

## Molecular fragment shape variation index applied to intramolecular interaction studies

Zoltan Antal · Paul G. Mezey

Received: 18 October 2011 / Accepted: 20 October 2011 / Published online: 12 November 2011  
© Springer Science+Business Media, LLC 2011

**Abstract** The fragment shape variation index approach is applied to intramolecular interactions involving C6 aromatic molecular fragments in the special case where the shape-modifying interactions are also caused primarily by other C6 aromatic fragments of the same molecule. This report is a part of a series of studies aimed at the detailed modeling of various components of intramolecular interactions among molecular fragments, including aromatic ring interactions, aromatic ring and non-aromatic conjugated and non-conjugated system interactions, and more general through-space and through-bond interactions. The ultimate purpose of these studies is a better understanding of the electron density shape modifying effects of intramolecular interactions.

**Keywords** Fragment shape variation index · Intramolecular interactions · Holographic electron density theorem · Common structural detail variations · Molecular fragmentation · Molecular similarity · Fragment similarity

---

Z. Antal  
Scientific Modeling and Simulation Laboratory (SMSL), Department of Chemistry,  
Memorial University of Newfoundland, 283 Prince Philip Drive, St. John's,  
NL A1B 3X7, Canada

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

P. G. Mezey  
Institute of Chemistry, Eötvös University of Budapest, Budapest, Hungary

## 1 Introduction

It is customary to associate chemical reactivity with various molecular fragments, and those molecular fragments which are expected to show only limited variations are expected to imply rather similar reactivities, at least in some common types of chemical processes. Reactivity, however, can be analyzed from several perspectives, for the more recent approaches the reader may consult ref. [1] and references therein. In some early theoretical studies on reactivity, especially those where local ranges of molecules have been known to have important roles in determining reactivity, some structural features and effects, called substituent effects have been studied within a common framework [2–6]. It is clear, that the variations in the shape of the electron density must reflect the variations in reactivity, since, according to the Hohenberg–Kohn theorem [7], all molecular properties are reflected in the molecular electron density. Furthermore, local shapes are also informative concerning the whole molecule, as this follows from the Holographic Electron Density Theorem [8]: any positive volume fragment of a molecular electron density cloud (in a non-degenerate ground state) contains the complete molecular information, that is, all information about the complete molecule. Hence, electron density shape analysis methods [9], even if applied to local shapes, can provide valid information about reactivity, as well as substituent effects.

The present study will focus on the electron density fragment shape variation index for actual fragments which are not functional groups in the conventional sense, but structural units which are often carriers of actual functional groups, so their shape variation often may be regarded as secondary, triggered by other groups more directly involved in determining reactivity. Nevertheless, in the actual examples used, involving primarily aromatic rings, a photochemical initiation reaction may have specific role, analogous to the role of conventional functional groups, as this has been discussed in some earlier studies [10–12].

The main purpose of this study is to investigate the shape variations of aromatic C6 ring fragments of molecules, where the influences from the rest of the molecules are also primarily provided by similar C6 aromatic ring structures. These results are expected to serve as a guide in later studies where the electron density shape variations of aromatic C6 systems are caused by far more diverse chemical groups. Those more general investigations will attempt to elucidate, among other questions, the roles of through-space and through-bond interactions, where the through-bond interactions are accomplished by molecular fragments with rather mobile electron density clouds, such as the aromatic C6 systems. In the latter case, the fragment shape variations are likely to have different character than in the present case where only the number and the mutual location of aromatic rings is a source of variations in the shapes of the electron densities of aromatic C6 systems.

When the interactions between aromatic rings are considered, polycyclic aromatic hydrocarbons, PAHs are likely candidates for studies. The experimental information and the range of biochemical effects of PAHs have been studied extensively, and the difficulties represented by conformational variations are very limited; the essentially planar structures of PAHs, with only limited deviations from planarity for a subclass of these structures is an advantage, both for the interpretation of some of the experimental results and in the actual computational studies using quantum chemical

or other molecular modeling methods. On the other hand, important insight can also be obtained if only a single aromatic C6 fragment is present in the molecule, when the local shape-modifying influences have entirely different sources, such as it is in the case of para-substituted styrene derivatives, the subject of a much earlier study on through-space and through-bond interactions, using the earlier computational quantum chemistry methods of somewhat limited power [2–6]. We may expect that some of the earlier conclusions can be refined if the more advanced quantum chemistry methods of today are applied.

In this work some aspects of the first of the two questions will be discussed, involving shape-modifying influences in C6 aromatic fragments if the sources of these influences are also C6 aromatic rings. Such C6 rings are special for several reasons; one aspect important in the present context is the fact that aromatic C6 rings are present in very large number of molecules, and they have a primary role to act as carriers of functional groups, otherwise, the C6 aromatic fragments are often considered rather stable structures, and within some conventional interpretation, their shapes are modified primarily through the influences affecting the pi bond system.

## 2 An application of the molecular fragment shape variation index

The earlier choice for Fragment Shape Variation Index [13], closely related to local Shape Dominance Indices, also referred to as local Shape Persistence Indices [12], has been introduced within a rather general context, applicable with respect to several possible shape descriptors, including simple internal coordinates referring to relative nuclear positions, as well as the more detailed shape descriptors of the Shape Group Methods, as applied to the molecular electron density clouds. In this report the latter methods, the Shape Group methods will be applied in order to provide the input information for the determination of the Fragment Shape Variation Index, and to follow the variations of the electron density shape of aromatic C6 rings, within various PAH molecules.

The steps employed in the introduction of the Shape Dominance Index [12] and the Fragment Shape Deviation Index [13] will not be repeated here, similarly, for the description of the Shape Group Methods the reader is also referred to the literature [9]. Here only the actual expressions for the Fragment Shape Deviation Index are given, in a form suitable for the applications in the present examples of polyaromatic hydrocarbon (PAH) electron densities.

In general, for a molecular electron density of a molecule A with nuclear configuration K the notation

$$\rho(\mathbf{r}) = \rho(\mathbf{r}, A, K) \quad (1)$$

is used, where  $\mathbf{r}$  is the spatial variable.

Using isosurfaces of the molecular electron density, also called Molecular Isodensity Surfaces, MIDCOs, the distributions of various local curvature properties define a family of homology groups of algebraic topology, also called “Shape Groups” within the molecular context. The ranks of these shape groups, the so-called Betti numbers,

generate a numerical shape code for each molecule or for each molecular fragment. These shape codes are given in the form of a matrix of integers  $M(a, b, X)$  for each molecule or fragment  $X$ , and for two species  $X_1$  and  $X_2$ , a direct numerical comparison, in terms of matches and mismatches of the corresponding matrix elements generate a similarity measure, denoted by

$$s_{SG}(X_1, X_2) = m[M(a, b, X_1), M(a, b, X_2)] / t, \quad (2)$$

where  $m[M(a, b, X_1), M(a, b, X_2)]$  is the number of entry matches in the two shape matrices and  $t$  is the total number of entries.

Specifically, for any pair of fragments  $F_1$  and  $F_2$  of any pair of molecules  $A_1$  and  $A_2$ , the possible values of the global and local electron-density shape-similarity measures  $s_{SG}(A_1, A_2)$  and  $s_{SG}(F_1, F_2)$  fall within the  $[0,1]$  interval:

$$0 \leq s_{SG}(A_1, A_2) \leq 1, \quad (3)$$

and

$$0 \leq s_{SG}(F_1, F_2) \leq 1. \quad (4)$$

Note that, in the present context the shape code matrices  $M(a, b, A)$  and  $M(a, b, F)$  take the roles of the general shape descriptors of ref. [13], as the global shape descriptor  $gsh(A)$  and the associated local shape descriptor  $lsh(F)$ , respectively. An important consistency condition of the general treatment is fulfilled: one can verify easily that these two shape descriptors,  $M(a, b, A)$  and  $M(a, b, F)$  are consistent with one another: if the local molecular fragment  $F$  is replaced with the complete molecule  $A$ , then one obtains

$$lsh(A) = gsh(A) = M(a, b, A) = M(a, b, F). \quad (5)$$

Then, applying again the general treatment of ref. [13] to the Shape Group approach, with the general comparison function taken as

$$\text{comp}(gsh(X_1), gsh(X_2)) = s_{SG}(X_1, X_2) = m[M(a, b, X_1), M(a, b, X_2)] / t, \quad (6)$$

for both cases,  $X = A$ , or  $X = F$ , where the subscript in the expression  $s_{SG}(X_1, X_2)$  refers to the Shape Group method.

For any pair of local fragments  $F_1$  and  $F_2$  of molecules  $A_1$  and  $A_2$ , the Shape Group induced local or *fragment shape variation index*  $fsvi_{SG}$ , denoted as  $fsvi_{SG}(F_1, F_2, A_1, A_2)$  is defined as

$$fsvi_{SG}(F_1, F_2, A_1, A_2) = s_{SG}(A_1, A_2) / [s_{SG}(F_1, F_2) + s_{SG}(A_1, A_2)]. \quad (7)$$

As in the case of the general shape descriptors  $gsh(A)$  and  $lsh(A)$ , the constraints on  $s_{SG}(F_1, F_2)$  and  $s_{SG}(A_1, A_2)$  imply that the Shape Group induced local or fragment

shape variation index  $FSVI_{SG}$  is also restricted to the  $[0,1]$  interval, that is, for any pair of fragments  $F_1$  and  $F_2$  of any pair of molecules  $A_1$  and  $A_2$ , the inequalities

$$0 \leq fsvi_{SG}(F_1, F_2, A_1, A_2) \leq 1 \quad (8)$$

hold.

A small local shape similarity of fragments  $F_1$  and  $F_2$  within molecules  $A_1$  and  $A_2$ , when taken in comparison to the global similarities of the molecules  $A_1$  and  $A_2$ , indicates a large shape variation for the fragments, and this is reflected in the fragment shape variation index  $FSVI_{SG}$  as given in terms of the results of the shape group analysis.

The evaluation of the fragment shape variation index for a set of different fragment types in a series of molecules can help in the identification of various molecular regions where the effect of changes in substituents have greater or lesser effect in modifying the influence of various intramolecular interactions.

### 3 Shape group results and fragment shape variation index computed for selected aromatic C6 ring fragments of 16 PAH molecules

A series of 16 PAH molecules, listed in Table 1, were studied earlier [10–12], and had their shape groups and shape code matrices computed, and their numerical similarity measures calculated. The main goal of those studies was to find correlations between their aquatic toxicity with respect to the plant *Lemna gibba*, and good correlations have been obtained between their shape descriptors and experimental toxicities. Those studies revealed some useful information concerning the actual mechanism of the photoinduced toxicity of these molecules within an aquatic environment, in terms of a photochemical sensitization step followed by an actual photochemical reaction.

The fragment similarities have been evaluated with reference to the central C6 ring fragment of the anthracene molecule, taking the most similar fragment form each molecule. Using the relevant raw data from refs. [10–12], reproduced in Table 1, the newly computed fragment shape variation indices of the present study are also listed in Table 1.

In general, if the value of the computed fragment shape variation index is greater than 0.5, then the chosen fragment, in our case the aromatic ring fragment C6 shows greater shape variation than the shape variation of the given molecule containing the aromatic ring fragment C6, that is, the local shape change is greater than the global shape change, when the shapes of the reference molecule anthracene and the actual other PAH molecule are compared. By contrast, a computed fragment shape variation index value smaller than 0.5 indicates that the actual aromatic C6 ring fragment shows smaller shape changes than the complete PAH molecule to which the C6 fragment belongs, when the local and global shapes of the anthracene molecule are used as reference in the comparisons.

The computed fragment shape variation index values indicate that both enhancement and partial suppression of fragment shape variation does occur, when compared to the global shape variation of the host molecules for these aromatic C6 fragments.

**Table 1** Evaluation of 16 PAH molecule C6 ring Fragment Shape Variation Index with respect to the central C6 ring of anthracene (taking the most similar ring in each molecule)

Name of PAH molecule	C6 ring similarity	Molecular similarity	Fragment shape variation index
Anthracene	1.000	1.000	0.500
Benzo( <i>a</i> )anthracene	0.568	0.422	0.426
Fluoranthene	0.386	0.513	0.571
Dibenzo[ <i>a,i</i> ]pyrene	0.425	0.410	0.491
Benzo[ <i>b</i> ]fluorene	0.640	0.436	0.405
Benzo[ <i>b</i> ]anthracene	0.640	0.412	0.392
Benzo[ <i>a</i> ]pyrene	0.471	0.404	0.462
Pyrene	0.435	0.514	0.542
Benzo[ <i>e</i> ]pyrene	0.429	0.409	0.488
Fluorene	0.413	0.663	0.616
Benzo( <i>g, h, i</i> )perylene	0.378	0.406	0.518
Coronene	0.350	0.403	0.535
Phenanthrene	0.358	0.606	0.629
Dibenz[ <i>a, h</i> ]anthracene	0.659	0.418	0.388
Triphenylene	0.342	0.417	0.549
Chrysene	0.355	0.411	0.537

Some general trends appear to emerge.

The largest values of the computed fragment shape variation index were obtained for phenanthrene and fluorene, two of the relatively small PAH molecules, whereas some of the largest PAH molecules show relatively small values for the computed fragment shape variation index. This finding is in good agreement with the notion that in larger aromatic systems the local shape variations are better balanced due to a more extensive delocalization, that is not as prominent in smaller PAH molecules.

It is interesting to note that the smallest fragment shape variation index values are obtained for two of the anthracene derivatives, dibenz[*a, h*]anthracene (0.388) and benzo[*b*]anthracene (0.392), where large fragment similarities, but relatively low global similarities have been obtained. Although the fact that both of these molecules are related to the reference molecule anthracene, may suggest otherwise, this fact, by itself does not imply large global similarities, yet, the closer relations of the bonding structures in such cases may provide good local similarities for the respective C6 fragments. The combinations of these factors produce low fragment shape variation indices, in spite of the fact, that higher global similarity, by itself, is likely to produce higher fragment shape variation indices.

The above findings indicate that the computation of fragment shape variation indices can reveal features of intramolecular interactions which are not obvious if one is concerned with global shapes and local shapes individually, without reference to their interrelations.

## 4 Summary

Molecular fragment shape variation measures have been applied within the context of the electron density shape characterization using the Shape Group methods. Actual computations have been carried out to evaluate the fragment shape variation measures for a set of 16 polyaromatic hydrocarbons, focusing on aromatic C6 fragments, where the shape-modifying influence of other molecular parts is also dominated by other aromatic rings, thereby limiting the types of influences and somewhat simplifying the analysis. These results provide reference and background for further studies, where a wider variety of fragments may produce enhanced shape changes, partly associated with through-space and through-bond interactions.

**Acknowledgements** 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

1. Toro-Labbé Alejandro (ed.), *Theoretical Aspects of Chemical Reactivity (Theoretical and Computational Chemistry)*, vol. 19, Elsevier, Amsterdam (2007)
2. W.F. Reynolds, T.A. Modro, P.G. Mezey, A theoretical investigation of the effect of positively charged substituents on product distribution in electrophilic aromatic substitution; evidence for a dominant field effect of the positive poles. *J. Chem. Soc. Perkin II* **1977**, 1066–1070 (1977)
3. P.G. Mezey, W.F. Reynolds, *Ab initio* calculations on 4-substituted benzoic Acids; a further theoretical investigation into the nature of substituent effects in aromatic derivatives. *Can. J. Chem.* **55**, 1567–1574 (1977)
4. W.F. Reynolds, P.G. Mezey, G.K. Hamer, *Ab initio* calculations on 4-substituted styrenes; a theoretical model for the separation and evaluation of field and resonance substituent parameters. *Can. J. Chem.* **55**, 522–529 (1977)
5. W.F. Reynolds, P.G. Mezey, W.J. Hehre, R.D. Topsom, R.W. Taft, The relationship between substituent effects on energy and on charge from *ab initio* molecular orbital theory. *J. Am. Chem. Soc.* **99**, 5821–5822 (1977)
6. W.F. Reynolds, T.A. Modro, P.G. Mezey, E. Skorupowa, A. Maron, An experimental and theoretical investigation of the unusual substituent effect of the vinyl group. *Can. J. Chem.* **58**, 412–417 (1980)
7. P. Hohenberg, W. Kohn, Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964)
8. P.G. Mezey, The holographic electron density theorem and quantum similarity measures. *Mol. Phys.* **96**, 169–178 (1999)
9. P.G. Mezey, *Shape in Chemistry: An Introduction to Molecular Shape and Topology* (VCH Publishers, New York, 1993)
10. P.G. Mezey, Z. Zimpel, P. Warburton, P.D. Walker, D.G. Irvine, D.G. Dixon, B. Greenberg, A high-resolution shape-fragment database for toxicological shape analysis of PAHs. *J. Chem. Inf. Comp. Sci.* **36**, 602–611 (1996)
11. P.G. Mezey, Z. Zimpel, P. Warburton, P.D. Walker, D.G. Irvine, X.-D. Huang, D.G. Dixon, B. Greenberg, Use of QShAR to model the photoinduced toxicity of PAHs: electron density shape features accurately predict toxicity. *Environ. Toxicol. Chem.* **17**, 1207–1215 (1998)
12. 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, New York, 2003)*, pp. 123–136
13. P.G. Mezey, *Molecular Fragment Shape Variation Index* (to be published)