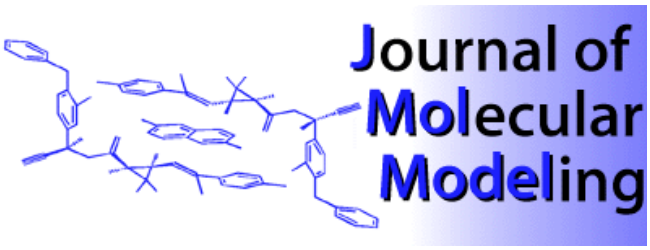
J. Mol. Model. **2000**, *6*, 150 – 157



FULL PAPER

Computer Aided Drug Design: Some Fundamental Aspects

Paul G. Mezey[1,2]

[1]Mathematical Chemistry Research Unit, Department of Chemistry and Department of Mathematics and Statistics, University of Saskatchewan, 110 Science Place, Saskatoon, SK CANADA S7N 5C9. Tel.: (306) 966 4661; Fax: (306) 966 4730. E-mail: mezey@sask.usask.ca

[2]Institute for Advanced Study, Collegium Budapest, 1014 Budapest, I. Szentharomsag u. 2, Hungary

Received: 25 November 1999/ Accepted: 20 January 2000/ Published: 28 February 2000

Abstract Understanding the molecular basis of drug action and exploring the chemical interactions involved in the complex processes of drug delivery are among the most important goals of contemporary drug design. The major recent advances in the detailed, mechanistic interpretation of molecular interactions, the global and local shape analysis of electron density clouds making up the actual fuzzy bodies of molecules, novel similarity and complementary measures, the detailed and accurate computational visualization techniques of molecular level „Computational Microscopy“, the advances in computer modeling of conformational processes and chemical reactions of drug molecules, the computer aided design of molecular templates fitting various receptor sites are among the powerful tools of computer aided drug discovery. In this contribution some of the latest advances are reviewed.

Keywords Theoretical drug design, Computational microscopy, Molecular electron density shape analysis, Similarity and complementary measures

Introduction

Computer aided drug design has matured into a scientifically still challenging, but industrially highly applicable field that has achieved considerable successes [1-4]. This field is an excellent example for the rapid introduction of novel, often seemingly abstract scientific methodologies and computational software advances into areas that were earlier dominated by experimental approaches and the conventional synthetic methods of chemistry. In this context, the early theoretical formulation and subsequent computer applications of molecular similarity represented important advances [5-10].

The biochemical processes involved in medicinal and pharmaceutical chemistry are usually very complex, and the detailed mechanisms of the microscopic chemical processes at the level of actual motions and rearrangements of individual molecular fragments are very seldom known throughout the entire process of drug action. Whereas the ultimate, detailed understanding of drug action would require the knowledge of such details, contemporary theoretical approaches to drug design usually set a more modest goal: the identification of the main active centers and the required

Dedicated to Professor Paul von Ragué Schleyer on the occasion of his 70th birthday

shape and electronic features of pharmacologically significant molecules. These features often have roles on several levels: at the level of actual, biochemical interactions with a specific target, generating the essence of the drug's activity, and also on the level of the processes required for the drug molecules to reach these targets. In all of these phases, the shapes of the molecules are of special importance.

A thorough approach to these problems very rapidly leads to some of the most interesting, and rather fundamental problems of contemporary biochemistry. On the actual, molecular level, all properties of molecules are fully determined by the electron density distribution, that is, by the electron density charge cloud surrounding the atomic nuclei present in the molecule. A molecule contains only positively charged atomic nuclei and the negatively charged electron density cloud; clearly, there is simply no other material present in a molecule that could possibly contain any information about the molecule. Furthermore, from the way the electronic charge cloud becomes dense in the vicinity of various nuclei it is possible to identify each nucleus; consequently, the electronic charge cloud alone already contains all information about the molecule. How and why a molecule behaves, interacts, and reacts the way it does are fully encoded in the shape of its fuzzy electron density cloud [11,12].

This realization has importance in highly applied areas, such as pharmaceutical chemistry and the study of drug action, but it also connects these applied fields to some of the fundamentals of theoretical chemistry, as it is well-founded in rigorous quantum mechanics. The essence of this idea has been formulated as the celebrated Hohenberg-Kohn theorem [13] of density functional theory [14], only seemingly removed from the medicinal aspects of chemistry. According to this theorem, the ground state electron density (in the non-degenerate cases) fully determines the energy of the molecule [13]. In fact, all other ground state properties of the molecule are also determined by the electron density. It is not surprising then, if in a series of similar molecules, various calculated electron density properties show excellent correlations with some biochemical activities.

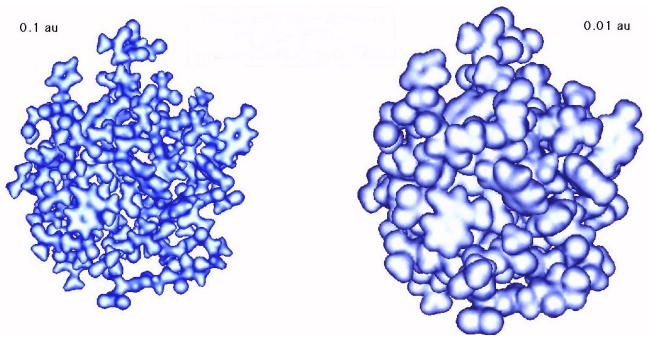


Figure 1 *Images of the* ab initio *quality 0.1 a.u. and 0.01 a.u. molecular isodensity contours (MIDCOs) of the Proto-Oncogene Tyrosine Kinese Protein 1ABL, computed by the MEDLA method.*

Computational approaches to drug design have their foundations in the correlations of computed and experimentally determined properties of molecules. The computed properties have a better chance of representing reality if they are based on a more rigorous theory; modern quantum chemistry represents such a theory, where many molecular properties can now be computed with an accuracy competing with or even surpassing the accuracy of experimental measurements [15]. Whereas the most accurate theoretical results are typically available mostly for small molecules, nevertheless, recent advances in the quantum chemistry of medium and large size molecules, including *ab initio* quality electron density computations for proteins [16-18] and the ADMA macromolecular density matrix approaches [19-22], suitable for the computation of forces acting on individual nuclei within a protein [22], have provided a new motivation for quantum chemistry computational approaches to drug design.

Theoretical chemistry basis of advanced computer modeling of molecules

Advanced molecular modeling approaches that are able to take advantage of the unprecedented development of computer hardware experienced in the recent decade are based on a thorough theoretical foundation provided by quantum chemistry. A well-established theoretical basis in such applied fields as drug research is no longer an unnecessary luxury; in fact, many theoretical results, which decades ago served only as guidelines or broad framework for the experiments in applied fields, are now becoming readily applicable for the experimentalists. Theoretical methods, implemented as computer programs [15,23-26], complement the experimental studies, in fact, these computer programs have become versatile instruments, tailor-made for the given application. Theory and experiment both benefit from this new scientific framework, since theories now can be tested easily against new experimental results, and experiments benefit from the computational techniques, which provide new tools for both interpretation and validation.

It is perhaps not surprising that advanced computational methods originally developed for small molecules are often employed for the study of pharmaceutically significant molecules. The transfer of methodology between different disciplines has accelerated in recent years, and the computational techniques of drug design increasingly rely on sophisticated and theoretically well established methodologies of computational quantum chemistry.

Many of the currently used quantum chemistry approaches are based on the determination of the molecular wavefunction, by solving (at least in an approximate sense) the fundamental equation of quantum chemistry, the molecular Schrödinger equation:

$$H\Psi = E\Psi \quad (1)$$

Here the standard notation is used, where H is the molecular Hamilton operator, that is in fact a complicated set of mathematical manipulations to be carried out on the molecular wavefunction Ψ (a rather complicated function), that converts Ψ into a multiple $E\Psi$ of itself, where the quantity E is the energy of the molecule. This description of the molecular Schrödinger equation is deceptively simple and hides the fact that the actual solution of this equation using sufficiently accurate approximate methods typically involves very complex computations, often including several millions of integrations. The development of a system of suitable computer programs for these calculations, such as the GAUSSIAN programs pioneered by Pople and coworkers [24], has taken many years by a large number of computational chemists. These computer programs are among the major achievements of modern theoretical chemistry.

In an LCAO (linear combinations of atomic orbitals) representation of the molecular wavefunction Ψ, the actual computation involves the determination of the relative contributions of various atomic orbitals to the molecular wavefunction Ψ.

If n is the total number of atomic orbitals, if r denotes the three-dimensional position vector variable, and if one denotes the i-th atomic orbital by $\varphi_i(r)$ (i = 1,2,...,n), then the set $\{\varphi_i(r)\}$ of functions is the molecular basis set for the given expansion of the molecular wavefunction Ψ. The relative weights of the contributions of the various basis functions determine the so-called density matrix P. Within the LCAO framework, most molecular properties can be calculated from the density matrix P, and most of the chemically relevant information present in the molecular wavefunction Ψ can also be represented by the density matrix P.

Using the density matrix and the molecular basis set $\{\varphi_i(r)\}$, an important property, the electronic density $\rho(r)$ of the molecule can be computed at any location r by the simple formula

$$\rho(r) = \sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij}\varphi_i(r)\varphi_j(r) \qquad (2)$$

Since the actual shape of the molecule is in fact the shape of the fuzzy electron density cloud, the computed electronic density has a primary role in molecular shape analysis, and the fuzzy „body“ of the electronic charge distribution is well described by the function $\rho(r)$.

The computational approach of the standard quantum chemistry molecular orbital framework, well represented by the programs included in the GAUSSIAN system [24], is suitable for the study of relatively small molecules. When using such *ab initio* computer programs, however, one faces severe difficulties if large molecules are the subjects of studies.

The usual implementation of the *ab initio* method for the generation of molecular wavefunctions suffers from a serious limitation: as the size of the molecules increases, the computer time requirements eventually grow with the fourth power of the molecular size (the number of electrons). For example, the computation of five times as large molecule requires 5^4, that is, 625 times more computer time. This rapid increase of time requirement practically excludes proteins and other large molecules from conventional quantum chemistry studies: such computations could require decades of computer time on the fastest computers, consequently, the conventional computational methods on the *ab initio* level are not applicable for truly large molecules.

It is possible, however, to circumvent these computational difficulties, using some of the more recently introduced methodologies involving additive, fuzzy electron density fragments (the AFDF approach [16-22]), where the electron density of the macromolecule is represented either numerically [16-18], or by a macromolecular density matrix P [19-22]. These methods are discussed briefly in the next section.

Additive, Fuzzy Electron Density Fragmentation (AFDF) methods for the computation of macromolecular electron densities and other molecular properties

Recently, a family of new techniques has been introduced based on the additive, fuzzy density fragmentation (AFDF) principle of molecular electron densities. The original version of these AFDF approaches, motivated by the early atomic charge computation scheme of Mulliken, and called the Mulliken-Mezey approach [19], is the basis of the numerical Molecular Electron Density Loge Assembler (MEDLA) method of Walker and Mezey [16-18] and the Adjustable Density Matrix Assembler method of Mezey [19-21]. The AFDF approach has been applied for a variety of large molecules, including proteins [16-18,21], and is also extended to the computation of macromolecular forces [22], that is, to the computation of the forces acting on individual nuclei within a macromolecule controlling both local vibrations and conformational motions, such as protein folding [22].

The general description of the essence of the AFDF scheme can be given using the concept of membership functions of nuclei within various molecular fragments. For this purpose, the family of all nuclei of the given molecule is subdivided into m mutually exclusive groups,

$$f_1, f_2, \ldots, f_k, \ldots f_m \qquad (3)$$

each such group f_k of atomic nuclei serving as a set of „anchor points“ for a fuzzy electron density $\rho^k(r)$ contribution of the corresponding fuzzy fragment F_k, one electron density function of a fragment for each of the m groups of nuclei:

$$\rho^1(r), \rho^2(r), \ldots, \rho^k(r), \ldots \rho^m(r) \qquad (4)$$

and

$$F_1, F_2, \ldots, F_k, \ldots F_m \qquad (5)$$

respectively.

If $m_k(i)$ denotes the membership function of atomic orbital $\varphi_i(r)$ in the set of orbitals centered on a nucleus of nuclear subfamily f_k of electron density fragment F_k, where

$$m_k(i) = \begin{cases} 1 & \text{if the atomic orbital } \varphi_i(r) \text{ is centered on any one of the nuclei of subfamily } f_k, \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

then the n × n AFDF density matrix P^k for the k-th electron density fragment F_k of the Mulliken-Mezey fuzzy fragmentation method is defined in terms of the matrix elements P^k_{ij} as

$$P^k_{ij} = 0.5\,[m_k(i) + m_k(j)]\,P_{ij} \quad (7)$$

More general fragmentation schemes have also been proposed [19] that can be expressed in terms of the membership function of Eq. (6) as

$$P^k_{ij} = [m_k(i)\,w_{ij} + m_k(j)\,w_{ji}]\,P_{ij} \quad (8)$$

In the general expression of equation (8), the weighting factors w_{ij} and w_{ji} are constrained:

$$w_{ij} + w_{ji} = 1 \quad (9)$$

This condition is required for additivity.

In the case of the original and simplest Mulliken-Mezey AFDF scheme, one takes

$$w_{ij} = w_{ji} = 0.5 \quad (10)$$

The AFDF scheme, briefly outlined above, provided the first method that could be applied successfully for the calculation of *ab initio* quality electron densities of natural products including taxol, and macromolecules such as proteins [16-22].

A theorem on the information content of molecular fragments and the „holographic“ property of molecular electron density clouds

In molecular modeling, the information content of various molecular models is an important, limiting aspect. Clearly, very little useful result can be hoped from an oversimplified model of poor information content. The molecular formula, and even the stereochemical structural formula of molecules can be characterized by a very limited set of numbers (for example, by 3N nuclear coordinated for a molecule of N nuclei), and it is, perhaps, too optimistic to hope that such a small set of numbers could efficiently convey the very complex nature and behavior of molecules involved in pharmaceutical activity. Whereas the nuclear arrangement of molecules provides a very useful insight, and in a complex manner implies strong constraints on the electron density, nevertheless, it is clear that it is the molecular electron density cloud that provides the complete information.

However, one may approach the problem of molecular information from a different perspective, and there is a valid question to raise: how much information is present in the electron density of local functional groups, or in more general terms, can one deduce conclusions concerning the complete molecule if one studies only a local molecular moiety, such as a molecular fragment?

A simple and perhaps surprising answer to this question is provided by a recently proven theorem, the Holographic Electron Density Theorem [27]. Some of the consequences of this theorem have fairly fundamental implications for molecular modeling [27-30]. According to this theorem, any small, positive volume local region of a complete, *boundaryless* electron density cloud of a molecule (in a nondegenerate electronic ground state) contains the complete information about the entire electron density, hence, any such local region of the electron density cloud completely determines all molecular properties [27]. This theorem represents an improvement on the Hohenberg-Kohn theorem [13], since there is no need for the complete electron density to determine the molecular energy and other properties, already a small piece of the electron density cloud is sufficient [27].

In other words, molecular fragments, such as individual functional groups are always influenced in a significant way by the rest of the molecule, to such an extent that the induced local changes in the fragment are fully characteristic to the molecule they belong, and the entire molecular information is contained within the fragment!

This result has fundamental implications in the context of molecular modeling involving a series of molecules with similar pharmacological activity. The holographic electron density theorem provides assurances that by studying local molecular regions, even regions that are perhaps not directly involved in a major way with the given type of pharmacological activity, it should be possible, at least in principle, to find correlations between local features of the electron density clouds and the experimentally determined descriptors of the pharmacologically relevant biochemical processes [28,30].

Shape analysis and similarity measures of molecular electron density clouds

Molecular similarity is a fundamental concept of central role in computer aided drug design [5-11,31-50]. The analysis of global and local shape of the electron density clouds can reveal important clues concerning a variety of molecular properties. Specifically, the global and local similarity of molecules can be expressed and numerically characterized by precisely defined shape-similarity measures based on the mathematical representation of computed molecular electron densities [7-9,11].

The detailed computer representations of electronic charge densities, now available for both small and large molecules, provide a systematic framework for modeling and understanding molecular behavior. These computational methods serve as a set of versatile tools for the determination of correlations between well-defined electron density shape features obtained by computations and the experimentally measured

levels of some specified biochemical activity. If such correlations are determined and verified, then electron density shape analysis can be used not only in an interpretative way but also in a predictive sense. By computing and analyzing the relevant shape features of the electron densities of a series of potentially active molecules, even if these molecules have never been synthesized, one is able to use such correlations for estimating their levels of biochemical activity.

Whereas such theoretical predictions are certainly not expected to replace eventual experiments, and usually serve only as guidelines, nevertheless, by identifying promising candidates within the series of new molecules, and also by identifying those molecules which are very unlikely to show significant levels of activities, the electron density based screening is a powerful selection tool. Indeed, the number of molecules that are suggested for the eventual, often rather expensive synthesis and subsequent experimental tests can be reduced substantially.

Among the molecular shape analysis techniques, those relying on the molecular electron density are of special significance. The formal description of the topological shape group methods (SGMs), as applied to the shape characterization and shape similarity analysis of molecular electron density clouds has been the subject of several detailed reviews [11,44-47]. Here we shall describe only a short summary of the methodology.

The starting point of such analysis is the computation of electron densities of the given molecules for a family of relevant nuclear arrangements. If for one such conformation the nuclear arrangement (or, as it is often called, the nuclear configuration) of a molecule M is denoted by K, and if an electron density threshold a is specified, then a *molecular isodensity contour*, MIDCO G(K,a) of M is, by definition, the collection of all those points r of the three-dimensional space where the electron density ρ(K,r) of the molecule M of the given conformation K is equal to the selected threshold value a:

$$G(K,a) = \{ r : \rho(K,r) = a \} \quad (11)$$

For each such MIDCO, the associated *density domain* DD(K,a) is defined as the collection of all points r of the three-dimensional space where the electron density ρ(K,r) of the selected molecule M is greater than or equal to the threshold a:

$$DD(K,a) = \{ r : \rho(K,r) = a \} \quad (12)$$

In fact, the MIDCO G(K,a) is the actual boundary surface of the corresponding density domain DD(K,a).

Density domains and MIDCO surfaces provide a simple quantum chemical definition for local molecular moieties usually involved in a given type of chemical reaction. According to this definition, an additive fuzzy electron density fragment (AFDF) associated with a family of nuclei f_k, is by definition a *quantum chemical functional group* [46,48,50] if there exists some density threshold value a such that a MIDCO G(K,a) of the given density threshold separates this nuclear family f_k from the rest of the nuclei of the molecule. One may regard the regions within this MIDCO G(K,a) of functional group electron density as the extent of "limited autonomy" of the given functional group within the molecule M.

Such a limited autonomy and separate identity of functional groups within molecules are well reflected in this definition, and the same principle of MIDCO analysis can be applied to the limited autonomy and identity of two molecules placed next to each other. If two different molecules are placed within a short distance of one another, then the existence of some MIDCOs separating the nuclei of the two molecules indicate the autonomy and separate identity of these molecules. When comparing molecule pairs in close proximity and functional groups within a single molecule, an important difference is the actual range of values for the density thresholds along the relevant MIDCOs, which are usually much smaller for two separate molecules.

One advantage of this quantum chemical definition of functional groups is the fact that the objects so defined are fully determined by the electron density itself, and the definition does not involve unreliable chemical intuition or individual bias.

The actual values of the thresholds a, which correspond to MIDCOs of functional groups, usually fall within a limited range of electron density. This range is called the "Functional Group Range", and several properties associated with this density range have been studied [11].

The essence of the topological shape analysis method of molecular electron densities, called the Shape Group Method, SGM, is the evaluation of local curvature properties of MIDCOs in order to characterize and classify shape [11]. In fact, the Shape Group Method is a general topological shape analysis technique of any almost everywhere continuously differentiable three-dimensional function; in the case of molecules, SGM is applied to complete, three-dimensional fuzzy molecular electron densities.

Whereas the general SGM method involves some of the special methods of topology [11], the main ideas of the method can be described in simple terms. In the application of SGM to molecular electron densities ρ(K,r), the local curvatures of a range of MIDCOs G(a) are compared to a range of reference curvatures b.

The algorithm of topological shape description using the simplest SGM approach consists of several steps:

(i) A range of electron density thresholds a and a range of reference curvature values b are selected.

(ii) 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.

In order to generate this partitioning, 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 matrices) at each point r to the reference curvature b. Each point r of G(K,a) is assigned to either a $D_0(b)$, or a $D_1(b)$, or a $D_2(b)$ curvature domain, if none, or

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 of values are considered.

Note that the three types of local curvature domains, $D_0(b)$, $D_1(b)$ or $D_2(b)$, indicate whether the MIDCO G(K,a) is convex, concave, or of the saddle type, respectively, relative to the actual curvature b.

(iii) The step of curvature domain identification is followed by a truncation of MIDCO surfaces for each (a,b) pair of parameters 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). Note that a truncated surface $G(K,a,\mu)$ is obtained for each (a,b) pair.

An important aspect of the method is the fact that 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. (This holds in all except some degenerate cases.)

(iv) The next step is the determination of the *shape groups* of the molecular electron density, that is, the zero-, one-, and two-dimensional algebraic homology groups of the truncated MIDCO surfaces. In molecular electron density analysis in three dimensions based on two-dimensional MIDCO surfaces, there is one family of shape groups for each of these dimensions of zero, one, and two, that is, there are three types of shape groups, one for each of the dimensions zero, one, and two. The zero-, one-, and two-dimensional shape groups are denoted by H^0_μ (a,b), H^1_μ (a,b), and H^2_μ (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. One should note that within each topological equivalence class of these surfaces, the shape groups are topological invariants, implying that these groups convey some essential information about the topology of molecular shape.

(v) In the next step simple, numerical descriptors of the Shape Groups are determined. The shape groups themselves can be characterized by numbers and among these numbers the Betti numbers are of special importance.

The zero-, one-, and two-dimensional Betti numbers are the ranks of the zero-, one-, and two-dimensional homology groups, H^0_μ (a,b), H^1_μ (a,b), and H^2_μ (a,b), respectively. The Betti numbers associated with the shape groups are denoted by b^0_μ (a,b), b^1_μ (a,b), and b^2_μ (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. The Betti numbers generate a set of numerical shape descriptors for the entire range of MIDCOs G(K,a) of the given molecule M (of a specific nuclear conformation K).

(vi) The final step is the construction of a numerical shape code from the computed Betti numbers.

In a typical application of the Shape Groups, the distribution of various values of Betti numbers b^p_μ (a,b) as a function of the density threshold a and curvature parameter b is described by various (a,b)-maps. In most practical applications, these maps can be formulated as matrices: the shape matrices $M^{(a,b)}$ are discretized versions of (a,b)-maps corresponding to the cases where only a finite number of thresholds are considered, for example, if n_a density threshold values a and n_b reference curvature values b are used.

The shape matrices $M^{(a,b)}$ can be regarded as numerical shape codes for the molecules, where the matrix elements can be listed either in the original matrix form or as a vector. Since the total number of elements in the shape code matrix $M^{(a,b)}$ is

$$t = n_a n_b \tag{13}$$

the corresponding shape vector is also t-dimensional.

By applying the same Shape Group Method to fuzzy molecular electron density fragments, to the AFDF components of molecular electron densities, the local shape analysis can reveal important trends which often correlate with biochemical activity. An essential aspect of the SGM and related approaches is the fact that they are equally applicable to the fuzzy electron density clouds of complete molecules and to local, fuzzy electron density fragments, hence these methods are suitable for ”zooming in” to the biochemically relevant local regions of molecules.

The shape codes, once determined, can be used for shape comparisons between molecules, or between molecular fragments. A new comparison does not require the recalculation of the shape groups of a molecule or molecular fragment already determined; for each new comparison only the shape codes are used and compared, that results in considerable savings of computer time. The shape codes of molecules and molecular fragments can be stored in various data banks and re-used when needed.

The comparison of shape codes can be formalized in terms of numerical similarity measures.

For any two molecules or molecular fragments A and B, the quantity $m[M^{(a,b),A}, M^{(a,b),B}]$ is defined as the number of matches between corresponding elements in the two shape code matrices $M^{(a,b),A}$ and $M^{(a,b),B}$, of objects A and B, respectively.

The shape-similarity measure s(A,B) defined as

$$s(A,B) = m[M^{(a,b),A}, M^{(a,b),B}] / t \tag{14}$$

expresses the similarity of the shapes of the two fuzzy electron density clouds A and B. Note that, in most applications the range of parameters a and b stretches over several orders of magnitudes and in such cases a logarithmic representation (log(a),log(b)) is used for the maps of these parameters [11].

It has been shown [11] that as a consequence of the nature of parameters a and b in the shape map representations, local shape complementarity can also be evaluated using essentially the same methodology. If the shape map (a,b) of one object, say object A is centrally inverted, resulting in a shape map (a*,b*), then the formal shape similarity measure

$$s^*(A,B) = m[M^{(a^*,b^*),A}, M^{(a,b),B}] / t \tag{15}$$

is in fact a shape complementary measure between objects A and B. This is a simple consequence of the fact that complementarity involves a match of low electron density of object A with high electron density of object B, furthermore, a negative local curvature of object A with a positive curvature of object B, and *vice versa*. This pairing is ensured by a central inversion of one of the (a,b) maps. Note that, in complementarity analysis based on such inverted shape maps, one no longer uses logarithmic representations.

Applications

The applications of detailed electron density shape analysis and similarity evaluation in the study of biochemical activities of various molecules are numerous. Here we shall provide only a brief summary of these results and some of the literature references involving applications of various relevant topological shape analysis methods [17,18,20,49,51-66].

Shape analysis studies of some macromolecules, including proteins, were the very first to derive high accuracy, quantum chemistry quality electron densities for proteins and for drug molecules such a taxol [17,18,20,55], and for the construction of efficient lipophilicity potentials [49,62]. Similarity approaches and an adaptation of the holographic electron density analysis provided excellent correlations between computed molecular features and experimentally measured optical activities in some chiral compounds [66]. The shape group methods have been applied successfully in a predictive sense to derive toxicity relations in a series of studies [61,63-64], and to analyze some prototype molecules in drug design applications [58-60].

Summary

Modern computer aided drug design approaches can take full advantage of the recent breakthroughs in computational quantum chemistry. Advances in the actual computation of electron densities of both small and large molecules provide the means to determine the source material for the molecular interpretation and various characterizations of molecular shapes. The novel methodologies for the shape analysis of both global and local electron density contributions are especially suitable for the study of correlations between shape features and experimentally measured biochemical activities. These correlations can be used in an interpretative manner if one is interested in studying the mechanistic aspects of drug action (these approaches, however, usually do require some knowledge about the actual receptors), or on a purely correlation basis for predictive purposes (this approach is applicable even if very little or no knowledge is available about the actual biochemical mechanism). In either way, the electron density approach to drug design is a promising field, motivated by the simple fact: the electron density contains all information about a given molecule.

References

1. Martin, Y. C. *Quantitative Drug Design: A Critical Introduction*; Marcel Dekker: New York, 1978.
2. Richards, W. G. *Quantum Pharmacology*; Butterworths: New York, 1983.
3. Franke, R. *Theoretical Drug Design Methods*; Elsevier: Amsterdam, 1984.
4. Dean, P. M. *Molecular Foundations of Drug-Receptor Interaction*; Cambridge University Press: Cambridge, 1987.
5. Carbó, R.; Leyda, L.; Arnau, M. *Int. J. Quantum Chem.* **1980**, *17*, 1185-1197.
6. Hodgkin, E. E.; Richards, W. G. *J. Chem. Soc., Chem. Commun.* **1986**, 1342-1344.
7. Mezey, P. G. *Int. J. Quantum Chem., Quant. Biol. Symp.* **1986**, *12*, 113-122.
8. Mezey, P. G. *J. Comput. Chem.* **1987**, *8*, 462-469.
9. Mezey, P. G. *Int. J. Quantum Chem., Quant. Biol. Symp.* **1987**, *14*, 127-132.
10. Johnson, M. A.; Maggiora, G. M., Eds. *Concepts and Applications of Molecular Similarity;* Wiley & Sons: New York, 1990.
11. Mezey, P. G. *Shape in Chemistry: An Introduction to Molecular Shape and Topology;* VCH Publishers: New York, 1993.
12. Kolossváry, I. *J. Math. Chem.* **1996**, *19*, 393-399.
13. Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864-871.
14. Levy, M. In *Recent Developments and Applications of Modern Density Functional Theory, Theoretical and Computational Chemistry,* Seminario, J. M., Ed.; Elsevier Science: Amsterdam, 1996; Vol 4, pp 3-24.
15. Schleyer, P. v. R.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; Schaefer, H. F., III; Schreiner. P. R., Eds.; *Encyclopedia of Computational Chemistry;* Wiley & Sons: Chichester, UK, 1998.
16. Walker, P. D.; Mezey, P. G. *J. Am. Chem. Soc.* **1994**, *116*, 12022-12032.
17. Walker, P. D.; Mezey, P. G. *Can. J. Chem.* **1994**, *72*, 2531-2536.
18. Walker, P. D.; Mezey, P. G. *J. Math. Chem.* **1995**, *17*, 203-234.
19. Mezey, P. G. *J. Math. Chem.* **1995**, *18*, 141-168.
20. Mezey, P. G. In *Computational Chemistry: Reviews and Current Trends;* Leszczynski, J., Ed., World Scientific: Singapore, 1996; Vol 1, pp 109-137.
21. Mezey, P. G. *Pharmaceutical News* **1997**, *4*, 29-34.
22. Mezey, P. G. *Int. J. Quantum Chem.* **1997**, *63*, 39-48.
23. Smith, V. H., Jr.; Schaefer, H. F., III; Morokuma, K., Eds. *Applied Quantum Chemistry. Proceedings of the Hawaii 1985 Nobel Laureate Symposium on Applied Quantum Chemistry;* D. Reidel: Dordrecht, The Netherlands, 1986.
24. Gaussian90; Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Schlegel, H. B.; Raghavachari, K.; Robb, M. A.; Binkley, J. S.; González, C.; DeFries, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.;

Baker, J.; Martin, R.; Kahn, L. R.; Stewart, J. J. P.; Topiol, S.; Pople, J. A. Gaussian Inc., Pittsburgh, 1990.
25. Lipkowitz, K. B.; Boyd, D. B., Eds. *Reviews in Computational Chemistry,* VCH Publisher: New York, 1990; Vol 1, chapters 1-8.
26. Leszczynski, J., Ed. *Computational Chemistry: Reviews and Current Trends;* World Scientific: Singapore, 1996.
27. Mezey, P. G. *Mol. Phys.* **1999**, *96*, 169-178.
28. Mezey, P. G. *J. Chem. Inf. Comp. Sci.* **1999**, *39*, 224-230.
29. Mezey, P. G. *J. Math. Chem.* **1998**, *23*, 65-84.
30. Mezey, P. G. In *QSARs in Environmental Toxicology - VIII. QSARs for Predicting Endocrine Disruption, Chemical Persistence and Effects,* Walker, J. D., Ed., SETAC Publishers: in press (accepted May 20, 1998).
31. Zachman, C.-D.; Heiden, M.; Schlenkrich, M.; Brickmann, J. *J. Comput. Chem.* **1992**, *13*, 76-84.
32. Zachman, C.-D.; Kast, S. M.; Sariban, A.; Brickmann, J. *J. Comput. Chem.* **1993**, *14*, 1290-1300.
33. Cooper, D. L.; Allan, N. L. In *Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches*; Carbó, R., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, l995; pp 31-55.
34. Sen, K., Ed. *Molecular Similarity, Topics in Current Chemistry;* Springer-Verlag: Heidelberg, 1995; Vol 173.
35. Dean, P. M., Ed. *Molecular Similarity in Drug Design;* Chapman & Hall - Blackie Publishers: Glasgow, 1995.
36. Carbó, R., Ed. *Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995.
37. Besalú, E.; Carbó, R.; Mestres, J.; Solà, M. *Top. Curr. Chem.* **l995**, *173*, 31-62.
38. Carbó-Dorca, R.; Mezey, P. G., Eds. *Advances in Molecular Similarity;* JAI Press: Greenwich, Connecticut, 1996; Vol 1.
39. Bywater, R. P. In *Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches;* Carbó, R., Ed.; Kluwer Academic Publisher: Dordrecht, The Netherlands, l995, pp 113-122.
40. Anzali, S.; Barnickel, G.; Krug, M.; Sadowski, J.; Wagener, M.; Gasteiger, J. In *Neural Networks in QSAR and Drug Design*; Devillers, J., Ed.; Academic Press: London, 1996; pp 209-222.
41. Mezey, P. G. *J. Math. Chem.* **1988**, *2*, 325-346.
42. Mezey, P. G. *J. Math. Chem.* **1988,** *2*, 299-323.
43. Mezey, P. G. *Top. Curr. Chem.,* **1995**, *173*, 63-84.
44. Mezey, P. G. In *Molecular Similarity in Drug Design*; Dean, P. M., Ed.; Chapman & Hall - Blackie Publishers: Glasgow, l995; pp 240-268.
45. Mezey, P. G. *Struct. Chem.* **1995**, *6*, 261-270.
46. Mezey, P. G. *Adv. Quantum Chem.* **1996**, *27*, 163-222.
47. Mezey, P. G. *Adv. Mol. Similarity* **1996**, *1*, 89-120.
48. Mezey, P. G.; Fukui, K.; Arimoto, S.; Taylor, K. *Int. J. Quantum Chem.* **1998**, *66*, 99-105.
49. Du, Q.; Mezey, P. G. *J. Comp.-Aided Mol. Des.* **1998**, *12*, 451-470.
50. Mezey, P. G. *Top. Curr. Chem.* **1999**, *203*, 167-186.
51. Arteca, G. A.; Jammal, V. B.; Mezey, P. G.; Yadav, J. S.; Hermsmeier, M. A.; Gund, T. M. *J. Mol. Graphics* **1988**, *6*, 45-53.
52. Arteca, G. A.; Hernández-Laguna, A.; Rández, J. J.; Smeyers, Y. G.; Mezey, P. G. *J. Comput. Chem.* **1991**, *12*, 705-716.
53. Rozas, I.; Arteca, G. A.; Mezey, P. G. *Int. J. Quantum Chem., Quant. Biol. Symp.* **1991**, *18*, 269-288.
54. Arteca, G. A.; Mezey, P. G. *Chem. Phys.* **1992**, *161*, 1-9.
55. Arteca, G. A.; Mezey, P. G. *Biopolymers* **1992**, *32*, 1609-1621.
56. Luo, X.; Arteca, G. A.; Zhang, C.; Mezey, P. G. *J. Organomet. Chem.* **1993**, *444*, 131-136.
57. Walker, P. D.; Mezey, P. G. *J. Comput. Chem.* **1995**, *16*, 1238-1249.
58. Walker, P. D.; Maggiora, G. M.; Johnson, M. A.; Petke, J. D.; Mezey, P. G. *J. Chem. Inf. Comp. Sci.* **1995,** *35*, 568-578.
59. Walker, P. D.; Mezey, P. G.; Maggiora, G. M.; Johnson, M. A.; Petke, J. D. *J. Comput. Chem.* **1995**, *16*, 1474-1482.
60. Heal, G. A.; Walker, P. D.; Ramek, M.; Mezey, P. G, *Can. J. Chem.* **1996**, *74*, 1660-1670.
61. Mezey, P. G.; Zimpel, Z.; Warburton, P.; Walker, P. D.; Irvine, D. G.; Dixon, D. G.; Greenberg, B. *J. Chem. Inf. Comp. Sci.* **1996**, *36*, 602-611.
62. Du, Q.; Arteca, G. A.; Mezey, P. G. *J. Comp.-Aided Mol. Des.* **1997**, *11*, 503-515.
63. Mezey, P. G.; Zimpel, Z.; Warburton, P.; Walker, P. D.; Irvine, D. G.; Huang, X-D.; Dixon, D. G.; Greenberg, B. M. *Environ. Toxicol. Chem.* **1998**, *17*, 1207-1215.
64. Mezey, P. G. In *Soil Chemistry and Ecosystem Health;* Huang, P.M., Ed.; SSSA Publishers: Pittsburgh, 1998; pp 21-43.
65. Badel-Chagnon, A.; Tuffery, P.; Hazout, S.; Mezey, P. G. *Mol. Eng.* **1999**, *8*, 267-288.
66. Mezey, P. G.; Ponec, R.; Amat, L.; Carbó-Dorca, R. *Enantiomers* **1999**, *4*, 371-378.