ORIGINAL PAPER

Fragment shape variation index for periodicity deficiency and gradual changes of internal coordinates along linear polymers

Eva Simon · Paul G. Mezey

Abstract   A special adaptation of the fragment shape variation index approach is applied to a series of rod-like structures obtained by replacing all C–C atom pairs by B–N atom pairs in linear polymers constructed by stacking adamantane units on top of each other. A special feature of these structures is a systematic, monotonic increase of the length of B–N bonds roughly parallel with the axis of the rods when moving from the B-terminus towards the N-terminus of the rods. All other changes in the unit cells of these short polymers are less important and less systematic, consequently, focusing on these B–N bonds of the given structures is especially suitable for the application of the fragment shape variation index approach used for a single type of internal coordinate shape descriptor. The results reveal additional trends which are easy to miss if one is restricted to a simple inspection of the data. The study also confirms that in special cases one may replace the detailed electron density shape descriptors with the far simpler bond length data as formal, but limited shape descriptors, whenever the overall structures and shapes show only minor changes, affecting primarily only one type of the internal coordinates.

E. Simon · 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

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

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

1 Introduction

The effect of intramolecular interactions on common local structural features in molecular families defined by some common properties, such as the presence of a given type of functional group, or a given type of monomeric unit in a family of polymers of various sizes, is often expected to show some systematic trends, characteristic of the changing global properties within the molecular family. These global properties may also correspond to some structural features, and the simpler and more systematic the global changes within the family, the simpler changes are expected for the local structural features as well. Finite length linear polymers which are formally periodic are excellent examples for the study of some systematic trends, since even such polymers only formally can be regarded as strictly periodic, because for finite systems the end-effects necessarily cause some deviations from perfect periodicity. Of course, if strict periodicity for short linear polymers would be present, there could be no trends to study, since then all unit cells would, indeed, show strictly identical features, an ideal case best approximated by the middle part of long, practically "infinitely long" linear polymers.

However, even formal, but less than strict periodicity also offers advantages, since in such cases the changes from formal "unit cell" to the next formal "unit cell" are expected to be small, and it is of some interest to study specific changes, where, indeed, some regularities in the extent of the deviations from strict periodicity are expected.

For such cases the methods of molecular similarity analysis, usually applied to actual series of molecules, can now be also applied within a single molecule: to the series of formal "unit cells", all slightly different, within a single linear polymer of relatively short length.

Molecular similarity analysis has a very extensive literature, initiated by the early works of Carbo and others [1–20], and depending on the focus of the similarity study, the quantum similarity approach, or the Shape Group approach [21–24], or some simpler, specialized approaches [24,25] may be the optimum choice.

The rather general approach of fragment shape variation index [25] was originally intended for applications with respect to the molecular shape codes within the Shape Groups approach [21–24]. These shape codes have been defined in terms of the Betti numbers, the ranks of the molecular shape groups (one-dimensional homology groups of algebraic topology) generated by a curvature-based truncation of the molecular isodensity contours [21], computed by some quantum chemistry method. However, in special cases, such as some minor shape variations within linear polymers, a far simpler approach is possible, circumventing the more involved computational steps required for the determination of the above shape codes based on the molecular electron density.

Besides the actual investigation of the questions of local and global deviations from ideal periodicity, we shall also demonstrate that a very simple, bond length

based approach also fits within the general fragment shape variation index formalism, hence, using the simplest of shape descriptors, such as bond lengths, can also be treated within the general framework designed for far more complex and detailed shape characterization approaches.

Our present goal is to study deviations from exact periodicity, and for this end the selected "vertical" B–N bond lengths in rod-like polymers are taken as exactly equal for an "ideal" reference polymer. In each example we shall use in the present demonstration, the actual reference will be either the longest or the average of such B–N bond length, and in this context the global and local shape changes will be defined in terms of deviations from this reference. We are not going to be concerned with the signs of these deviations; since there is a strict monotonicity of bond length changes along the rods for all calculated "non-ideal" polymers in this sequence, the important additional information is the size of the deviation from periodicity. Note that, for each length of these polymers, there are four different structures, depending on the structure of the N-terminus of the rods, where in the terminal, zig-zagged B_3N_3 rings the three N atoms, with H atoms attached, are the furthest from the center of the molecules. We shall study in detail the four different three-adamantane polymers, that is, the four polymers each with six B_3N_3 rings each imitating the carbon skeleton of cyclohexane, where one finds different possible orientations of the N–H bonds of the N-terminus for the four different such structures. In each molecule there are three "vertical' bond sequences between the B_3N_3 rings as one moves from the N-terminal to the B-terminal, and in principle, we should be concerned with all three. Note, however, that if one denotes the N-terminal orientations of the H atoms as e for equatorial orientation and a for axial orientation, then the four structures are eee, aaa, aea, and eae, and from these the first two, eee and aaa, each has identical vertical bond sequences, whereas in each of the last two, aea and eae, there are two identical and one unique vertical bond sequence. In each case, we are focusing on the vertical sequence showing the largest deviation from periodicity.

Here we shall give only a brief summary of the defining equations of the fragment shape variation index approach, for more detail the reader is referred to the literature references [21–25].

The deviation from periodicity for any pair of subsequent fragments F_i and F_{i+1} along the actual, imperfectly periodic polymeric molecule P, with respect to some ideally periodic reference polymer P_{ideal}, is specified by the fragment similarity measure $s_{pf}(F_i, F_{i+1})$,

$$s_{pf}(F_i, F_{i+1}) = \left[v(P_{ideal}) - v(F_i, F_{i+1}) \right] / v(P_{ideal}), \quad (1)$$

where $v(P_{ideal})$ is the value of some structural parameter, exactly repeated within the ideally periodic model polymer P_{ideal}, and $v(F_i, F_{i+1})$ is the corresponding value in the actual, non-ideal polymer P, with respect to subsequent fragments F_i and F_{i+1}. As expected for the general similarity measures, the possible values fall within the [0,1] interval:

$$1 \leq s_{pf}(F_i, F_{i+1}) \leq 1. \quad (2)$$

The global similarity measure $s_{pg}(P, P_{ideal})$ can also be taken with respect to the ideal, perfectly periodic model polymer P_{ideal}, providing values which fall within the [0,1] interval:

$$1 \le s_{pg}(P, P_{ideal}) \le 1. \tag{3}$$

The simplest choice for $s_{pg}(P, P_{ideal})$ is the average of $s_{pf}(F_i, F_{i+1})$ values.

Following the general treatment for fragment shape variation index of ref. [25], the above $s_{pf}(F_i, F_{i+1})$ and $s_{pg}(P, P_{ideal})$ quantities generate the "deficient periodicity" induced local or *fragment shape variation index* $FSVI_{DP}$, denoted as $fsvi_{DP}(F_i, F_{i+1}, P, P_{ideal})$, and defined as

$$fsvi_{DP}(F_i, F_{i+1}, P, P_{ideal}) = s_{pg}(P, P_{ideal}) / \left[s_{pf}(F_i, F_{i+1}) + s_{pg}(P, P_{ideal}) \right]. \tag{4}$$

In agreement with the properties of the general shape descriptors $gsh(A)$ and $lsh(A)$ discussed in ref. [25], the constraints on $s_{pf}(F_i, F_{i+1})$ and $s_{pg}(P, P_{ideal})$ ensure that the "deficient periodicity" induced fragment shape variation index $FSVI_{DP}$ must have values also restricted to the [0,1] interval:

$$0 \le fsvi_{DP}(F_i, F_{i+1}, P, P_{ideal}) \le 1. \tag{5}$$

The "deficient periodicity" induced fragment shape variation index $FSVI_{DP}$ gives numerical indications for the extent by which various local ranges violate strict periodicity within the polymer P.

2 The case of imperfectly periodic B–N rod structures

In order to illustrate the general ideas of the applications of fragment shape invariance index to study the deficiencies of periodicity in terms comparisons of geometrical parameters of formal unit cells, we use the example of the four $H_3(B_3N_3)_6H_3$ Boron-Nitrogen nanoneedles.

These molecules, together with several related structures, were the subject of a detailed computational study, with emphasis on the evaluation of their stability and optimum geometry [26]. In that computational study [26] both shorter and longer polymers were included, specifically, the series $H_3(B_3N_3)_nH_3$, $n = 2 - 10$. The quantum chemical study involved the DFT computational method at the B3LYP/6-31G(d,p) level, focusing on the internal regions and the peripheral regions of these nanoneedles. For the analysis of some characteristic patterns, an electron density analysis has also been applied, using a series of molecular isodensity contours (MIDCO's). For each member of the $H_3(B_3N_3)_nH_3$ sequence, $n = 2 - 10$, the geometry optimizations have lead to four distinct structures, differing in the alignment of the three terminal N-H bonds. Using the specifications axial (a) and equatorial (e) for these alignments, the four possible combinations are eee, aaa, aea, and eae. For each length of the poly-

mers, $n = 2 - 10$, the four conformations have been geometry optimized, leading to proper energy minima (as verified by having only real vibrational frequencies).

Although the level of geometry optimization of the earlier study was sufficient to verify the stability and main structural features of these polymers, here our needs are somewhat more demanding concerning small deviations from symmetry. In order to avoid the need to deal with small deviations from the expected symmetry, in addition to the small deviations from periodicity, further readjustment of internal coordinates was required. Beyond the geometry optimization presented in reference [26], where no symmetry restriction has been applied, some additional steps were needed for the purposes of the present study. Accordingly, we have performed one more step to bring the molecular geometry results to the necessary symmetry, as required for the evaluation of deviations from periodicity. The resulting, slightly modified bond length values for the four polymers $H_3(B_3N_3)_6H_3$, are listed in Table 1.

When investigating the deviations from periodicity and using the actual realization of the fragment shape variation index introduced above, the following considerations apply for the four molecules of overall formula $H_3(B_3N_3)_6H_3$. The six layers of B_3N_3 units correspond to five "vertical" B–N bond lengths, and to four deviations from periodicity, and these deviations are the very quantities we intend to compare. The local fragment similarity measure in this case is defined with respect to the strictly periodic, ideal polymer, where the maximum of the geometry optimized vertical B–N bond lengths is taken as reference when the deviations from periodicity is computed.

Based on the symmetrized vertical B–N bond length data presented in Table 1, for each consecutive B–N bond pair, the similarity measure with respect to the ideal periodicity are calculated, where the general value v in Eq. (1) is replaced with the bond length values d:

$$s_{pf}(F_i, F_{i+1}) = \left[d(P_{ideal}) - d(F_i, F_{i+1}) \right] / d(P_{ideal}). \qquad (6)$$

Based on these data, and focusing on the fragments with the largest deviations from periodicity, the following fragment similarity measure values $s_{pf}(F_1, F_2)$ are obtained for the four molecules eee, aaa, aea, eae:

$$s_{pf,eee}(F_1, F_2) = 0.980,$$
$$s_{pf,aaa}(F_1, F_2) = 0.996,$$
$$s_{pf,aea}(F_1, F_2) = 0.990,$$
$$s_{pf,eae}(F_1, F_2) = 0.986.$$

Similarly, the four global similarity measure values $s_{pg}(P, P_{ideal})$ are calculated as

$$s_{pg,eee}(P, P_{ideal}) = 0.994,$$
$$s_{pg,aaa}(P, P_{ideal}) = 0.998,$$
$$s_{pg,aea}(P, P_{ideal}) = 0.996,$$
$$s_{pg,eae}(P, P_{ideal}) = 0.995.$$

These data, in turn, generate the following values for the fragment shape variation indices $fsvi_{DP}$ (F_i, F_{i+1}, P, P_{ideal}) defined by Eq. (4):

$$fsvi_{DP} (F_1, F_2, P_{eee}, P_{ideal}) = 0.504$$
$$fsvi_{DP} (F_1, F_2, P_{aaa}, P_{ideal}) = 0.501$$
$$fsvi_{DP} (F_1, F_2, P_{aea}, P_{ideal}) = 0.502$$
$$fsvi_{DP} (F_1, F_2, P_{eae}, P_{ideal}) = 0.502$$

Each of the four calculated fragment shape variation indices for the terminal fragment, for $i = 1$, shows a slight, but clear enhancement of shape variations, as defined in terms of deviations from perfect periodicity. This can also be interpreted as various levels

Table 1 B–N "vertical" bond lengths (Å) between layers of the four conformers, (eee), (aaa), (aea), and (eae), of the polymer $H_3(B_3N_3)_6H_3$, re-optimized for symmetry

(eee) $H_3(B_3N_3)_6H_3$	e	e	e
$d_{BN,1}$ (N–H end)	1.473	1.473	1.473
$d_{BN,2}$	1.443	1.443	1.443
$d_{BN,3}$	1.439	1.439	1.439
$d_{BN,4}$	1.437	1.437	1.437
$d_{BN,5}$ (B–H end)	1.435	1.435	1.435
(aaa) $H_3(B_3N_3)_6H_3$	a	a	a
$d_{BN,1}$ (N–H end)	1.447	1.447	1.447
$d_{BN,2}$	1.441	1.441	1.441
$d_{BN,3}$	1.439	1.439	1.439
$d_{BN,4}$	1.438	1.438	1.438
$d_{BN,5}$ (B–H end)	1.436	1.436	1.436
(aea) $H_3(B_3N_3)_6H_3$	a	e	a
$d_{BN,1}$ (N–H end)	1.457	1.452	1.457
$d_{BN,2}$	1.443	1.440	1.443
$d_{BN,3}$	1.439	1.440	1.439
$d_{BN,4}$	1.438	1.438	1.438
$d_{BN,5}$ (B–H end)	1.436	1.436	1.436
(eae) $H_3(B_3N_3)_6H_3$	e	a	e
$d_{BN,1}$ (N–H end)	1.461	1.467	1.461
$d_{BN,2}$	1.441	1.446	1.441
$d_{BN,3}$	1.440	1.439	1.440
$d_{BN,4}$	1.438	1.438	1.438
$d_{BN,5}$ (B–H end)	1.435	1.436	1.435

Significant deviation from periodicity, in fact, monotonicity of bond length changes is found in all cases

of enhancement as compared to other fragments involved in the ideal, perfectly periodic polymer, however, the enhancement is relatively minor. This representation of the results is also suitable to guard against exaggerated conclusions, clearly showing the relatively minor deviations of the index values from the case of

$$\text{fsvi}_{DP} \left(F_i, \ F_{i+1}, \ P, \ P_{ideal} \right) = 0.500, \tag{7}$$

corresponding to the "neutral", or "no enhancement" possibility.

3 Summary

Molecular fragment shape variation measures have been applied to test the levels of deviations from exact periodicity for the special case of rod-like structures with a systematic, monotonic bond length variation along the rods. The numerical results suggest that some conclusions based on such deviations must be treated with caution, since by formulating the results within a more general framework, the impact of these deviations shows less significance. Nevertheless, the results also demonstrate that simple shape descriptors, such as internal coordinates, can be used for the study of the local and global aspects of intramolecular interactions, without invoking the fundamental properties of electron densities, regarded as the information carrier of molecular properties, both globally [27] and locally [28].

References

1. R. Carbó, L. Leyda, M. Arnau, Int. J. Quantum Chem. **17**, 1185 (1980)
2. E.E. Hodgkin, W.G. Richards, J. Chem. Soc. Chem. Commun. **1986**, 1342 (1986)
3. R. Carbó, Ll Domingo, Int. J. Quantum Chem. **32**, 517 (1987)
4. S. Leicester, R. Bywater, J.L. Finney, J. Mol. Graph. **6**, 104 (1988)
5. M.A. Johnson, G.M. Maggiora (eds.), *Concepts and Applications of Molecular Similarity* (Wiley, New York, 1990)
6. R. Ponec, M. Strnad, J. Chem. Inf. Comput. Sci. **32**, 693 (1992)
7. G. Sello, J. Chem. Inf. Comput. Sci. **32**, 713 (1992)
8. P.G. Mezey (l995) in *Density Domain Bonding Topology and Molecular Similarity Measures*, ed. by K. Sen. Topics in Current Chemistry, Molecular Similarity, vol. 173 (Springer, Heidelberg), pp. 63–83
9. B.B. Stefanov, J. Cioslowski, J. Comput. Chem. **16**, 1394 (1995)
10. P.M. Dean (ed.), *Molecular Similarity in Drug Design* (Chapman & Hall-Blackie Publishers, Glasgow, 1995)
11. P.G. Mezey, Molecular similarity measures of conformational changes and electron density deformations. Adv. Mol. Similarity **1**, 89–120 (1996)
12. 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)
13. 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)
14. 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)

15. 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)
16. 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)
17. 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)
18. 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)
19. 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)
20. 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
21. P.G. Mezey, *Shape in Chemistry: An Introduction to Molecular Shape and Topology* (VCH Publishers, New York, 1993)
22. 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. Comput. Sci. **36**, 602–611 (1996)
23. 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)
24. P.G. Mezey, in *Quantitative Shape: Activity Relations (QShAR), Molecular Shape Analysis, Charge Cloud Holography, and Computational Microscopy*, ed. by Walker J.D. Quantitative Structure-Activity Relationships for Pollution Prevention, Toxicity Screening, Risk Assessment, and Web Applications. SETAC (Society of Environmental Toxicity and Chemistry). Publ., (SETAC Press, 2003), pp. 123–136
25. P.G. Mezey, Molecular fragment shape variation index for functional groups and the holographic properties of electron density. J. Math. Chem. doi:10.1007/s10910-011-9938-9
26. E. Simon, P.G. Mezey, Imperfect periodicity and systematic changes of some structural features along linear polymers: the case of rod-like boron-nitrogen nanostructures. Theor. Chem. Acc. (to be published)
27. P. Hohenberg, W. Kohn, Inhomogeneous electron gas. Phys. Rev. **136**, B864–B871 (1964)
28. P.G. Mezey, The holographic electron density theorem and quantum similarity measures. Mol. Phys. **96**, 169–178 (1999)