0.5300	**I**	0.5112	−0.41237	22.67	0.00000
C(45)	5.4495	5.4486	5.4536		5.4527	**IV**	5.4527	−37.51104	58.19	0.00175
H(46)	0.9941	0.9954	0.9829		0.9837	**IV**	0.9837	−0.63340	43.07	0.00013
H(47)	0.9880	0.9891	0.9840		0.9864	**IV**	0.9864	−0.63068	45.83	0.00007
H(48)	0.9851	0.9872	0.9827		0.9858	**IV**	0.9858	−0.63045	45.56	0.00008
O(49)			9.2655	9.3017		**II**	9.2655	−75.33416	114.30	0.00010
H(50)	0.3253		0.3289	0.3293		**II**	0.3289	−0.31196	11.37	−0.00099
C(51)				5.7736		**III**	5.7736	−37.71143	57.76	0.00137
H(52)				1.0625		**III**	1.0625	−0.66319	46.02	0.00003
H(53)				1.0875		**III**	1.0875	−0.66971	48.99	−0.00085
H(54)				1.0721		**III**	1.0721	−0.66157	50.82	0.00011
C(55)				5.8098		**III**	5.8098	−37.72067	50.80	0.00183
H(56)				1.0509		**III**	1.0509	−0.66734	46.24	0.00014
H(57)				1.0784		**III**	1.0784	−0.67468	46.63	0.00002
C(58)		5.7717		5.8090		**III**	5.8090	−37.69094	53.41	0.00230
H(59)		1.0451		1.0798		**III**	1.0798	−0.67528	47.02	0.00002
H(60)		1.0317		1.0675		**III**	1.0675	−0.66689	49.83	0.00010
C(61)	6.0231	6.0139		6.0177		**III**	6.0177	−37.86771	61.43	0.00015
C(62)		5.9632		5.9591		**III**	5.9591	−37.82078	77.01	0.00065
H(63)		1.0316		1.0592		**III**	1.0592	−0.66183	49.47	0.00011
C(64)		5.9786		5.9589		**I**	5.9786	−37.81338	80.36	0.00028
H(65)		1.0287		1.0493		**I**	1.0287	−0.64767	48.21	0.00007
C(66)	5.9897	5.9885		5.9546		**I**	5.9885	−37.81426	79.48	0.00068
H(67)		1.0399		1.0495		**I**	1.0399	−0.65170	49.50	0.00008
C(68)		5.9805		5.9583		**I**	5.9805	−37.81344	80.82	−0.00030
H(69)		1.0351		1.0491		**I**	1.0351	−0.64991	49.08	0.00008
C(70)		5.9640		5.9644		**III**	5.9644	−37.82527	75.78	0.00029
H(71)		1.0368		1.0523		**III**	1.0523	−0.66146	47.55	0.00000
sums							254.0199	−1510.68084	3821.54	0.02308

[a] Fragments are labeled as in Figure 2. Vol 1 is the integrated volume of the atom bounded by the zero-flux surfaces intersecting the outer 0.001 au isodensity envelope. The "frag. used" column indicate which fragment was used to provide the atomic properties for the corresponding atom in PEO, i.e., which fragment the particular atom used in the reconstruction (abbreviated "rec" in the legend) comes from. $E(\Omega)$ is the integrated atomic energy, and $L(\Omega)$ is the integrated atomic Laplacian. All values in au.

Intramolecular Basis Set Nonsuperposition Error (BSNSE). The use of truncated basis sets in SCF calculations leads to an (*inter*molecular) basis set superposition error (BSSE).[70−72] When the energies of monomers **A** and **B** are calculated separately, the effective basis set used is smaller than the one used to estimate the energy of the complex **AB**. This results in an artificial lowering of the energy of the complex. To correct for this error, the counterpoise (CP) method works best and is theoretically justified.[72]

Only recently has attention been drawn to an *intra*molecular BSSE which can result in significant errors in the estimation of conformational barriers.[73] In the fragment approach a type of intramolecular BSSE arises, which is opposite to the BSSE in the complex **AB**. The fragments within the intact molecule are in close spatial vicinity (they constitute one system) and therefore basis functions centered on atoms belonging to neighboring fragments overlap. This overlap is particularly important for atoms closely located in space, and is important for atoms that lie on the edges interfacing two fragments, e.g., atoms participating in bonding severed in the fragmentation. When reconstructing the energy of the whole molecule from isolated fragment molecules, one is reconstructing the energy of the total system from the integrated energies of subsystems, each of which is described by a smaller basis set than the same subsystem within the intact molecule. Since the energy of a fragment molecule is expected to be higher than the energy of the same fragment molecule computed with the basis set of the whole intact molecule, the reconstructed energy will be higher than that computed in an SCF calculation of the intact molecule. In other words, the reconstructed molecule is effectively described by a smaller basis set than the set used in a direct SCF calculation of the intact molecule. This leads to a *destabilization* of the reconstructed molecule (as opposed to the usual artificial stabilization of a complex **AB**). For these reasons, it is proposed hereby to give this error the designation "basis set nonsuperposition error (BSNSE)" to distinguish it from BSSE which arises from the *superposition* of basis functions and result into an overall *stabilization*.

To minimize the BSNSE one has to include in the density reconstruction atoms which are far from the edges of the fragment, edges that were formed in the fragmentation. This will diminish the effect of the missing overlap of basis functions centered on atoms in the fragment with the tails of basis functions centered on atoms of its surrounding, a surrounding present in the intact molecule but absent in the case of the fragment. The same strategy will also minimize the error resulting from zero-flux surfaces mismatches, i.e., imperfect transferability errors. The use larger basis sets is another means of reducing BSNSE which is rooted in the basis set truncation.

In summary to improve the accuracy of the fragment approach one can use larger basis sets, larger fragments, and tighter integration criteria. The computational resources will determine the cost-versus-accuracy balance.

Theoretical Basis of the Transferability in the Fragment Methodology. The "locality" of the atomic and functional group properties is a consequence of the "near-sightedness" of the one-electron reduced density matrix $\Gamma^{(1)}$.[74] "Near-sightedness" of $\Gamma^{(1)}$ means that the matrix elements between two points in space decay relatively quickly with the distance between them. At the Hartree−Fock (single determinant) level, $\Gamma^{(1)}$ determines the higher-order density matrixes as well as the state function. Therefore, $\Gamma^{(1)}$ determines all the measurable properties and their respective real space densities, including the total energy and its components.[75] As a result, the properties of an atom in a molecule are dependent only on its immediate neighborhood, and are almost completely independent of remote parts of the molecule. In other words, the near sightedness of the density matrix provides the physical basis for the functional group transferability underlying much of experimental and theoretical chemistry. This near sightedness is exploited in this paper since the density of an atom or a group of atoms can be modeled to a very high degree of accuracy in fragments from which they are carved out and assembled to yield an excellent approximation to the density of the intact molecule.

Future Developments: Similarity Measures of a Pharmacophoric Fragment in Two Different Molecules

The success of the fragment approach is a manifestation of the transferability of functional groups. The concepts of transferability and similarity are related: If two pieces of matter are similar they are transferable, and vice versa. Transferability is a quantitative expression of similarity. The transferability of a functional group's properties from one molecule to another parallels the transferability of its density and of its bounding zero-flux surfaces.[34]

The visual similarity of the density of a common moiety embedded in two different molecules is manifested in the transferability of the properties of this moiety. For instance, Figure 4 reveals the transferability of the charges of equivalent atoms (e.g., the phenolic ring and its surrounding) in three entirely different molecules: PEO, enkephalins, and quaternized or neutral morphine. The transferability of any one property—such as the atomic charges—is a partial quantitative expression of the visual similarity that one can easily see from the envelope plots of the density [$\rho(\mathbf{r})$] or of its Laplacian [$\nabla^2\rho(\mathbf{r})$] in the phenolic region (Figure 5) or from the contour plots in the plane of the same region in morphine (Figure 6a) and PEO (Figure 6b). The pronounced similarity exhibited in the phenolic region fades as one moves toward the bottom-right sides of the plots in Figure 6a,b as clearly seen from the superimposition of the contour plots in Figure 6c.

A direct measure of the similarity of the density of *a part* of a molecule is desirable, since this will allow one to zoom-in on the similarity of a crucial (e.g. pharmacophoric) moiety between two or more molecules, leaving out parts that are not necessary for the biological activity. In Figure 5, even though the envelope plots display a strong similarity at the phenolic end, they show much dissimilarity at the opposite side of the molecules since the chemical nature of the groups are different there. The same is true in the corresponding contour plots (Figure 6a−c). Suppose that only the phenolic moiety was known to be important for the biological activity of the series, and we wish to obtain an overall similarity measure for the density of this moiety in morphine and in PEO. Carbó and coworkers[76] proposed the now well-known similarity index (eq 4), largely used in quantitative-structure-to-activity-relationships(QSAR) studies:[77]

$$R_{\mathbf{AB}} = \frac{\int \rho_{\mathbf{A}}\rho_{\mathbf{B}}\, d\nu}{\sqrt{\int \rho_{\mathbf{A}}^2\, d\nu \int \rho_{\mathbf{B}}^2\, d\nu}} \qquad (4)$$

where $\rho_{\mathbf{A}}$ and $\rho_{\mathbf{B}}$ stand for the density of molecules **A** and **B**, respectively, and the integral is over the *entire space*. The index is a normalized cross-correlation coefficient that can take values $0 \le R_{\mathbf{AB}} \le 1$, where 1 indicates a complete similarity and 0 a

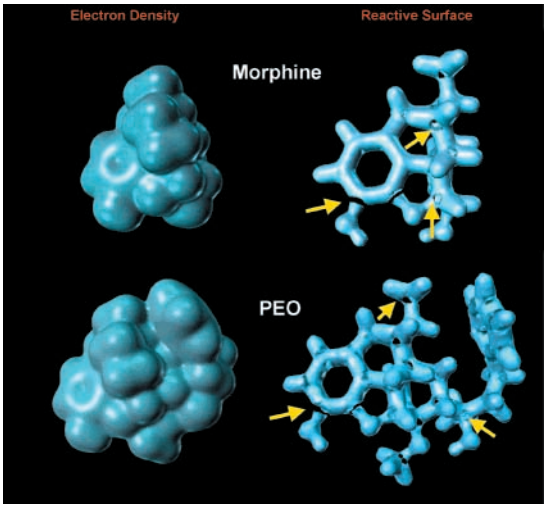


Figure 5. Electron density 0.002 au envelopes of morphine and PEO (left), and their zero-Laplacian (reactive surface) envelope (right). The Laplacian of the density $[\nabla^2\rho(\mathbf{r})]$ measures the local curvature of the electron density in all of its three dimensions. Where the Laplacian is positive, the density is locally depleted and the electron density is expanded relative to its average distribution; where it is negative the density is locally concentrated and the density is tightly bound and compressed above its average distribution. A local charge concentration is a Lewis base (nucleophile), while a local charge depletion is a Lewis acid (electrophile). The Laplacian reproduces the spherical shell structure of isolated atoms. The zeroes in the Laplacian are spherical node envelopes bounding regions of density depletion or concentration. The last shell of charge concentration followed by charge depletion (extending to infinity) is called the valence shell charge concentration (VSCC). When an atom is involved in bonding the spherical symmetry of the VSCC is broken and it may become punctured. A chemical reaction corresponds to the combination of a "lump" in the VSCC of the base with a "hole" in the VSCC of the acid. Some regions of charge depletion (holes) are indicated by the yellow arrows in the figure representing the reactive surfaces. Similar regions of the morphine and of the PEO molecules have similar regions of local charge depletion and concentration. The reactive surface of PEO serves as a "complementary template" for the receptor: where there is a region of charge depletion (holes) in the reactive surface of PEO there should be a region of charge concentration (lumps) in the receptor pocket, and vice versa.

complete dissimilarity of the two density functions. Carbó-type similarity indexes are particularly useful when the relative disposition of the atoms is important as is the case for example in optical isomers. This type of index is also useful to compare observed and calculated densities using real-space-correlation coefficient in X-ray protein crystallographic refinement.[78]

Another definition of the similarity index was proposed later:[79]

$$R_{\mathbf{AB}} = \frac{2\int \rho_{\mathbf{A}}\rho_{\mathbf{B}}\,\mathrm{d}\nu}{\int \rho_{\mathbf{A}}^2\,\mathrm{d}\nu + \int \rho_{\mathbf{B}}^2\,\mathrm{d}\nu} \tag{5}$$

It has been argued that eq 5 is sensitive to the magnitudes of the two densities (as opposed to comparing only their shapes in eq 4) since that, if $\rho_{\mathbf{B}} = k^2\rho_{\mathbf{A}}$, where k is a constant, eq 4 will indicate a perfect similarity.[79,80] This is true of course, but it is shown in the Appendix that $\rho_{\mathbf{B}} = k^2\rho_{\mathbf{A}}$ is only allowed for the trivial case of $k = \pm 1$. In other words, there can be no density which is a "scaled" version of another density. Furthermore, eq 4 has a direct statistical meaning lost if one uses eq 5. Thus, eq 4 will be used in the following discussion, but the same ideas could be extended to eq 5 should a specific use arises.

Similarity indices (eqs 4 and 5) have been used so far to compare the *entire* densities of two molecules. Because this usage does not allow one to select particular regions *within* the two molecules to be compared, the indices have been most useful when the molecules are roughly of the same size and have most of the atoms in common with relatively minor substitutions. If one is interested in comparing a particular grouping of atoms of pharmacological importance (a fragment) embedded in very different molecules where the remainder of each molecule is biologically irrelevant, the mode of integration over all space in eqs 4 and 5 is not likely to yield meaningful or useful results. In other words, the use of similarity indices to compare densities of full molecules "dilutes" the similarity (or dissimilarity) of a pharmacologically important moiety. We now return to our hypothetical case of comparing the density of the phenolic moiety in morphine to the one in PEO. If we use eq 4 or eq 5 to compare the densities of the entire molecules, the similarity of the phenolic part will be offset by the dissimilarity of the density of the flexible side chain and of the density in the Diels−Alder adduct region in PEO, regions that are absent (or very different) in morphine. The comparison of the full molecule density cannot possibly recover the similarity of the phenolic region in both molecules that one can see in Figure 5 or Figure 6 or by comparing the atomic properties of corresponding atoms in the two molecules. Alternatively, one can extract only the density fragments to be compared at their respective zero-flux surfaces. Thus instead of comparing the density of morphine and PEO over all space, one can compare the density over the space occupied by the fragment of interest extracted from the rest of the molecule at the zero-flux surface, since these surfaces determine the maximally transferable proper open systems within a total system and will, hence, maximize the similarity of the fragment in different total systems. In two dimensions, this is represented by Figure 6d, which shows the superposition of only the phenolic region of morphine and PEO, each extracted from its respective molecule at the zero-flux surface. As another illustration of proper open systems, three-dimensional density fragments representing two amino acid residues extracted from a polypeptide at their bounding zero-flux amidic surfaces are shown in Figure 7. A data bank of predetermined side chains densities at their common conformations and separated from peptides moulds at the R|Cα zero-flux surface can be used in real-space crystallographic refinement[78,81] and in locating errors[81] during model building of protein densities from X-ray experiments.

Thus, instead of integration over all space, the integration in eqs 4 and 5 can be restricted to the volume enclosed between a certain outer isodensity envelope, say the 0.001 au, and its intersection with the zero-flux surfaces separating the fragment of interest Ω from the rest of the molecule. The fragment may consist of a single atom or a grouping of atoms. The similarity index equation is therefore rewritten to reflect the bounded integration domain:

$$R_{\Omega\cup\Omega'} = \frac{\int_{\Omega\cup\Omega'} \rho_{\Omega}\rho_{\Omega'}\,\mathrm{d}\nu}{\sqrt{\int_{\Omega} \rho_{\Omega}^2\,\mathrm{d}\nu \int_{\Omega'} \rho_{\Omega'}^2\,\mathrm{d}\nu}} \tag{6}$$

where ∪ is necessary to emphasize that the integration is carried out only over the volume necessary to enclose both fragments, and no longer over the entire space. Of course, as in the case of entire molecular densities, there exist an alignment problem. Several excellent methods have been devised to optimize the alignment of densities, the optimum alignment being the one

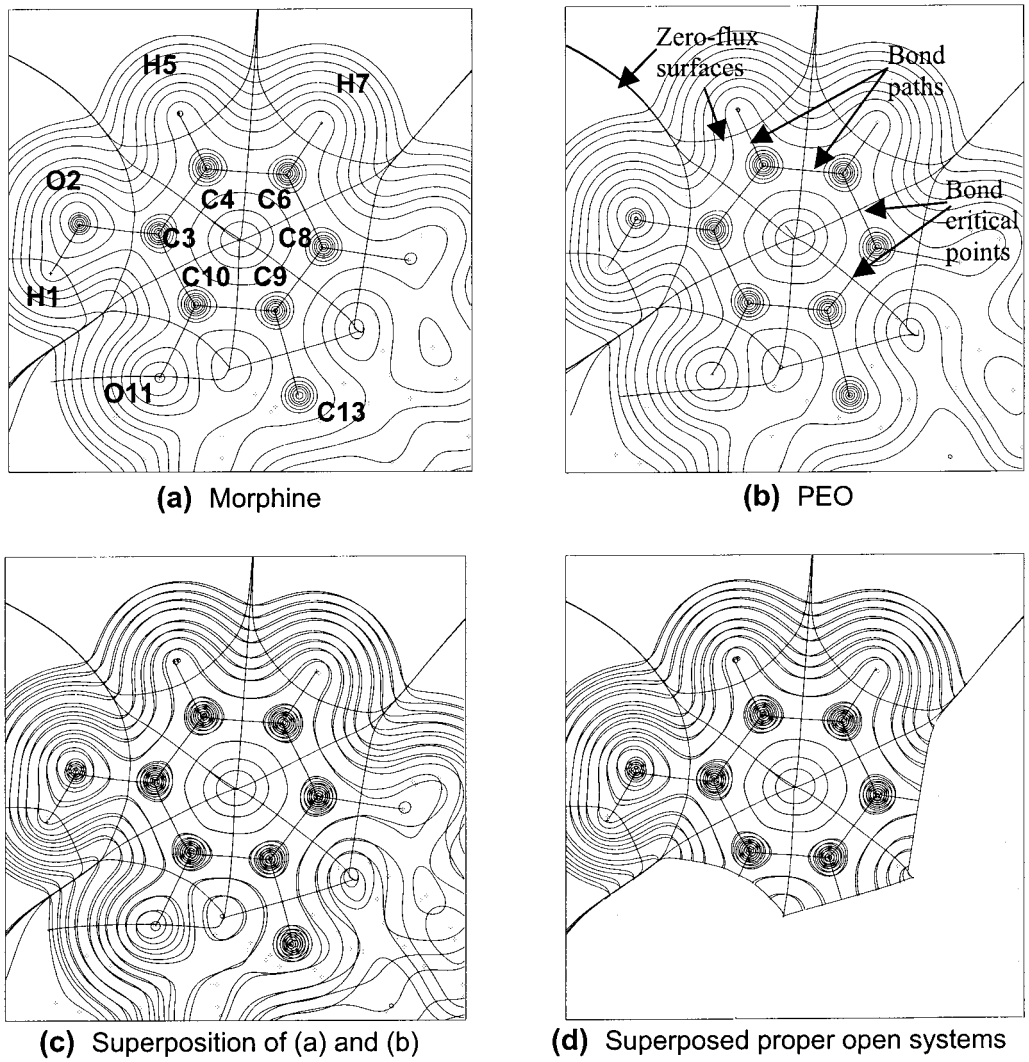


Figure 6. Contour maps of the electron density in the plane of the phenolic region of (a) quaternized morphine and (b) quaternized PEO. The density increases from the outermost 0.001au isodensity contour in steps of 2×10^n, 4×10^n, 8×10^n au with n starting at -3 and increasing in steps of unity. The lines connecting the nuclei are the bond paths, and the lines delimiting each atom are the intersection of the respective interatomic surface with the plane of the drawing. Crosses not linked by bond paths are the projections of the nuclear positions of atoms out of the plane of the drawing. Since the region is not perfectly planar, bond paths and zero-flux interatomic surfaces are allowed to leave the plane and their projections to be plotted. This accounts for artifacts like the line crossing through the density of O(11). (c) Is a superposition of the two density maps of (a) morphine and (b) PEO. It shows the remarkable degree of similarity, contour by contour, between the densities of the phenolic region of morphine and PEO. The similarity is most striking as one goes to the top left of and decreases as we move to the bottom right. (d) Is the superposition of the two densities of the phenolic region extracted from their respective surroundings and compared as proper open systems in isolation. The superposition of the two molecular densities is terminated at their bounding zero-flux surface (lower surface in the plot). The fragment in (d) consists of the atoms: H(1), O(2), C(3), C(4), H(5), C(6), H(7), C(8), C(9), and C(10). The rest of the molecular space has been deleted from the zero-flux surface on.

maximizing the similarity index $R_{\mathbf{AB}}$. (See, for example, refs 82−85 and the literature cited therein.) The extension of similarity maximization algorithms to deal with regions of space bounded by an isodensity envelope intersecting with a zero-flux surface should be straightforward.

The indices based on the density, while having the virtue of being based on an experimental observable, can give exaggerated weight to small mismatches in regions where the density is extremely high, a situation that may arise in proximity of the nuclei when the geometries of the molecules to be compared is slightly different. Cooper's group proposed momentum-space similarity indices which give more weight to the outer chemically interesting regions in the molecule.[86,87] A more natural measure of the reactivity of a molecule in *real space* is provided by the Laplacian of the charge density $\nabla^2\rho(\mathbf{r})$.[1] The Laplacian is an excellent descriptor of chemical reactivity since it indicates regions in the molecule where the charge is locally concentrated (nucleophilic regions), $\nabla^2\rho(\mathbf{r}) < 0$, and regions where it is locally depleted (electrophilic regions), $\nabla^2\rho(\mathbf{r}) > 0$. In plots of the zero-Laplacian envelope—the outer zero-Laplacian envelope being the molecule's "reactive surface"—holes are the sites of prone to nucleophilic attacks (Lewis acids) and lumps are the sites prone to act as nucleophiles (Lewis bases).[1] (See Figure 5.) Thus, the Laplacian is an excellent candidate to provide the basis for a similarity measure that reflect the reactivity of the molecule without being biased by small mismatches at the nuclear cusps. The integral of the Laplacian over a proper open quantum system vanishes, and it does so for the total system, however the product of two Laplacian distributions does not necessary vanish when integrated over all space or over the space of a proper open system. Therefore, one can use the Laplacian product to obtain a "reactive similarity index" over the region of interest in two molecules. One can compare regions bounded by the reactive surface between fragments bounded

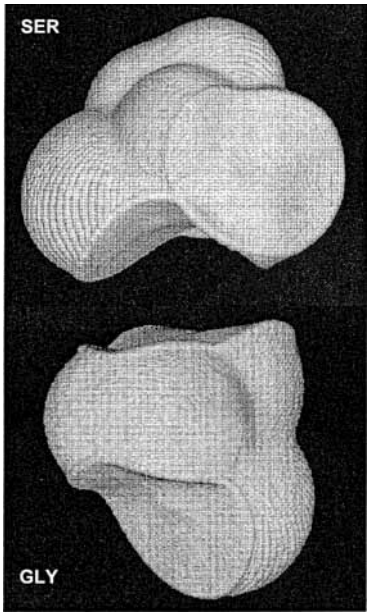


Figure 7. The density of two amino acid residues determined as proper open systems carved-out of their polypeptide. The two residues are represented by the intersection of the van der Waals surface with the two amidic surfaces. The residue at the top is a serinyl residue, |NHCH-(CH2OH)C(═O)| with the −HN| surface facing the viewer and leading to the previous amino acid residue in the α-helix. The −(O═)C| amidic surface of the serinyl group face downward to the −HN| amidic surface of the next amino acid residue in the helix, the residue at the bottom (glycyl). The two residues face each other with complementary surfaces, being the two sides of one and the same surface separated here only for clarity: We have severed an amide bond. These are two examples of proper open systems represented by their density in real three-dimensional space.

by a zero-flux surfaces. This can be expressed as

$$R_{\Omega\cup\Omega'} = \frac{\int_{\Omega\cup\Omega'} \nabla^2\rho_\Omega \nabla^2\rho_{\Omega'}\, d\nu}{\sqrt{\int_\Omega (\nabla^2\rho_\Omega)^2\, d\nu \int_{\Omega'} (\nabla^2\rho_{\Omega'})^2\, d\nu}} \qquad (7)$$

Equations 6 and 7 lead one to an operational definition of a pharmacophore as *one (or more) properly oriented proper open quantum system(s) in a molecule responsible for eliciting a pharmacological action, each bounded by a zero-flux surface, a surface within which the similarity ($R_{\Omega\cup\Omega'}$) and the transferability of the density is maximized in a series of related molecules*.

This definition underlines the possibility of studying a pharmacophore in a model system much smaller than the original drug molecule.

Conclusions

The feasibility of the theoretical reconstruction of large complicated molecules is established and shown to provide an excellent approximation to the density and to the properties of the intact molecule. The success of the method depends the simultaneous transferability as well as the additivity of the atomic properties, conditions that are closely met when the fragments are determined as proper open systems according to the quantum theory of atoms-in-molecules (QT-AIM). Transferability is a manifestation of the short-range nature of the first-order reduced density matrix. The concepts of transferability and similarity are intimately interrelated. The integration over selected molecular regions bounded by zero-flux surfaces is proposed to extend the use of Carbó-type similarity indices to parts of interest within a molecule as opposed to the entire molecule. This mode of integration will eliminates the unnecessary noise introduced by the remainder of the molecules that may have little biological relevance. The Laplacian of the density can provide a "reactive similarity index" free from the dominance of nuclear cusps in the density.

Three different conformations for PEO have been found to bear a striking geometrical similarity to three corresponding enkephalin conformations. These three geometrical pairs are candidates for the biologically-active conformations at the opioid receptors.

The present paper can be of value in extending the computational arm to large complex molecule with potential implications in nanotechnology and in drug design.

Acknowledgment. The author thanks Prof. Richard F. W. Bader for his support, for generous computer time, and for numerous suggestions. Prof. Jesús Hernández-Trujillo is acknowledged for constructive criticism. Financial support was provided by the Chemistry Department, McMaster University. This work was presented at the American Chemical Society, 222nd National Meeting, August 26-30, 2001, Chicago, IL; paper COMP 94.

Appendix

The Impossibility of a "Scaled" Electron Density. The density is derived from the state function and $\rho_{\mathbf{B}} = k^2\rho_{\mathbf{A}}$ implies $\Psi_{\mathbf{B}} = \pm k\Psi_{\mathbf{A}}$; in other words, the two densities $\rho_{\mathbf{B}}$ and $\rho_{\mathbf{A}}$ are derivable from two state functions differing only by a constant factor and hence describing the same state. The scaled density $k^2\rho'_{\mathbf{A}}$ integrates to k^2N electrons rather than N, a situation with no physical consequences since it entails division by a proper normalization factor.[34] For example, the expectation value of a multiplicative operator $\hat{\Omega}$ is

$$\langle\hat{O}\rangle = \frac{\int \hat{O}\rho_{\mathbf{B}}(\mathbf{r})\, d\mathbf{r}}{\int d\mathbf{r} \int d\tau\, \Psi_{\mathbf{B}}{}^*\Psi_{\mathbf{B}}} = \frac{\int \hat{O}k^2\rho_{\mathbf{A}}(\mathbf{r})\, d\mathbf{r}}{\int d\mathbf{r} \int d\tau\, (\pm k\Psi_{\mathbf{A}})^*(\pm k\Psi_{\mathbf{A}})} = \frac{\int \hat{O}\rho_{\mathbf{A}}(\mathbf{r})\, d\mathbf{r}}{\int d\mathbf{r} \int d\tau\, \Psi_{\mathbf{A}}{}^*\Psi_{\mathbf{A}}} \qquad (A.1)$$

Equation A.1 states that the same expectation value for the ground-state property O can equally be obtained from either densities, $\rho_{\mathbf{B}}$ or $\rho_{\mathbf{A}}$. The Hohenberg−Kohn theorem[9,10]—that the ground-state electron density uniquely define any ground-state property—completes the proof that $\rho_{\mathbf{B}} = k^2\rho_{\mathbf{A}}$ is only allowed for the trivial case $k = \pm 1$. In other words, there can be no density which is equal to another density multiplied by a constant other than one. (See ref. 34.)

Supporting Information Available: A table of some integrated atomic properties for a few representative atoms of PEO, and for each of the four fragment molecules defined in the text. Properties included: atomic electron population, atomic charge, atomic energy, atomic volume integrated up to the 0.001 au envelope, integrated atomic Laplacian, the components and magnitude of the atomic dipolar polarization with the origin at the nucleus, the eigenvalues of the diagonalized quadrupole tensor and its magnitude. Wave functions in Gaussian 94 format of PEO, the four fragment molecules, as well as neutral and

quaterized morphine also available. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, U.K., 1990.

(2) Wieberg, K. B.; Bader, R. F. W.; Lau, C. D. H. *J. Am. Chem. Soc.* **1987**, *109*, 1001.

(3) Keith, T. A.; Bader, R. F. W. *Chem. Phys. Lett.* **1992**, *194*, 1.

(4) Keith, T. A.; Bader, R. F. W. *Chem. Phys. Lett.* **1993**, *210*, 223.

(5) Keith, T. A.; Bader, R. F. W. *Int. J. Quantum Chem.* **1996**, *60*, 373.

(6) Bader, R. F. W.; Keith, T. A.; Gough, K. M.; Laidig, K. E. *Mol. Phys.* **1992**, *75*, 1167.

(7) Bader, R. F. W.; Keith, T. A. *J. Mol. Struct. (THEOCHEM)* **1991**, *234*, 75.

(8) Matta, C. F.; Bader, R. F. W. In preparation.

(9) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

(10) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1989.

(11) Matta, C. F.; Bader, R. F. W. *Proteins: Struct., Funct., Genet.* **2000**, *40*, 310.

(12) Popelier, P. L. A.; Bader, R. F. W. *J. Phys. Chem.* **1994**, *98*, 4473.

(13) Bader, R. F. W. *Theor. Chem. Acc.* **2001**, *105*, 276.

(14) Bader, R. F. W. *Phys. Rev.* **1994**, *B49*, 13348.

(15) Bader, R. F. W.; Popelier, P. L. A. *Int. J. Quantum Chem.* **1993**, *45*, 189.

(16) Biegler-König, F. W.; Nguyen-Dang, T. T.; Tal, Y.; Bader, R. F. W.; Duke, A. J. *J. Phys. B: At. Mol. Phys.* **1981**, *14*, 2739.

(17) Biegler-König, F. W.; Bader, R. F. W.; Tang, T.-H. *J. Comput. Chem.* **1982**, *13*, 317.

(18) Biegler-König, F. W. *J. Comput. Chem.* **2000**, *21*, 1040.

(19) Koritsanszky, T. S.; Coppens, P. *Chem. Rev.* **2001**, *101*, 1583.

(20) Coppens, P. *X-ray Charge Densities and Chemical Bonding*; Oxford University Press: New York, 1997.

(21) Spackman, M. A. *R. Soc. Chem. Ann. Rep. Section C* **1998**, *94*, 177.

(22) Massa, L.; Huang, L.; Karle, J. *Int. J. Quantum. Chem: Quantum Chem. Symp.* **1995**, *29*, 371.

(23) Huang, L.; Massa, L.; Karle, J. *Int. J. Quantum. Chem: Quantum Chem. Symp.* **1996**, *30*, 479.

(24) Pichon-Pesme, V.; Lecomte, C.; Lachekar, H. *J. Phys. Chem.* **1995**, *99*, 9.

(25) Wiest, R.; Pichon-Pesme, V.; Bénard, M.; Lecomte, C. *J. Phys. Chem.* **1994**, *98*, 1351.

(26) Pichon-Pesme, V.; Lecomte, C.; Wiest, R.; Bénard, M. *J. Am. Chem. Soc.* **1992**, *114*, 2713.

(27) Hansen, N. K.; Coppens, P. *Acta Crystallogr.* **1978**, *A34*, 909.

(28) Jelsch, C.; Pichon-Pesme, V.; Lecomte, C.; Aubry, A. *Acta Crystallogr.* **1998**, *D54*, 1306.

(29) Jelsch, C.; Teeter, M. M.; Lamzin, V.; Pichon-Pesme, V.; Blessing, R. H.; Lecomte, C. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3171.

(30) Housset, D.; Benabicha, F.; Pichon-Pesme, V.; Jelsch, C.; Maierhofer, A.; David, S.; Fontecilla-Camps, J. C.; Lecomte, C. *Acta Crystallogr.* **2000**, *D56*, 151.

(31) Chang, C.; Bader, R. F. W. *J. Phys. Chem.* **1992**, *96*, 1654.

(32) Martín, F. J. *Theoretical Synthesis of Macromolecules from Transferable Functional Groups*. Ph.D. Thesis, McMaster University, Hamilton, Ontario, Canada, 2001.

(33) Martín, F. J.; Bader, R. F. W. To be published.

(34) Bader, R. F. W.; Martín, F. J. *Can. J. Chem.* **1998**, *76*, 284.

(35) Breneman, C. M.; Thompson, T. R.; Rhem, M.; Dung, M. *Computers and Chemistry* **1995**, *19*, 161.

(36) Breneman, C. M.; Weber, L. W. In *NATO ASI Series: The Application of Charge Density Research to Chemistry and Drug Design*; Jeffrey, G. A., Piniella, J. F. E., Eds.; Plenum Press: New York, 1991.

(37) Breneman, C. M.; Rhem, M. *J. Comput. Chem.* **1997**, *18*, 182.

(38) Hernández-Trujillo, J.; Bader, R. F. W. *J. Chem. Phys.* **2001**. Accepted for publication.

(39) Yang, W. *Phys. Rev. Lett.* **1991**, *66*, 1438.

(40) Yang, W. *Phys. Rev. A* **1991**, *44*, 7823.

(41) Yang, W. *J. Mol. Struct. (THEOCHEM)* **1992**, *255*, 461.

(42) Yang, W.; Lee, T.-S. *J. Chem. Phys.* **1995**, *103*, 5674.

(43) Scuseria, G. E. *J. Phys. Chem. A* **1999**, *103*, 4782.

(44) Strain, M. C.; Scuseria, G. E.; Frisch, M. J. *Science* **1996**, *271*, 51.

(45) Bradbury, A. F.; Smyth, D. G.; Snell, C. R. *Nature* **1976**, *260*, 165.

(46) Feinberg, A. P.; Creese, I.; Snyder, S. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 4215.

(47) Snyder, S. H. *Sci. Am.* **1977**, *236*, 44.

(48) Schiller, P. W. In *The Peptides: Analysis, Synthesis, Biology*; Udenfriend, S., Meienhofer, J. E., Eds. Academic Press: New York, 1984; Vol. 6, p 219.

(49) Reisine, T.; Pasternak, G. In *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*; Hardman, J. G., Limbird, L. E., Molinoff, P. B., Ruddon, R. W., Gilman, A. G. E., Eds.; McGraw-Hill: New York, 1996; p 527.

(50) Kaufman, J. J.; Semo, N. M.; Koski, W. S. *J. Med. Chem.* **1975**, *18*, 647.

(51) Michel, A. G.; Proulx, M.; Evrard, G.; Norberg, B.; Milchert, E. *Can. J. Chem.* **1988**, *66*, 2498.

(52) Bader, R. F. W. *J. Phys. Chem. A* **1998**, *102*, 7314.

(53) Fulmor, W.; Lancaster, J. E.; Morton, G. O.; Brown, J. J.; Howell, C. F.; Nora, C. T.; Hardy, R. A., Jr. *J. Am. Chem. Soc.* **1967**, *89*, 3322.

(54) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. GAUSSIAN94, revision B.3; Gaussian, Inc.: Pittsburgh, PA, 1995.

(55) AIMPAC program suite can be downloaded free of charge from Internet at http://www.chemistry.mcmaster.ca/aimpac/.

(56) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*, 2nd ed.; Gaussian, Inc.: Pittsburgh, PA, 1996.

(57) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications: New York, 1989.

(58) Levine, I. N. *Quantum Chemistry*. Prentice Hall: New Jersy, 1991.

(59) Loew, G. H.; Villar, H. O.; Cometta, C.; Perez, J. *Int. J. Quantum Chem.: Quantum Biol. Symp.* **1991**, *18*, 165.

(60) Cambridge Crystallographic Data Centre (CCDC), Cambridge, U.K., 2001.

(61) HyperChem, version 5.01; HyperCube, Inc.: Waterloo, Canada, 1996.

(62) Aubry, A.; Birlirakis, N.; Sakarellos-Daitsiotis, M.; Sakarellos, C.; Marraud, M. *Biopolymers* **1989**, *28*, 27.

(63) Tóth, G.; Watts, C. R.; Murphy, R. F.; Lovas, S. *Proteins: Struct., Funct., Genet.* **2001**, *43*, 373.

(64) Bader, R. F. W.; Matta, C. F. *Int. J. Quantum. Chem.* **2001**, *85*, 592.

(65) Popelier, P. L. A. *Mol. Phys.* **1996**, *87*, 1196.

(66) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83*, 233.

(67) Stone, A. J.; Alderton, M. *Mol. Phys.* **1985**, *56*, 1047.

(68) Uberuaga, B. P.; Batista, E. R.; Jónsson, H. *J. Chem. Phys.* **1999**, *111*, 10664.

(69) Popelier, P. L. A. *Theor. Chem. Acc.* **2001**, *105*, 393.

(70) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(71) Liu, B.; McLean, A. D. *J. Chem. Phys.* **1973**, *59*, 4557.

(72) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873.

(73) Senent, M. L.; Wilson, S. *Int. J. Quantum Chem.* **2001**, *82*, 282.

(74) Bader, R. F. W.; Becker, P. *Chem. Phys. Lett.* **1988**, *148*, 452.

(75) Bader, R. F. W. *Int. J. Quantum Chem.* **1995**, *56*, 409.

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

(77) Carbó-Dorca, R.; Robert, D.; Amat, Ll.; Gironé, X.; Besalú, E. *Molecular Quantum Similarity in QSAR and Drug Design*; Springer: Berlin, 2000.

(78) Kjeldgaard, M. In *Structure and Dynamics of Biomolecules: Neutron and Synchrotron Radiation for Condensed Matter Studies*; Fanchon, E., Geissler, E., Hodeau, J.-L., Regnard, J.-R., Timmins, P. A. E., Eds.; Oxford University Press: New York, 2000.

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

(80) Richards, W. G. *Pure Appl. Chem.* **1988**, *60*, 277.

(81) Jones, T. A.; Zou, J.-Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr.* **1991**, *A47*, 110.

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

(83) Chae, C. H.; Oh, D. G.; Shin, W. *J. Comput. Chem.* **2001**, *22*, 888.

(84) McMahon, A. J.; King, P. M. *J. Comput. Chem.* **1997**, *18*, 151.

(85) Nissink, J. W. M.; Verdonk, M. L.; Kroon, J.; Mietzner, T.; Klebe, G. *J. Comput. Chem.* **1997**, *18*, 638.

(86) Cooper, D. L.; Allan, N. L. *J. Comput.-Aided Mol. Design* **1989**, *3*, 253.

(87) Measures, P. T.; Mort, K. A.; Allan, N. L.; Cooper, D. L. *J. Comput.-Aided Mol. Design* **1995**, *9*, 331.

(88) Lideman, R. H.; Merenda, P. F.; Gold, R. Z. *Introduction to Bivariate and Multivariate Analysis*; Scott, Foresman and Co.: Glenview, IL, 1980.