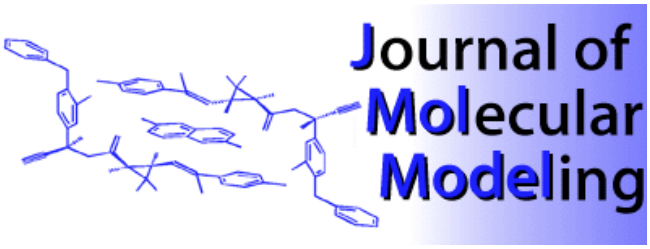
J. Mol. Model. **1999**, *5*, 116 – 124



FULL PAPER

An Optimisation of Asymmetry Evaluation of Molecules within the Folding-Unfolding Method

Socrates Alikhanidi and Victor Kuz'min

Department of Organic Chemistry, I. I. Mechnikov Odessa State University, 2 Dvoryanskaya St., Odessa 270100, Ukraine. Tel: 380-482-225127; Fax: 380-482-652012. E-mail: victor@farlep.net

Received: 23 March 1999/ Accepted: 21 April 1999/ Published: 18 June 1999

Abstract This article concerns the study of the folding-unfolding (Continuous Symmetry Measure, CSM) method of Avnir et al. (Zabrodsky, H.; Peleg, S.; Avnir, D. *J. Am. Chem. Soc.* **1992**, *114*, 7843) for quantitative evaluation of the asymmetry of molecular objects. It is shown that a series of modifications to the folding-unfolding method are expedient. An efficient solution for optimisation problems in the folding-unfolding method has been proposed. It allows a significant speed up of the calculations and provides better results. Differences in the behaviours of the original and optimised folding-unfolding methods for various molecular structures are investigated.

Keywords Methods for QSPR, Mathematical chemistry, Quantitative asymmetry

Introduction

Symmetry as one of the main characteristics of material objects causes many molecular properties. An analysis of the molecular asymmetry is often necessary for quantum-mechanical calculations, investigation of spectral properties of substances, stereochemistry of reactions and so on. Several molecular modelling programs (Spartan [1], Gaussian [2], SYMMAPS [3] etc.) allow point groups of molecular symmetry to be determined. However, for some "structure-activity" correlations it is necessary to know the quantitative asymmetry measure of molecular objects with respect to the specific symmetry elements. Here, the set of methods and programs is limited. Universal approaches for asymmetry measurement have been proposed by the authors of the folding-unfolding (CSM) method [4] and, for quantifying dissymmetry, the Dissymmetry Function method [5]. The first is considered in this article (for examples of the practical applicability of the folding-unfolding method, see ref. [6] and references therein).

Previous publications [4,7,8] are of a fundamental character and include good mathematical grounds of the main folding / unfolding principles, but there has been no detailed discussion of the practical realisation of this method.

Using the folding-unfolding method, we have concluded that modification of several steps of the original algorithm is expedient either for the method to work adequately for all models or for the accurate definition of respective steps.

Two types of optimisation tasks should be resolved. The first (combinatorial) has not been discussed previously, despite its importance, since it limits the speed of the algorithm and often requires rigorous calculations. As compared to the exhaustive algorithm, the proposed procedure is more efficient (Appendix A); it considerably decreases the vol-

Correspondence to: V. Kuz'min

ume of the calculations. The second optimisation task (minimisation of function from several variables) has been reviewed previously. However, only numerical optimisation has been proposed for 3D structures. We have solved the given problem analytically, leading to the following advantage as compared with numerical optimisation: quick detection of all extrema including the global minimum.

The original folding-unfolding method. Ways for optimisation

The essence of the discussed method is the following: the asymmetry measure S(G) of a point object P_b with respect to symmetry operation G is defined as:

$$S(G) = \frac{1}{a}\sum_{b=1}^{a}\left\|P_b - \ddot{P}_b\right\|^2 \qquad (1)$$

where a is total number of points of the object, $\ddot{P}_b$ are the corresponding points of the nearest G-symmetric object, whose determination is the task of folding-unfolding algorithm.

An attempt of application of the algorithm of the folding-unfolding method for symmetry evaluation in the case of equilateral triangle (length of side is 1.54 Å, like cyclopropane) with respect to rotation axis C_5 is given below. The object is not scaled as ref. [4] requires (so that the maximum distance of any point to the centroid is 1) since step 7 of the following algorithm allows the application of an extended set of normalisation options, *e.g.* from ref. [9]. Everywhere we mean the 3D models.

The algorithm [4,7] consists of the following steps:

1. Allocation of unequivalent points (atoms) of the object in groups. The authors of ref. [7] suggested searching for the topologically unequivalent points according to an analysis of the molecular graph isomorphism.

Example. All points are of the same type and they are located in one group (Figures 1a and 1b).

Addition. Differences in atom weights are not considered in the original algorithm. We propose to distinguish atoms by their nature and we use either the difference of the atom's chemical nature or the method of canonical labelling of molecular graphs CANON [10], which discriminates atoms easily according to the specificity of the task. It is possible to apply algorithms with more hard atom discrimination for structures with high topological symmetry, e.g. fullerenes or frame structures [11].

2. Select a position in space for the symmetry element G passing through the origin coinciding with the centre of mass. There is no instruction concerning initial orientation of the symmetry element, although this information is a very important part of the algorithm.

Concrete definition. We propose the three main inertial axes of the object be put consistently in conformity to the initial position of the symmetry element. It is proved by the fact that when an object has any symmetry axis, then this axis coincides with one of the main inertial axes [5,9]. The main inertial axes are unambiguously determined for objects, except spherical tops (three main inertia moments are the same) or cylindrical tops (two main inertia moments are same) only. Further detailed discussion is given in ref. [9].

Example. Let in both variants (Figures $1a_1$ and $1b_1$), the symmetry element be perpendicular to the triangle plane and pass through its centre. This position of the rotation axis coincides with one of the main inertia axes of the equilateral triangle.

3. Select a division of the points of the object into sets of n points (n is the symmetry number, in our example 5; n corresponds to the symmetry element order, but is not less than 2). The most general instruction is the following: all groups of points are divided into sets with numbers of points that are divisors of n. If a set contains a single point, that point is multiplied n times; if a set contains two points, each is multiplied $n/2$ times; . . .; if a set contains n points, nothing is multiplied.

Example, original algorithm. There is a single way only: each point is multiplied five times after location in different sets (upper symbol means the multiplicability factor): $\left\{P_1^{\mathrm{I}}, P_1^{\mathrm{II}}, P_1^{\mathrm{III}}, P_1^{\mathrm{IV}}, P_1^{\mathrm{V}}\right\}$, $\left\{P_2^{\mathrm{I}}, P_2^{\mathrm{II}}, P_2^{\mathrm{III}}, P_2^{\mathrm{IV}}, P_2^{\mathrm{V}}\right\}$, $\left\{P_3^{\mathrm{I}}, P_3^{\mathrm{II}}, P_3^{\mathrm{III}}, P_3^{\mathrm{IV}}, P_3^{\mathrm{V}}\right\}$ (Figure $1a_1$).

Modification. We propose that the following set formation is expedient: all groups of points are divided into sets, where each point of a group can appear 0, 1, 2, …, n times in each set (each set must have exactly n points, the point order in sets is important). This process corresponds to "sampling with replacement" (the combinatorial mathematics term).

Example, modified algorithm. The following division can be used (task of the selection of the best sets is resolved by the procedure given in Appendix A, or by the exhaustive approach): $\left\{P_1^{\mathrm{I}}, P_1^{\mathrm{II}}, P_3^{\mathrm{I}}, P_2^{\mathrm{I}}, P_2^{\mathrm{II}}\right\}$ (Figure $1b_1$).

4. Fold each set of points by applying to each point of the ordered set the G symmetry operation (j-1) times, where j is the order number of this point in the set.

Example, original algorithm. Obtained points are: $\left\{\dot{P}_1^{\mathrm{I}}, \dot{P}_1^{\mathrm{II}}, \dot{P}_1^{\mathrm{III}}, \dot{P}_1^{\mathrm{IV}}, \dot{P}_1^{\mathrm{V}}\right\}$, $\left\{\dot{P}_2^{\mathrm{I}}, \dot{P}_2^{\mathrm{II}}, \dot{P}_2^{\mathrm{III}}, \dot{P}_2^{\mathrm{IV}}, \dot{P}_2^{\mathrm{V}}\right\}$, $\left\{\dot{P}_3^{\mathrm{I}}, \dot{P}_3^{\mathrm{II}}, \dot{P}_3^{\mathrm{III}}, \dot{P}_3^{\mathrm{IV}}, \dot{P}_3^{\mathrm{V}}\right\}$ (Figure $1a_2$, where the first set is shown only).

Example, modified algorithm. Obtained points are: $\left\{\dot{P}_1^{\mathrm{I}}, \dot{P}_1^{\mathrm{II}}, \dot{P}_3^{\mathrm{I}}, \dot{P}_2^{\mathrm{I}}, \dot{P}_2^{\mathrm{II}}\right\}$ (Figure $1b_2$).

5. Average each set of folded points. From each set the single point multiplied n times is obtained.

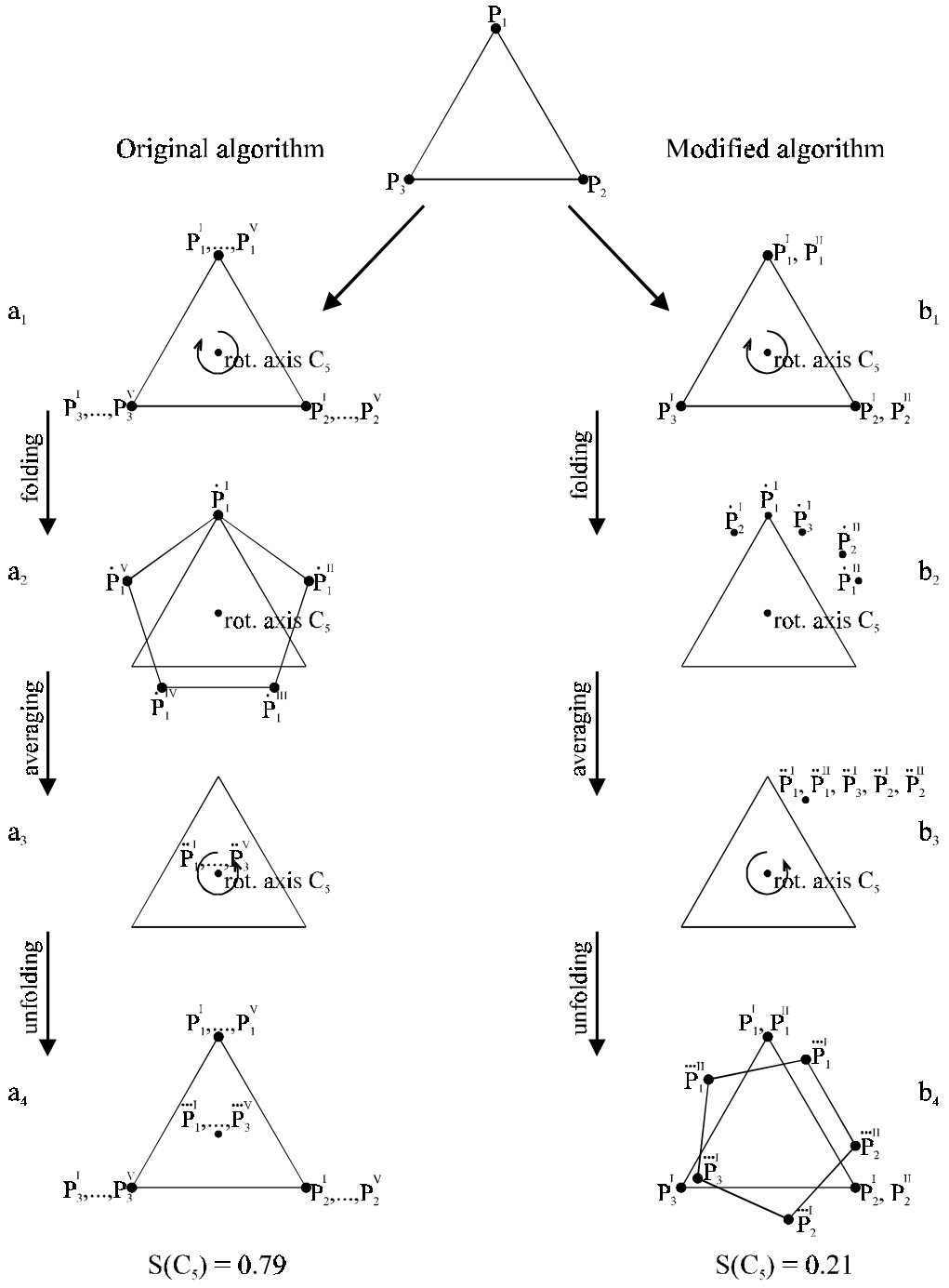


Figure 1 *Scheme of the determination of asymmetry measure of the equilateral triangle with respect to the rotation symmetry axis C_5 using the original (a_1-a_4) and the optimised (b_1-b_4) folding-unfolding methods. Subfigures a_2 and b_2 contain the folded points $\dot{P}_x^y$ of the original triangle. Subfigures a_4 and b_4 contain the unfolded points $\ddot{P}_x^y$ forming the nearest C_3-symmetric object to the original triangle.*

Example, original algorithm. Obtained points are located in the centre: $\{\ddot{P}_1^{I},\ddot{P}_1^{II},\ddot{P}_1^{III},\ddot{P}_1^{IV},\ddot{P}_1^{V}\}$, $\{\ddot{P}_2^{I},\ddot{P}_2^{II},\ddot{P}_2^{III},\ddot{P}_2^{IV},\ddot{P}_2^{V}\}$, $\{\ddot{P}_3^{I},\ddot{P}_3^{II},\ddot{P}_3^{III},\ddot{P}_3^{IV},\ddot{P}_3^{V}\}$ (Figure 1a$_3$).

Example, modified algorithm. Obtained points are: $\{\ddot{P}_1^{I},\ddot{P}_1^{II},\ddot{P}_3^{I},\ddot{P}_2^{I},\ddot{P}_2^{II}\}$ (Figure 1b$_3$).

6. Unfold each set of points by applying the G^{-1} symmetry operation (reverse operation to G; $(G^{-1})^{-1}=G$) to each averaged point from the ordered set (j-1) times, where j is the order number of this averaged point in the set. This process forms a figure with the G symmetry element.

Example, original algorithm. Obtained points are located in the centre: $\{\ddot{\ddot{P}}_1^{I},\ddot{\ddot{P}}_1^{II},\ddot{\ddot{P}}_1^{III},\ddot{\ddot{P}}_1^{IV},\ddot{\ddot{P}}_1^{V}\}$, $\{\ddot{\ddot{P}}_2^{I},\ddot{\ddot{P}}_2^{II},\ddot{\ddot{P}}_2^{III},\ddot{\ddot{P}}_2^{IV},\ddot{\ddot{P}}_2^{V}\}$, $\{\ddot{\ddot{P}}_3^{I},\ddot{\ddot{P}}_3^{II},\ddot{\ddot{P}}_3^{III},\ddot{\ddot{P}}_3^{IV},\ddot{\ddot{P}}_3^{V}\}$ (Figure 1a$_4$).

Example, modified algorithm. Obtained points are: $\{\ddot{\ddot{P}}_1^{I},\ddot{\ddot{P}}_1^{II},\ddot{\ddot{P}}_3^{I},\ddot{\ddot{P}}_2^{I},\ddot{\ddot{P}}_2^{II}\}$ (Figure 1b$_4$).

7. Difference between configurations obtained on steps **3** and **6** (i.e., distance between initial object and its symmetric analogue) is defined according to:

$$S = \frac{\sum_{k=1}^{M}\sum_{j=1}^{N_k'}\left\| {}_j^k P - {}_j^k \ddot{\ddot{P}} \right\|^2}{\sum_{k=1}^{M} N_k'} \qquad (2)$$

where S is the asymmetry value at the current division of object into sets (see step **3**) and at the current orientation of symmetry element, M is the number of groups of points (see step **2**), N_k' is the number of points in group k after possible multiplications of some points ($N_k' \geq N_k$, N_k' is multiple of n, N_k is number of points in group k).

Example, original algorithm. For current orientation of C_5 axis: $S(C_5) = 0.79$.

Modification. In previous reports [4,7,8] all objects were considered with unweighted points but simplification is often unacceptable for real tasks. We propose the following more general variant of expression (2):

$$S = \frac{\sum_{k=1}^{M} w_k \sum_{j=1}^{N_k'}\left\| {}_j^k P - {}_j^k \ddot{\ddot{P}} \right\|^2}{\sum_{k=1}^{M} w_k N_k'} \qquad (3)$$

where w_k is the weight factor of group k. Several types of normalisations of the value S (3) have recently been suggested [9].

Example, modified algorithm. For current orientation of C_5 axis: $S(C_5) = 0.21$.

8. Minimise the asymmetry value S by repeating steps **2-7** with all possible divisions of points into sets and for all possible orientations of the symmetry element.

Concrete definition. The consecutive scheme for realisation of the given step was not described [4,7,8]. The following iterative circle for the algorithm is proposed:

a) selection of the initial orientation for symmetry element;

b) determination of the best sets combination (with minimal value S) for the current orientation of the symmetry element;

c) determination of the best orientation for the symmetry element for the current sets combination;

d) return to the step **b)** in case of reduction of value S.

Comment to step b). It is a very complicated combinatorial problem since the large number of equivalent atoms require numerous calculations. Previous reports have not discussed this problem. Therefore, we suppose that an exhaustive procedure has been used. In Appendix A, the accelerated tool is demonstrated..

Comment to step c). The authors of ref. [4] have demonstrated that the necessary condition for the optimal position of the symmetry element is when it passes through the object centroid (this was also checked and applied for the weighted objects). Attempts to solve the problem of optimisation of symmetry element orientation for the current sets combination were undertaken [4]. This problem was solved analyti-

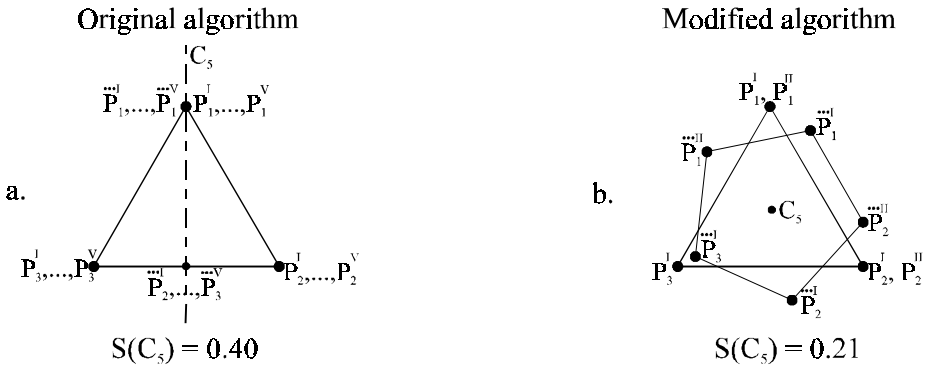


Figure 2 *The final results of the process shown in Figure 1 after checking all possible divisions of points into sets and all initial positions of the rotation axis C_5. It is shown that the asymmetry measure is less for the modified algorithm (b) and its "nearest" symmetric object is closer and preferable.*

cally for 2D objects, but only numerical methods have been proposed for 3D. However, most of the real tasks concern 3D objects, but the best result can be missed with such a numerical approach. We have resolved the given problem analytically for the general 3D weighted objects (see Appendix B).

Example, original algorithm. The best orientation of symmetry element C_5 is shown in Figure 2a. $S(C_5) = 0.40$.

Example, modified algorithm. The best orientation of symmetry element C_5 and the best division of points into sets are shown in Figure 2b. $S(C_5) = 0.21$.

Discussion. It is seen that the asymmetry measure is less for the modified algorithm (b) and its "nearest" symmetric object is closer and preferable. This difference is the result of modifying step **3** of the folding-unfolding algorithm.

Results and discussion

The efficiency of each optimisation procedure is demonstrated individually by comparison of the original and the optimised folding-unfolding algorithms. In the following, the original method indicates the folding-unfolding method, where step **3** of the above-mentioned algorithm is without modification, the combinatorial problem is solved by the trivial exhaustive approach, and the orientation of the symmetry element is optimised by the numerical approach. In contrast, the optimised technique means the folding-unfolding method, where step **3** of the above-mentioned algorithm is modified as described in modification of step **3**, the combinatorial problem is solved by the procedure described in Appendix A, and the orientation of the symmetry element is optimised by the analytical approach from Appendix B. Atoms are distinguished by their chemical nature in both folding-unfolding methods.

I. Modification of the step 3 of folding-unfolding algorithm leads to the closer "nearest symmetric shape" and, as a result, the asymmetry value decreases.

An example is shown in Figure 1 and 2 and other examples are listed in Table 1. Since the given modification affects cases in which the symmetry number is more than 2 (i.e., all elements, except C_2, S_1, S_2), the difference in asymmetry values only occurs for these cases. Thus, the modification is more efficient when the symmetry number is larger.

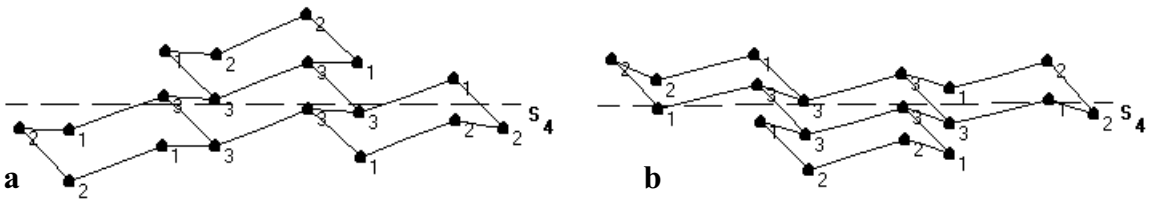
II. Optimisation of the combinatorial problem (see Appendix A). This modification allows the acceleration of calculations of the results as compared with the exhaustive approach.

Since solution of the given combinatorial problem limits the speed of calculation of the entire algorithm, we can analyse just the whole time of the calculation. In Table 1, the calculation time for several molecular structures by both the original and the optimised folding-unfolding methods are shown.

Thus, a larger number of equivalent atoms requires much more calculation time, but the proposed optimisation procedure (see Appendix A) significantly accelerates the calculation (by a 1000 times, approximately, for objects with a large number of equivalent points) as compared with the exhaustive approach of the original folding-unfolding algorithm. The shortest time of calculation as compared with the original algorithm is for the symmetry elements S_1, C_2 (and, in fact, S_2). It is caused by the application of a special, very fast, combinatorial algorithm of the *P* combinatorial class for these symmetry elements, which accelerates the calculation exceptionally - in $>>10^5$ times.

III. Optimisation of problem of the best orientation of symmetry element (see Appendix B). This allows us to avoid missing some minima of asymmetry values.

An example, where the better solution could be missed (case b) with the original, numerical approach to searching for the optimal orientation of the symmetry element, is given in the Figure 3. The optimised analytical approach of Appendix B overcame this problem (case a).



$S(S_4) = 2.1014$ $S(S_4) = 2.1125$

Figure 3 *Example of the efficiency of the analytical solution (a) of the problem of the symmetry element orientation compared with the numerical solution (b), which allows to avoid the missing of some minima of asymmetry values S. There are the* trans-trans-trans-*perhydrotriphenylene, S_4 symmetry element, the topological location of atoms in groups based on the CANON [10] algorithm. The atom labels denote the group numbers, where atoms are located. For clarity, the single modified step **3** of the folding-unfolding algorithm has been used in both cases.*

Table 1 *The analysis of the difference in the evaluated unweighted asymmetry S values (numerators in cells having fractions) and the calculation time (denominators in cells having fractions, in seconds) between the original (M1) and optimised (M2) folding-unfolding methods. The absence of data means too long calculation time (>10 hours by Pentium 66 processor).* N_{max} *is the maximal number of equivalent atoms in molecules. The discrimination of atoms is based on their chemical nature*

Structure	N_{max}	symm. point group	method	C_2 axis	S_1 axis	C_3 axis	S_4 axis
milk acid	6	C_1	M1	.335/ 6 s	.175/ 3 s	.743/ 6 s	1.209/ 8 s
			M2	.335/ 4 s	.175/ 5 s	.567/ 7 s	1.149/ 5 s
glucose	12	C_1	M1	.503/ 575 s	.213/ 500 s	.876/ 540 s	1.170/ 2205 s
			M2	.503/ 6 s	.213/ 6 s	.789/ 9 s	1.141/ 27 s
phenanthrene	14	C_{2V}	M1	.000/ 4500 s	.000/ 4850 s	–	–
			M2	.000/ 2 s	.000/ 2 s	1.26/ 450 s	1.320/ 125 s
tetrahelicene	18	C_2	M1	–	–	–	–
			M2	.000/ 8 s	.150/ 7 s	.860/ 8905 s	.955/ 4535 s
fullerene C_{76}	76	D_2	M1	–	–	–	–
			M2	.000/ 12 s	.151/ 36 s	–	–

Conclusion

The optimised algorithm of the folding-unfolding method runs more exactly and more quickly due to the proposed two optimisation procedures and some modifications.

The optimised and original folding-unfolding (CSM) methods, together with methods for the overall chirality measurement [12] and the Dissymmetry Function [5] were included into the program DisFact [13]. It allows the calculation of different asymmetries for various types of molecular structures.

Acknowledgment We are grateful to Prof. D. Avnir (Hebrew University of Jerusalem, Israel) and Dr. Sci. S. S. Tratch (Moscow State University, Russia) for useful discussion of the folding-unfolding method. Financial support from INTAS foundation through the project INTAS-UA 97-17 30 is gratefully acknowledged.

References and notes

1. Spartan software. Wavefunction, Inc, 18401 Von Karmen Ave., Suite 370, Irvine, CA 92612 USA. *http://www.wavefun.com*
2. Gaussian software. Gaussian, Inc, Carnegie Office park, Building 6, Suite 230, Carnegie, PA 15106 USA. *http://www.gaussian.com*
3. SymApps software. Bio-Rad Laboratories, Sadtler Division, U.S. Headquarters, 3316 Spring Garden Street, Philadelphia, PA 19104-2596 USA. *http://www.softshell.com*
4. Zabrodsky, H.; Peleg, S.; Avnir, D. *J. Am. Chem. Soc*. **1992**, *114*, 7843.
5. Kuz'min, V. E.; Stel'makh, I. B.; Bekker, M. B.; Pozigun, D. V. *J. Phys. Org. Chem.* **1992**, *5*, 295.
6. Pinto, Y.; Salomon Y.; Avnir, D. *J. Math. Chem. Soc*. **1998**, *23*, 13.
7. Zabrodsky, H.; Avnir, D. *J. Am. Chem. Soc*. **1995**, *117*, 462.
8. Zabrodsky, H.; Peleg, S.; Avnir, D. *J. Am. Chem. Soc*. **1993**, *115*, 8278.
9. Alikhanidi, S. E.; Kuz'min, V. E. *Zh. Strukt. Khim.* **1998**, *39*, 543.
10. Ugi, I.; Dugundij, J.; Kopp, R.; Marquarding, D. *Perspectives in Theoretical Stereochemistry*; Springer-Verlag: Berlin, Heidelberg, 1984.
11. Laidbocur, T.; Cabrol-Bass, D.; Ivanciuc, O. *J. Chem. Inf. Comput. Sci*. **1996**, *36*, 811.
12. Alikhanidi, S. E.; Kuz'min, V. E. *Reports of NAS of Ukraine*, **1999**, *3*, 138.
13. DisFact software for quantitative analysis of chirality and specific asymmetry of molecular structures: version 3.10 of 1999 year, authors are S. Alikhanidi and V. Kuz'min from Odessa State University of Ukraine. For free demo version and more information, *http://www.chat.ru/~theochem/disfact.htm*
14. Hu, T. C. *Integer Programming and Network Flows*; Addison-Wesley: Menlo Park, California, London and Don Mills, Ontario, 1970.
15. Christofides, N. *Graph Theory. An Algorithmic Approach*; Academic: New York, London and San Francisco, 1975.
16. Balas, E. *Oper. Res.* **1965**, *13*, 517.

Appendix A: Determination of the optimal combination of sets of points for current orientation of the symmetry element

The expression numerator (3) corresponds to the non-normalised asymmetry measure. Therefore, it is possible to be limited by optimisation of this numerator only, since optimisation of entire fraction (3) requires solution of an extremely difficult task of nonlinear programming.

The overall task for the whole object can be divided into M (number of groups of points) smaller tasks (for each group individually) since the construction of sets from points of different types (groups) is forbidden. Now we work with each group $k \in \overline{[1, M]}$ individually.

Let us consider that we have generated all possible sets of points (their number is N_{row}) according to the requirement of step **3** of folding-unfolding algorithm (e.g. for modified step **3**, $N_{row} = \sum_{i=1}^{\min(n,N)} C_N = \sum_{i=1}^{\min(n,N)} \frac{N!}{i!(N-i)!}$ where N is number of points in group k). The problem we want to solve is, how could we find the such combination of sets containing all points of the group (one more constraint may require each point to be located in one set only), which has the lowest asymmetry value?

The trivial exhaustive algorithm solves this problem in 2^{Nrow} iterations. Since very often N_{row} is >100, this approach requires too many calculations (>>10^{30} iterations) for the exact solution.

We formulated this problem as a Boolean programming task [14] to minimise:

$$S_{\text{nomin.}} = \sum_{l=1}^{N_{row}} \Delta_l x_l \qquad (4)$$

under constraints:

$$\sum_{l=1}^{N_{row}} a_{jl} x_l = 1, \qquad j = 1,\ldots,N, \qquad (5)$$

or

$$\sum_{l=1}^{N_{row}} a_{jl} x_l \geq 1, \qquad j = 1,\ldots,N, \qquad (6)$$

$$1 \geq x_l(\text{integer}) \geq 0, \quad l = 1,\ldots,N_{row},$$

where N is the number of points in group k; N_{row} is the number of all available sets of points in group k according to the requirement of step **3** of the folding-unfolding algorithm; Δ_l is the nonnormalised asymmetry value of set l ($\Delta_l = \sum_{i=1}^{n} \left\| P_i - \ddot{\ddot{P}}_i \right\|^2$ - over all points of set l where i is number of point of set); x_l is the Boolean variable, which is equal either 1 or 0 that is either the set l is taken into account in the solution or not; parameter $a_{jl} = 1$ if set l includes point j, otherwise $a_{jl} = 0$; n is the symmetry number of given symmetry element.

Constraint (5) corresponds to the original folding-unfolding algorithm (each point must be only in one set of the optimal solution) and is also accepted in its optimised version, but the alternative constraint (6) (each point must be at least in one set of the optimal solution) can also be useful for some tasks. The choice of constraints depends on the user.

Generally, the Boolean programming tasks (the particular case of integer programming) are the difficult to solve *NP*-full tasks [14], i.e.,. methods for solution have exponential working time from the dimension of task (N_{row} in our case). In this connection, attempts to separate the particular cases for solving the polynomial (tasks of *P*-class) are expedient.

For the case of n=2 and constraint (5), we succeeded in formulating the given combinatorial problem as the task of "minimal cost matching in general graph" [15]. This solution algorithm requires a polynomial working time. For the formulation of the task the full graph ($2N$ vertices) should be described. The weight of each edge of the graph depends on the end vertices x and y ($x, y \in \overline{[1, 2N]}$) of that edge by the following: (first one true item only)

a) ∞, if x=y; b) $\left\| {}^{k}_{x}P - {}^{k}_{y}\ddot{\ddot{P}} \right\|^2$, if x <=N and y <=N; c) $\left\| {}^{k}_{x}P - {}^{k}_{x}\ddot{\ddot{P}} \right\|^2$, if $y - N = x$; d) $\left\| {}^{k}_{y}P - {}^{k}_{y}\ddot{\ddot{P}} \right\|^2$, if $x - N = y$; e) 0, if x >N and y >N; f) ∞ else.

Application of this *P*-solvable task accelerates the computation extremely and allows us to calculate exactly the structures with numerous equivalent points: e.g. fullerenes with hundreds of atoms (N_{row} is near ten thousand and over).

The additive algorithm of Balas [14,16] has been used for solution of the general case of task (4) (*NP*-full task). While the asymptotic time for this algorithm is the same as for the exhaustive algorithm (2^{Nrow} iterations), usually the first algorithm works much more rapidly.

The above-mentioned additive algorithm of Balas and the algorithm for solving the task of minimal cost matching in general graph are too bulky to be described here, but are well explained in the corresponding references [14,15,16].

Appendix B: Determination of the optimal orientation of symmetry element for the current combination of sets of points

The proposed structure of the original algorithm [4] (folding, averaging, unfolding and following "comparison" (step 7 of folding-unfolding algorithm) of the obtained points with initial points) has been probably described for the visual performance of the search process for the asymmetry measure. From the computing point of view, another structure (identical to first by result) is more acceptable: folding, averaging and "comparison" of the obtained points with folded ones. Then, for the current combination of sets, the numerator of

value S (3) (denominator is fixed in this procedure) can be rewritten as the following:

$$S_{nomin.} = \sum_{k=1}^{m} w_k \sum_{i=1}^{n} \left\| \frac{\sum_{i\_=1}^{n} \dot{P}_{i\_}^{k}}{n} - \dot{P}_i^k \right\|^2 \qquad (7)$$

where *m* is the total number of sets into which the object is divided; *i* and *i_* are different variables meaning numbers of points in set *k*.

The problem we want to solve is, how could we find the best orientation of the symmetry axis (for current combination of sets)?

For this purpose let us consider that the current orientation of the symmetry element is put along Z crossbar axis and the centre of mass coincides with the origin. Then the *x*, *y*, *z* components of the folded points are the following:

$$\begin{aligned} \dot{P}_{xj}^{\ k} &= \overline{P}_{xj}^{\ k} c_j + \overline{P}_{yj}^{\ k} s_j \\ \dot{P}_{yj}^{\ k} &= -\overline{P}_{xj}^{\ k} s_j + \overline{P}_{yj}^{\ k} c_j \\ \dot{P}_{zj}^{\ k} &= \overline{P}_{zj}^{\ k} k_j \end{aligned} \qquad (8)$$

where:

$$\begin{aligned} c_j &= \cos\left(2\frac{\pi(j-1)}{n'}\right), \\ s_j &= \sin\left(2\frac{\pi(j-1)}{n'}\right), \\ k_j &= \begin{cases} (-1)^{j-1} \ for \ S_{n'}, \\ 1 \ for \ C_{n'}; \end{cases} \end{aligned} \qquad (9)$$

where j=1,2,..., n is number of a point in set k, n' is the symmetry element order (*i.e.* lower symbol of the symmetry element name); points $\overline{P}_j^k$ of rotated object are defined as:

$$\overline{P}_j^k = AP_j^k \qquad (10)$$

where rotation matrix *A* is:

$$A = \begin{pmatrix} 0 & \frac{-Z}{\sqrt{Y^2+Z^2}} & \frac{Y}{\sqrt{Y^2+Z^2}} \\ \frac{Y^2+Z^2}{\sqrt{Y^2+Z^2}} & \frac{-XY}{\sqrt{Y^2+Z^2}} & \frac{-XZ}{\sqrt{Y^2+Z^2}} \\ X & Y & Z \end{pmatrix} \qquad (11)$$

where *X*, *Y*, *Z* are directional cosines determining orientation of the given symmetry element. For creation of the rotation matrix *A*, the symmetry element has been put along the Z crossbar axis and other matrix elements have been chosen under constraints of the rotation matrix, i.e., all their vectors (crossbar axes) must be orthonormal (det *A*=1).

Now we try to find the best *X*, *Y*, *Z* values analytically. After substitution of equation (8) - (11) into equation (7) and further bulky simplifications, we can obtain the following expression:

$$\begin{aligned} S_{nomin.} = {} & (A11 - A6 + A1 - A10)Y^2 + \\ & + (A12 - A9 + A1 - A10)Z^2 + (A13 - A2)XY + \\ & + (A14 - A3)XZ + (A15 - A7)YZ + \\ & - A8X - A5Y - A4Z + (A6 + A9 + A10) \end{aligned} \qquad (12)$$

where:

$$\begin{aligned} A1 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} [(a1_i^k)^2 + (a4_i^k)^2], \\ A2 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} 2[a1_i^k a2_i^k + a4_i^k a5_i^k], \\ A3 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} 2[a1_i^k a3_i^k + a4_i^k a6_i^k], \\ A4 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} 2[a2_i^k a4_i^k - a5_i^k a1_i^k], \\ A5 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} 2[a1_i^k a6_i^k - a4_i^k a3_i^k], \\ A6 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} [(a2_i^k)^2 + (a5_i^k)^2], \\ A7 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} 2[a2_i^k a3_i^k + a5_i^k a6_i^k], \\ A8 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} 2[a5_i^k a3_i^k - a2_i^k a6_i^k], \\ A9 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} [(a3_i^k)^2 + (a6_i^k)^2], \end{aligned} \qquad (13)$$

$$\begin{aligned} A10 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} (a7_i^k)^2, \\ A11 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} (a8_i^k)^2, \\ A12 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} (a9_i^k)^2, \\ A13 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} 2a7_i^k a8_i^k, \\ A14 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} 2a7_i^k a9_i^k, \\ A15 &= \sum_{k=1}^{m} w_k \sum_{i=1}^{n} 2a9_i^k a8_i^k; \end{aligned}$$

and:

$$
\begin{aligned}
a1_i^k &= \frac{\sum_{i\_=1}^{n} x_{i\_}^k c_{i\_}}{n} - x_i^k c_i, \quad a6_i^k = \frac{\sum_{i\_=1}^{n} z_{i\_}^k s_{i\_}}{n} - z_i^k s_i, \\
a2_i^k &= \frac{\sum_{i\_=1}^{n} y_{i\_}^k c_{i\_}}{n} - y_i^k c_i, \quad a7_i^k = \frac{\sum_{i\_=1}^{n} x_{i\_}^k k_{i\_}}{n} - x_i^k k_i, \\
a3_i^k &= \frac{\sum_{i\_=1}^{n} z_{i\_}^k c_{i\_}}{n} - z_i^k c_i, \quad a8_i^k = \frac{\sum_{i\_=1}^{n} y_{i\_}^k k_{i\_}}{n} - y_i^k k_i, \\
a4_i^k &= \frac{\sum_{i\_=1}^{n} x_{i\_}^k s_{i\_}}{n} - x_i^k s_i, \quad a9_i^k = \frac{\sum_{i\_=1}^{n} z_{i\_}^k k_{i\_}}{n} - z_i^k k_i; \\
a5_i^k &= \frac{\sum_{i\_=1}^{n} y_{i\_}^k s_{i\_}}{n} - y_i^k s_i,
\end{aligned}
\tag{14}
$$

All extrema of expression (12) in respect with *X*, *Y*, *Z* can be found analytically using Lagrange multipliers under constraint $X^2+Y^2+Z^2=1$ (values *A*1-*A*15 are constants). The global minimum found corresponds to directional cosines of the optimal orientation of the given symmetry element for the current combination of sets of points.