Journal of Mathematical Chemistry 13(1993)59–70

Dynamic shape analysis of molecules in restricted domains of a configuration space

Paul G. Mezey
Mathematical Chemistry Research Unit,
Department of Chemistry and Department of Mathematics,
University of Saskatchewan, Saskatoon, Canada S7N 0W0

A method is presented for the description of dynamic shape properties of molecular configurations confined to a domain of the metric nuclear configuration space M. The method is based on the shape groups of formal three-dimensional bodies associated with a continuum of all shape representations (for example, all isodensity surfaces of molecular electron distributions) occurring within a selected domain F of M. The bodies of the first type are minimal 3D bodies which contain all isodensity surfaces $G(a, K)$ of a given density value a which occur for configurations K within domain F, whereas bodies of the second type are maximal 3D bodies contained within all isodensity surfaces $G(a, K)$ within domain F. The boundary surfaces of these bodies lead to the concepts of envelope shape and core shape for families of nuclear configurations. The differences of the shape groups as well as the volumes of these bodies provide a dynamic shape as well as dynamic size characterization of nonrigid molecules. Special emphasis is given to configurational domains defined by various energy bounds (level sets), symmetry properties (such as chirality), and the invariance domains of deformations preserving chemical identity (catchment regions).

1. Introduction

The analysis of shapes of molecules is of fundamental importance in both theoretical and applied chemistry. Much of the current research is focused on various applications in drug design, where both relatively small organic compounds as well as macromolecules such as biopolymers are of interest [1–10]. Molecular shape is not a static concept, since molecules themselves are dynamic, nonrigid entities. Although shape characterization of static molecular models is a valuable approach, nevertheless, models of shape description which take into account the nonrigidity of molecules provide more insight. The dynamic shape space model DSS [11] has been proposed as a systematic framework for a dynamic shape characterization of molecules, formulated within the nuclear configuration space model. In this general model, the 3D shape variations of molecules is monitored within the space of internal configurations M (a metric space with elements of all the possible relative nuclear arrangements for a given overall stoichiometry). Shape is regarded as a function of the changes of nuclear configurations K, as well as a function of some parameters of the chosen shape representation P. For example, a shape representation P may be chosen as an electronic isodensity contour $G(a, K)$ for a given contour density value a and nuclear configuration K. Within the shape group technique [12,13], the local curvature domains of the contour surface $G(a, K)$ are identified, where the local extreme curvatures are compared to a reference curvature b. In this case, the parameters of the shape representation P (the contour surface with curvature domains identified) are the density value a and the reference curvature value b. In a more general case, there may be more numerous parameters, for example, if the local curvatures of $G(a, K)$ are compared to an oriented reference ellipsoid T, then in addition all the density value a, the three characteristic curvatures b_1, b_2, and b_3 (the reciprocals of the half axes of the ellipsoid) may serve as parameters [14]. Different patterns on $G(a, K)$, hence different shape representations P, can be generated, for example, by mapping the ranges of the differences between local canonical curvatures, or the ranges of the local density gradient vector components (oriented shape representation) or the ranges of the length of the gradient vector on the contour surface $G(a, K)$. The shape group technique can be applied to these patterns as well.

In general, the dynamic shape space D is the product space $M \otimes \boldsymbol{P}$, where $\boldsymbol{P}$ is a vector space spanned by the parameters $\boldsymbol{p}$ of shape representation $P = P(\boldsymbol{p})$ [11]. (Note that M is not in general a vector space, that justifies the distinction in the notation.) The actual shape descriptor W of the shape representation P can be chosen as a topological descriptor, for example, as the shape groups [12–14] (homology groups of molecular surfaces truncated according to a criterion defined by the parameters $\boldsymbol{p}$ of shape representation $P = P(\boldsymbol{p})$). According to the GSTE principle (geometrical similarity as topological equivalence [15]), two molecular contour surfaces $G(a, K)$ and $G'(a', K')$ are (P, W)-similar if the W descriptors of their shape representation P are topologically equivalent.

For example, if the shape groups of surfaces G are selected as shape descriptors,

$$W = g(G, \omega', h), \tag{1}$$

then for each dimension h and type ω' of domains truncated from G (in the typical, nondegenerate cases) there is only a finite number of different shape groups within the entire dynamic shape space D. This allows one to reduce the continuum problem of infinitely many, geometrically different shapes to a finite set of different topological shape types $\tau_{i,(P,W)}$,

$$\tau_{i,(P,W)} = \tau_{i,(P,g(G,\omega',h))}. \tag{2}$$

If the index i of the ith shape type $\tau_{i,(P,g(G,\omega',h))}$ is assigned to the corresponding abstract shape group $g_i(G, \omega', h)$, then this latter can be referred to as the ith reference group. The invariance domains $A_{i,j,}(g_i(G, \omega', h))$ of each reference group within the dynamic shape space D of elements $\boldsymbol{d}$ are defined as

$$A_{i,j,}(g_i(G,\omega',h)) = \{\boldsymbol{d} : \boldsymbol{d} \in D,\ \ g(G(\boldsymbol{d}),\omega',h) = g_i(G,\omega',h)\}, \tag{3}$$

where index j refers to the jth maximum connected component $A_{i,j,}(g_i(G, \omega', h))$ of the dynamic shape space (DSS) invariance set $A_i(g_i(G, \omega', h))$ of the ith reference group $g_i(G, \omega', h)$ of shape type $\tau_{i,(P,g(G,\omega',h))}$:

$$A_i(g_i(G,\omega',h)) = \bigcup_j A_{i,j,}(g_i(G,\omega',h)). \tag{4}$$

The pair composed from a given shape group $g_i(G, \omega', h)$ and its invariance set $A_i(g_i(G, \omega', h))$ within the dynamic shape space D is regarded as the *dynamic shape group*:

$$(g_i, A_i)_{(G,\omega',h)} = (g_i(G,\omega',h),\ \ A_i(g_i(G,\omega',h))). \tag{5}$$

The shape invariance domains $A_i(g_i(G, \omega', h))$ of the dynamic shape space D provide a characterization of the dynamic shape properties of molecular conformations and arrangements. In the present study, we shall approach the problem of dynamic shape of molecules from a different perspective; we shall follow a procedure that is formally the inverse of the above technique. Instead of defining domains in the dynamic shape space D (hence, for any fixed parameter $\boldsymbol{p}$, shape invariance domains in the nuclear configuration space M) in terms of families of 3D objects, we shall base our method on domains in the nuclear configuration space as tools to define various 3D objects. The topological description of these 3D objects will provide the means to characterize the dynamic shape properties of all configurations belonging to the given configuration space domain.

2. Unconstrained and accessibility constrained external envelope and internal core surfaces of molecular contour surfaces for families of nuclear configurations

Consider a subset F of the nuclear configuration space, $F \subset M$. At this point, we shall not make any restriction of F; however, it is useful to select a potential energy hypersurface $E(K)$ above M and think of F either as a connected level set for some energy bound A, or as a catchment region representing a stable molecular species of the corresponding electronic state [16]. Next, consider a molecular contour surface G, for example, an electrostatic isopotential surface or an isodensity contour surface $G(a, K)$, or a fused sphere van der Waals surface (VDWS) for a selected family of atomic radii. For illustration, we shall explicitly discuss the case of isodensity contours $G(a, K)$; however, the general treatment is equally applicable for any closed molecular surface, where the parameter a defining the contour for a given nuclear configuration K can be a single scalar or it can represent a vector

$a = \{a_1, a_2, \ldots, a_k\}$ containing various specifications for the surface (for example, the set $\{a_1, a_2, \ldots, a_k\}$ of VDW atomic radii).

We define six 3D bodies, $B^{\mathrm{uc,e}}(a, F)$, $B^{\mathrm{uc,i}}(a, F)$, $B^{\mathrm{ac,e}}(a, F)$, $B^{\mathrm{ac,i}}(a, F)$, $B^{\mathrm{sac,e}}(a, F)$, and $B^{\mathrm{sac,i}}(a, F)$, and their boundary surfaces $G^{\mathrm{uc,e}}(a, F)$, $G^{\mathrm{uc,i}}(a, F)$, $G^{\mathrm{ac,e}}(a, F)$, $G^{\mathrm{ac,i}}(a, F)$, $G^{\mathrm{sac,e}}(a, F)$, and $G^{\mathrm{sac,i}}(a, F)$, respectively. The first of these surfaces, the unconstrained exterior envelope surface $G^{\mathrm{uc,e}}(a, F)$, is the boundary surface of the 3D body

$$B^{\mathrm{uc,e}}(a, F) = \min\{B : \forall K \in F, G(a, K) \subset B\}, \tag{6}$$

$$G^{\mathrm{uc,e}}(a, F) = \Delta\, B^{\mathrm{uc,e}}(a, F), \tag{7}$$

whereas the unconstrained interior core surface $G^{\mathrm{uc,i}}(a, F)$ is the boundary surface of the 3D body

$$B^{\mathrm{uc,i}}(a, F) = \max\{B : \forall K \in F, B \subset B(a, K)\}, \tag{8}$$

$$G^{\mathrm{uc,i}}(a, F) = \Delta\, B^{\mathrm{uc,i}}(a, F). \tag{9}$$

Here, as well as in the following expressions, min and max refer to minimum and maximum 3D volumes $V(B)$ of set B, where the $B(a, K)$ bodies and their respective boundary surfaces $G(a, K)$ are taken for all nuclear configurations within domain F, and the symbol Δ denotes the boundary operation.

The $G(a, K)$ contours for all nuclear configurations of domain F fit within the unconstrained exterior envelope surface $G^{\mathrm{uc,e}}(a, F)$. Nevertheless, it is possible that along a path p of a configuration change within F the $G(a, K_1)$ and $G(a, K_2)$ contours of two configurations K_1 and K_2, which differ only infinitesimally according to the metric of the configuration space M, can be enclosed within the minimum volume envelope surface $G^{\mathrm{uc,e}}(a, F)$ only in very different ways, implying a discontinuity in the embedding arrangements as the nuclear configurations K change continuously along the path p within domain F. Consequently, it is meaningful to consider the following accessibility constraints:

(i) a restriction that for any two nuclear configurations K_1 and K_2 of domain F there must exist a path p within F such that along p, the contour surface $G(a, K)$ can be continuously deformed from $G(a, K_1)$ to $G(a, K_2)$ within an envelope surface $G^{\mathrm{ac,e}}(a, F)$, where the superscript ac refers to this accessibility constraint.

A stronger accessibility constraint is given by the following choice:

(ii) a restriction that for any two nuclear configurations K_1 and K_2 of domain F along *all* paths p interconnecting K_1 and K_2 within F, the contour surface $G(a, K)$ can be continuously deformed from $G(a, K_1)$ to $G(a, K_2)$ within an envelope surface $G^{\mathrm{sac,e}}(a, F)$, where the superscript sac refers to this strong accessibility constraint.

One may formulate these constraints by considering continuous functions

$$f : p(K_1, K_2) \to G(S), \tag{10}$$

where $p(K_1, K_2)$ is a path in configuration space from K_1 to K_2 and $G(S)$ denotes the set of continuous closed 2D surfaces embedded in a subset S of the 3D Euclidean space E^3. Then,

$$\begin{aligned} &B^{\mathrm{ac,e}}(a, F) \\ &\quad = \min\{B : \forall K \in F, G(a, K) \subset B; \\ &\qquad\qquad \forall K_1, K_2 \in F, \exists p(K_1, K_2) \subset F : \exists f : p(K_1, K_2) \to G(F)\}, \end{aligned} \tag{11}$$

$$G^{\mathrm{ac,e}}(a, F) = \Delta\, B^{\mathrm{ac,e}}(a, F) \tag{12}$$

and

$$\begin{aligned} &B^{\mathrm{sac,e}}(a, F) \\ &\quad = \min\{B : \forall K \in F, G(a, K) \subset B; \\ &\qquad\qquad \forall K_1, K_2 \in F, \forall p(K_1, K_2) \subset F : \exists f : p(K_1, K_2) \to G(F)\}, \end{aligned} \tag{13}$$

$$G^{\mathrm{sac,e}}(a, F) = \Delta\, B^{\mathrm{sac,e}}(a, F). \tag{14}$$

The analogous conditions for the interior core surfaces can be formulated by requiring that

(a) the maximal $B^{\mathrm{ac,i}}(a, F)$ and $B^{\mathrm{sac,i}}(a, F)$ bodies are contained within all contour surfaces associated with the configuration space domain F, and that

(b) between any two configurations K_1 and K_2 of set F there exists at least one path $p(K_1, K_2)$ (case ac), or all paths $p(K_1, K_2)$ are such (case sac) that the embeddings of the $G(a, K)$ contours along the paths change continuously within the complements $E^3 \backslash B^{\mathrm{ac,i}}(a, F)$ and $E^3 \backslash B^{\mathrm{sac,i}}(a, F)$ of the maximal $B^{\mathrm{ac,i}}(a, F)$ and $B^{\mathrm{sac,i}}(a, F)$ bodies, respectively. That is, none of the contours along continuous deformation paths $p(K_1, K_2)$ has a common point with $B^{\mathrm{ac,i}}(a, F)$ and $B^{\mathrm{sac,i}}(a, F)$, respectively.

The following formal definitions can be given for the interior core bodies and core surfaces:

$$\begin{aligned} B^{\mathrm{ac,i}}(a, F) = \max\{&B : \forall K \in F, B \subset G(a, K); \forall K_1, K_2 \in F, \\ &\exists p(K_1, K_2) \subset F : \exists f : p(K_1, K_2) \to G(E^3 \backslash F)\}, \end{aligned} \tag{15}$$

$$G^{\mathrm{ac,i}}(a, F) = \Delta\, B^{\mathrm{ac,i}}(a, F) \tag{16}$$

and

$$B^{\mathrm{sac,i}}(a, F) = \max\{B : \forall K \in F, B \subset G(a, K); \forall K_1, K_2 \in F,$$
$$\forall p(K_1, K_2) \subset F : \exists f : p(K_1, K_2) \rightarrow G(E^3 \backslash F)\}, \tag{17}$$

$$G^{\mathrm{sac,i}}(a, F) = \Delta\, B^{\mathrm{sac,i}}(a, F). \tag{18}$$

The unconstrained exterior envelope surfaces $G^{\mathrm{uc,e}}(a, F)$ and interior core surfaces $G^{\mathrm{uc,i}}(a, F)$ provide a description of shape constraints present within the chosen configuration domain F. Since within any open subset F of the nuclear configuration space M some configurational variation is allowed, these envelope and core surfaces do represent some of the dynamic features of molecular shape. However, in practical applications it is also important to find out whether a given configuration change may or may not occur if some shape constraints are imposed; for example, this is the case for conformational changes of various biologically active molecules within enzyme cavities. Consequently, in such cases the constrained external envelope surfaces $G^{\mathrm{ac,e}}(a, F)$ and $G^{\mathrm{sac,e}}(a, F)$ are of more relevance.

3. Some properties of envelope and core surfaces of families of nuclear configurations

The following relations among the supporting bodies of various exterior envelope surfaces are simple consequences of their definitions:

$$B^{\mathrm{uc,e}}(a, F) \subset B^{\mathrm{ac,e}}(a, F) \subset B^{\mathrm{sac,e}}(a, F). \tag{19}$$

Similarly, for the supporting bodies of the internal core surfaces, the relations

$$B^{\mathrm{uc,i}}(a, F) \supset B^{\mathrm{ac,i}}(a, F) \supset B^{\mathrm{sac,i}}(a, F) \tag{20}$$

hold, where the sense of the inclusion relations is reversed.

Since for any domain F

$$B^{\mathrm{uc,i}}(a, F) \subset B^{\mathrm{uc,i}}(a, F) \tag{21}$$

also holds, one may combine the above two sequences (19) and (20) of inclusion relations into a single sequence

$$B^{\mathrm{sac,i}}(a, F) \subset B^{\mathrm{ac,i}}(a, F) \subset B^{\mathrm{uc,i}}(a, F) \subset B^{\mathrm{uc,e}}(a, F)$$
$$\subset B^{\mathrm{ac,e}}(a, F) \subset B^{\mathrm{sac,e}}(a, F). \tag{22}$$

If for two conformational domains F_1 and F_2 the condition

$$F_1 \subset F_2 \tag{23}$$

applies, then

$$B^{\mathrm{uc,e}}(a, F_1) \subset B^{\mathrm{uc,e}}(a, F_2), \tag{24}$$

$$B^{\mathrm{ac,e}}(a, F_1) \subset B^{\mathrm{ac,e}}(a, F_2), \tag{25}$$

$$B^{\mathrm{sac,e}}(a, F_1) \subset B^{\mathrm{sac,e}}(a, F_2), \tag{26}$$

$$B^{\mathrm{uc,i}}(a, F_1) \supset B^{\mathrm{uc,i}}(a, F_2), \tag{27}$$

$$B^{\mathrm{ac,i}}(a, F_1) \supset B^{\mathrm{ac,i}}(a, F_2), \tag{28}$$

$$B^{\mathrm{sac,i}}(a, F_1) \supset B^{\mathrm{sac,i}}(a, F_2) \tag{29}$$

follow, again, as simple consequences of the definitions.

In the special case of isodensity contours $G(a, K)$, their supporting bodies $B(a, K)$ satisfy the relation

$$B(a_1, K) \subset B(a_2, K) \tag{30}$$

for any fixed nuclear configuration K, if

$$a_1 \geq a_2. \tag{31}$$

Consequently, if $a_1 \geq a_2$, then the following relations also hold for any configuration space domain F:

$$B^{\mathrm{uc,e}}(a_1, F) \subset B^{\mathrm{uc,e}}(a_2, F), \tag{32}$$

$$B^{\mathrm{ac,e}}(a_1, F) \subset B^{\mathrm{ac,e}}(a_2, F), \tag{33}$$

$$B^{\mathrm{sac,e}}(a_1, F) \subset B^{\mathrm{sac,e}}(a_2, F), \tag{34}$$

$$B^{\mathrm{uc,i}}(a_1, F) \subset B^{\mathrm{uc,i}}(a_2, F), \tag{35}$$

$$B^{\mathrm{ac,i}}(a_1, F) \subset B^{\mathrm{ac,i}}(a_2, F), \tag{36}$$

$$B^{\mathrm{sac,i}}(a_1, F) \subset B^{\mathrm{sac,i}}(a_2, F). \tag{37}$$

One should notice that for configuration space inclusion relations of configuration families F, the supporting bodies of envelope surfaces and core surfaces have inclusion relations of opposite directions (relations (24)–(26) and (27)–(29)), whereas for charge density changes in a contour value a, the supporting bodies of envelope surfaces and core surfaces, the sense of inclusion relations agree ((32)–(37)).

For simultaneous variations in the size of configuration domain F, in the value(s) of contour parameter(s) a, and in the accessibility constraints uc, ac, and sac, one can use a simple combination of the above families of inclusion relations.

4. The shape and size of molecular configuration of a family F: envelope shapes and core shapes, envelope sizes and core sizes

The above sets provide the means for the introduction of the concept of the shape of a family F of molecular configurations.

DEFINITION 1

The shapes of envelope surfaces $G^{\mathrm{uc,e}}(a, F)$, $G^{\mathrm{ac,e}}(a, F)$, and $G^{\mathrm{sac,e}}(a, F)$ are the unconstrained, accessibility constrained and strong accessibility constrained envelope shapes, respectively, of the configurational family F.

If no further qualification is given, the term "envelope shape of F" refers to the unconstrained case.

DEFINITION 2

The shapes of core surfaces $G^{\mathrm{uc,i}}(a, F)$, $G^{\mathrm{ac,i}}(a, F)$, and $G^{\mathrm{sac,i}}(a, F)$ are the unconstrained, accessibility constrained and strong accessibility constrained core shapes, respectively, of the configurational family F.

If no further qualification is given, the term "core shape of F" refers to the unconstrained case.

The shape characterization of unconstrained and constrained envelope shapes and the core shapes of configurational domains F can be carried out by the standard shape group methods [12–14] using local curvature properties. The pattern of local curvature domains (e.g. those of locally convex, concave, or saddle-type domains relative to a reference curvature b) of various envelope surfaces $G^{\mathrm{uc,e}}(a, F)$, $G^{\mathrm{ac,e}}(a, F)$, $G^{\mathrm{sac,e}}(a, F)$ and core surfaces $G^{\mathrm{uc,i}}(a, F)$, $G^{\mathrm{ac,i}}(a, F)$, and $G^{\mathrm{sac,i}}(a, F)$ can be described topologically by their shape groups (homology groups of objects obtained from these surfaces after cutting our domains of a given local curvature type). Since the surfaces characterize the shapes of an entire configurational domain F of nuclear arrangements, this approach allows one to use methods originally developed for static shape characterization also for dynamic shape characterization of molecules, providing an alternative to the earlier dynamic shape space technique [11].

For a given configurational domain F, the difference between the envelope shape and the core shape is of particular importance. These differences can be represented by the hollow bodies

$$B^{\mathrm{uc,d}}(a, F) = B^{\mathrm{uc,e}}(a, F) \backslash B^{\mathrm{uc,i}}(a, F), \tag{38}$$

$$B^{\mathrm{ac,d}}(a, F) = B^{\mathrm{ac,e}}(a, F) \backslash B^{\mathrm{ac,i}}(a, F), \tag{39}$$

and

$$B^{\mathrm{sac,d}}(a, F) = B^{\mathrm{sac,e}}(a, F) \backslash B^{\mathrm{sac,i}}(a, F), \tag{40}$$

for the three cases uc, ac, and sac, respectively.

The above bodies, together with those defined in eqs. (6), (8), (11), (13), (15), and (17), provide the means for the introduction of the concept of the envelope size, core size, and size range of the family F of molecular configurations (with respect to the given shape representation P). The volumes of the various interior and exterior bodies satisfy the following relations:

$$\begin{aligned} V(B^{\mathrm{sac,i}}(a, F)) &\le V(B^{\mathrm{ac,i}}(a, F)) \le V(B^{\mathrm{uc,i}}(a, F)) \le V(B^{\mathrm{uc,e}}(a, F)) \\ &\le V(B^{\mathrm{ac,e}}(a, F)) \le V(B^{\mathrm{sac,e}}(a, F)), \end{aligned} \tag{41}$$

and can be used for size characterization. The quantities

$$\Delta V^{\mathrm{uc}}(a, F)) = V(B^{\mathrm{uc,d}}(a, F))/V(B^{\mathrm{uc,e}}(a, F)), \tag{42}$$

$$\Delta V^{\mathrm{ac}}(a, F)) = V(B^{\mathrm{ac,d}}(a, F))/V(B^{\mathrm{ac,e}}(a, F)) \tag{43}$$

and

$$\Delta V^{\mathrm{sac}}(a, F)) = V(B^{\mathrm{sac,d}}(a, F))/V(B^{\mathrm{sac,e}}(a, F)) \tag{44}$$

characterize the unconstrained, accessibility constrains, and strong accessibility constrained size range of the family F of molecular configurations, where the symbol Δ stands for "difference", and is not to be confused with the boundary operation.

One may consider an entire range of shape representations $P(\boldsymbol{p})$ for a range R of choices for parameters $\boldsymbol{p}$. For example, one may consider an interval $R = [a_1, a_2]$ of electron density values and the associated family of isodensity contours $G(a, K)$ for any fixed nuclear configuration K. If the maximum and minimum volumes in definitions (6), (8), (11), (13), (15), and (17) are selected by considering all choices of configurations from F and all contour density values within the range $[a_1, a_2]$, then by repeating the above treatment one obtains a dynamic shape description for the entire range $[a_1, a_2]$ of isodensity surfaces. For example, definition (6) can be extended for a range $[a_1, a_2]$ of electronic density values, leading to a formal 3D body,

$$B^{\mathrm{uc,e}}([a_1, a_2], F) = \min\{B : \forall K \in F, \forall a \in [a_1, a_2], G(a, K) \subset B\}, \tag{45}$$

and its boundary

$$G^{\mathrm{uc,e}}([a_1, a_2], F) = \Delta\, B^{\mathrm{uc,e}}([a_1, a_2], F), \tag{46}$$

representing an unconstrained envelope surface for configuration family F and electron density range $[a_1, a_2]$. The analogous treatment applies to all other bodies, surfaces, and volumes discussed above.

5. Comments on shapes and sizes of special molecular configuration families F

If a potential energy hypersurface $E(K)$ is selected, then one important choice for family F of nuclear configurations is that of a connected level set $F(A)$ of the space M, that is, a connected subset of configurations which fall below the energy bound A. In this case, the dependence of shapes and sizes of various envelopes and cores on the energy bound A is of particular interest. If the change of the energy bound is confined between two successive critical levels [17] of the chosen potential energy hypersurface $E(K)$, then one expects a continuous change with A in all ac and sac properties defined above; however, the change in the uc properties is not necessarily continuous even between critical levels. At a critical energy level, all uc, ac, and sac properties may change discontinuously. By regarding a given ac or sac property, and by considering all those objects equivalent which interconvert into one another continuously, one obtains a set of equivalence classes. A monotonic change of the energy bound A implies a partial order for these equivalence classes along the given potential energy hypersurface $E(K)$, where this partial order leads to a semilattice structure analogous to those of the fundamental groups of reaction mechanisms in level sets of $E(K)$ [18, 19]. This connection between algebraic structures for reactivity and algebraic structures for dynamic shape variations will be described in detail elsewhere [20].

Another natural choice for F is that of various catchment regions $C(\lambda, i)$ of the potential energy hypersurface $E(K)$ ([21], see also ref. [16]). Within the Topology Program [16] of reinterpretation of fundamental chemical concepts in terms of topology, chemical species are represented by catchment regions [21]. Hence, if one makes the choice

$$F = C(\lambda, i) \tag{47}$$

for a selected catchment region (where λ is the index of the unique critical point in $C(\lambda, i)$ and i is a serial index), then the associated uc, ac, and sac envelope and core shapes and sizes are regarded as those of the chemical species $C(\lambda, i)$, with respect to the given shape representation P. In particular, if the catchment region is that of a potential energy minimum, $F = C(0, i)$, then one obtains the envelope and core shapes and the envelope and core sizes of a stable molecule $C(0, i)$, with respect to the given shape representation P. Similarly, if the catchment region is that of a potential energy saddle point of index 1, $F = C(1, i)$, then one obtains the envelope and core shapes and the envelope and core sizes of a transition structure $C(1, i)$ (transition "state", according to a common but misleading terminology).

A global analysis of molecular symmetry [22], in particular that of molecular chirality, can be based on the configuration space distribution of nuclear arrangements of specified symmetry properties. In particular, chiral nuclear arrangements are found in various chirality domains of space M, separated by subsets of achiral nuclear configurations [23]. A dynamic shape characterization of chiral nuclear arrangements can be given by taking F as one of the chirality domains, or as an energy constrained subset of a chirality domain. The latter approach can be regarded as a combination of the energy level set and chirality domain constraints.

The dynamic shape and dynamic size properties of arrangements occurring in chemical reactions can also be characterized by the methods described above. Set F can be chosen as the family of nuclear configurations occurring along a reaction channel. Many approximations are possible; for example, one may take F as a $(3N-6)$-dimensional tube of nuclear configurations defined by the following condition: F contains each configuration K along a formal reaction path p and all additional configurations K' which are A-accessible from at least one configuration K of path p, where A-accessibility means access that requires an energy change less than A. The extreme choice of $A = 0$ gives the formal reaction path p itself.

An alternative choice for F is the union of the catchment regions of all stable species and transition structures participating in the reaction

$$F = \bigcup_{i \in r} C(1, i), \tag{48}$$

where r is the family of indices involved in the reaction. This, as well as the previous choices, provide a systematic approach for the analysis of dynamic molecular shape and size variations in chemical reactions.

Acknowledgement

It is a pleasure to acknowledge stimulating discussions with Professor Jacques-Emile Dubois of the Institute of Topology and System Dynamics, Paris, France, Dr. Jean Maruani, CMOA, CNRS, Paris, France, and Professor A. Rassat, ENS, Paris, France, during a sabbatical visit of the author. Financial support for this work was provided in part by the Computational Chemistry Unit of the Upjohn Laboratories, Kalamazoo, Michigan, and by both strategic and operating research grants from the Natural Sciences and Engineering Research Council of Canada.

References

[1] Y.C. Martin, *Quantitative Drug Design: A Critical Introduction* (Dekker, New York, 1978).
[2] M. Karplus and J.A. McCammon, Annu. Rev. Biochem. 53(1983)263.
[3] P. De Santis, S. Morosetti and A. Palleschi, Biopolymers 22(1983)37.
[4] W.G. Richards, *Quantum Pharmacology* (Butterworths, London, 1983).
[5] J.S. Richardson, Methods Enzymol. 115(1985)359.

[6] M.N. Liebman, C.A. Venanzi and H. Weinstein, Biopolymers 24(1985)1721.

[7] T. Kikuchi, G. Némethy and H.A. Scheraga, J. Comput. Chem. 7(1986)67.

[8] P.M. Dean, *Molecular Foundations of Drug–Receptor Interaction* (Cambridge University Press, New York, 1987).

[9] M.L. Connolly, Visual Comput. 3(1987)72.

[10] O. Tapia, H. Eklund and C.I. Brändén, Molecular electronic and structural aspects of the catalytic mechanism of alcohol dyhydrogenase, in: *Steric Aspects of Biomolecular Interactions*, ed. G. Náray-Szabó and K. Simon (CRC Press, Boca Raton, 1987).

[11] P.G. Mezey, J. Math. Chem. 2(1988)299.

[12] P.G. Mezey, Int. J. Quant. Chem. Quant. Biol. Symp. 12(1986)113.

[13] P.G. Mezey, J. Comput. Chem. 8(1987)462.

[14] P.G. Mezey, J. Math. Chem. 2(1988)325.

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

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

[17] P.G. Mezey, Theor. Chim. Acta 60(1981)97.

[18] P.G. Mezey, Theor. Chim. Acta 67(1985)43.

[19] P.G. Mezey, Theor. Chim. Acta 67(1985)91.

[20] P.G. Mezey, to be published.

[21] P.G. Mezey, Theor. Chim. Acta 58(1981)309.

[22] P.G. Mezey, J. Am. Chem. Soc. 112(1990)3791.

[23] P.G. Mezey, A global approach to molecular chirality, in: *New Developments in Molecular Chirality*, ed. P.G. Mezey (Kluwer, Dordrecht, 1991).