ORIGINAL PAPER

Molecular fragment shape variation index for functional groups and the holographic properties of electron density

Paul G. Mezey

Received: 17 October 2011 / Accepted: 20 October 2011 / Published online: 4 November 2011 © Springer Science+Business Media, LLC 2011

Abstract For a specific functional group, considered as a molecular fragment, the rest of the molecule produces a range of interactions which influence various properties of the functional group. Considering a family of molecules with the "same" functional group, the range of variations in properties determines the range of chemical reactivity of the functional group, and a similar conclusion is valid for more general molecular fragments. By the application of conventional as well as more advanced indices of fragment properties, including local electron density shape characterization, various shape variation indices can be introduced for fragments, and their relations to the holographic properties of electron densities can be examined.

Keywords Molecular fragments · Molecular similarity · Fragment similarity · Fragment shape variations · Holographic electron density theorem · Transition structures · Electronic excited states · Reactivity · Conformational variations

1 Introduction

The theoretical interpretation and modeling of chemical reactivity have been important goals already in the early days of the post-Pauling era of Quantum Chemistry, and this

P. G. Mezey (🖂)

P. G. Mezey

Department of Chemistry, Scientific Modeling and Simulation Laboratory (SMSL), Memorial University of Newfoundland, 283 Prince Philip Drive, St. John's, NL A1B 3X7, Canada e-mail: paul.mezey@gmail.com

Department of Physics and Physical Oceanography, Memorial University of Newfoundland, 283 Prince Philip Drive, St. John's, NL A1B 3X7, Canada

field continues to be the focal point of current research activity [1-5]. Besides traditional reactivity approaches and reaction dynamics studies, explicit considerations of large regions of potential energy surfaces [6-12], as well as the study of local and fragment electron densities [13-34] have provided clues to some questions of reactivity, as well as to novel methods for the construction of macromolecular representations exploiting fuzzy fragment densities. In particular, the study of quantum chemical functional groups [22] has provided links between earlier and more advanced approaches to the interpretation of some chemical interactions relevant to reactivity.

A quantum chemical functional group has been defined [22] as a family of nuclei from a given molecule together with the associated fuzzy electron density fragment obtained by the fragment density matrix approach proposed in refs. [21,24], if in the complete molecule there exists a molecular isodensity contour (MIDCO), such that this contour separates the given nuclear family from the rest of the nuclei of the molecule. A MIDCO has been defined as

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

$$\rho(\mathbf{r}) = \rho(\mathbf{r}, \mathbf{A}, \mathbf{K}),\tag{2}$$

where $\rho(\mathbf{r})$ is the short notation for the electronic density $\rho(\mathbf{r}, A, K)$ of molecule A with nuclear configuration K. Note that $\rho(\mathbf{r}, A, K)$ is a function of position vector \mathbf{r} , and in Eq. (1) a is the density threshold value for the isocontour.

The presence of such a MIDCO, that separates the given nuclear family from the rest of the nuclei of the molecule, indicates some degree of local "autonomy" of this nuclear family within the molecule. A similar situation, with more obvious "autonomy", is more clearly recognizable for the complete family of all nuclei of one molecule placed in the proximity of some other molecules only weakly interacting with the first molecule. Obviously, the nuclei of the first molecule do possess some degree of autonomy in relation to any other neighbor molecule, also indicated by the fact that there must exist some MIDCO of the combined system that separates the nuclei of the first molecule from the nuclei of the second molecule. Returning now to the problem of a functional group within a single molecule, indeed, for most functional groups regarded as such by organic chemists, the above condition of MIDCO separation is fulfilled even within a single molecule.

It is clear, however, that this functional group "autonomy" is limited. In fact, one such limitation follows from a consequence of the Hohenberg–Kohn Theorem [35] of density functional theory, and from an extension of the closed and bounded set results of Riess and Munch [36] to open and boundaryless molecules. The actual fragment result, applicable to fragments of real, boundaryless molecules is the Holographic Electron Density Theorem [37], stating that in a non-degenerate ground state, any positive volume part of the molecular electron density contains the complete information about the entire molecule. This implies that no two different molecules can have exactly identical local electron densities for their functional groups, that is, complete "autonomy", without any influence on the functional group electron density from the rest of the molecule is impossible. In fact, more is true: the actual differences in the local electron densities of functional groups completely determine the rest of the

molecule, that is, the affiliation of each and every functional group. Functional groups of the same chemical formula are all different if they belong to different molecules, and their differences fully distinguish the complete molecules they belong.

This observation raises the following question: is it possible to provide quantitative indices for these influences on functional groups? Whereas electron density properties appear as the natural tools to answer this question, following an earlier study of molecular fragment "Site Dominance Indices" [38], we can apply similar ideas to define "fragment shape persistence indices" and more importantly, "fragment shape variation indices" in terms of general similarity measures. Although such general similarity measures are useful and are often easier to connect to traditional chemistry concepts, nevertheless, in some of the applications we may return to the specific case of electron density shape quantification using Shape Groups [23,39– 48]. Both the general and the electron density shape similarity measures appear useful, and are expected to become tools which fit well within the tool-kits of *density functional theory* and *quantum similarity measures* of Carbo and related approaches [49–68].

The above problems also have strong connections to latent properties of molecules, that is, to properties not actually exhibited, but triggered by some interaction (which actually changes the molecule); after the interaction has taken place, this formerly only latent property also becomes one that is exhibited, now by the new, modified system. In fact, in any molecular process, the changes can be interpreted in terms of latent and exhibited properties: as the molecular process is completed, some latent properties become exhibited, and (with reference to the inverse process) some exhibited properties become latent. In this context, the Holographic Electron Density Theorem for Latent Molecular Properties becomes relevant [69].

2 Molecular fragment and functional group shape variation indices

Although for detailed shape characterization the so-called Shape Group Methods [39–46] (based on the algebraic-topological homology group approach applied to various curvature-based truncations of molecular isodensity contour surfaces) have provided good correlations between electron density shape and molecular activity [39–46], in some alternative applications simpler approaches are still useful, hence, the following discussion on local shape variation indices will be presented in a general way, applicable to several types of shape descriptors.

Following the notations of ref. [38], where the concept of molecular fragment "Site Dominance Index" has been introduced, we are going to discuss the question, how much shape variation can occur in a functional group or a molecular fragment of a given stoichiometry and bond-connectivity, when the effects relative to the shape changes of the surrounding molecular parts are considered? In fact, these indices may be viewed as addressing the following question: compared to the global shape changes? The actual formulation will ensure the fulfillment of a normalization condition, that is, the numerical characterization by this index will provide values falling within the [0,1] interval.

For some molecule A and molecular fragment F of molecule A, we shall consider some shape descriptors, and distinguish the global shape descriptor gsh(A), applicable for any molecule A, and an associated local shape descriptor lsh(F), applicable for any fragment F of any molecule A, where we also assume that these two shape descriptors are consistent with one another, that is, if the local molecular fragment F is replaced with the complete molecule A, then one obtains

$$lsh(A) = gsh(A).$$
 (3)

Such shape descriptors may be based on some internal coordinates with reference to the atomic nuclei, or they may be chosen as the shape descriptors of the Shape Group Method, as applied to the molecular electron-density clouds, as long as they fulfill the above requirements.

In addition, we will consider global and local shape-similarity measures, denoted by gshs(A1, A2) and lshs(F1, F2), involving molecules A1 and A2, and their fragments F1 and F2, respectively.

Specifically, based on the above global and local shape descriptors gsh(A) and lsh(A), we consider global and local shape-similarity measures, gshs(A1, A2) and lshs(F1, F2), defined in terms of some appropriate method for comparisons, suitable for the shape descriptors gsh(A) and lsh(A) used. In order to formalize the above, we consider a general comparison function, denoted by comp(gsh(A1), gsh(A2)), leading to the following expressions:

$$gshs(A1, A2) = comp(gsh(A1), gsh(A2)),$$
(4)

and, by analogy

$$lshs(F1, F2) = comp(lsh(F1), lsh(F2)).$$
(5)

Note that, both the global and local shape-similarity measures are restricted to the [0,1] unit interval:

$$0 \le \operatorname{gshs}(A1, A2) \le 1,\tag{6}$$

for any pair of molecules A1 and A2, and

$$0 \le \operatorname{lshs}(F1, F2) \le 1,\tag{7}$$

for any pair of fragments F1 and F2. Note that the value 0 practically never occurs, except in the case of some rather simplistic shape descriptors.

The above general concepts were used earlier [38] to define Site Dominance Indices for molecular fragments, where the emphasis was placed on measuring similarity enhancement by local domains of molecules. Since in the present study we are interested in the range of shape changes of functional groups and more general molecular fragments, we are not going to follow the earlier approach.

Here we deal with a different question that requires somewhat different tools, specifically, we are going to introduce the following fragment shape variation index.

Consider a pair of local fragments F1 and F2 of molecules A1 and A2, where, typically, the local fragments F1 and F2 are functional groups of the same stoichiometry and bond connectivity. The local or *fragment shape variation index* FSVI, denoted as fsvi(F1, F2, A1, A2) is defined as

$$fsvi(F1, F2, A1, A2) = gshs(A1, A2)/[lshs(F1, F2) + gshs(A1, A2)].$$
 (8)

As a consequence of the restrictions on the possible values of gshs(A1, A2) and lshs(F1, F2), specified by inequalities (6) and (7), the possible values of the fragment shape variation index fsvi(F1, F2, A1, A2) are also restricted to the [0,1] interval:

$$0 \le \text{fsvi}(\text{F1}, \text{F2}, \text{A1}, \text{A2}) \le 1,$$
 (9)

for any pair of fragments F1 and F2 of any pair of molecules A1 and A2.

3 Some comments on applications of fragment shape variation indices

The fragment shape variation index fsvi(F1, F2, A1, A2), as expressed with respect to the actual global and local shape-similarity measures gshs(A1, A2) and lshs(F1, F2), and ultimately by the global and local shape descriptors gsh(A) and lsh(A), provides information on the variability of the local shape of the fragment (functional group) F, as represented by F1 and F2, when the two molecules, A1 and A2 are compared. The smaller the local similarity of F1 and F2 within molecules A1 and A2, taken in the context of the global similarities of the molecules A1 and A2, the greater the variability of the local shapes of the fragments (functional group).

In turn, a greater variability value for fsvi(F1, F2, A1, A2) indicates a greater influence of the molecules A1 ad A2, as whole molecules, on the fragment shapes.

In some sense, a large numerical value for the fragment shape variation index fsvi(F1, F2, A1, A2) may indicate a local concentration of shape changes in the local region, that may also indicate a greater change in reactivity.

The actual scaling with respect to the global shape similarity measure ensures that a value of 0.5 for the fragment shape variation index,

$$fsvi(F1, F2, A1, A2) = 0.5$$
 (10)

represents the special case when the local or fragment shape changes and the global, molecular shape changes are characterized by equal measures. Note, however, that this occurrence is determined by the actual shape similarity measures used, for example, a shape similarity measure based on some simple, bond length or bond angle internal coordinates may provide the number 0.5, indicating a "tie" between the extents of local and global shape changes, whereas using some more sophsticaded and detailed shape measures, such as those based on the Shape Groul methods, a different numerical value can be obtained, and a formal "tie" of value 0.5 for the fragment shape

variation index may be associated with a different pair of molecules. Consequently, when quoting such fragment shape variation index values, the actual shape similarity measures used must also be specified.

In some forthcoming studies we shall attempt to provide a direct connection between electron density predictions based on the Holographic Electron Density Theorem and the numerical indices for such fsvi(F1, F2, A1, A2) results for a family of simple molecules with common functional groups.

4 Summary

Molecular fragment shape variation indices have been introduced with the purpose of providing simple, numerical indications on the differences in the influence of the rest of the molecule on the shape of a given type of molecular fragment, typically, functional group. The connections of this approach to the Holographic Electron Density Theorem, specifically, to the version of this theorem applicable to latent molecular properties, have been noted.

Acknowledgments This work was supported by a Discovery Research Grant from the Natural Sciences and Engineering Research Council of Canada, and by the Canada Research Chair and Canada Foundation for Innovation projects at the Memorial University of Newfoundland, Canada.

References

- K. Fukui, in *Chemical Reactivity Theory: Its Pragmatism and Beyond*. Pure and Applied Chemistry, vol. 54 (Pergamon Press, London, UK, 1982) pp. 1825–1836,
- 2. A. Toro-Labbé (ed.), *Theoretical Aspects of Chemical Reactivity (Theoretical and Computational Chemistry)*, vol. 19 (Elsevier, Amsterdam, 2007)
- 3. F.L. Pilar, *Elementary Quantum Chemistry* (McGraw-Hill, New York, 1968)
- 4. R.D. Levine, Molecular Reaction Dynamics (Cambridge University Press, Cambridge, 2009)
- 5. F.F. Crim, Reaction dynamics: surrounded by complications. Nat. Chem. 3, 344–345 (2011)
- P.G. Mezey, Reaction topology of excited state potential energy hypersurfaces. Can. J. Chem. 61, 956–961 (1983). (volume dedicated to Prof. H. Gunning)
- P.G. Mezey, The topology of energy hypersurfaces. II. Reaction topology in euclidean spaces. Theor. Chim. Acta 63, 9–33 (1983)
- P.G. Mezey, The metric properties of the reduced nuclear configuration space. Int. J. Quantum Chem. 26, 983–985 (1984)
- 9. P.G. Mezey, The topology of energy hypersurfaces. III. The fundamental group of reaction mechanisms on potential energy hypersurfaces. Theor. Chim. Acta 67, 43–61 (1985)
- P.G. Mezey, The topology of energy hypersurfaces. IV. Generator sets for the fundamental group of reaction mechanism and the complete set of reaction paths. Theor. Chim. Acta 67, 91–113 (1985)
- P.G. Mezey, The topology of energy hypersurfaces. V. Potential defying chemical species: a global analysis of vibrational stabilization and destabilization on potential energy hypersurfaces. Theor. Chim. Acta 67, 115–136 (1985)
- 12. P.G. Mezey, Potential Energy Hypersurfaces (Elsevier, Amsterdam, 1987)
- P.D. Walker, P.G. Mezey, Molecular electron density Lego approach to molecule building. J. Am. Chem. Soc. 115, 12423–12430 (1993)
- P.D. Walker, P.G. Mezey, Ab initio quality electron densities for proteins: a MEDLA approach. J. Am. Chem. Soc. 116, 12022–12032 (1994)
- P.D. Walker, P.G. Mezey, Realistic, detailed images of proteins and tertiary structure elements: ab initio quality electron density calculations for bovine insulin. Can. J. Chem. 72, 2531–2536 (1994)

- P.D. Walker, P.G. Mezey, Towards similarity measures for macromolecular bodies: MEDLA test calculations for substituted benzene systems. J. Comput. Chem. 16, 1238–1249 (1995)
- P.D. Walker, 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, 203–234 (1995)
- P.G. Mezey, Z. Zimpel, P. Warburton, P.D. Walker, D.G. Irvine, D.G. Dixon, B. Greenberg, A highresolution shape-fragment database for toxicological shape analysis of PAHs. J. Chem. Inf. Comput. Sci. 36, 602–611 (1996)
- P.G. Mezey, P.D. Walker, Fuzzy molecular fragments in drug research. Drug Discov. Today (Elsevier Trend J.) 2, 6–11 (1997)
- 20. P.G. Mezey, Computational microscopy: pictures of proteins. Pharm. News 4, 29-34 (1997)
- P.G. Mezey, Macromolecular density matrices and electron densities with adjustable nuclear geometries. J. Math. Chem. 18, 141–168 (1995)
- 22. P.G. Mezey, Functional groups in quantum chemistry. Adv. Quantum Chem. 27, 163-222 (1996)
- P.G. Mezey, in *Local Shape Analysis of Macromolecular Electron Densities*, ed. by J. Leszczynski Computational Chemistry: Reviews and Current Trends (World Scientific Publ., Singapore, 1996), pp. 109–137
- P.G. Mezey, Quantum similarity measures and Löwdin's transform for approximate density Matrices and Macromolecular forces. Int. J. Quantum Chem. 63, 39–48 (1997)
- 25. P.G. Mezey, Quantum chemistry of macromolecular shape. Int. Rev. Phys. Chem. 16, 361-388 (1997)
- P.G. Mezey, A crystallographic structure refinement approach using ab initio quality additive fuzzy density fragments. Adv. Mol. Struct. Res. 4, 115–149 (1998)
- P.G. Mezey, Computational aspects of combinatorial quantum chemistry. J. Comput. Methods Sci. Eng. (JCMSE) 1, 99–106 (2001)
- T.E. Exner, P.G. Mezey, A comparison of nonlinear transformation methods for electron density approximation. J. Phys. Chem. A 106, 5504–5509 (2002)
- T.E. Exner, P.G. Mezey, Ab initio quality electrostatic potentials for proteins: an application of the ADMA approach. J. Phys. Chem. A 106, 11791–11800 (2002)
- T.E. Exner, P.G. Mezey, Ab initio quality properties for macromolecules using the ADMA approach. J. Comput. Chem. 24, 1980–1986 (2003)
- T.E. Exner, P.G. Mezey, The field-adapted ADMA approach: introducing point charges. J. Phys. Chem. 108, 4301–4309 (2004)
- Zs. Szekeres, T.E. Exner, P.G. Mezey, Fuzzy fragment selection strategies, basis set dependence, and HF: DFT comparisons in the applications of the ADMA method of macromolecular quantum chemistry. Int. J. Quantum Chem. 104, 847–860 (2005)
- T.E. Exner, P.G. Mezey, Evaluation of the field-adapted ADMA approach: absolute and relative energies of crambin derivatives. Phys. Chem. Chem. Phys. 7, 4061–4069 (2005)
- Zs. Szekeres, P.G. Mezey, P. Surjan, Diagonalization-free initial guess to SCF calculations for large molecules. Chem. Phys. Lett. 424, 420–424 (2006)
- 35. P. Hohenberg, W. Kohn, Inhomogeneous electron gas. Phys. Rev. 136, B864–B871 (1964)
- J. Riess, W. Münch, The theorem of Hohenberg and Kohn for subdomains of a quantum system. Theor. Chim. Acta 58, 295–300 (1981)
- P.G. Mezey, The holographic electron density theorem and quantum similarity measures. Mol. Phys. 96, 169–178 (1999)
- P.G. Mezey, in *Quantitative Shape: Activity Relations (QShAR), Molecular Shape Analysis, Charge Cloud Holography, and Computational Microscopy*, ed. by J.D. Walker Quantitative Structure-Activity Relationships for Pollution Prevention, Toxicity Screening, Risk Assessment, and Web Applications (SETAC (Society of Environmental Toxicity and Chemistry) Publ., SETAC Press, Fairfax, 2003), pp. 123–136
- P.G. Mezey, Shape in Chemistry: An Introduction to Molecular Shape and Topology, vol. 1 (VCH Publishers, New York, 1993)
- P.G. Mezey, Group theory of electrostatic potentials: a tool for quantum chemical drug design. Int. J. Quantum Chem. Quant. Biol. Symp. 12, 113–122 (1986)
- 41. P.G. Mezey, Tying knots around chiral centres: chirality polynomials and conformational invariants for molecules. J. Am. Chem. Soc. **108**, 3976–3984 (1986)
- 42. P.G. Mezey, The shape of molecular charge distributions: group theory without symmetry. J. Comput. Chem. **8**, 462–469 (1987)

- P.G. Mezey, Group theory of shapes of asymmetric biomolecules. Int. J. Quantum Chem. Quant. Biol. Symp. 14, 127–132 (1987)
- 44. P.G. Mezey, Shape analysis of macromolecular electron densities. Struct. Chem. 6, 261–270 (1955)
- P.G. Mezey, in *Shape Analysis*, ed. by P.v.R. Schleyer, N.L. Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F. Schaefer III, P.R. Schreiner Encyclopedia of Computational Chemistry, Vol. 4 (Wiley, Chichester, 1998), pp. 2582–2589
- 46. P.G. Mezey, Quantum chemistry of macromolecular shape. Int. Rev. Phys. Chem. 16, 361–388 (1997)
- P.G. Mezey, Holographic electron density shape theorem and its role in drug design and toxicological risk assessment. J. Chem. Inf. Comput. Sci. 39, 224–230 (1999)
- A. Jokic, Z. Zimpel, P.M. Huang, P.G. Mezey, Molecular shape analysis of a maillard reaction intermediate. SAR QSAR Environ. Res. 12, 297–307 (2001)
- 49. R. Carbó, L. Leyda, M. Arnau, Int. J. Quantum Chem. 17, 1185 (1980)
- 50. E.E. Hodgkin, W.G. Richards, J. Chem. Soc. Chem. Commun. 1986, 1342 (1986)
- 51. R. Carbó, Ll. Domingo, Int. J. Quantum Chem. 32, 517 (1987)
- 52. S. Leicester, R. Bywater, J.L. Finney, J. Mol. Graph. 6, 104 (1988)
- M.A. Johnson, G.M. Maggiora (eds.), Concepts and Applications of Molecular Similarity (Wiley, New York, 1990)
- 54. R. Ponec, M. Strnad, J. Chem. Inf. Comput. Sci. 32, 693 (1992)
- 55. G. Sello, J. Chem. Inf. Comput. Sci. 32, 713 (1992)
- P.G. Mezey, in *Density Domain Bonding Topology and Molecular Similarity Measures*, ed. by K. Sen Topics in Current Chemistry, Molecular Similarity (Springer, Heidelberg, 1995), pp. 63–83
- 57. B.B. Stefanov, J. Cioslowski, J. Comput. Chem. 16, 1394 (1995)
- P.M. Dean (ed.), *Molecular Similarity in Drug Design*, vol. 173 (Chapman & Hall Blackie Publishers, Glasgow, 1995)
- P.G. Mezey, Molecular similarity measures of conformational changes and electron density deformations. Adv. Mol. Similarity 1, 89–120 (1996)
- P. Constans, L. Amat, R. Carbó-Dorca, Toward a global maximization of the molecular similarity function: Superposition of two molecules. J. Comput. Chem. 18, 826–846 (1997)
- L. Amat, R. Carbó-Dorca, R. Ponec, Molecular quantum similarity measures as an alternative to log P values in QSAR studies. J. Comput. Chem. 19, 1575–1583 (1998)
- R. Ponec, L. Amat, R. Carbó-Dorca, Molecular basis of quantitative structure-properties relationships (QSPR): A quantum similarity approach. J. Comput. Aided Mol. Des. 13, 259–270 (1999)
- L. Amat, E. Besalú, R. Carbó-Dorca, R. Ponec, Identification of active molecular sites using quantum-self-similarity measures. J. Chem. Inf. Comput. Sci. 41, 978–991 (2001)
- A.G. Saliner, D. Robert, X. Gironés, R. Carbó-Dorca, Structure-toxicity relationships of polycyclic aromatic hydrocarbons using molecular quantum similarity. J. Comput. Aided Mol. Des. 15, 67–80 (2001)
- X. Gironés, R. Carbó-Dorca, R. Ponec, Molecular basis of LFER. Modeling of the electronic substituent effect using fragment quantum self-similarity measures. J. Chem. Inf. Comput. Sci. 43, 2033–2038 (2003)
- P. Bultinck, T. Kuppens, X. Gironés, R. Carbó-Dorca, Quantum similarity superposition algorithm (QSSA): a consistent scheme for molecular alignment and molecular similarity based on quantum chemistry. J. Chem. Inf. Comput. Sci. 43, 1143–1150 (2003)
- P. Bultinck, R. Ponec, R. Carbó-Dorca, Aromaticity in linear polyacenes: generalized population analysis and molecular quantum similarity approach. J. Comput. Chem. 28, 152–160 (2007)
- R. Carbó-Dorca, A. Gallegos, in *Quantum Similarity and Quantum Quantitative Structure-Properties Relationships (QQSPR)*, ed. by Robert A. Meyers. Encyclopedia of Complexity and Systems Science (2009), pp. 7422–7480
- P.G. Mezey, The holographic principle for latent molecular properties. J. Math. Chem. 30, 299– 303 (2001)