ELSEVIER

Journal of Molecular Structure: THEOCHEM 725 (2005) 145–149

THEO CHEM

www.elsevier.com/locate/theochem

Pathway of the conformational transitions in flexible molecules

T. Körtvélyesi[a,*], L.L. Stachó[b], Gy. Dömötör[a], B. Jójárt[c], M.I. Bán[a]

[a]*Department of Physical Chemistry, University of Szeged, H-6701 Szeged, P.O. Box 105, Hungary*
[b]*Bolyai Institute for Mathematics, University of Szeged, Hungary*
[c]*Department of Pharmacodynamics and Pharmacokinetics, University of Szeged, Hungary*

Received 30 October 2004; accepted 28 February 2005
Available online 13 June 2005

Abstract

The pathways of the conformational transitions in flexible molecular systems were studied by the **DDRP** (Dynamically Determined Reaction Path) method implemented in TINKER molecular modelling package. Our first model systems were conformational transitions in small organic molecules (determination of rotational transitions, boat-chair transitions, etc.). The method was found to be very effective in finding such transitions and the location of the transition state. In comparison with other methods (modified Elber and Czerminski (**EC**) method implemented also in TINKER), similar results were obtained. For the conformational change of di- and small peptide fragments also acceptable paths and barriers were obtained. To determine the conformational transition in large peptides (i.e. folding), the effectivity of the procedure demands further studies.

Keywords: Conformational transition; Molecular mechanics; Reaction path; Search for global path

1. Introduction

To find the stationary points (equilibrium geometries and transition states) on the potential energy surface (**PES**) is a central problem in the theoretical study of chemical reactions. It is also important to know the reaction paths (**RPs**) between the reactant(s) and product(s) through one (or more) transition states. In some cases, there are possibilities for the formation of several products in different pathways. The **RP**, therefore, is a sequence of steepest descent paths (**SDP**) that joins the minima pertaining to the stable states of reactants and products through saddle points (**SPs**). Fukui introduced the concept of the intrinsic reaction coordinate (**IRC**), which is an **SDP** in mass-weighted Cartesian coordinates, assuming that the transitions from reactant to the product(s) take place infinitely slowly [1]. Though the indirect or global methods e.g. the Elber–Karplus (**EK**) method [2] and its improvements [3–7] and the 'dynamically defined reaction path' (**DDRP**) method [8–18] to find **RP**(s) are expensive related to local **RP** following methods, their use may have adequate ground because exploring the network of reaction routes, connecting points and coupling branches of various ramifying **RP**s can be successful only by global methods. The stability of global methods is also an important aspect in their use. **DDRP** method [8–18] was proved to be stable not only in small molecular systems [15–18] but for larger systems [19], too. We investigated the possibility of the determination of transition paths at small and large flexible systems with energy functions determined by molecular mechanics (**MM**) procedures.

Throughout this paper we focus on calculations of Fukui type **RP**s [1] in conformational transitions of flexible molecules. Given a system $A_1,\ldots,A_N$ of atoms, and denoting by $U(x_1, y_1, z_1,\ldots,x_N, y_N, z_N)$ its energy by Born–Oppenheimer approximation, when the atoms A_k are located at the points with Cartesian coordinates (x_k,y_k,z_k), $(k=1, 2,\ldots,N)$, respectively, a piecewise differentiable curve c: $[0,1]\to\mathbf{R}^{3N}$ is called a **RP** for the system $(A_1,\ldots,A_N)$ if the starting and end points ($c(0)$, respectively, $c(1)$) are locations of minima of the function U and the differential vector $dc(s)/ds$ is parallel to the mass weighted gradient $\nabla U(c(s))$ at every point $c(s)$ of differentiability of c while we have stationary points (**SP**s or local minima or maxima) at the breaking points of c. It is a remarkable fact

* Corresponding author. Tel.: +36 62 544612; fax: +36 62 544652.
E-mail address: kortve@chem.u-szeged.hu (T. Körtvélyesi).

0166-1280/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.theochem.2005.02.089

[14] that the flow F of the vector field of the (mass weighted) negative gradient $-\nabla U$ governs almost every piecewise analytic curve (polygons in particular) in some reaction paths under rather weak and realistic mathematical conditions on the function U. By the flow F of $-\nabla U$, we mean the mapping $F(t, x_1, y_1, z_1,\ldots,x_N, y_N, z_N) \equiv$ [solution of $d\mathbf{x}(\tau)/d\tau = -\nabla U(\mathbf{x}(\tau))$ with initial value $\mathbf{x}(0)=(x_1, y_1, z_1,\ldots,x_N, y_N, z_N)$ at time $\tau=t$].

This theoretical observation can be utilized in the following idealized algorithm. Starting from a polygon p: $[0,1]\rightarrow\mathbf{R}^{3N}$ (a set of consecutive line segments in the 3N-dimensional space parametrized with the values between 0 and 1), compute the curves

$$c_t(s) \equiv F(t, p(s)) \quad (t \geq 0, 0 \geq s \leq 1)$$

and form their arc-length proportional reparametrizations c_t^*: $[0,1]\rightarrow\mathbf{R}^{3N}$. The uniform limit of the curves results a **RP**. There is a **RP** c^*: $[0,1]\rightarrow\mathbf{R}^{3N}$ of $(A_1,\ldots,A_N)$ such that

$$c_t^* \rightarrow c^* \text{ as } t \rightarrow \infty$$

Also, generally, curves of the type $C_t^{\varphi}(s) \equiv F(\varphi(t,s), p(s))$ converge to a **RP** if φ is smooth and $\varphi(t,s)\rightarrow\infty$ for fixed s as $t\rightarrow\infty$.

The idealized procedure described, provides correct **RP**s between two predescribed local minima of U (corresponding to reactants and products of the reactions), even if there is no so-called minimum energy **RP** (a path passing through 'valley bottoms' of the graph of U) between reactants and products.

The ideal algorithm was realized in computer calculations as follows. Usually we are given good approximations $\mathbf{x}^0 \equiv (\xi_1^0, \eta_1^0, \zeta_1^0, \ldots, \xi_N^0, \eta_N^0, \zeta_N^0)$ and $\mathbf{x}^1 \equiv (\xi_1^1, \eta_1^1, \zeta_1^1, \ldots, \xi_N^1, \eta_N^1, \zeta_N^1)$ for the coordinates of reactants and products. Then for the starting polygon p: $[0,1]\rightarrow\mathbf{R}^{3N}$ we take the straight line segment

$$p(s) := (1-s)\mathbf{x}^0 + s\,\mathbf{x}^1$$

During the procedure we shall represent curves simply by a sequence of its points according to the ordering by parametrization such that consecutive points should be approximately in a given distance $\varepsilon>0$. A curve $[0,1]\rightarrow\mathbf{R}^{3N}$ can be represented by any sequence $\mathbf{y}^k \equiv (\xi_1^k, \eta_1^k, \zeta_1^k, \ldots, \xi_N^k, \eta_N^k, \zeta_N^k)$ $(k=1,\ldots,m)$ such that there exist parameters $0=s^1 < s^2 < \ldots < s^{m-1} < c^m = 1$ and $\mathbf{y}^k = c(s^k)$ $(k=1,\ldots,m)$ and $\text{dist}_M(\mathbf{y}^{k-1}, \mathbf{y}^k)$ $(\varepsilon/2, \varepsilon)$ $(k=2,\ldots,m)$. dist_M stands for the standard mass-weighted distance of points in the configuration space $\mathbf{R}^{3N}$. Any sequences of points $x^1,\ldots,x^{\ell}\in\mathbf{R}^{3N}$ obtained were homogenized by means of the following procedure. This homogenization begins with a root mean square (**RMS**) fitting of the consecutive points. We find the translate $\tilde{x}^2$ of $\mathbf{x}^2$ being the closest in dist_M to $\mathbf{x}^1$. Then we find the translate $\tilde{x}^3$ of $\mathbf{x}^3$ being the closest in dist_M to $\mathbf{x}^2$ and so on up to establishing the translate $\tilde{x}^{\ell}$ of $\mathbf{x}^{\ell}$ as the closest point in dist_M to $\mathbf{x}^{\ell-1}$. Having constructed the RMS-fit sequence $\tilde{x}^1 \equiv \mathbf{x}^1, \tilde{x}^2, \ldots, \tilde{x}^{\ell}$, by inserting new points between $\tilde{x}^k$ and $\tilde{x}^{k+1}$ if $\text{dist}_M(\tilde{x}^k, \tilde{x}^{k+1}) > 2\varepsilon$ and by deleting $\tilde{x}^k$ if $\text{dist}_M(\tilde{x}^{k-1}, \tilde{x}^k) + \text{dist}_M(\tilde{x}^k, \tilde{x}^{k+1}) < \varepsilon$ we finally construct an admissible **RMS**-fit sequence $\mathbf{y}^1,\ldots,\mathbf{y}^m$ which stands close to $\mathbf{x}^1,\ldots,\mathbf{x}^{\ell}$ and consist of interpolates of points from $\mathbf{x}^1,\ldots,\mathbf{x}^{\ell}$ with the ε-distance homogenity property dist_M $(\mathbf{x}^{k-1},\mathbf{x}^k) \in (\varepsilon/2,\varepsilon)$ $(k=2,\ldots,m)$. Both ε distance homogenity and **RMS**-fitting are of high practical importance in case of large molecules. There are widely used algorithms where curves as tentative approximations of reaction paths are represented by a given number of nearly equidistant points in the configuration space $\mathbf{R}^{3N}$. However, if the number of atoms is large, we cannot predict easily the number representing the points for a true approximation. If it is too small, e.g. smaller than the number of stationary points the real **RP** passes through, then the result obtained may be completely irrelevant even if the corresponding energy diagram is seemingly credible. The final conditions or optimal parameters for ε and σ are difficult to set automatically. We suggest these parameters should be chosen on the basis by experience on systems with different number of atoms. This is an unavoidable numerical problem for all pathway finding algorithms if no specific information is available on the potential energy surface (**PES**).

For a given admissible (and **RMS** fit) representation $\mathbf{y}_k \approx c_t^*(s_k)$ $(k=1,\ldots,m_t)$ of the curve c_t^* in the ideal procedure, we apply the simple DDRP-shift with time step $\sigma>0$

$$\mathbf{x}^k = \mathbf{y}^k - \sigma\nabla U(\mathbf{y}^k) \quad (k = 1, \ldots m_t)$$

Some examples of the conformational transitions in small (**3D**) chemical systems and the successful application of **DDRP** are reported first time in this article.

2. Calculations

DDRP algorithm was implemented into TINKER version 4.0 molecular modelling package [20]. The implementation ensured the calculations with all of the force fields and solvation models available in the package. The conformations were built up by PCMODEL [21] and optimized by the Newton–Raphson method contained in TINKER using a 0.001 kcal mol^{-1} Å^{-1} convergence criteria using the force field of path calculations. In some cases, we used implicit solvation model GB/SA [22] to compare our results obtained in gas-phase and in solute. The optimized geometries of the two conformations were used as input for the further calculations. The energies of the points on the **RP** obtained were depicted against the reaction coordinates (**RC**). The latter values were obtained for the kth point of the last **RP** sequence by the expression $\Sigma_i^k(\Sigma_j^N((x_j^i - x_{j-1}^i)^2 + (y_j^i - y_{j-1}^i)^2 + (z_j^i - z_{j-1}^i)^2)/N)^{1/2}$ normalized by $\Sigma_i^n(\Sigma_j^N((x_j^i - x_{j-1}^i)^2 + (y_j^i - y_{j-1}^i)^2 + (z_j^i - z_{j-1}^i)^2)/N)^{1/2}$, where N is the number of atoms and n is the number of points on the curve of the last **RP**. The calculated **RP**s obtained by the DDRP method is compared with those of the **EC** method [3] which is a modified version

of the original **EK** method [2]. This algorithm has been implemented in the program PATH in TINKER and the results are also described. In the calculation with PATH 50 points were generated with a tolerance of 0.001 kcal/mol/A.

3. Results and discussion

Some examples of the calculations performed by the **DDRP** method and the **PATH** calculations are described. By these calculations we show the **RP**s between stable conformers as between anti- and gauche butanes, anti- and gauche-2,2,5,5-tetramethyl-hexanes, between chair and twist cyclohexane, between C_{7ax} and C_{7eq} of Ac–Ala–$NHCH_3$ and between extended and α-helical Ac–$(Ala)_4$–$NHCH_3$. The calculation of the transition between two conformations of these small molecules supports the effectiveness of the **DDRP** method.

3.1. Conformational transition between anti- and gauche butanes

The minima of anti- and gauche conformations of butane was minimized in gas phase. The energies were 3.1 and 3.9 kcal/mol applied MM3 force field in gas phase. σ and ε values in **DDRP** were 0.01 and 0.01, respectively. The results after 1000 iterations were depicted in Fig. 1a (conformations of the molecule in point 1, 35, 60 and 79 on the path) and b. The results obtained by using **DDRP** method is very similar to that of **EC** method.

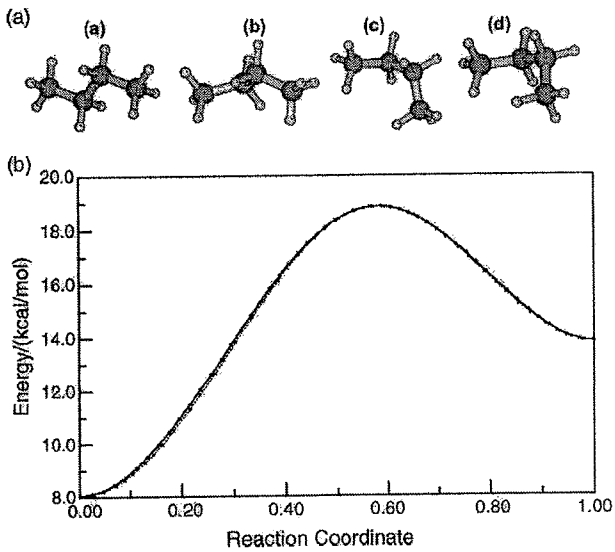

Fig. 1. (a) Conformational transition between anti- and gauche butanes (Points 1 (**a**), 35 (**b**), 60 (**c**) and 79 (**d**) on the intrinsic reaction path). (b) Energy profile of the conformational transition between anti- and gauche butanes (line with legend x: **EC** method, line with ●: **DDRP** method).

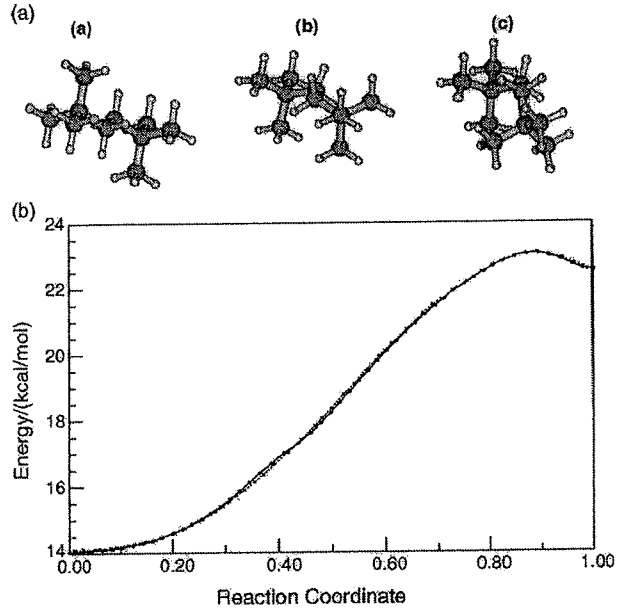

Fig. 2. (a) Conformational transition between anti- and gauche-2,2,5,5-tetramethyl-hexanes (Points 1 (**a**), 15 (**b**) and 30 (**c**) on the intrinsic reaction path). (b) Energy profile of the conformational transition between anti- and gauche-2,2,5,5-tetramethyl-hexanes (line with legend x: **EC** method, line with ●: **DDRP** method).

3.2. Conformational transition between anti- and gauche-2,2,5,5-tetramethyl-hexanes

The minima of anti- and gauche conformations of 2,2,5,5-tetramethyl-hexane was minimized in gas phase.

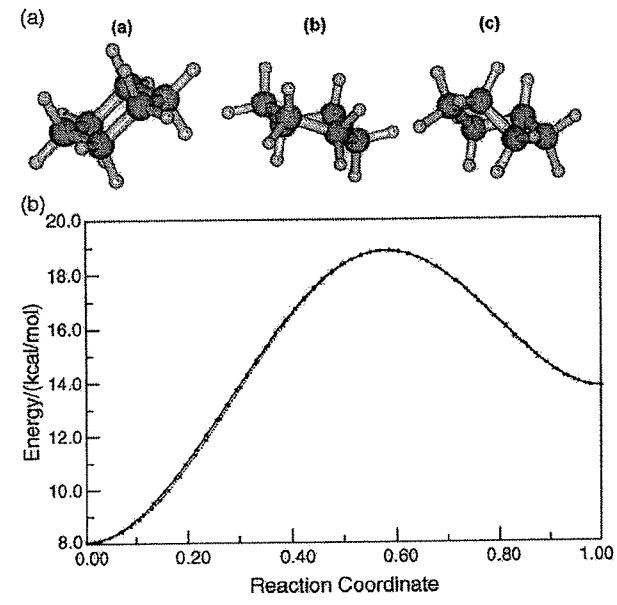

Fig. 3. (a) Conformational transition between two conformations of cyclohexane (chair to twist) (Points 1 (**a**), 21 (**b**) and 46 (**c**) on the intrinsic reaction path). (b) Energy profile of the conformational transition between two conformations of cyclohexane (chair to twist) (line with legend x: **EC** method, line with ●: **DDRP** method).

The energies were 14.1 and 22.5 kcal/mol applied MM3 force field in gas phase. σ and ε values in **DDRP** were 0.005 and 0.025, respectively. 1000 iteration steps were chosen. At larger iteration steps, the character of the energy profile did not changed, they are fitted to each other. The change in geometry is described in Fig. 2a (Point 1, 15 and 30). The energy profile of the reaction path and the structures in the minima and maxima is depicted in Fig. 2b. A barrier height of 23.1 kcal/mol was obtained by means of the path calculations using both **DDRP** and **EC** method.

3.3. Transition between two conformations of cyclohexane (chair to twist)

The minima of the chair and twist conformation of cyclohexane were calculated by using MM3 force field (8.0 and 13.8 kcal/mol). σ and ε values were 0.01 and 0.01, respectively, the number of iterations was 1000. The conformations on the reaction paths (points 1, 21 and 46 at **DDRP**) and the energy profiles obtained by **DDRP** and **EC** methods are described in Fig. 3a and b. The results of the two methods were found to be practically the same.

3.4. Transition between two conformations of C_{7ax} and C_{7eq} of Ac–Ala–NHCH$_3$

MM3PRO force field was applied in the calculation of the paths between C_{7ax} and C_{7eq} of Ac–Ala–NHCH$_3$. The initial energies of the reactant and product were 7.0 and 5.5 kcal/mol (in gas phase). σ and ε values were 0.1 and 0.02, respectively, the number of iterations was 5000. The largest barrier height was found at the 35th point (13.8 kcal/mol). The results obtained by **DDRP** and **EC** methods were different. The calculations were repeated by GB/SA solvation model, too, where the energies of the reactant and product were 9.8 and 8.4 kcal/mol, respectively. Structures obtained by DDRP in points 0, 50, 100 and 159 are described in Fig. 4a. The energy profiles obtained by **DDRP** and **EC** are depicted in Fig. 4b. As it can be seen, **DDRP** and **EC** results are different. **EC** method predicts path only with some breaks, the curve obtained by **DDRP** is continuous.

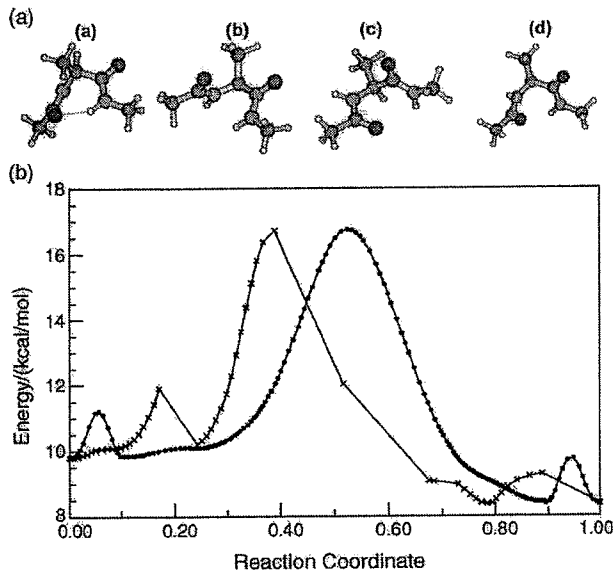

Fig. 4. (a) Conformational transition between two conformations of C_{7ax} and C_{7eq} of Ac–Ala–NHCH$_3$ (Points 1 (**a**), 50 (**b**), 100 (c.) and 159 (**d**) on the intrinsic reaction path). (b) Energy profile of the conformational transition between two conformations of two conformations of C_{7ax} and C_{7eq} of Ac–Ala–NHCH$_3$ (line with legend x: **EC** method, line with ●: **DDRP** method).

3.5. Transition between extended and helical structure of Ac–(Ala)$_4$–NHCH$_3$

AMBER94 force field with GB/SA implicit solvation model was applied in the calculation of the path between

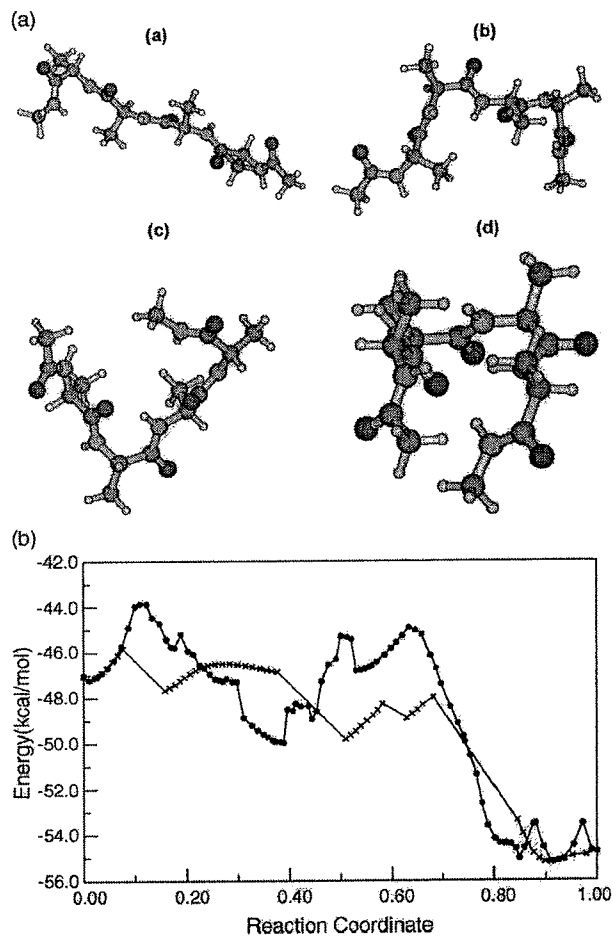

Fig. 5. (a) Conformational transition between extended and helical structure of Ac–(Ala)$_4$–NHCH$_3$ (Points 1 (**a**), 30 (**b**), 60 (**c**) and 94 (**d**) on the intrinsic reaction path). (b) Energy profile of the conformational transition between extended and helical structure of Ac–(Ala)$_4$–NHCH$_3$ (line with legend x: **EC** method, line with ●: **DDRP** method).

extended and α-helical structure of Ac–$(Ala)_4$–$NHCH_3$. σ and ε values were 0.1 and 0.1, respectively, the number of iteration was 50,000. The conformations on the reaction paths (points 0, 30, 60 and 94 at **DDRP**—where points 0 and 89 are the initial structures) and the energy profiles (obtained by the two methods) are described in Fig. 5a and b, respectively. The result of **EC** calculations shows some break on the curve, the curve obtained by **DDRP** is continuous.

4. Conclusion

DDRP algorithm was applied for conformational changes in small molecular systems and the results were compared to that of obtained by EC method implemented in TINKER. Results of the calculations support that the conformation change of small molecules are similar in **EC** and **DDRP** methods. At slightly larger and more flexible small molecules (e.g. small peptide models) with more rotational degree of freedom, the **RP**s obtained by these methods are different and one or more breaks were found at the results of **EC** method. **DDRP RP**s are continous.

Acknowledgements

This work is supported by the Hungarian National Science Foundation, grant no. OTKA T 032190.

References

[1] K. Fukui, Acc. Chem. Res. 14 (1981) 363.
[2] R. Elber, M. Karplus, Chem. Phys. Lett. 139 (1987) 375.
[3] R. Czerminski, R. Elber, Proc. Natl Acad. Sci. 86 (1989) 6963.
[4] R. Czerminski, R. Elber, J. Chem. Phys. 92 (1990) 5580.
[5] R. Czerminski, R. Elber, Int. J. Quantum Chem.: Quantum Chem. Symp. 24 (1990) 167.
[6] S.S.L. Chiu, J.J.W. McDouall, I.H. Hillier, J. Chem. Soc. Faraday Trans. 90 (1994) 1575.
[7] R. Olender, R. Elber, J. Mol. Struct. (THEOCHEM) 398 (1997) 63.
[8] L.L. Stachó, M.I. Bán, Theor. Chim. Acta 83 (1992) 433.
[9] L.L. Stachó, M.I. Bán, J. Math. Chem. 11 (1992) 405.
[10] L.L. Stachó, M.I. Bán, Theor. Chim. Acta 84 (1993) 535.
[11] L.L. Stachó, M.I. Bán, Comput. Chem. 17 (1993) 21.
[12] Gy. Dömötör, L.L. Stachó, M.I. Bán, J. Comput. Chem. 14 (1993) 1491.
[13] M.I. Bán, Gy. Dömötör, L.L. Stachó, J. Mol. Struct. (THEOCHEM) 311 (1994) 29.
[14] L.L. Stachó, M.I. Bán, Gy Dömötör, J. Mol. Struct. (THEOCHEM) 337 (1995) 99.
[15] L.L. Stachó, M.I. Bán, J. Math. Chem. 17 (1995) 337.
[16] Gy. Dömötör, M.I. Bán, L.L. Stachó, J. Comput. Chem. 17 (1996) 289.
[17] Gy. Dömötör, M.I. Bán, L.L. Stachó, J. Mol. Struct. (THEOCHEM) 455 (1998) 219.
[18] L.L. Stachó, Gy Dömötör, M.I. Bán, Chem. Phys. Lett. 317 (1999) 328.
[19] P.R. Surján, M. Kállay, Gy. Dömötör, L.L. Stachó, M.I. Bán, in: H. Kuzmány, J. Fink, M. Mehring, S. Roth (Eds.), Molecular Nanostructures. Proceedings of the IWEP'97, Kirchberg, Austria, World Scientific, Singapore, 1998, p. 147.
[20] J.W. Ponder, TINKER Version 4.0, Software Tools for Molecular Design, 2003.
[21] PCMODEL, Version 7.0, Molecular Modeling Software, Serena Software, Box 3076, Bloomington, IN 47402-3076, USA.
[22] W.C. Still, A. Tempczyk, R.C. Hawley, T. Hendrikson, J. Am. Chem. Soc. 112 (1990) 6127.