Holographic Electron Density Shape Theorem and Its Role in Drug Design and Toxicological Risk Assessment

Paul G. Mezey*

Mathematical Chemistry Research Unit, Department of Chemistry and Department of Mathematics and Statistics, University of Saskatchewan, 110 Science Place, Saskatoon, SK, Canada, S7N 5C9

Received April 5, 1998

Each complete, boundaryless molecular electron density is fully determined by any nonzero volume piece of the electron density cloud. This inherent feature of molecules, called the "holographic" property of molecular electron densities, provides a strong foundation for the local, quantum chemical shape analysis of various functional groups, pharmacophores, and other local molecular moieties. A proof is presented for the relevant molecular shape theorem, the "holographic electron density shape theorem", and the role of this theorem in quantum chemical, quantitative shape−activity relations (QShAR) is discussed. The quantum chemical methods of molecular shape analysis can be extended to ab initio quality electron densities of macromolecules, such as proteins, as well as to local molecular moieties, such as functional groups or pharmacophores, based on the transferability and additivity of local, fuzzy density fragments and the associated local density matrixes within the framework of the ADMA (Adjustable Density Matrix Assembler) approach. In addition to new results on chemical bonding and the development of macromolecular force methods, the new methodologies are also applicable to QShAR studies in computer-aided drug discovery and in toxicological risk assessment.

1. INTRODUCTION

One problem of importance in correlating molecular structure information with biochemical effects is the extent to which local molecular regions determine the properties of single molecules. The very concept of chemical functional groups has apparently evolved from classification efforts where various types of chemical reactions are classified according to the presence of certain reactive local molecular regions. Although it is well understood that within a family of molecules containing the same functional group not all these functional groups react exactly the same way, one nevertheless expects at least similar chemical behavior from these functional groups in most of these molecules. For example, in a family of different carbonyl compounds, no two carbonyl groups are exactly alike, and various degrees of differences are expected as a consequence of the influence of their different surroundings. Nevertheless, most of these compounds are expected to show similar reactivities in some typical reactions involving carbonyl groups.

Most investigations of this problem have focused on the question of how and to what degree does the surrounding local environment of a functional group influence the chemical properties; for example, the reactivities of these functional groups. Because the carbonyl groups themselves are somewhat different in these different compounds, it is also of interest to study the reverse problem of how well these local differences of functional groups of a given type can reflect the differences in their surroundings, and, ultimately, the differences in the complete molecules they belong.

Both of these questions have fundamental implications concerning the predictability of chemical and biochemical effects of molecules from information obtained for some of their local regions. Such information is valuable in drug design, in toxicological risk assessment, and in many other fields where local structure−property−activity correlations are used.

In this study we shall focus on the second question of the predictability of complete molecular properties from local electron density information, using some of the tools of density functional theory[1−5] and mappings onto spheres of various dimensions.[5,6] Our focus will be on electron density shape analysis,[7] and on the relations between the shape of local regions, such as quantum chemical functional groups,[8] and the shape analysis of complete molecules, which now may include rather large systems, such as proteins.[9−13] In this context, the recent developments of macromolecular electron density computation methodologies, including the macromolecular density matrix approach ADMA (Adjustable Density Matrix Assembler) provide the necessary tools.[11,12]

When addressing the local shape and global shape connection, we shall rely on a result that for a somewhat limited model of electron densities confined to a finite region[4] has been known for some time. More recently, a more general result has been proven for the molecular case of fuzzy, boundaryless electron densities.[5] This result resembles some characteristic properties of holographic plates: any small, nonzero area piece of a holographic plate contains, in principle, the entire information necessary to reconstruct the entire three-dimensional image. Electron densities have a similar property: in principle, any nonzero volume electron density fragment of a molecule contains the complete electron density information about the entire molecule. In

* E-mail mezey@sask.usask.ca.

10.1021/ci980072y CCC: $18.00
Published on Web 09/25/1998

the present study, we shall focus on the aspects of shape of electron densities, specifically, on the predictability of the shape of the electron density cloud of the complete molecule from the shape of a local region, for example, of a functional group. A proof of the relevant result, the "Holographic Electron Density Shape Theorem" will be given, and some consequences of this theorem to applied shape analysis in drug design[14] and in toxicological risk assessment[15−17] will be discussed.

2. SOME ELEMENTS OF MOLECULAR SHAPE ANALYSIS AND ELECTRON DENSITY FUNCTIONALS

There are extensive reviews on the topological techniques of molecular electron density shape analysis in the literature,[7,8] and here only a brief summary will be given.

Denote the nuclear arrangement (or, as it is often called, the nuclear configuration) of a molecule M by *K*, and an electron density threshold by *a*, then a *molecular isodensity contour*, MIDCO $G(K,a)$, of M is the collection of all those points **r** of the three-dimensional space where the electron density $\rho(K,\mathbf{r})$ of the molecule M of conformation *K* is equal to the threshold *a*:

$$G(K,a) = \{\mathbf{r}: \rho(K,\mathbf{r}) = a\} \tag{1}$$

The associated *density domain* $DD(K,a)$ is the collection of all points **r** where the electron density $\rho(K,\mathbf{r})$ of molecule M is greater than or equal to the threshold *a*:

$$DD(K,a) = \{\mathbf{r}: \rho(K,\mathbf{r}) \geq a\} \tag{2}$$

The MIDCO $G(K,a)$ is the boundary surface of the density domain $DD(K,a)$.

An additive fuzzy electron density fragment[8] (AFDF) associated with a family of nuclei f_k, is by definition a *quantum chemical functional group* [8] if there exists some density threshold *a* such that this nuclear family f_k is separated from the rest of the nuclei of the molecule by a MIDCO $G(K,a)$ of the given density threshold. The regions within these MIDCOs of functional group electron densities mark the extent of "limited autonomy" of functional groups within molecules. The density thresholds of these MIDCOs fall within a typical range, called the "Functional Group Range", discussed in detail in the literature.[7,8] These functional groups are fully determined by the electron density, and no chemical intuition or chemical bias is involved.

The limited autonomy and separate identity of functional groups within molecules are reflected in this definition. One might compare the case of functional groups to the case of two different molecules placed within a short distance of one another, where the existence of some MIDCOs separating the nuclei of the two molecules indicate the autonomy and separate identity of these molecules.

A topological shape analysis technique of complete, three-dimensional fuzzy electron densities, or of any almost everywhere twice continuously differentiable three-dimensional function, is the Shape Group Method[7] (SGM).

If the SGM is applied to molecular electron densities, then the local curvatures of a range of MIDCOs are compared with a range of reference curvatures.[7] The process involves several steps. First, a range of electron density thresholds *a* and a range of reference curvature values *b* are selected. For each MIDCO $G(K,a)$ of density threshold *a* within the specified range $G(K,a)$ is partitioned into local curvature domains relative to each reference curvature *b* within the corresponding range. There are three types of local curvature domains, $D_0(b)$, $D_1(b)$ or $D_2(b)$, indicating whether the MIDCO $G(K,a)$ is convex, concave, or of the saddle type, respectively, relative to the actual curvature b.

To generate this classification, the local curvature at each point **r** of the MIDCO surface $G(K,a)$ is characterized by a local curvature matrix called the local Hessian matrix, by comparing the local canonical curvatures (the eigenvalues of the local Hessian matrixes) at each point **r** to the reference curvature *b*. Each point **r** of $G(K,a)$ is assigned to a $D_0(b)$, a $D_1(b)$, or a $D_2(b)$ curvature domain, if none, one, or two, respectively, of the eigenvalues of the local Hessian matrix at point **r** are smaller than the curvature parameter *b*. In practical computations, only a finite number of (a,b) pairs are considered.

This step is followed by a truncation of MIDCO surfaces by removing all curvature domains $D_\mu(b)$ of a specified type μ (in most application the type $\mu = 2$) from the MIDCO $G(K,a)$, for each (a,b) pair of parameters. A truncated surface $G(K,a,\mu)$ is obtained for each (a,b) pair. For the whole range of parameter values *a* and *b* of each molecule M, only a finite number of topologically different truncated surfaces are obtained, in all except some degenerate cases.

In the next step, the *shape groups* of the molecular electron density, that is, the zero-, one-, and two-dimensional algebraic homology groups of the truncated surfaces, are computed. Consequently, for electron density analysis in three dimensions, using two-dimensional MIDCO surfaces, there are three types of shape groups, one for each of the dimensions zero, one, and two. Within each topological equivalence class of these surfaces, the shape groups are topological invariants.

The zero-, one-, and two-dimensional Betti numbers are the ranks of these zero-, one-, and two-dimensional homology groups, respectively. The zero-, one-, and two-dimensional shape groups are denoted by $H^0_\mu(a,b)$, $H^1_\mu(a,b)$, and $H^2_\mu(a,b)$, respectively, where the letter H refers to the fact that these groups are the homology groups of truncated MIDCO surfaces. In these notations, the dimensions 0, 1, and 2, the truncation type μ, the electron density threshold *a*, and the reference curvature parameter *b* are also specified.

The Betti numbers associated with the shape groups are denoted by $b^0_\mu(a,b)$, $b^1_\mu(a,b)$, and $b^2_\mu(a,b)$, where, similarly to the shape groups, the dimensions 0, 1, and 2, the two parameters *a* and *b*, as well as the truncation type μ are specified. For the entire range of MIDCOs $G(K,a)$ of the molecule M, the Betti numbers generate a set of numerical shape descriptors.

In most applications, the distribution of various values of Betti numbers $b^p_\mu(a,b)$ as a function of the density threshold *a* and curvature parameter *b* is described by various (a,b) maps. The shape matrixes $\mathbf{M}^{(a,b)}$ are discretized versions of (a,b) maps obtained if a finite number, n_a and n_b, of *a* and *b* values are used, respectively. These matrixes are numerical shape codes for molecules, where the total number of elements in the shape code matrix $\mathbf{M}^{(a,b)}$ is

$$t = n_a n_b \tag{3}$$

Define $m[\mathbf{M}^{(a,b),A}, \mathbf{M}^{(a,b),B}]$ as the number of matches between corresponding elements in the two shape code matrixes $\mathbf{M}^{(a,b),A}$ and $\mathbf{M}^{(a,b),B}$ of the two molecules or molecular fragments A and B. Based on the shape codes, the shape similarity between two molecules or molecular regions A and B can be expressed using the shape-similarity measures (A,B) defined as

$$s(\mathrm{A,B}) = m[\mathbf{M}^{(a,b),A}, \mathbf{M}^{(a,b),B}]/t \tag{4}$$

For the derivation of some of the results discussed in this report, it is advantageous to formulate specific shape aspects of local molecular regions in terms of density functional theory. According to the fundamental Hohenberg−Kohn theorem[3] of density functional theory, the nondegenerate ground-state electron density $\rho(\mathbf{r})$ of a molecule of n electrons in a local spin-independent external potential V, expressed in a spin-averaged form as

$$\rho(\mathbf{r}) = n\sum_{s_1}\ldots\sum_{s_n}\int\ldots\int|\Psi(\mathbf{r},s_1,\mathbf{r}_2,s_2,\ldots\mathbf{r}_n,s_n)|^2\,\mathrm{d}^3\mathbf{r}_2\ldots\mathrm{d}^3\mathbf{r}_n \tag{5}$$

fully determines the electronic energy E as well as other properties of the molecule.

In more precise terms, the Hohenberg−Kohn theorem[3] establishes that a nondegenerate ground-state electron density $\rho(\mathbf{r})$ determines the Hamiltonian H of the system within an additive constant (two identical densities must belong to two Hamiltonians that are identical up to a constant), consequently, the electron density $\rho(\mathbf{r})$ also determines all ground state and all excited-state properties of the system, where for a molecule M, the Hamiltonian H can be expressed as

$$H = \sum_{i=1}^{n} V(\mathbf{r}_i) + T + V_{ee} \tag{6}$$

in terms of the usual kinetic energy operator T, the electron−electron repulsion operator V_{ee}, and potential $V(\mathbf{r})$,

$$V(\mathbf{r}) = \sum_{i=1}^{n} V(\mathbf{r}_i) \tag{7}$$

a sum of one-electron potentials, where $V(\mathbf{r}_i)$ is the electron−nuclear attraction operator for the interaction of the set of nuclei with the ith electron of the molecule.

The "constrained search" approach, as described by Levy,[2] leads to an elegant proof of the Hohenberg−Kohn theorem. However, as discussed in an earlier study,[5] a simple consideration of the information content of a molecule also leads to the same conclusion:

1. A molecule contains only a set of nuclei and an electron density cloud.

2. Because there is no other material present to encode information, all information concerning the static properties of the molecule must be contained in the nuclear and electron distributions.

3. The location and atomic numbers of the nuclei are fully determined by the nondegenerate ground-state electron density $\rho(\mathbf{r})$ of the molecule.

4. Consequently, the nondegenerate ground-state electron density $\rho(\mathbf{r})$ contains all information concerning all static properties of the molecule, including its ground-state energy and any other molecular properties.

The first application of the Hohenberg−Kohn theorem to a local electron density problem was to a finite subsystem of another, larger, but still finite and bounded system.[4] However, according to the fundamental quantum chemical properties of electron distributions and obeying the Heisenberg uncertainty relation, molecular electron densities do not have boundaries. Hence, in a strict sense, they are not confined to a finite volume. These fundamental properties were taken into account in a later study,[5] where the local electron densities of subsystems as parts of complete, boundaryless electron densities of entire molecules were studied in terms of the Hohenberg−Kohn theorem. This latter study[5] used a four-dimensional representation of molecular electron densities [where the first three dimensions are those of the ordinary three-space E^3 and the fourth dimension represents the electron density values $\rho(\mathbf{r})$], and a compactification method that maps all points of the ordinary three-dimensional space E^3 to a manifold S^3 embedded in a four-dimensional Euclidean space E.[4] Using these tools, it was possible to use the theorem of analytic continuation of functions almost everywhere analytic on a compact manifold to establish the "holographic properties" of molecular electron densities represented on a compact manifold S^3, and the equivalent "holographic properties" of complete, boundaryless molecular electron densities no longer confined to any finite, bounded region of the ordinary three-dimensional space:

If $\rho'_d(\mathbf{r}')$ denotes the nondegenerate ground-state electron density over a subset d of manifold S^3, $S^3 \supset d$, where d has nonzero volume, then the ground-state electron density $\rho'(\mathbf{r}')$ over the entire manifold S^3 is uniquely determined by the local ground-state electron density $\rho'_d(\mathbf{r}')$ over the subdomain d.

The proof of this result[5] was based on a special, four-dimensional representation of the electron density. However, as a consequence of the properties of the compactification assigning points of the ordinary space E^3 to the manifold S^3, the conclusions apply to the ordinary, three-dimensional density functions $\rho(\mathbf{r})$.

As one consequence of this holographic property, exact transferability of electron density "pieces" with boundaries is necessarily very limited, being restricted to corresponding local electron densities in identical molecules and to local electron densities related by symmetry.

The local electron density fully determines the complete electron density; that is, if the local electron density is exact, the determination is also exact. If, however, the local electron density is not exact, only some approximation, then, of course, the extrapolation to the complete molecule, and to molecular properties, will also be only approximate (just as if the holographic plate or the recording of the holographic image is imprecise, as it always is in practical holography, then the image reconstruction is only approximate). For this reason, in approximate calculations, as most calculations are, a larger piece of the electron density is likely to give better property correlations.

The holographic property has interesting consequences for local shape analysis of molecular electron densities, as discussed in the next section.

3. THE HOLOGRAPHIC ELECTRON DENSITY SHAPE THEOREM

Chemists often base their conclusions regarding properties of entire molecules on the presence and interactions of local molecular regions. In much of the recent studies on structure−activity correlations, the focus is often placed on assigning roles to local regions, such as formal functional groups, or combinations of functional groups, or moieties regarded as pharmacophores. The local shape analysis of such regions, and the detection and possible interpretation of their subtle differences are among the tasks in contemporary computational chemistry contributions to drug design efforts and toxicological risk assessment.[14−17] The assumption underlying these efforts is the expectation that certain well-selected local regions of larger molecules hold the key to the explanation of their various degrees of activities. Small differences in these local regions and their local surroundings can have a significant effect on the overall reactivity of molecules.

In fact, a fundamental justification of this assumption can be found in the holographic properties of electron densities. The following result is of importance in molecular shape analysis, that warrants special emphasis.

Holographic Electron Density Shape Theorem. The nondegenerate ground-state electron density $\rho_D(\mathbf{r})$ within a nonzero volume local region D of the ordinary three-space E^3 fully determines all shape groups $H^0_\mu(a,b)$, $H^1_\mu(a,b)$, and $H^2_\mu(a,b)$ of the complete molecule for the entire range of electron density thresholds a and reference curvature parameter b.

The proof follows directly from the holographic property of electron densities:

(i) any nondegenerate ground-state electron density $\rho_D(\mathbf{r})$ within a nonzero volume local region D of the ordinary three-space E^3 fully determines the complete, boundaryless electron density $\rho(\mathbf{r})$ of the entire molecule, and in turn,

(ii) all shape groups $H^0_\mu(a,b)$, $H^1_\mu(a,b)$, and $H^2_\mu(a,b)$ of the complete molecule are fully determined by the complete electron density $\rho(\mathbf{r})$.

Because a molecule contains nothing else than a boundaryless electron density cloud and a family of nuclei, where all information on the nuclear distribution is fully contained in the electron density $\rho(\mathbf{r})$, the electron density itself carries all information concerning the static properties of the molecule. Because the shape groups $H^0_\mu(a,b)$, $H^1_\mu(a,b)$, and $H^2_\mu(a,b)$ provide a detailed description of the electronic density, it is not surprising that excellent correlations can be obtained between the shape codes $\mathbf{M}^{(a,b)}$ and shape similarity measures s(A,B), derived from the shape groups and various, experimentally observed molecular properties,[14] including toxicities[15−17] within complex biological organisms. The fact that a local region of the electron density contains sufficient information to determine the shape of the entire molecule indicates that it is indeed justified to use local regions for the interpretation of chemical and biochemical properties. This observation is reassuring with regard to the roles of local molecular regions, even if it does not appear computationally practical to actually reconstruct a complete electron density of a large molecule from a single, local piece of this density.

One may rely on a few fundamental facts to formally deduce some general biochemical and toxicological consequences of the holographic electron density property.

(i) The local electron density determines the complete electron density, that in turn determines all properties of any single molecule. Simply, there is no other material in a molecule but electron density and nuclei, and the nuclear distribution is fully described by the electron density. Because there nothing else exists, no other material in a molecule that could contain any information on any property of a single molecule, electron density must determine all poperties of a single molecule. The local electron density determines the complete electron density, hence, every property of the single molecule.

(ii) In particular, the local electron density determines the differences in biochemical activities because the differences in biochemical activities, when compared with any fixed reference, are single molecule properties. For example, the differences between the activities of A versus X, B versus X, C versus X, ... and D versus X, where X is any reference molecule of some given activity, are due entirely to the differences among the individual molecules A, B, C, ... and D; that is, these differences are single molecule properties! As such, they are fully determined by the molecular electron densities, hence, by the local electron densities.

(iii) The same is true if X is a receptor, and the difference of interactions of A, B, C, ... and D, with the common receptor X, are compared. These differences are also single molecule properties, fully determined by the local electron densities, as long as X is fixed.

(iv) The differences in molecular flexibilities, regarded as important in biochemical complementarity, are also determined by the electron density; hence, by the local electron densities.

4. REGIONAL ELECTRON DENSITIES, FUNCTIONAL GROUP ELECTRON DENSITIES, AND HOST−GUEST INTERACTION DENSITIES

Whereas the chemically most important local regions are those usually associated with functional groups, the AFDF approach can be used for the study of larger regions or local molecular moieties that do not correspond to any functional group. With respect to the holographic electron density property, it is a natural expectation that a larger local region including several nuclei contains the information concerning the rest of the molecule in a more accessible form than a small domain at some distance from the nuclei. For this reason, if local domains of the molecule are used for correlations with biochemical activity, it is natural to expect that larger domains provide better correlations. We shall see in the next section that this is not necessarily the case.

In the context of host−guest interactions,[18] the shape analysis of regional electron densities of interacting molecule pairs has been proposed as a tool for characterizing these interactions. Biochemical host−guest interactions are typically accompanied by changes of electron densities as manifested in the composite electron density of the host−guest system when compared with the electron densities of the noninteracting molecules. Similarity and dissimilarity measures of the electron densities of interacting and noninteracting host and guest molecules provide quantum chemical information on these interactions.

The AFDF principle is applicable to molecular regions or to regions in the composite of two, interacting molecules. These regions are described by fuzzy electron densities analogous to densities of complete molecules. The local, topological shape analysis of regions is based on the same principles as the shape analysis of complete, individual molecules. For large molecular regions and for regions of interactions of molecule pairs, the "region isodensity contour" (RIDCO) surface terminology is used,[18] by analogy with "molecular isodensity contour " (MIDCO) surfaces.

A molecule M or a pair of interacting molecules can be regarded as a composite of several regions,

$$R_1, R_2,..., R_r \tag{8}$$

where individual regions might involve several functional groups, and where the existence of an isodensity contour separating the nuclei of a region from the other nuclei is not required. Nevertheless, AFDF method is applicable for the generation of the fuzzy regional electron density contributions[18]

$$\rho_{R1}(\mathbf{r}), \rho_{R2}(\mathbf{r}),... \rho_{Ri}(\mathbf{r}),... \rho_{Rr}(\mathbf{r}) \tag{9}$$

of these regions R_1, R_2,..., R_r. These regional densities, in turn, collectively reproduce the macromolecular electron density $\rho_M(\mathbf{r})$:

$$\rho_M(\mathbf{r}) = \rho_{R1}(\mathbf{r}) + \rho_{R2}(\mathbf{r}) +... + \rho_{Ri}(\mathbf{r}) +... + \rho_{Rr-1}(\mathbf{r}) + \rho_{Rr}(\mathbf{r}) \tag{10}$$

If $R = R_r$ denotes the actual region studied, and M′ denotes the rest of the molecule regarded as the composite of all the other regions R_1, R_2,..., R_{r-1}, then the electron density $\rho_{M'}(\mathbf{r})$ for the region M′ complementing R is described as a composition of the fuzzy regional densities of regions R_1, R_2,..., R_{r-1}:

$$\rho_{M'}(\mathbf{r}) = \rho_{R1}(\mathbf{r}) + \rho_{R2}(\mathbf{r}) +... + \rho_{Ri}(\mathbf{r}) +... + \rho_{Rr-1}(\mathbf{r}) \tag{11}$$

The shape group analysis, the determination of shape codes, and shape similarity measures of the isodensity surfaces of these regional densities are generated by the usual technique.

Some of the interactions involve the low density regions of the interfacing, peripheral ranges of the two molecules. Many of the self-interactions of folded macromolecules also involve the low density regions, where the contribution to chemical bonding and to the actual conformation is less directional than ordinary chemical bonds; also, these contributions are less localized, not being located primarily between two nuclei. In proteins, these contributions to the folding pattern are assumed to have important role, analogous to a "low-density glue" (LDG) bonding, that is superimposed on the conventional, directional, and more localized bonding contributions.[19]

The electron density domains of regional interactions are expected to provide local shape information that correlates with the energy, strength, and the feasibility of these interactions. In addition, the holographic electron density property implies that the local electron densities provide symmetry and chirality information about the complete molecule, a fact that has interesting consequences with respect to generalized chirality and symmetry deficiency properties of molecules.[20]

5. THE ROLE OF LOCAL REGIONS OF ELECTRON DENSITIES IN THE PREDICTION OF BIOCHEMICAL EFFECTS

The simplest version of the general AFDF approach,[11] the numerical MEDLA technique (Molecular Electron Density "Loge" Assembler method) was used in the first computation of ab initio quality electron densities for macromolecules, such as various proteins including the HIV Protease of 1564 atoms.[9,10,13] The AFDF methods have also provided practical tools for the computation of ab initio quality electron density fragments. These fragments are "tailor-made" for any given molecular environment, reflecting, in fact, all the local interactions within this environment including the local forces acting on the nuclei,[11,12] as well as interactions between various local domains of the electron density. For the range of interactions represented, the reader may consult the relevant references.[11,12] This AFDF approach has been used for the study of correlations between local shape features and biochemical properties,[14−17] including toxicities of various molecular series.[15−17]

Electron density fragmentation methods are applicable to the study of the combined effects of local and global shape features of molecular electron densities. In some instances the biochemical effect is a result of several factors, and in such cases, no single shape feature can be expected to correlate well with the experimental results. Here we shall use the example of a toxicological study[17] involving the aquatic species *L. gibba* (duck weed) and a series of polyaromatic hydrocarbons (PAHs). These PAH molecules are known to undergo photochemical activation and subsequent chemical reactions in the environment, triggering a series of biochemical processes resulting in various degrees of toxicity in aquatic species. Although the mechanism of the chain of all these complex processes is far from being completely understood, molecules of high degree of similarity are expected to have similar toxic effects; the extreme case of identical molecules is clearly one where identical toxic effects are expected. Hence, a similarity analysis of molecular shapes, using shape similarity measures correlated with experimentally observed toxicities, is expected to provide predictions for the toxic effects of molecules from the same general molecular family.

There is considerable current interest in the toxicological risk assessment involving the photoinduced toxicity of PAHs. Photoinduced toxicity of PAHs is initiated by photosensitization reactions usually resulting in the generation of singlet oxygen, followed by one or several photomodification steps, for example, photooxidation or photolysis, leading to more toxic substances.

To illustrate some of the interplay between local and global shape features of electron densities in correlation with experimentally measured biochemical effects, we shall use the example of a study that has given excellent correlations,[17] and in the present context, it also provides some computational evidence of the role of the holographic property. Using the SGM and the associated shape similarity measures, detailed electron density shape analysis was carried out for

16 PAH molecules. These PAHs had been experimentally tested previously for *L. gibba* photoinduced toxicity, where the plant growth was measured in terms of doubling of leaves directly related to the weight gain of *L. gibba*.

The shape similarities of the (*a*,*b*) maps of the complete molecules did not show useful correlations with the experimental toxicities, neither did the one-ring fragments, nor the two-ring fragments, among the several local shape features tested for correlations. However, the combination of the single ring and global, complete molecule similarities correlated well with toxicity. The heuristic formula

$$Y = c_0/(1 - c_1 X_1 - c_2 X_2) \tag{12}$$

where X_1 and X_2 are the one-ring and whole molecule similarities, respectively, and the quantities c_0, c_1 , and c_2 are the regression coefficients, gave excellent theoretical toxicities. The relation between the experimental and theoretical toxicities, based on the relation in eq 12 was found as

$$Y_{\text{exper}} = 1.000\, Y_{\text{theor}} - 0.075 \tag{13}$$

with a correlation coefficient of 0.96.

It is interesting to note that, the combination of the one-ring and two-ring similarities also correlated well with toxicity. A heuristic formula of the same structure as eq 12, but where X_1 and X_2 were the one-ring and two-ring similarities, respectively, also gave good theoretical toxicities. In this case, the relation between the experimental and theoretical toxicities was found as

$$Y_{\text{exper}} = 0.927\, Y_{\text{theor}} + 40.153 \tag{14}$$

with a correlation coefficient of 0.88.

In the context of the holographic property of electron densities, it is noteworthy that none of the single, local region shape similarity measures actually evaluated did show a high degree of correlation with experimental toxicities, although some underlying trends were evident. This result, by itself, does not mean that no such indicative shape feature of these local regions exist, but if such features are present, we have either missed them, or they are too subtle to be detected at the resolution of the (*a*,*b*) shape maps used. The complete molecule shape similarity measures did not correlate well with the experiments either; the information content of these global shape maps is apparently too rich and the shape features that do correlate with experiments are buried in other information not relevant to toxicity, if a sample of only 16 PAHs is used.

However, the combination of a local (one-ring) and a global (complete molecule) similarity measures provided excellent correlations (0.96), which is justified by the mechanistic hypothesis that among the many factors influencing overall toxicity, a global shape feature is probably relevant to the photosensitization step involving the extended conjugated system of PAHs, whereas a local shape feature is important in the local reaction leading to photomodification.

With respect to the holographic property, it is significant that the correlations are only somewhat lower if in the combined shape features the complete molecule is replaced by a two-ring fragment (0.88). Apparently, the shape information about the relevant aspect of the global electron density can be successfully recovered from the shape similarity analysis of the two-ring fragments. As the holographic property implies, any small, nonzero volume subsystem contains the full information. However, if the subsystem is very small, it may be very hard to identify and extract this information. Apparently, these two-ring fragments are already large enough, and at the level of resolution of the (*a*,*b*) maps used the two-ring fragments are already suitable to reproduce reasonably well the information otherwise provided by the complete molecules. Whereas in the context of toxicities studied the single-ring fragments appear too small to have practical use as a single source of shape information concerning the complete molecules, the two-ring fragments in the actual set of PAH molecules already appear as reliable "holographic" carriers of the complete molecule shape information.

ACKNOWLEDGMENT

This study was supported by research grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the Canadian Network of Toxicology Centers (CNTC).

REFERENCES AND NOTES

(1) Kolossváry, I. Information theory and electron density. *J. Math. Chem.* **1996**, *19*, 393−399.
(2) Levy, M. Elementary Concepts in Density Functional Theory. In *Recent Developments and Applications of Modern Density Functional Theory, Theoretical and Computational Chemistry, Vol. 4*; Seminario, J. M., Ed.; Elsevier Science B. V.: Amsterdam, 1996; pp 3−24.
(3) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864−B871.
(4) Riess, J.; Münch, W. The Theorem of Hohenberg and Kohn for Subdomains of a Quantum System. *Theor. Chim. Acta* **1981**, *58*, 295−300.
(5) Mezey, P. G. The Holographic Electron Density Theorem and Quantum Similarity Measures. accepted for publication in *Mol. Phys.*
(6) Mezey, P. G. *Potential Energy Hypersurfaces*; Elsevier: Amsterdam, 1987.
(7) Mezey, P. G. *Shape in Chemistry: An Introduction to Molecular Shape and Topology*; VCH: New York, 1993.
(8) Mezey, P. G. Functional Groups in Quantum Chemistry. *Adv. Quantum Chem.* **1996**, *27*, 163−222.
(9) Walker, P. D.; Mezey, P. G. Ab initio Quality Electron Densities for Proteins: A MEDLA Approach. *J. Am. Chem. Soc.* **1994**, *116*, 12022−12032.
(10) Walker, P. D.; Mezey, P. G. A New Computational Microscope for Molecules: High-Resolution MEDLA Images of Taxol and HIV-1 Protease, Using Additive Electron Density Fragmentation Principles and Fuzzy Set Methods. *J. Math. Chem.* **1995**, *17*, 203−234.
(11) Mezey, P. G. Macromolecular Density Matrixes and Electron Densities with Adjustable Nuclear Geometries. *J. Math. Chem.* **1995**, *18*, 141−168.
(12) Mezey, P. G. Quantum Similarity Measures and Löwdin's Transform for Approximate Density Matrixes and Macromolecular Forces. *Int. J. Quantum Chem.* **1997**, *63*, 39−48.
(13) Mezey, P. G. Computational Microscopy: Pictures of Proteins. *Pharm. News* **1997**, *4*, 29−34.
(14) Walker, P. D.; Maggiora, G. M.; Johnson, M. A.; Petke, J. D.; Mezey, P. G.Shape Group Analysis of Molecular Similarity: Shape Similarity of Six-Membered Aromatic Ring Systems, *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 568−578.
(15) Mezey, P. G.; Zimpel, Z.; Warburton, P.; Walker, P. D.; Irvine, D. G.; Dixon, D. G.; Greenberg, B. A High-Resolution Shape-Fragment Database for Toxicological Shape Analysis of PAHs. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 602−611.
(16) Mezey, P. G. Molecular Structure − Reactivity − Toxicity Relationships. In *Soil Chemistry and Ecosystem Health*; Huang, P. M., Ed.; SSSA: Pittsburgh, PA, 1998; pp 21−43.

(17) Mezey, P. G.; Zimpel, Z.; Warburton, P.; Walker, P. D.; Irvine, D. G.; Huang, X.-D.; Dixon, D. G.; Greenberg, B. Use of QShAR to Model the Photoinduced Toxicity of PAHs: Electron Density Shape Features Accurately Predict Toxicity. *Environ. Toxicol. Chem.* **1998**, *17*, 1207−1215.
(18) Mezey, P. G. Molecular Similarity and Host−Guest Interactions. In *Pauling's Legacy: Modern Modelling of Chemical Bonding*; Maksic, Z.; Orville-Thomas, W. J., Eds.; Elsevier Science: Amsterdam, 1998 (in press).
(19) Mezey, P. G. Chemical Bonding in Proteins and Other Macromolecules. In *Pauling's Legacy: Modern Modelling of Chemical Bonding*, Maksic, Z.; Orville-Thomas, W. J., Eds.; Elsevier Science: Amsterdam, 1998 (in press).
(20) Mezey, P. G. Generalized Chirality and Symmetry Deficiency. *J. Math. Chem.* **1998**, *23*, 65−84.

CI980072Y