Journal of Mathematical Chemistry, Vol. 40, No. 2, August 2006 (© 2006)
DOI: 10.1007/s10910-005-9019-z

Large-scale chirality measures and general symmetry deficiency measures for functional group polyhedra of proteins

Lijie Wang

Scientific Modeling and Simulation Laboratory, Departments of Chemistry and Physics and Physical Oceanography, Memorial University of Newfoundland, St. John's, NL, Canada A1B 3X7

Liqin Wang

Chemistry Department, Harbin Normal University, Heilongjiang 150080, China

Shigeru Arimoto

Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, SK, Canada S7N 5C9

Paul G. Mezey*

Scientific Modeling and Simulation Laboratory, Departments of Chemistry and Physics and Physical Oceanography, Memorial University of Newfoundland, St. John's, NL, Canada, A1B 3X7
Institute for Advanced Study, Collegium Budapest Szentháromság u. 2 H-1014 Budapest, Hungary
E-mail: pmezey@mun.ca

Received 6 June 2005; revised 28 July 2005; published online 17 March 2006

Functional Group Polyhedra provide a simplified representation of the most essential spatial features of macromolecules, especially, of globular proteins. Since the functional group polyhedron model focuses on large scale features, the chirality and other symmetry deficiency measures of these molecules, when adapted to these polyhedra, should also be based on the characterization of large scale shape features. Two new approaches for the evaluation of such symmetry deficiency and chirality measures are presented.

KEY WORDS: macromolecular chirality measures, protein symmetry deficiency, functional groups, polyhedral models, biochemical function and activity

*Corresponding author.

1. Introduction

The concepts of approximate symmetry, "almost" symmetric molecular conformations, "almost" achiral nuclear arrangements, the degree of chirality, and in a more general sense, the degree of symmetry deficiency have been applied to molecules in several different ways (for a selection of approaches, see references [1–24]). By exploiting the non-geometrical features of topology, some of the approaches have used the freedom afforded by topology to handle approximate symmetry [25–28]. Measures describing the degree of molecular symmetry, or in an alternative approach, the degree of symmetry deficiency, and in particular, the degree of chirality, are tools focusing on specific aspects of molecular shape.

Molecular shape analysis focuses on the electron density distributions of molecules. The electron density forms a fuzzy cloud surrounding the nuclei of the molecule, and this cloud fades away gradually to zero density value as the distance from the nearest nucleus converges to infinity. For large enough distances from the nuclei, the decrease of electron density follows an exponential function, however, within the chemically more important regions closer to the nuclei, the functional representation of electron density can become very complex. For relatively small molecules, up to a few hundred atoms, the pattern of the electronic density can be studied easily in terms of isodensity contour surfaces, however, for macromolecules, this approach becomes cumbersome, since these patterns become too complex for simple analysis. Nevertheless, even for macromolecules, the electron density carries all molecular information, and it is the basis for shape analysis.

One of the essential elements of the study of molecular electron densities is the identification of reactive regions. Often, these reactive regions are characterized as functional groups, and the quantum chemical definition of functional groups is based on the patterns of the electron density cloud [29–31]. The properties of quantum chemically defined functional groups have been studied in terms of fuzzy set models [32], also in terms of relations between local shape and global shape [33], where the holographic properties of electron densities [34] have been exploited. Alternative approaches [35,36] have been based on the potential energy hypersurface model [37,38] of conformational changes and chemical reactions.

For macromolecules, especially for proteins, alternative, simplified representations have also been proposed. One such approach, emphasizing the distribution of the biochemically essential functional groups within the protein molecule, is the Functional Group Polyhedron model [39,40]. This model has also led to a family of similarity measures between proteins, based on the biochemically important functional groups. The quantum chemical definition of functional groups [29–31] is based on electron density and it has a natural connection to molecular shape analysis, consequently, this definition is also a natural

basis for the functional group polyhedron model. Nevertheless, one may use other, possibly simpler criteria for the construction of a valid a functional group polyhedron model.

2. General symmetry deficiency measures

First we shall consider a rather general approach to symmetry deficiency measures, that is applicable to fields beyond chemistry, and is formulated in terms of general set theory. Only a brief review of the definions is given without formal theorems and proofs, for more detail the reader is referred to the literature [19,20,22].

A set is an R-*set* if it has the symmetry element R.

B is an R-*subset* of a set A if B is a subset of A, and B is an R-set.

B is a *maximum volume* R-*subset* of a set A if B is an R-subset of A and if volume V(B) is maximum among all R-subsets of A. (note that, while B is not necessarily unique, V(B) is).

C is an R-*superset* of a set A if C is a superset of A (A is a subset of C), and C is an R-set.

C is a *minimum volume* R-*superset* of a set A if C is an R-superset of A and if volume V(C) is minimum among all R-supersets of A. (note that, while C is not necessarily unique, V(C) is).

The *internal* R-*symmetry deficiency measure*, ISD(A,R) of a set A is

$$\mathrm{ISD(A,R)} = 1 - \mathrm{V(B)/V(A)}, \tag{1}$$

where B is a maximum volume R-subset of A.

The *external* R-*symmetry deficiency measure*, ESD(A,R) of a set A is

$$\mathrm{ESD(A,R)} = 1 - \mathrm{V(A)/V(C)}, \tag{2}$$

where C is a minimum volume R-superset of A.

3. Density domains and functional groups

The quantum chemical definition of functional groups is based on the concepts of isodensity contours and density domains.

A *molecular isodensity contour surface,* MIDCO G(K,a) of nuclear configuration K and density threshold a is defined as

$$\mathrm{G(K,a)} = \{r : \mathbf{r}(K, r) = a\}, \tag{3}$$

that is, as the collection of all points r of the 3D space where the electronic density **r**(K,r) is equal to the threshold value a.

Note that one application of the MIDCOs is the Shape Group approach developed for the detailed shape analysis of molecular electron densities, making up the fuzzy bodies of molecules. The *shape groups* are algebraic groups, not related to point symmetry groups, although the presence of symmetry may influence the shape groups. The shape groups are the algebraic-topological *homology groups* of truncated objects, where the truncation is determined by local shape properties. In most applications of shape groups, the local shape properties are specified in terms of shape domains: for example, in terms of the locally convex, concave, or saddle-type regions of MIDCO's, relative to tangent planes, or tangent spheres.

A *density domain* DD(K,a) is a domain enclosed by a MIDCO C(K,a), that is

$$DD(K,a) = \{ \mathbf{r}:\rho(K,\mathbf{r}) > a\}. \tag{4}$$

The density domain approach has been suggested for a quantum chemical representation of formal functional groups. Consider a single connected density domain DD(K,a) and the nuclei enclosed by it. The very fact that this subset of the nuclei of the molecule is separated from the rest of the nuclei by the boundary G(K,a) of the density domain DD(K,a) indicates that these nuclei, together with the local electronic density cloud surrounding them, represent a sub-entity of the molecule, with individual identity.

It is natural to regard this density domain DD(K,a) as a representative of a *formal functional group.*

4. Functional group polyhedra

Take a set of functional groups from a molecule M, and define a unique point for each. Take the convex hull of these points, this defines a functional group polyhedron FGP(M) for the molecule M.

This polyhedron provides a simplified representation of some of the chemically significant, essential geometrical aspects of the molecule M.

The symmetry and chirality properties of an FGP(M) polyhedron can be studied directly, leading to a "low resolution" description, and an FGP(M) polyhedron can be compared directly to another FGP(M′) polyhedron. Whereas the source of the original input information is a quantum chemistry study of the molecule, some of the tools for the purposes of the analysis of functional group polyhedra are provided by topology and discrete mathematics.

5. Application of general symmetry deficiency measures for functional group polyhedra

Take the functional group polyhedron FGP(M) of a molecule M, and consider the general internal and external symmetry deficiency measures ISD(A,R) and ESD(A,R) discussed in section 2. By taking A as the convex functional group polyhedron FGP(M) of molecule M, these symmetry deficiency measures can be evaluated relatively easily for any symmetry element R, since the very fact of dealing with convex polyhedra allows one to use efficient algorithms for applying and testing various symmetry operations. The internal and external symmetry deficiency measures

$$\mathrm{ISD(FGP(M),R)} \tag{5}$$

and

$$\mathrm{ESD(FGP(M),R)} \tag{6}$$

provide a low resolution description of the approximate symmetry of the distribution of the essential functional groups of the macromolecule M.

We shall be especially concerned with a particular type of symmetry deficiency, chirality.

An internal chirality measure ICH(FGP(M)) and an external chirality measure ECH(FGP(M)) of the functional group polyhedron FGP(M) can be defined as follows:

$$\mathrm{ICH(FGP(M))} = \min\{\mathrm{ISD(FGP(M),R)},\ \mathrm{R} = \sigma, \mathrm{S_{2n}},\ \mathrm{n} = 2, 3, 4, \ldots\} \tag{7}$$

and

$$\mathrm{ECH(FGP(M))} = \min\{\mathrm{ESD(FGP(M),R)},\ \mathrm{R} = \sigma, \mathrm{S_{2n}},\ \mathrm{n} = 2, 3, 4, \ldots\}, \tag{8}$$

respectively, where symmetry element σ is a reflection plane, whereas symmetry elements S_{2n}, $n = 2, 3, 4, \ldots$ correspond to rotation-reflection symmetry operators of even fold. If present, any one of these symmetry elements implies achirality for the functional group polyhedron FGP(M).

Note that even if only one of the symmetry elements listed in these definitions shows only a small degree of symmetry deficiency, then the structure of the functional group polyhedron is "almost" achiral, and this justifies taking the minimum of symmetry deficiency measures in definitions (7) and (8).

Also note that for most macromolecules it is sufficient to consider n values smaller than 10, since higher n values for symmetry elements of improper rotations would assume higher levels of angular regularities in macromolecules than it is commonly possible.

6. A symmetry-biased dissimilarity measure for functional group polyhedra

Consider two macromolecules M and M′, and assume that they are sufficiently similar in a chemical sense, so that they have either the same lists of essential functional groups, or the two lists can be matched up in a reasonable way. Whereas the above considerations are not strictly necessary for the forthcoming discussion, and the symmetry-biased dissimilarity measures can be introduced for two fundamentally different molecules and functional group polyhedra FGP(M) and FGP(M′), the dissimilarity measures to be introduced are likely to be more useful if there is at least some chemical commonality between the two molecules. A typical example where this commonality is present, if one considers two different conformations of the same macromolecule.

Consider a set **R** of k symmetry elements,

$$\mathbf{R} = \{R_i, i = 1, 2, \ldots, k\} \tag{9}$$

and the associated internal symmetry deficiency measures

$$ESD(FGP(M),\ R_i),\ \ i = 1, 2, \ldots, k \tag{10}$$

and

$$ESD(FGP(M'), R_i),\ \ i = 1, 2, \ldots, k, \tag{11}$$

respectively, where we focus on the external symmetry deficiency measures.

With respect to the given set of symmetry elements **R**, a symmetry-biased external dissimilarity measure $SBS_{ext}(M,M',FGP, \mathbf{R})$ of the two functional group polyhedra FGP(M) and FGP(M′) can be defined as

$$SBS_{ext}(M,M',FGP, \mathbf{R}) = [\Sigma_{i=1,k}[ESD(FGP(M), R_i) - ESD(FGP(M'), R_i)]^2]^{1/2}. \tag{12}$$

For any two macromolecules M and M′, the quantity $SBS_{ext}(M,M',FGP, \mathbf{R})$ expresses how different the approximate symmetries of the two functional group polyhedra are, wu=ith respect to the given set of symmetry elements.

Note that a symmetry-biased internal dissimilarity measure $SBS_{int}(M,M',FGP, \mathbf{R})$ can be defined entirely analogously, simply by replacing the external symmetry deficiency measures $ESD(FGP(M), R_i)$ with the internal symmetry deficiency measures $ISD(FGP(M), R_i)$ in the above definition (12).

In the comparisons of proteins of similar biochemical functions, or in the comparisons of drug molecules of comparable effects, or simply, in the comparisons of different conformations of the same protein or those of the same drug molecule, an important aspect of their dissimilarity is the different degree of their approximate symmetry. This dissimilarity measure provides a tool for a numerical characterization of their differences.

7. Comments on the choice of vertices of functional group polyhedra and preferences for internal or external symmetry deficiency measures

There is a considerable freedom in the potential choices for defining vertices of the functional group polyhedra FGP(M). A simplistic choice is the nuclear position of the most relevant atom of the given functional group, that might be, for example, the proton of a carboxyl group. Alternatively, one may consider a MIDCO of a particular density threshold a of the local electron density cloud of the functional group, and chose the point of this MIDCO furthest away from the center of mass of the macromolecule. This might be a reasonable choice if the peripheral regions of the macromolecule are considered important. Yet another alternative is the choice of a point of the MIDCO that is falling on the line drawn through the two nuclei involved in the bond most likely to break when the functional group undergoes its typical chemical reaction. Since there are at least two such points, in most cases one should chose the one nearest to the center of the broken bond. This latter choice is one that is likely to be most relevant to the actual chemical role of the given functional group, but further, alternative choices are also possible.

In many instances the external symmetry deficiency measures ESD(FGP(M), R) and the external chirality measures ECH(FGP(M)) are the chemically more relevant. This is the case for most smaller molecules where the expected chemical or biochemical activity is initiated on the peripheral regions of the molecule.

On the other hand, for proteins which function as enzymes where the cavity region has special significance, if the chosen functional groups are restricted to those present in the cavity region, then the interior of the functional group polyhedron FGP(M) is relevant. Consequently, in such cases the internal symmetry deficiency measures ISD(FGP(M),R) and the internal chirality measures ICH(FGP(M)) are of higher importance.

Acknowledgments

The research support received from Natural Sciences and Engineering Research Council of Canada, the Canada Foundation for Innovation, the Canada Research Chairs Program, and from the Memorial University of Newfoundland are gratefully acknowledged.

References

[1] A.I. Kitaigorodski, *Organic Chemical Crystallography* (Consultants Bureau, New York, 1961) Chapter 4, pp. 75–100.
[2] A. Rassat, Un critere de classement des sytemes chiraux de points a partir de la distance au sens de Haussdorf, Compt. Rend. Acad. Sci. (Paris) II 299 (1984) 53–55.
[3] G. Gilat, Chiral interactions in biomolecules, Chem. Phys. Lett. 121 (1985) 13–16.

[4] L.D. Barron, Symmetry and molecular chirality, Chem. Soc. Rev. 15 (1986) 189–223.
[5] P.G. Mezey, Tying knots around chiral centres: chirality polynomials and conformational invariants for molecules, J. Am. Chem. Soc.108 (1986) 3976–3984.
[6] P.G. Mezey, Symmetry and periodicity of potential surfaces, Theor. Chim. Acta 73 (1986) 221–228.
[7] G. Gilat and J. Phys. A, Chiral coefficient – a measure of the amount of structural chrality, Math. Gen. 22 (1989) L545–L550.
[8] G. Gilat, On the quantification of symmetry in nature, Found. Phys. Lett. 3 (1990) 189–196.
[9] Y. Hel-Or, S. Peleg and D. Avnir, Two-dimensional rotational dynamic chirality and a chirality scale, Langmuir 6 (1990) 1691–1695.
[10] P.G. Mezey (ed.), *New Developments in Molecular Chirality* (Kluwer, Dordrecht, 1991).
[11] T.P.E. Auf der Heyde, A.B. Buda and K. Mislow, Desymmetrization and the degree of chirality, J. Math. Chem. 6 (1991) 255–265.
[12] A.B. Buda and K. Mislow, On geometric measures of chirality, J. Mol. Struct. (Theochem). 232 (1991) 1–12.
[13] A.Y. Meyer and D. Avnir, The relation between molecular shape and molecular rotations in chiral halogenated alkanes, Struct. Chem. 2 (1991) 475–478.
[14] P.G. Mezey, The degree of similarity of three-dimensional bodies; applications to molecular shapes, J. Math. Chem. 7 (1991) 39–49.
[15] P.G. Mezey, On the allowed symmetries of all distorted forms of conformers, molecules, and transition structures, Can. J. Chem. 70 (1992) 343–347.
[16] P.G. Mezey, Shape in chemistry : an introduction to molecular shape and topology, (VCH Publishers, New York, 1993).
[17] P.G. Mezey, Quantum chemistry of macromolecular shape, Internat. Rev. Phys. Chem. 16 (1997) 361–388.
[18] P.G. Mezey, The proof of the metric properties of a fuzzy chirality measure of molecular electron density clouds, J. Mol. Struct. Theochem. 455 (1998) 183–190.
[19] P.G. Mezey, Generalized chirality and symmetry deficiency, J. Math. Chem. 23 (1998) 65–84.
[20] P.G. Mezey, Mislow's label paradox, chirality-preserving conformational changes, and related chirality measures, chirality 10 (1998) 173–179.
[21] P.G. Mezey, Theory of biological homochirality: chirality, symmetry deficiency, and electron-cloud holography in the shape analysis of biomolecules, In: *Advances in BioChirality*, eds. G. Palyi, C. Zucchi and L. Caglioti (Elsevier Sci. Publ., Amsterdam, The Netherlands, 1999) pp. 35–46.
[22] P.G. Mezey, The role of imperfect symmetry in nature, art, mathematics, and chemistry: approximate symmetry and symmetry deficiency measures, J. Int. Soc. Interdisc. Study Symmetry 2004 (2004) 166–169.
[23] P.G. Mezey, The theory of chirality induction and chirality reduction in biomolecules, In: *Progress in Biological Chirality*, eds. G. Palyi, C. Zucchi and L. Cagliotti (Elsevier, Oxford, GB, 2004) Chapter 17, pp. 209 – 219.
[24] P.G. Mezey, Molecular similarity, quantum topology, and shape, In: *Computational Medicinal Chemistry and Drug Discovery*, eds. Patric Bultinck, Jan P. Tollenaere, Hans De Winter and Wilfried Langenaeker (Marcel Decker Inc., New York, 2004) pp. 345–364.
[25] P.G. Mezey, The topological model of non-rigid molecules and reaction mechanisms, In: *Symmetries and Properties of Non-Rigid Molecules: A Comprehensive Survey*, eds. J. Maruani and J. Serre (Elsevier Sci. Publ. Co., Amsterdam, 1983) pp. 335–353.
[26] P.G. Mezey, Topological theory of molecular conformations, In: *Structure and Dynamics of Molecular Systems*, eds. R. Daudel, J.-P. Korb, J.-P. Lemaistre, and J. Maruani (Reidel, Dordrecht, 1985) Vol. I., pp. 41–56.

[27] P.G. Mezey and J. Maruani, The concept of "syntopy": a continuous extension of the symmetry concept for quasi-symmetric structures using fuzzy-set theory, Mol. Phys. 69 (l990) 97–113.
[28] P.G. Mezey and J. Maruani, The fundamental syntopy of quasi-symmetric systems: geometric criteria and the underlying syntopy of a nuclear configuration space, Int. J. Quantum Chem. 45 (1993) 177–187.
[29] P.G. Mezey, Quantum chemical shape: new density domain relations for the topology of molecular bodies, functional groups, and chemical bonding, Can. J. Chem. 72 (1994) 928–935 (Special issue dedicated to Prof. J.C. Polanyi).
[30] P.G. Mezey, Functional groups in quantum chemistry, Adv. Quantum Chem. 27 (1996) 163–222.
[31] P.G. Mezey, Local electron densities and functional groups in quantum chemistry, In: *Topics in Current Chemistry, Correlation and Localization*, ed. P.R. Surjan (Springer-Verlag, Berlin, Heidelberg, New York, 1999) Vol. 203, pp. 167–186.
[32] G.M. Maggiora and P.G. Mezey, A fuzzy set Approach to functional group comparisons based on an asymmetric similarity measure, Int. J. Quant. Chem. 74 (1999) 503–514.
[33] 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.
[34] P.G. Mezey, The holographic electron density theorem and quantum similarity measures, Mol. Phys. 96 (1999) 169–178.
[35] J.-E. Dubois and P.G. Mezey, Relations among functional groups within a stoichiometry: a nuclear configuration space approach, Int. J. Quantum Chem. 43 (1992) 647–658.
[36] J.-E. Dubois and P. G. Mezey, A functional group database: a charge density – DARC approach, Molec. Eng. 8 (1999) 251–265.
[37] P.G. Mezey, Manifold theory of multidimensional potential surfaces, Int. J. Quantum Chem., Quant. Biol. Symp. 8 (1981) 185–196.
[38] P.G. Mezey, Potential energy hypersurfaces (Elsevier, Amsterdam, 1987).
[39] P.G. Mezey, K. Fukui, S. Arimoto and K.Taylor, Polyhedral shapes of functional group distributions in biomolecules and related similarity measures, Int. J. Quantum Chem. 66 (1998) 99–105.
[40] P.G. Mezey, K. Fukui and S. Arimoto, A treatment of small deformations of polyhedral shapes of functional group distributions in biomolecules, Int. J. Quant. Chem. 76 (2000) 756–761.