Theor Chim Acta (1995) 92: 333–338

Theoretica Chimica Acta

Three properties of relative shape envelopes of molecular electron density contours

Paul G. Mezey

Mathematical Chemistry Research Unit, Department of Chemistry and Department of Mathematics and Statistics, 110 Science Place, University of Saskatchewan, Saskatoon, SK, Canada S7N 5C9

Received January 17, 1995/Final version received March 16, 1995/Accepted May 11, 1995

Summary. The relative shapes of molecular electron density contour surfaces (MIDCO's), and various molecular shape constraints in solvent–solute interactions, in external electromagnetic fields and within enzyme cavities, are representable by electron density T-hulls, introduced earlier. Three general properties of T-hulls are proven, serving as the justification of a recently proposed computational scheme of molecular similarity measures.

Key words: Relative shape envelopes – Electronic densities – Solvent – solute interactions – T-hulls

1 Introduction

The concept of α-hull has been introduced by Edelsbrunner et al. [1] as a generalization of convexity. The T-hull, introduced recently [2], can be regarded as a generalization of the α-hull, hence, as a further generalization of the convex hull. The chemical relevance of T-hulls lies in their role as tools for shape analysis of electronic densities [3], as the basis of molecular similarity measures, and as mathematically precise representations of solvent contact surfaces of molecules [4]. By the introduction of the MEDLA method for *ab initio* quality electron density computations for proteins and other large molecules [5–7], the role of computational shape analysis methods designed for molecular applications [8] is expected to increase.

In Ref. [1], the introduction of two-dimensional α-hulls has been based on the concept of *generalized disc of radius* $1/\alpha$, defined as a disc of radius $1/\alpha$ if $\alpha > 0$, the complement of a disc of radius $-1/\alpha$ if $\alpha < 0$, and a half-plane if $\alpha = 0$. The *α-hull* $\langle S \rangle_\alpha$ of a point set S in the plane has been defined as the intersection of all closed generalized discs of radius $1/\alpha$ which contain S.

Following the description in [3], the three-dimensional case is entirely analogous. A *generalized ball of radius* $1/\alpha$ is defined as a ball of radius $1/\alpha$ if $\alpha > 0$, as the complement of a ball of radius $-1/\alpha$ if $\alpha < 0$, and as a half-space if $\alpha = 0$. The *α-hull* $\langle S \rangle_\alpha$ of a finite point set S in a 3D Euclidean space is defined as the intersection of all closed generalized balls of radius $1/\alpha$ which contain S.

The α-hull $\langle S \rangle_\alpha$ of S is a "curvature-biased" shape representation of S, using the specific curvature value α. For a finite point set S (for example, for the collection of nuclei in a specific configuration) and for a sufficiently small negative value of α, the α-hull $\langle S\rangle_\alpha$ of S is the finite point set S itself. In the special case of $\alpha = 0$, the α-hull $\langle S\rangle_\alpha$ of set S is the ordinary convex hull $\langle S\rangle$ of S. According to the usual convention, the empty intersection is regarded as the entire space, consequently, the α-hull of any set S exists for any α value.

The T-hull of both discrete point sets and continua has been introduced [2] as a generalization of the convex hull with respect to a reference object T. Within the chemical context, a shape characterization of the molecular electronic density of a molecule A in terms of its T-hull defined by an electronic density contour of another molecule B serves as a direct shape comparison of molecules A and B, and also as a "B-biased" shape representation of molecule A.

Following the original definition [2], consider an arbitrary, bounded and closed, three-dimensional set T, and regard it as a reference object. Using T-hulls, the shapes of various other objects S are described relative to the reference object T.

If a reference object T (for example, a molecular isodensity contour surface, MIDCO, of a molecule B) is selected, then any set obtained by translation and rotation of T is called a *version* of T. Some motions may be excluded, for example, the test object T may be required to fulfill some orientation constraints; in such cases a version of T is a set obtained from T by translation.

The *T-hull* $\langle S\rangle_T$ of a point set S has been defined [2] as the intersection of all rotated and translated versions of T which contain set S. If no version T_v of T contains S then the T-hull of S is the empty intersection, interpreted as the full space. Consequently, the T-hull $\langle S\rangle_T$ exists for every set S and for every reference object T. Evidently, the T-hull of a set S depends on the shapes of both objects, S and T, more specifically, on the relative shapes of S and T.

In some applications, for example, in solvent contact surface analysis, the closure $\mathrm{clos}(E^3\backslash T)$ of the relative complement $E^3\backslash T$ of T is required. Following the notation used in [2–4], the expression $-T$ stands for the closure of the relative complement of T:

$$-T = \mathrm{clos}(E^3\backslash T). \tag{1}$$

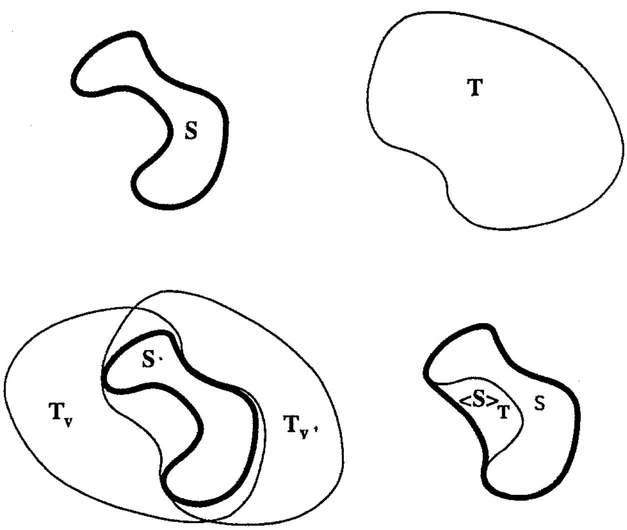

Fig. 1. A two-dimensional example of an object S, reference object T, the inclusion relations of S for two versions, T_v and $T_{v'}$, of T, and the actual T-hull $\langle S\rangle_T$ of S

By analogy with α-hulls of negative α values, the set $(-T)$ can also be chosen as a reference object.

In order to aid the visualization of properties of T-hulls, a two-dimensional example is given in Fig. 1.

2 Three identities for T-hulls

The first identity we prove is a simple generalization of an elementary property of convex hulls: for any set S the convex hull $\langle\langle S\rangle\rangle$ of the convex hull $\langle S\rangle$ is the convex hull $\langle S\rangle$, i.e.,

$$\langle\langle S\rangle\rangle = \langle S\rangle. \tag{2}$$

Clearly, the convex hull is already convex.

For T-hulls the analogous relation applies. The theorem and its proof given below are valid in all finite dimensions n.

Theorem 1. *For any set S and reference set T, the T-hull $\langle\langle S\rangle_T\rangle_T$ of the T-hull $\langle S\rangle_T$ is the T-hull $\langle S\rangle_T$:*

$$\langle\langle S\rangle_T\rangle_T = \langle S\rangle_T. \tag{3}$$

Proof. According to the definition of T-hulls, $\langle S\rangle_T$ contains S, hence, each version T_v of T that contains $\langle S\rangle_T$ also contains S:

$$T_v \supset \langle S\rangle_T \quad \Rightarrow \quad T_v \supset S. \tag{4}$$

Let us denote the family of all such versions T_v by V_1. The intersection of all sets in V_1 is $\langle\langle S\rangle_T\rangle_T$.

We show now that each version $T_{v'}$ of T that contains S also contains $\langle S\rangle_T$. Since $T_{v'} \supset S$, this version $T_{v'}$ must occur in the intersection defining $\langle S\rangle_T$, consequently, $T_{v'} \supset \langle S\rangle_T$:

$$T_{v'} \supset S \quad \Rightarrow \quad T_{v'} \supset \langle S\rangle_T. \tag{5}$$

Let us denote the family of all such versions $T_{v'}$ by V_2. The intersection of all sets in V_2 is $\langle S\rangle_T$.

Since the two implications (4) and (5) are inverses of each other, the two sets V_1 and V_2 must agree:

$$V_1 = V_2 = V. \tag{6}$$

The intersection of all sets in V is both $\langle\langle S\rangle_T\rangle_T$ and $\langle S\rangle_T$, consequently,

$$\langle\langle S\rangle_T\rangle_T = \langle S\rangle_T. \quad \square \tag{7}$$

The assertion of this theorem corresponds to the rhyme "the T-hull of the T-hull is the T-hull".

The T-hull $\langle S\rangle_T$ itself can be used as a reference set. In such a case, a different rhyme applies: "the T-hull–hull is the T-hull". We prove this below. This theorem and its proof are also valid in all finite dimensions n.

Theorem 2. *For any set S and reference set T, the T-hull $\langle S\rangle_T$ of set S is the $\langle S\rangle_T$-hull $\langle S\rangle_{\langle S\rangle_T}$ of S, obtained with the T-hull $\langle S\rangle_T$ as reference set:*

$$\langle S\rangle_{\langle S\rangle_T} = \langle S\rangle_T. \tag{8}$$

Proof. By the definition of T-hulls, $\langle S\rangle_{\langle S\rangle_T}$ is the intersection of all versions $(\langle S\rangle_T)_{v'}$ of $\langle S\rangle_T$ which contain S.
(i) First we show that $\langle S\rangle_{\langle S\rangle_T}$ is an intersection of some versions $T_{v'''}$ of T which contain S.

By the definition of T-hulls, set $\langle S\rangle_T$ is the intersection of all versions T_v of T which contain S:

$$\langle S\rangle_T = \bigcap\nolimits_v T_v. \tag{9}$$

Consequently, each version $(\langle S\rangle_T)_{v'}$ which contains S is also an intersection of some versions $T_{v''}$ of T which contain S. Since $\langle S\rangle_{\langle S\rangle_T}$ is the intersection of all versions $(\langle S\rangle_T)_{v'}$ which contain S, the set $\langle S\rangle_{\langle S\rangle_T}$ must be the intersection of some versions $T_{v'''}$ of T which contain S:

$$\langle S\rangle_{\langle S\rangle_T} = \bigcap\nolimits_{v'}(\langle S\rangle_T)_{v'} = \bigcap\nolimits_{v'}(\bigcap\nolimits_{v''}T_{v''})_{v'} = \bigcap\nolimits_{v'''}T_{v'''}. \tag{10}$$

(ii) We show that the families of sets T_v and $T_{v'''}$ in the intersections of Eqs. (9) and (10) are the same.

(a) Since $\langle S\rangle_T$ itself can be chosen as a version $(\langle S\rangle_T)_{v'}$ which contains S, for each version T_v of Eq. (9),

$$T_v \supset \langle S\rangle_T \supset \langle S\rangle_{\langle S\rangle_T}, \tag{11}$$

must hold. Hence, each T_v of Eq. (9) is one version of T that contains $\langle S\rangle_{\langle S\rangle_T}$. Consequently, for each version T_v of Eq. (9) there must exist a version $T_{v'''}$ of the intersection in the far right of Eq. (10) such that

$$T_{v'''} = T_v. \tag{12}$$

(b) Since for each version $T_{v'''}$ of the intersection in Eq. (10)

$$T_{v'''} \supset \langle S\rangle_{\langle S\rangle_T} \supset S \tag{13}$$

must hold, each $T_{v'''}$ of Eq. (10) is one version of T that contains S. Consequently, for each version $T_{v'''}$ of Eq. (10) there must exist a version T_v of the intersection in Eq. (9) such that

$$T_v = T_{v'''}. \tag{14}$$

Consequently, the intersections $\bigcap_v T_v$ and $\bigcap_{v'''} T_{v'''}$ are the same, hence

$$\langle S\rangle_{\langle S\rangle_T} = \langle S\rangle_T. \quad \square \tag{15}$$

Another important property of T-hulls is a formal "shape quantization" effect, whenever a set S is transformed into its T-hull. This "shape quantization" is based on the following simple result.

Theorem 3. *For any set S, reference set T, and set S' fulfilling the condition*

$$\langle S\rangle_T \supset S' \supset S, \tag{16}$$

the two T-hulls $\langle S\rangle_T$ and $\langle S'\rangle_T$ are the same:

$$\langle S'\rangle_T = \langle S\rangle_T. \tag{17}$$

Proof. By the definition of T-hulls, $\langle S\rangle_T$ is the intersection of all versions T_v of T which contain S. Let us denote the family of all these T_v versions by V. For each of these versions,

$$T_v \supset \langle S\rangle_T \supset S'. \tag{18}$$

Consequently, each version T_v from the family V participates in the intersection of versions of T containing S' and defining $\langle S' \rangle_T$. There exists no additional version $T_{v'}$ containing S' and not present in the family V, since if a version $T_{v'}$ contains S' then according to relation (16) it must also contain S, hence $T_{v'}$ must be present in family V. Consequently, $\langle S' \rangle_T$ is the intersection of all sets in family V, hence $\langle S' \rangle_T = \langle S \rangle_T$. □

Theorem 3 and its proof are valid in all finite dimensions n.

This result implies that for an entire continuum of sets S', where the condition $\langle S \rangle_T \supset S' \supset S$ holds, the T-hulls are invariant. By a continuous change of set S into $\langle S \rangle_T$, all intermediate sets S' have the constant T-hull $\langle S \rangle_T$, as long as none of these sets S' "hangs out" from the T-hull $\langle S \rangle_T$ of the initial set S.

Note that Theorem 3 can be regarded as a generalization of Theorem 1: by taking $S' = \langle S \rangle_T$ in Theorem 3, the statement of Theorem 1 follows.

3 Comments and closing remarks

For molecular shape analysis problems with some orientation constraints, for example, if an external electric field is applied on polar molecules, the oriented T-hull approach has been proposed [2, 3]. In such cases, only those versions T_v of the reference set T are included in the intersections which fulfill the appropriate orientation constraints. For example, using the most severe orientation restriction by disallowing rotation, only translated versions of the reference set T are used in the intersections.

Alternatively, one may include reflected versions of the reference set T besides the translated and rotated versions; for chiral reference sets this implies that a larger family of versions is considered in the intersections [2, 3].

For any of these alternatives, Theorems 1–3 apply, with the same proofs as given above, where in each case the versions of the reference set T from the restricted or enlarged families are used throughout.

If both objects S and T are selected as molecular isodensity contour surfaces (MIDCO's), then the T-hulls can be regarded as "relative shape envelopes" of molecular electron density contours. Usually, T-hulls show less shape detail than the original MIDCO S, and the T-hulls of two different molecules are often more similar than their individual MIDCO's themselves. This suggests a shape classification by T-hulls, where MIDCO's of two different molecules are regarded T-similar if the T-hulls of the two MIDCO's show equivalent shape features, for example, common shape groups [3]. Note that common shape groups for the T-hulls are possible even if the shape groups of the two MIDCO' do not agree.

Within a chemical context, T-hulls have been proposed for modeling *solvent contact surfaces* in the shape analysis of solvent–solute interactions [4]. In this model, the T-hull $\langle S \rangle_T$ of a solute electron density level set S is generated with respect to a reference object T where $(-T)$ is taken as an electron density level set of the solvent molecule. Theorem 3 implies that if the solute S undergoes some limited shape change in a conformational process and takes up a new form S', then the solvent contact surface $\langle S' \rangle_T = \langle S \rangle_T$ remains invariant as long as $\langle S \rangle_T \supset S'$. *The entire continuum of conformational changes and the associated electron density shape changes* within the range $\langle S \rangle_T \supset S' \supset S$ belong to the single, constant T-hull $\langle S \rangle_T$, i.e., to the *single, constant solvent contact surface* $\langle S \rangle_T$.

Acknowledgements. This work was supported by an operating research grant from the Natural Sciences and Engineering Research Council of Canada.

References

1. Edelsbrunner H, Kirkpatrick DG, Seidel R (1983) IEEE Trans Inform Theor IT-29:551
2. Mezey PG (1991) J Math Chem 8: 91
3. Mezey PG (1993) Shape in Chemistry; an introduction to molecular shape and topology; VCH Publ., New York
4. Mezey PG (1995) Theor Chim Acta 91:67
5. Walker PD, Mezey PG (1993) J Amer Chem Soc 115:12423
6. Walker PD, Mezey PG (1994) J Amer Chem Soc 116:12022
7. Walker PD, Mezey PG (1994) Can J Chem 72:2531
8. Mezey PG (1995) in: Molecular similarity in drug design, PM Dean (ed) Chapman and Hall–Blackie Publishers, Glasgow, UK