A uniqueness theorem of molecular recognition

Paul G. Mezey

*Institute for Advanced Study, Collegium Budapest, Szentháromság u. 2, Budapest, Hungary and
Mathematical Chemistry Research Unit, Department of Chemistry and Department of Mathematics and
Statistics, University of Saskatchewan, Saskatoon, SK, Canada S7N 5C9*
E-mail: mezey@sask.usask.ca

Dedicated to the 80th birthday of Professor Frank Harary

Phrased in terms of electron density deformations due to molecular interactions, an optimality condition, and the fundamental holographic properties of molecular electron densities, it is shown that molecular recognition is necessarily unique. A simple proof is given and the connections of this result with the Duality Principle of Molecular Recognition and related Selectivity Measures for molecular recognition are discussed.

KEY WORDS: supramolecular holographic electron density theorem, uniqueness of molecular interactions, molecular recognition, molecular reactivity, biochemical function and activity

1.   Introduction

In some interpretations of the wide-ranging phenomena of molecular recognition, it is assumed that some fundamental aspects of the theory of pattern recognition are applicable. For example, the pattern recognition methods employed for the identification of hydrogen-bonded structural features and related geometrical features in proteins, themselves primary examples of molecular recognition, may mimic some aspects of the actual recognition processes involved in the molecular world [1,2]. It is well understood that the shapes of molecular electron density clouds are the primary players in molecular recognition, and the mutual shape conditions, especially shape similarity and shape complementarity of molecular electron densities are responsible for the degree and specificity of molecular recognition.

The quantum chemical analysis of molecular shapes, involving the complete electron density clouds in a variety of possible nuclear conformations has a long history. Motivated in part by higher-dimensional shape analysis approaches applied to potential energy hypersurfaces [3], some of the early approaches have involved the shape group methods and related techniques [4–11], based on a differential and algebraic topological study of an infinite sequence of nested molecular isodensity contour surfaces

(MIDCO's). The shape groups are homology groups of curvature-based truncations of complete families of MIDCO surfaces.

Alternative molecular shape analysis methods have been introduced by combining discrete mathematical methods pioneered by Harary as applied to various molecular descriptors [12–14].

The specificity aspect of molecular recognition involves molecular similarity and dissimilarity. Based on the shape of the electron density cloud, various methods developed for the assessment of the degree of molecular similarity, including several approaches to the assessment of the degree of molecular chirality, have been studied from a topological perspective [15–20]. A detailed account of the early approaches is found in reference [21]. A major change of direction has become possible by the introduction of the first practically applicable *ab initio* quality macromolecular quantum chemistry techniques, based on the additive fuzzy density fragmentation (AFDF) approach and the development of explicit macromolecular density matrix methods [22–27]. In particular, the adjustable density matrix assembler (ADMA) technique [26,27] is suitable to generate *ab initio* quality macromolecular density matrices for molecules as large as proteins, leading to the computation of all macromolecular properties accessible if a density matrix and the associated atomic orbital basis set is available. The calculated properties include the forces acting on individual atomic nuclei of the protein; hence, a practical, *ab initio* quality quantum chemistry computational method is available to study the protein folding problem.

These developments have also opened the way to the detailed shape analysis of both global and local ranges of macromolecules [28–48]. It is often, but not always sufficient to focus the analysis on a specific local range of a macromolecule, and a local analysis has become possible without loosing the validity of the quantum chemistry approach.

2.   Information content and the relation between global and local molecular information

Density functional theory has been established as a valid framework for molecular physics [49]. Although many of the approximations involved in the actual computational applications of the theory still involve approaches justified only by their practical success in reproducing many (but not all) physical properties of molecules, the theoretical framework provides intuitively transparent and appealing approaches. One fundamental relation of this theory, the Hohenberg–Kohn theorem [50] states that a non-degenerate ground state electron density of any molecule fully determines the energy (and through the Hamiltonian) all other ground state properties of the molecule. That is, the information about non-degenerate ground state properties of the molecule is fully contained within such electron density.

For artificial molecular models, restricted to a bounded volume of the space and having finite boundaries, Riess and Münch [51] have shown that any positive subvolume of such finite and bounded electron density can be extended uniquely to the density in

the bounded volume with finite boundaries. In a later development, Mezey has established the Holographic Electron Density Property [52–56], valid for real, boundaryless molecules: any positive volume part of the non-degenerate ground state electron density of any molecule is already fully sufficient to determine the energy and all other ground state properties of the molecule. That is, the information about non-degenerate ground state properties of real, boundaryless molecules is fully contained within any positive volume part of the electron density.

Subject to the condition of nondegeneracy for the ground state electron density, it is now established that the information in the part contains all the information about the whole.

3. The Supramolecular Holographic Electron Density Theorem in the context of molecular recognition

Consider a supramolecular object, for example, the interacting pair ED of an enzyme E and a drug molecule D. We shall assume that the electron distributions of all these molecules are characterized by non-degenerate ground states. The entire supramolecular object can be regarded as a single entity. Although the very fact that these two molecules are able to form an identifiable single object implies that some degree of mutual recognition manifested in the interaction of the two molecules takes place, nevertheless, it is reasonable to refer to the new supramolecular object as a combination of two, originally independent molecules. The recognition process itself can be regarded as a change from the non-interactive states of the two molecules to the interacting supramolecular entity. Consequently, the information about the recognition itself must be contained in the change of the electron density as the independent molecules combine and form a single supramolecular object.

Consider a nonzero volume part P' of the electron density of independent molecule E, for example, a spherical volume about a specific nucleus X of molecule E. According to the holographic electron density theorem, this volume P' contains all the information about independent molecule E. Consider now the same nonzero volume part P (for example, the spherical volume of the same radius about the same nucleus X) in the supramolecular object ED. Although this volume was originally specified for the independent molecule E, nevertheless, according to the Holographic Electron Density theorem as applied to the entire supramolecular object, this volume P now contains all the information about the supramolecular object ED. This result, although rather obvious based on the original holographic electron density theorem, is of special significance and it appears worthwhile to emphasize it as a fundamental aspect of supramolecular chemistry:

The Supramolecular Holographic Electron Density Theorem (The Supramolecular Holographic Electron Density Theorem). If P is a nonzero volume within a molecular component E in a supramolecular assembly ED, where the electron density of the supramolecular object ED is characterized by a non-degenerate ground state, than the

electron density in this volume P contains all information about the entire supramolecular object ED, specifically, all information about all other molecular components D as they occur within the supramolecular object ED.

4.   The duality principle of molecular recognition

It is rather evident that if a molecule A recognizes another molecule B, for example, by temporarily forming a complex, or by undergoing a specific chemical reaction, then the interaction process involved in the recognition affects both molecules, although possibly to a different degree. The actual recognition is a process characteristic to the given *pair* of molecules. Whereas the roles of the two molecules are seldom symmetric, nevertheless, the recognition is mutual, even if the associated changes in the two molecules and the degrees of selectivities concerning the interactions with this and other potential partners are different. The process of molecular recognition is characterized by a (possibly asymmetric) *duality,* where the roles of the recognizer molecule A and the recognized molecule B can be interchanged.

The changes involved in the recognition process can be analyzed using the electron density shape analysis methods.

Consider the following electron densities:

ρ_A the electron density of independent molecule A;

ρ_B the electron density of independent molecule B;

ρ_{AB} the electron density of interacting molecule pair AB;

$\rho_{A(AB)}$ the electron density of fragment A within interacting molecule pair AB (as obtained by the AFDF density fragmentation process applied to molecule pair AB);

$\rho_{B(AB)}$ the electron density of fragment B within interacting molecule pair AB (as obtained by the AFDF density fragmentation process applied to molecule pair AB).

By carrying out a shape group analysis on these densities and calculating their shape similarities, the similarities $s(\rho_A, \rho_{A(AB)})$ and $s(\rho_B, \rho_{B(AB)})$ are of special importance.

If

$$s(\rho_A, \rho_{A(AB)}) < s(\rho_B, \rho_{B(AB)}) \tag{1}$$

then molecular component A is affected by a greater degree in the recognition process, since independent molecule A is less similar to the interacting molecule A, when compared to the change in the case of molecule B. (By re-assigning labels A and B, the roles can be reversed.) Nevertheless, these similarities are practically never perfect, that is

$$s(\rho_B, \rho_{B(AB)}) = 1 \tag{2}$$

almost never happens, implying that the recognition process introduces at least some changes in the electron densities of all molecules involved.

For example, the shapes of fluorine atom moieties within molecules show remarkable resistance to change, even if these very atoms may cause relatively large electron

density deformations in the interacting partners. Thus, in such recognition processes, the asymmetry is considerable.

Whereas some asymmetry of recognition as expressed by the inequality (1) is typical, in the case of a self-recognition process between two identical molecules along a symmetric pathway, and in some other, presumably rare instances, perfect duality, as measured by the shape group similarity measure, is possible:

$$s(\rho_A, \rho_{A(AB)}) = s(\rho_B, \rho_{B(AB)}). \tag{3}$$

The aspect of selectivity of recognition can also be treated in terms of shape group similarity measures.

Consider a family **A** of two molecules,

$$\mathbf{A} = \{A_1, A_2\} \tag{4}$$

in the role of recognizer, and a set **B** of k molecules to be recognized:

$$\mathbf{B} = \{B_1, \ldots, B_i, \ldots, B_k\}. \tag{5}$$

With respect to family B, the selectivity of recognition of molecule B_i by molecule A_1 can be characterized by the quantity $t(A_1, B_i)$

$$t(A_1, B_i) = \min_{k, k \neq i} \left\{ \mathrm{abs} \left[s(\rho_{A_1}, \rho_{A(AB_i)}) - s(\rho_{A_1}, \rho_{A(AB_k)}) \right] \right\}. \tag{6}$$

The greater this number, the greater is the smallest difference between the changes of electron densities caused by the interaction between A_1, and B_i, on the one hand, and the next most similar change between A_1, and any of the B_k molecules excluding B_i. That is, the greater this number $t(A_1, B_i)$, the greater the selectivity of molecule A_1 recognizing molecule B_i from the given family B. If

$$t(A_1, B_i) > t(A_2, B_i), \tag{7}$$

then molecule A_1 is more selective in recognizing molecule B_i from the molecular family B than molecule A_2.

5. The Uniqueness Theorem of Molecular Recognition

The recognition responses of a small child are different when he recognizes his mother or the doctor who has given him a shot already, and even in the case of his mother approaching him in full view or surprises him from behind, the recognition response is different. Similarly, the range of responses in the recognition processes of a given molecule with various partner molecules and in various mutual arrangements are also expected to be different. It is well understood that a given molecule can recognize a large number of other molecules, however, what appears to be unique is the response in each different case. Here we shall make a precise statement concerning the uniqueness of molecular recognition in the case of each molecule pair, for each different mutual arrangement.

Since the molecular electron densities have primary roles in molecular recognition, it is natural to characterize the recognition process in terms of electron densities.

Consider an A, B pair of molecules involved in the recognition process and the electron densities ρ_A, ρ_B, ρ_{AB}, $\rho_{A(AB)}$, and $\rho_{B(AB)}$ discussed in the previous section. The difference densities $\Delta\rho_{A(AB)}$, and $\Delta\rho_{B(AB)}$ are defined as

$$\Delta\rho_{A(AB)} = \rho_{A(AB)} - \rho_A, \tag{8}$$

and

$$\Delta\rho_{B(AB)} = \rho_{B(AB)} - \rho_B, \tag{9}$$

respectively. These two quantities can be regarded as the individual electron density responses of molecules A and B to the interaction associated with their mutual recognition process. Whereas the holographic theorem applies to densities ρ_A and ρ_B, it does not directly apply to fuzzy density fragments $\rho_{A(AB)}$ and $\rho_{B(AB)}$. However, it does apply to the supramolecular object AB, hence to density ρ_{AB}. Consequently, the difference densities $\Delta\rho_{AB\backslash A}$, and $\Delta\rho_{AB\backslash B}$, defined as

$$\Delta\rho_{AB\backslash A} = \rho_{AB} - \rho_A, \tag{10}$$

and

$$\Delta\rho_{AB\backslash B} = \rho_{AB} - \rho_B, \tag{11}$$

respectively, also have the holographic property, as long as molecule A is specified for $\Delta\rho_{AB\backslash A}$, and molecule B is specified for $\Delta\rho_{AB\backslash B}$, and an optimality condition is also specified for the relative geometrical placements of the interacting and non-interacting molecules.

With reference to the interacting molecule pair AB, take a positive volume P. Since the holographic theorem applies for the corresponding non-degenerate ground electronic states, the difference densities $\Delta\rho_{AB\backslash A}$ and $\Delta\rho_{AB\backslash B}$, restricted to the volume P, fully determine the complete difference densities $\Delta\rho_{AB\backslash A}$ and $\Delta\rho_{AB\backslash B}$, respectively. Consequently, both the local and the global electron density responses involved in A recognizing B and B recognizing A are unique. This proves the following:

Uniqueness Theorem of Molecular Recognition. Molecular recognition, as monitored by changes of electron densities in any positive volume P, is necessarily unique, characteristic to the given molecule pair with the given mutual arrangement.

6. Conclusions

Whereas the main statement of this study, the Uniqueness Theorem of Molecular Recognition is a plausible result on intuitive grounds, nevertheless, the proof presented here provides justification for regarding the detailed analysis of molecular recognition processes as molecular fingerprinting, sufficient, in principle, for unambiguous identification of molecules.

Acknowledgements

References

[1] W. Kabsch and C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.

[2] W. Kabsch and C. Sander, On the use of sequence homologies to predict protein structure: identical pentapeptides can have completely different conformations, Proc. Nat. Acad. Sci. U.S.A. 81 (1984) 1075–1078.

[3] P.G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987).

[4] P.G. Mezey, Group theory of electrostatic potentials: a tool for quantum chemical drug design, Internat. J. Quantum Chem., Quant. Biol. Symp. 12 (1986) 113–122.

[5] P.G. Mezey, Tying knots around chiral centres: chirality polynomials and conformational invariants for molecules, J. Amer. Chem. Soc. 108 (l986) 3976–3984.

[6] P.G. Mezey, The shape of molecular charge distributions: group theory without symmetry, J. Comput. Chem. 8 (1987) 462–469.

[7] P.G. Mezey, Group theory of shapes of asymmetric biomolecules, Internat. J. Quantum Chem., Quant. Biol. Symp. 14 (1987) 127–132.

[8] P.G. Mezey, Shape group studies of molecular similarity: shape groups and shape graphs of molecular contour surfaces, J. Math. Chem. 2 (1988) 299–323.

[9] P.G. Mezey, Global and local relative convexity and oriented relative convexity; Application to molecular shapes in external fields, J. Math. Chem. 2 (1988) 325–346.

[10] P.G. Mezey, The degree of similarity of three-dimensional bodies; Applications to molecular shapes, J. Math. Chem. 7 (1991) 39–49.

[11] P.G. Mezey, Similarity analysis in two and three dimensions using lattice animals and polycubes, J. Math. Chem. 11 (1992) 27–45.

[12] F. Harary and P.G. Mezey, Graphical shapes: seeing graphs of chemical curves and molecular surfaces, J. Math. Chem. 2 (1988) 377–389.

[13] F. Harary and P.G. Mezey, Similarity and complexity of the shapes of square-cell configurations, Theor. Chim. Acta 79 (1991) 379–387.

[14] F. Harary and P.G. Mezey, Chiral and achiral square-cell configurations and the degree of chirality, in: *New Developments in Molecular Chirality*, ed. P.G. Mezey (Kluwer Academic, Dordrecht, 1991) pp. 241–256.

[15] P.G. Mezey, A global approach to molecular chirality, in: *New Developments in Molecular Chirality*, ed. P.G. Mezey (Kluwer Academic, Dordrecht, 1991) pp. 257–289.

[16] P.G. Mezey, Three-dimensional topological aspects of molecular similarity, in: *Concepts and Applications of Molecular Similarity*, eds. M.A. Johnson and G.M. Maggiora (Wiley, New York, 1990) pp. 321–368.

[17] P.G. Mezey, Molecular surfaces, in: *Reviews in Computational Chemistry*, eds. K.B. Lipkowitz and D.B. Boyd (VCH, New York, l990) chapter 7, pp. 265–294.

[18] P.G. Mezey, Topological shape analysis of chain molecules: an application of the GSTE principle, J. Math. Chem. 12 (1993) 365–373.

[19] P.G. Mezey, Dynamic shape analysis of molecules in restricted domains of a configuration space, J. Math. Chem. 13 (1993) 59–70.

[20] P.G. Mezey, Dynamic shape analysis of biomolecules using topological shape codes, in: *The Role of Computational Models and Theories in Biotechnology*, ed. J. Bertran (Kluwer Academic, Dordrecht, 1992) pp. 83–104.

[21] P.G. Mezey, *Shape in Chemistry: an Introduction to Molecular Shape and Topology* (VCH, New York, 1993).

[22] P.D. Walker and P.G. Mezey, Molecular electron density lego approach to molecule building, J. Amer. Chem. Soc. 115 (1993) 12423–12430.

[23] P.D. Walker and P.G. Mezey, Ab initio quality electron densities for proteins: a MEDLA approach, J. Amer. Chem. Soc. 116 (1994) 12022–12032.

[24] P.D. Walker and P.G. Mezey, Realistic, detailed images of proteins and tertiary structure elements: ab initio quality electron density calculations for Bovine insulin, Canad. J. Chem. 72 (1994) 2531–2536.

[25] P.D. Walker and P.G. Mezey, 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. 17 (1995) 203–234.

[26] P.G. Mezey, Macromolecular density matrices and electron densities with adjustable nuclear geometries, J. Math. Chem. 18 (1995) 141–168.

[27] P.G. Mezey, Quantum similarity measures and Löwdin's transform for approximate density matrices and macromolecular forces, Internat. J. Quantum Chem. 63 (1997) 39–48.

[28] P.G. Mezey, Quantum chemical shape: new density domain relations for the topology of molecular bodies, functional groups, and chemical bonding, Canad. J. Chem. 72 (1994) 928–935 (special issue dedicated to Prof. J.C. Polanyi).

[29] P.G. Mezey, Iterated similarity sequences and shape ID numbers for molecules, J. Chem. Inf. Comput. Sci. 34 (1994) 244–247.

[30] P.G. Mezey, Shape analysis of macromolecular electron densities, Structural Chem. 6 (1995) 261–270.

[31] P.G. Mezey, Molecular similarity measures for assessing reactivity, in: *Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches*, ed. R. Carbó (Kluwer Academic, Dordrecht, The Netherlands, l995) pp. 57–76.

[32] P.G. Mezey, Methods of molecular shape-similarity analysis and topological shape design, in: *Molecular Similarity in Drug Design*, ed. P.M. Dean (Chapman & Hall/Blackie, Glasgow, UK, l995) pp. 241–268.

[33] P.G. Mezey, Density domain bonding topology and molecular similarity measures, in: Topics in Current Chemistry, Vol. 173, Molecular Similarity, ed. K. Sen (Springer, Heidelberg, l995) pp. 63–83.

[34] P.D. Walker, P.G. Mezey, G.M. Maggiora, M.A. Johnson and J.D. Petke, Application of the shape group method to conformational processes: shape and conjugation changes in the conformers of 2-phenyl pyrimidine, J. Comput. Chem. 16 (1995) 1474–1482.

[35] P.D. Walker, G.M. Maggiora, M.A. Johnson, J.D. Petke and P.G. Mezey, Shape group analysis of molecular similarity: shape similarity of six-membered aromatic ring systems, J. Chem. Inf. Comput. Sci. 35 (1995) 568–578.

[36] P.G. Mezey, Local shape analysis of macromolecular electron densities, in: *Computational Chemistry: Reviews and Current Trends*, Vol. 1, ed. J. Leszczynski (World Scientific, Singapore, 1996) pp. 109–137.

[37] P.G. Mezey, Molecular similarity measures of conformational changes and electron density deformations, Advances in Molecular Similarity 1 (1996) 89–120.

[38] Z. Zimpel and P.G. Mezey, A topological analysis of molecular shape and structure, Internat. J. Quantum Chem. 59 (1996) 379–390.

[39] P.G. Mezey, Functional groups in quantum chemistry, Advances in Quantum Chemistry 27 (1996) 163–222.

[40] P.G. Mezey, Descriptors of molecular shape in 3D, in: *From Chemical Topology to Three-Dimensional Geometry*, ed. A.T. Balaban (Plenum, New York, 1997) pp. 25–42.

[41] P.G. Mezey, Fuzzy measures of molecular shape and size, in: *Fuzzy Logic in Chemistry*, ed. D.H. Rouvray (Academic Press, San Diego, 1997) pp. 139–223.

[42] P.G. Mezey, Quantum chemistry of macromolecular shape, Internat. Rev. Phys. Chem. 16 (1997) 361–388.

[43] P.G. Mezey, Shape in quantum chemistry, in: *Conceptual Trends in Quantum Chemistry*, Vol. 3, eds. J.-L. Calais and E.S. Kryachko (Kluwer Academic, Dordrecht, The Netherlands, 1997) pp. 519–550.

[44] P.G. Mezey, Shape analysis, in: *Encyclopedia of Computational Chemistry*, Vol. 4, eds. P.V.R. Schleyer, N.L. Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F. Schaefer III and P.R. Schreiner (Wiley, Chichester, UK, 1998) pp. 2582–2589.

[45] P.G. Mezey, Combinatorial aspects of biomolecular shape analysis, Bolyai Soc. Math. Stud. 7 (1999) 323–332.

[46] P.G. Mezey, K. Fukui, S. Arimoto and K.Taylor, Polyhedral shapes of functional group distributions in biomolecules and related similarity measures, Internat. J. Quantum Chem. 66 (1998) 99–105.

[47] P.G. Mezey, Molecular similarity and host-guest interactions, Theor. Comput. Chem. 6 (1999) 593–612, in: *Pauling's Legacy: Modern Modelling of the Chemical Bond*, eds. Z. Maksic and W.J. Orville-Thomas (Elsevier Science, Amsterdam, The Netherlands, 1999) chapter 23, pp. 593–612.

[48] P.G. Mezey, Topological methods of molecular shape analysis: continuum models and discretization, in: DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 51 (2000) pp. 267–278.

[49] M. Levy, Elementary concepts in density functional theory, in: *Recent Developments and Applications of Modern Density Functional Theory*, Theoretical and Computational Chemistry, Vol. 4, ed. J.M. Seminario (Elsevier Science, Amsterdam, 1996) pp. 3–24.

[50] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, Phys. Rev. 136 (1964) B864–B865.

[51] J. Riess and W. Münch, The theorem of Hohenberg and Kohn for subdomains of a quantum system, Theor. Chim. Acta 58 (1981) 295–300.

[52] P.G. Mezey, Generalized chirality and symmetry deficiency, J. Math. Chem. 23 (1998) 65–84.

[53] P.G. Mezey, The holographic electron density theorem and quantum similarity measures, Mol. Phys. 96 (1999) 169–178.

[54] P.G. Mezey, Holographic electron density shape theorem and its role in drug design and toxicological risk assessment, J. Chem. Inf. Comp. Sci. 39 (1999) 224–230.

[55] P.G. Mezey, The holographic principle for latent molecular properties, J. Math. Chem. 30 (2001) 299–303.

[56] P.G. Mezey, The holographic electron density theorem and some of its consequences, in: *Computational Chemistry Approaches to Molecular Similarity*, ed. R. Carbó-Dorca (Kluwer Academic/Plenum, New York) in press (accepted May 29, 2000).