Theor Chim Acta (1995) 91: 67–71

Theoretica Chimica Acta

The *T*-hull approach to shape analysis

Paul G. Mezey

Mathematical Chemistry Research Unit, Department of Chemistry and Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, SK, Canada S7N 0W0

Received August 10, 1994/Accepted October 10, 1994

Summary. The T-hull of both discrete point sets and continua is a generalization of the convex hull with respect to a reference object T. Various T-hulls serve as tools for shape analysis and as a basis for the introduction of a family of shape similarity measures. Some potential applications of these measures to shape problems arising in the natural sciences and a specific chemical example are discussed. The method of T-hulls is applicable for the description of molecular shape effects in solvent–solute interactions, in external electromagnetic fields and within enzyme cavities.

Key words: T-hull – Shape analysis – Solvent–solute interactions

1 Introduction

Various generalizations of the concept of convexity are useful tools for shape analysis. One such generalization is the T-hull [1] that provides both absolute and relative shape characterizations. The T-hull of a bounded set A can be regarded as a generalization of the α-hull introduced earlier by Edelsbrunner et al. [2].

The approach of T-hulls is motivated by prospective applications to a family of shape problems of chemistry and molecular physics. The central shape problem in these fields is the characterization of molecular shapes, in isolation, and with respect to some external shape constraints [3]. Some related problems are:

(i) the description of molecular shapes and similarity in external electromagnetic fields;
(ii) the similarity and complementarity of shapes of interacting molecules;
(iii) the analysis of biochemical interactions within enzyme cavities.

For simplicity, and in view of the chemical applications, we shall focus our discussion on three-dimensional T-hull problems, although it is clear that all the relevant methods have straightforward generalizations to arbitrary finite dimensions.

2 *T*-hulls, oriented *T*-hulls, and negative *T*-hulls

Consider an arbitrary, closed, three-dimensional set T, and regard it as a reference object. The shape of various other objects S will be described relative to this reference object. In some instances, the closure $\mathrm{clos}(E^3\backslash T)$ of the relative complement $E^3\backslash T$ of T is needed. For the sake of simplicity in the notation, we shall write $-T$ for the closure of the relative complement of T, that is,

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

The set $T' = -T$ can also be chosen as a reference object.

A set obtained by translation and rotation of T is called a *version* of T. In some instances, the test object T is subject to 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 [1] as the intersection of all rotated and translated versions of T which contain S. If no version of T contains S then the T-hull of S is the empty intersection, interpreted as the full space. Clearly, the T-hull of a set S depends on the shapes of both objects, S and T. The T-hull of S exhibits properties that are dependent on the relative shapes of S and T.

First, we shall describe the connection between the concepts of T-hull and α-hull. Following the original reference for the two-dimensional case of α-hulls [2], a *generalized disc of radius* $1/\alpha$ is 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 is the intersection of all closed generalized discs of radius $1/\alpha$ which contain S. If $\alpha = 0$, then the α-hull $\langle S\rangle_\alpha$ of S is the ordinary convex hull $\langle S\rangle$ of S. If one regards the empty intersection as the entire space, then for any set S the α-hull exists for any α value. For a finite point set S and a sufficiently small negative value of α the α-hull $\langle S\rangle_\alpha$ of S is the finite point set S itself.

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

The T-hull is, indeed, a simple generalization of the α-hull. If the reference object T is a sphere of radius $1/\alpha$, then the T-hull of S is the α-hull of S. If T is a sphere of radius $1/\alpha$, then the $(-T)$-hull of a set S is the α'-hull of S, where $\alpha' = -\alpha$.

Consider now orientation constraints: take a reference set T of some fixed orientation with respect to the Cartesian axes of the three-dimensional Euclidean space. The *oriented T-hull* $\langle S\rangle_{T,o}$ of a point set S has been defined [1] as the intersection of all *translated* versions of the oriented set T which contain point set S.

Consider a reference set T. Neither the T-hull $\langle S\rangle_T$ nor the oriented T-hull $\langle S\rangle_{T,o}$ of a point set S is necessarily connected, even for a connected reference set T. Let $k(\langle S\rangle_T)$ and $k(\langle S\rangle_{T,o})$ denote the number of maximum connected components of $\langle S\rangle_T$ and the number of maximum connected components of $\langle S\rangle_{T,o}$, respectively. These numbers provide information on the shape compatibility of sets S and T.

Select the reference object T as an ellipsoid. An ellipsoid is achiral, that is, the mirror image $T^\diamond$ of T is superimposable on T by translations and rotations in the three-dimensional space. Hence, by allowing *reflection* of an ellipsoid in addition to translation and rotation when generating versions of T, one does not obtain a different condition for the T-hull $\langle S\rangle_T$ of a point set S.

If, however, the reference object T is chiral, then a different generalization of T-hull is obtained if reflected versions of T are also included in the intersection, in addition to rotated and translated versions.

The $T, T^{\diamond}$-*hull* $\langle S\rangle_{T,T^{\diamond}}$ of a point set S has been defined [1] as the intersection of all rotated, translated, and reflected versions of T which contain S.

The following inclusion relations hold for the point set S, its $T, T^{\diamond}$-hull $\langle S\rangle_{T,T^{\diamond}}$, its T-hull $\langle S\rangle_T$, and its oriented T-hull $\langle S\rangle_{T,\mathrm{o}}$:

$$S \subset \langle S\rangle_{T,T^{\diamond}} \subset \langle S\rangle_T \subset \langle S\rangle_{T,\mathrm{o}}. \tag{2}$$

Furthermore, the following equalities can be proven [4]:

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

3 Similarity measures based on T-hulls

Several similarity measures have been proposed for applications in the natural sciences, based on some topological and geometrical properties of objects [3]. In principle, all these similarity measures can be applied to the T-hulls of the original objects, expressing their similarities "biased" by their shape relations to the reference object T. In the last section of this note, some applications of this approach will be outlined.

Consider a similarity measure $s(G_1, G_2)$ of objects G_1 and G_2. The corresponding similarity measures of various T-hulls,

$$s(\langle G_1\rangle_T, \langle G_2\rangle_T), \tag{4}$$

$$s(\langle G_1\rangle_{T,\mathrm{o}}, \langle G_2\rangle_{T,\mathrm{o}}), \tag{5}$$

and

$$s(\langle G_1\rangle_{T,T^{\diamond}}, \langle G_2\rangle_{T,T^{\diamond}}), \tag{6}$$

express the similarities of G_1 and G_2 with respect to the reference object T as "shape standard".

Another similarity measure is obtained by direct comparisons of volumes of the various sets obtained. If $V(A)$ is the volume of object A, then the ratios

$$r(\langle G_1\rangle_T) = (V(\langle G_1\rangle_T) - V(G_1))/V(G_1), \tag{7}$$

$$r(\langle G_1\rangle_{T,\mathrm{o}}) = (V(\langle G_1\rangle_{T,\mathrm{o}}) - V(G_1))/V(G_1), \tag{8}$$

and

$$r(\langle G_1\rangle_{T,T^{\diamond}}) = (V(\langle G_1\rangle_{T,T^{\diamond}}) - V(G_1))/V(G_1) \tag{9}$$

express the relative volume increase of the T-hulls as compared to the original object. This relative increase itself can be used as a similarity criterion; for two objects, G_1 and G_2, the smaller the difference between their respective r values $r(\langle G_1\rangle_T)$ and $r(\langle G_2\rangle_T)$,

$$\Delta r(G_1, G_2)_T = |r(\langle G_1\rangle_T) - r(\langle G_2\rangle_T)|, \tag{10}$$

the more similar their relative space demand in an environment dominated by the shape conditions specified by the reference object T. The same considerations apply to the oriented and reflected T-hulls. The significance of these similarity measures will be illustrated by a chemical example in the next section of this note.

4 Construction of molecular accessibility and contact surfaces using negative *T*-hulls

One chemical problem of current interest is the characterization of solvent–solute interactions. Some aspects of these interactions can be described by formal

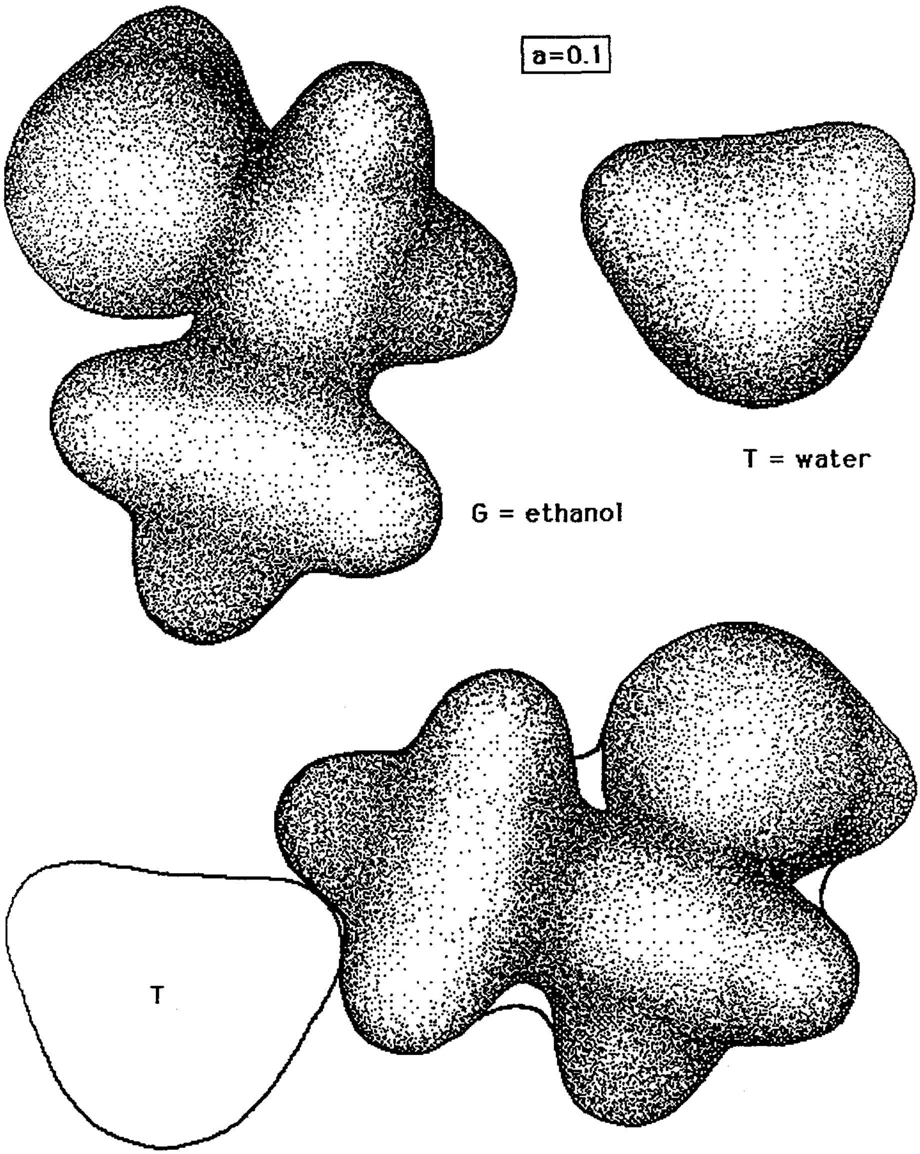

Fig. 1. The 0.1 a.u. (atomic unit) molecular isodensity contours (MIDCOs) of ethanol and water. Taking the body enclosed by the water MIDCO as reference object *T*, the ($-T$)-hull of the ethanol MIDCO *G*, shown in the lower part of the figure, illustrates the concept of water accessibility (contact) surface of the ethanol–water solute–solvent interaction

molecular surfaces, representing the closest approach of one molecule along the formal surface of another. In a simple, semiclassical model the molecules are represented by their molecular isodensity contour (MIDCO) surfaces $G(a)$ at the electronic density value $a = 0.001$ a.u. (atomic units). The associated formal molecular bodies are the corresponding $a = 0.001$ a.u. threshold level sets of the electronic density. Consider a single, fixed solute molecule A and a solvent molecule B, where molecule B is allowed to change its position. The solvent contact surface $SCS(A, B)$ of molecule A dissolved in solvent B is the boundary of the set composed from all points of the molecular body A and all additional points of the three-dimensional space not accessible by molecule B. In actual molecular calculations exact determination of these contours is seldom attempted, and an approximate Monte Carlo fitting procedure is used instead.

These molecular contact surfaces can be defined as negative T-hulls. If T is the formal body of the solvent molecule B and S is a formal molecular surface of the solute A, usually represented as an isodensity contour G [3], then the solvent contact surface is the $(-T)$-hull of S.

A spirited example is shown in Fig. 1, where the solute molecule $A = S$ is ethyl alcohol, CH_3CH_2OH, and the solvent $B = T$ is water, H_2O.

If the reference object is taken as a MIDCO of the solvent molecule, then the similarity measures $s(\langle G_1 \rangle_T, \langle G_2 \rangle_T)$ of two solute MIDCOs G_1 and G_2 express the similarity of their actual shape requirements within the solvent. The oriented case $s(\langle G_1 \rangle_{T,o}, \langle G_2 \rangle_{T,o})$ corresponds to solutions where an external orientation constraint applies, such as an external field, or local ordering such as those found in liquid crystals. The case involving reflection, $s(\langle G_1 \rangle_{T,T\diamond}, \langle G_2 \rangle_{T,T\diamond})$, is applicable to chiral solvent molecules which may easily undergo configuration inversion.

The volume-based similarity measures $\Delta r(G_1, G_2)_T$ are applicable to solute–solvent interaction problems where the actual shape features are of secondary importance, but the relative space requirements are of interest. These similarity measures are likely to have relations to similarities in solvation energies of various solutes.

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

References

1. Mezey PG (1991) J Math Chem 8:91
2. Edelsbrunner H, Kirkpatrick DG, Seidel R (1983) IEEE Trans Inform Theor IT-29:551
3. Mezey PG (1993) Shape in chemistry; an introduction to molecular shape and topology. VCH Publ., New York
4. Mezey PG (to be published)