



A MULTIDIMENSIONAL DRIVER FOR QUANTUM CHEMISTRY PROGRAM MOPAC

MARTIN ČERNOHORSKÝ, MICHAL KUTÝ and JAROSLAV KOČA*

Department of Organic Chemistry, Faculty of Science, Masaryk University, 611 37 Brno, Czech Republic

(Received 19 July 1995; in revised form 19 January 1996)

Abstract—A method has been developed that is able to control multidimensional linear driving along potential energy hypersurfaces. All the internal coordinates, bond lengths, angles, and dihedral angles may be driven. The method serves for providing fast starting inspection of reaction and conformational pathways, as well as for the estimation of transition states. The method has been implemented as a part of the semiempirical quantum chemistry program MOPAC. It is activated by the key word DRIVER. The method has been tested on the Diels–Alder reaction of ethene and butadiene. Also some examples on *t*-butyl halides are discussed. Copyright © 1996 Elsevier Science Ltd

Key words: Semiempirical method, energy surface, reaction path, Diels–Alder reaction of ethene and butadiene, *t*-butyl halide.

1. INTRODUCTION

A chemical reaction may conveniently be described as a pathway along a multidimensional potential energy hypersurface, PES. According to Eyring and Polanyi (1931), the pathway describes a geometrical movement leading from one minimum (educts) to another minimum (products). The most important coordinate to describe the movement is mathematically well defined, and known as the intrinsic reaction coordinate (IRC) concept of Fukui (Fukui, 1970). Two categories of points are of key importance in the IRC concept; minima and transition states (TS) both are critical points of index zero and one (Mezey, 1987), respectively. Although the developments of computational chemistry during the last decade have made exploration of PES a practical reality, a method that would generally solve the question of finding reaction paths is still missing. The problem has, in fact, three independent levels:

- search for minima and their geometry optimization;
- search for transition states and their optimization; and
- mapping of the reaction pathway itself.

While several effective methods are implemented in quantum chemistry and molecular mechanics packages to solve the first level, the methods to solve the second and third levels are not so successful. For reviews see Kolossvary & Guida (1993); Muller (1980). It seems that the search for TS is a key problem.

One of the possibilities to generate a first approximation of TS is travelling along the PES. This approach, called also driver technology, has for several years been tested in our group on conformational PES. The theoretical background is based on formal expression of interconversions along PES by a graph-theory model (Koča, 1991a,b). As we would like to extend our effort also to reaction PES study, we decided to employ semiempirical quantum chemistry software. The reason why, in the first stage of our work, *ab initio* calculations have been avoided is that the work is, principally, just methodological, and *ab initio* methods are still very time consuming to describe a pathway for realistic chemical systems. Moreover, the method presented could principally be used also within an *ab initio* framework.

We decided to extend the widely used semiempirical quantum chemistry package MOPAC and to develop a new subroutine, which is able to control a multidimensional driving. This work should also be considered in the framework of non-empirical synthesis design, as the PES are a very general theoretical basis for chemical synthesis itself.

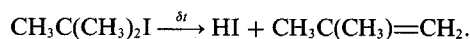
2. MOTIVATION

There are two tools in MOPAC to study reaction pathways. The dynamic reaction coordinate, DRC, option, which allows for calculation of reaction coordinates based on the conservation energy rule. It means that when potential energy is decreased, kinetic energy must consequently increase. The IRC option makes it possible to follow a reaction pathway leading from TS to two neighbouring minima, educts and products. A more realistic DRC is also

*Author for correspondence; E-mail: jkoca@chemi.muni.cz.

substantially more time consuming. Moreover, and this is the principal disadvantage of both IRC and DRC, in order to obtain sufficient results one should have optimized TS before calculation. It means that one needs at least an estimation of TS, which may then be optimized. There is a special tool to search for TS in MOPAC, which is called SADDLE. It makes the estimation based on the geometry of the starting and final structure (Dewar *et al.*, 1984). But, this tool is very sensitive to input geometries, and it is not general. Another possibility how to search for TS, which is implemented in MOPAC, is *one-dimensional driving*. It means continuous movement of one internal coordinate, which is expected to be the reaction coordinate, from educts to products by a positive step. After each step of driving the structure is fully optimized except for the driven coordinate. This is the basic principle on which the recently developed software DAISY (Koča & Carlsen, 1992) and CICADA (Koča, 1994) is based. However, CICADA has been developed to study conformational PES, which are principally simpler than the reaction ones. Also the estimation of the conformational reaction coordinate is clearer. For reaction PES the one-dimensional driving may completely fail (Muller, 1980). A more "secure" method, implemented in MOPAC, is two-dimensional *grid*, which produces a two-dimensional cut of PES. If the two coordinates are well selected, i.e. the real reaction coordinate is a combination of the two selected internal coordinates, then the picture produced clearly describes not only estimated geometry of TS, but also the reaction path itself. However, the grid calculation needs to calculate $n \cdot m$ points on PES (n steps for one coordinate, m for the other one), which may be time consuming. Moreover, the grid method has several substantial pitfalls. The basic one is that the grid will only work in such a case when the problem is well defined by two coordinates. If it is not so, then wells on PES will keep geometries unrealistic, and the surface calculated will also be unrealistic. It is illustrated in Fig. 1.

The following reaction (pyrrolytic elimination, E_i) was suggested:



A minimized geometry of the educt with two selected internal coordinates (the distances C–I and H–I, cf. Fig. 1) was submitted to the grid procedure. The following approach was used to create the surface:

- (i) one internal coordinate [bond H–I, see Fig. 1(a)] has been changed step by step (size of the step is determined by the key word STEP1 = 0.1 Å) within a pre-defined range (this range is determined by the size and number of steps, key word POINT1 = 16 defines the number of steps);
- (ii) the second coordinate (C–I) has afterwards been changed by only one step (key word STEP2 = 0.1 Å);

(iii) first coordinate has again been changed using the same size and number of steps but in the opposite direction; and

(iv) the same procedure (i–iii) is repeated several times (14 cycles in total, key word POINT2 = 14). Schematic representation of the grid method is shown in Fig. 1(c) (arrows indicate directions of coordinate changes).

Energy diagrams of the two-dimensional cut of the AM1 energy surface show the pathological discrepancies on the surface. The educt (a) is not converted to the expected product, but jumps into the valley (b) and the pathway then leads to the shallow valley of an unexpected product $\text{CH}_2=\text{C}(\text{CH}_3)\text{CH}_2\text{I} + \text{H}_2$ (d) via high energy structure (c). It is seen that the surface is not as smooth as one would expect. The reason is that the real changes take place in a different part of the configurational space than that which is expressed by the two selected coordinates.

Because of the above disadvantages we decided to develop a multidimensional driver, which will extend MOPAC tools to search TS and reaction pathway, and which will make the tools more efficient.

3. MULTIDIMENSIONAL DRIVER

The subroutine *driver* has been written for multidimensional driving along PES. It makes it possible to change geometry to simulate travelling along PES in any linear direction. The direction is defined by any set of internal coordinates. Once defined as driven, the internal coordinate cannot be optimized. It is always fixed on a certain value defined by the driver. The driver could also be used for conformational studies. Problems may appear around the dihedral angle value of 180 degrees, where the sign changes. This may cause inconsistency.

Driver is written as an additional part of the program package MOPAC. It requires MOPAC subroutines, and a small change in the MOPAC source itself is also needed. Driver is activated by the key word DRIVER, which has to be used together with DEBUG. It requires a file called *name.drv*, where the parameters are saved. Parameters describe the direction of driving. Each line corresponds with one driven internal coordinate. It is written as follows:

number_of_atom start limit step {A|B|D}

where *number_of_atom* is the reference number of the atom in the normal MOPAC input file, where the driven internal coordinate is defined; *start* and *limit* are the starting and final values of the driven coordinate, respectively; and *step* is the step for the driven coordinate. Step may be positive as well as negative. This is an extension to the standard MOPAC rules, where only positive steps are allowed. The letters at the end of the line inform, whether the

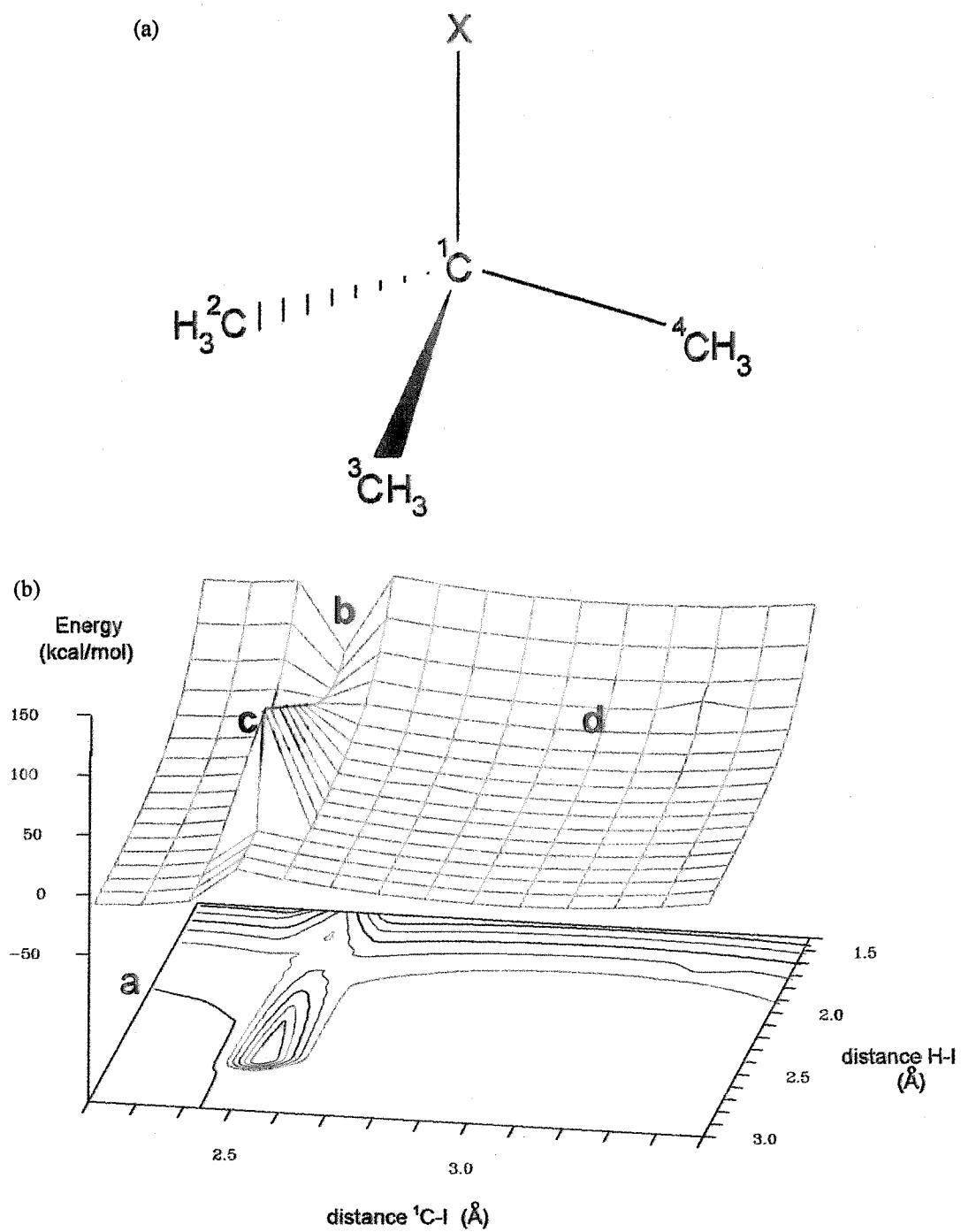


Fig. 1(a) and (b).

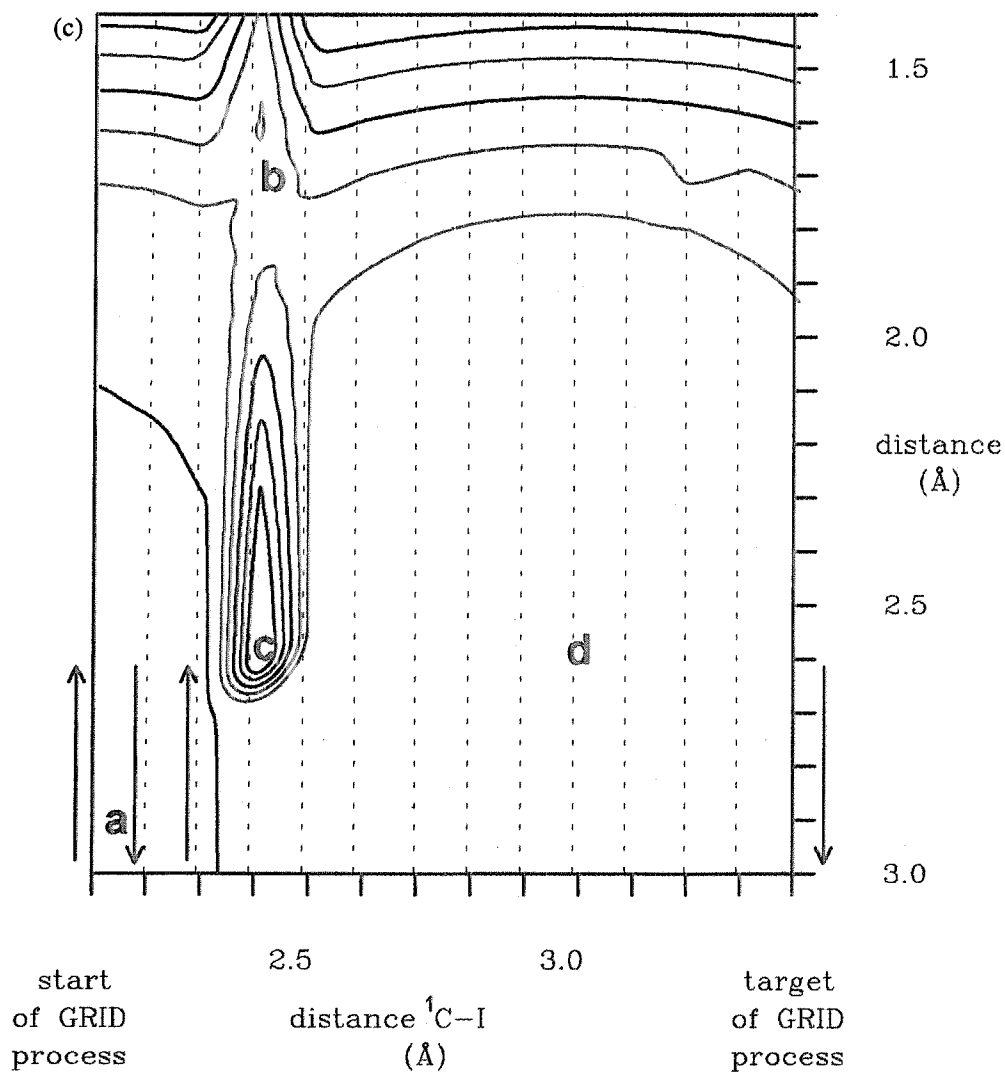


Fig. 1. Two-dimensional cut of AM1 energy hypersurface of the pyrrolytic elimination of *t*-butyl iodide [parts b and c generated by the graphic software Gnuplot (Kelley & Williams, 1986)]. (a) Atom labelling ($X = \text{I, F}$). (b) Three-dimensional energy diagram. (c) Contour representation of (b) (contours by step 14 kcal/mol); arrows indicate direction of driving within the standard MOPAC grid procedure. The educt **a** is not converted to the expected product, but jumps into the valley **b** and the pathway then leads to the shallow valley of an unexpected product $\text{CH}_2=\text{C}(\text{CH}_3)\text{CH}_2\text{I} + \text{H}_2$ **d** via high-energy structure **c**. For further explanation see the text.

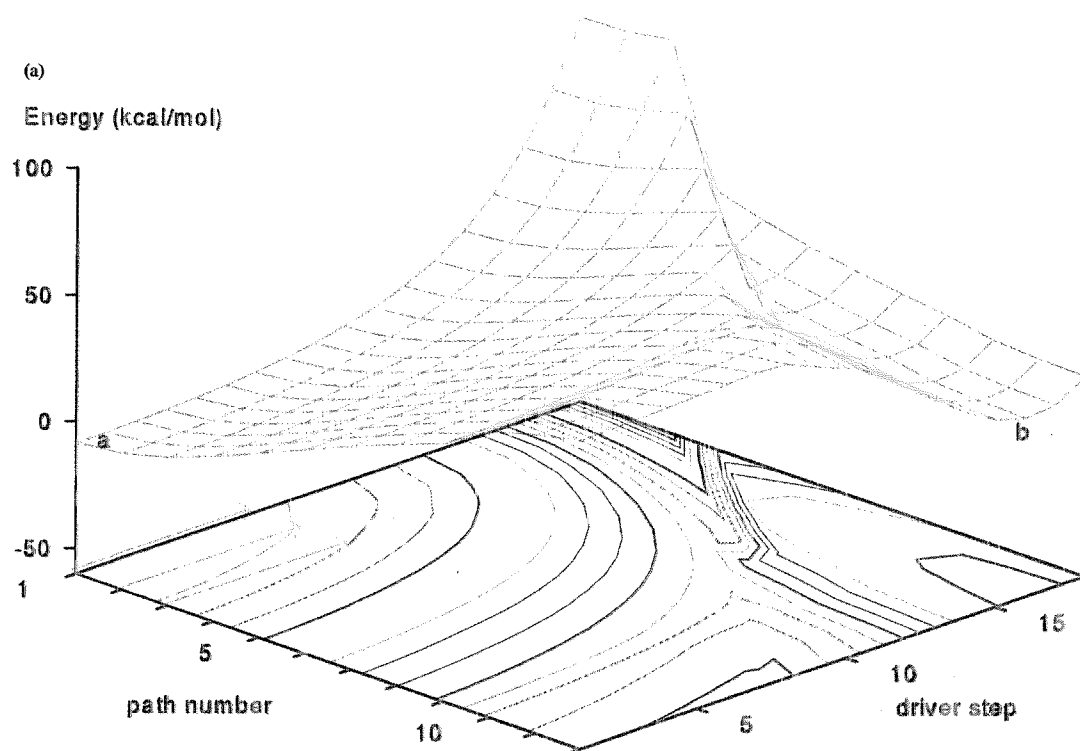


Fig. 2(a).

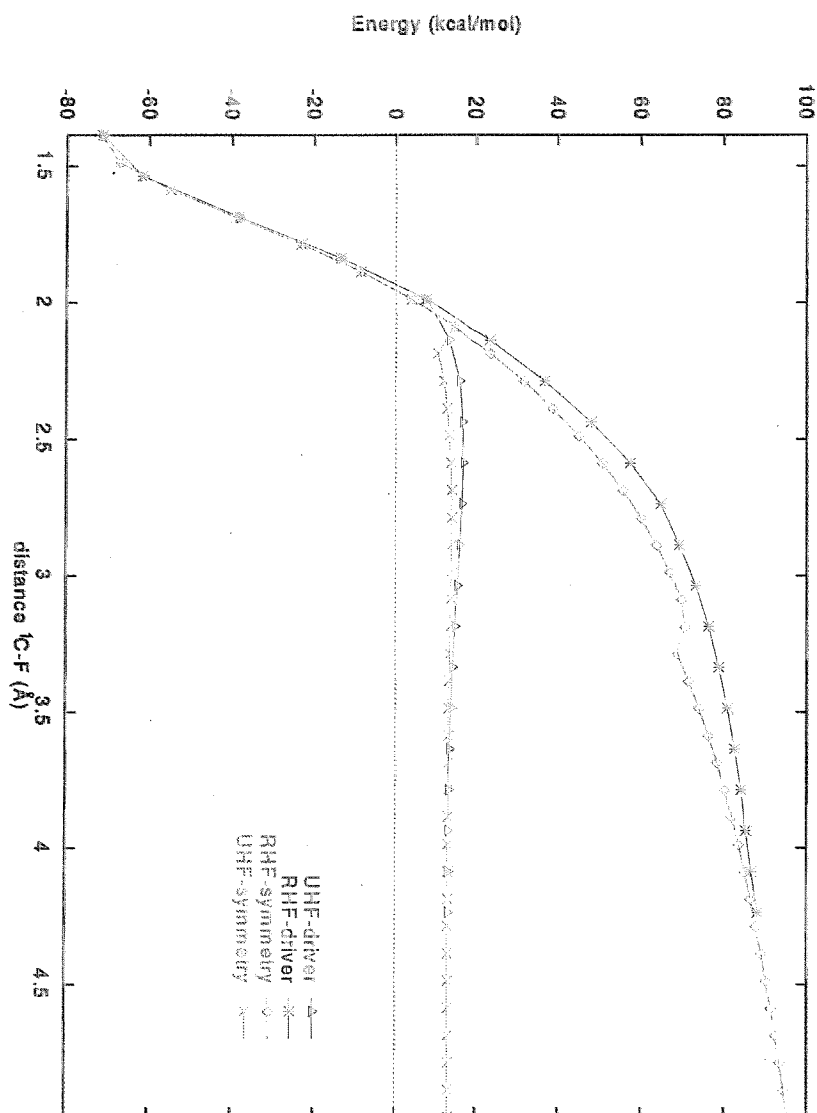


Fig. 4. Energy profile obtained by AM1 method combined with SYMMETRY and DRIVER key words of MOPAC for dissociation of ${}^1\text{C}-\text{F}$ bond of *t*-butyl fluoride. The curves marked "symmetry" show energy calculated by one-dimensional driving of the ${}^1\text{C}-\text{F}$ bond and SYMMETRY key word. The curves marked "driver" have been obtained by a multidimensional driving method when the coordinates ${}^4\text{C}-{}^1\text{C}-\text{F}$, ${}^3\text{C}-{}^1\text{C}-\text{F}$, ${}^2\text{C}-{}^1\text{C}-\text{F}$ (bond angles), and ${}^1\text{C}-\text{F}$ (distance) have simultaneously been driven using the new DRIVER procedure. For notation see Fig. 1(a).

Table 1. A sequence of dihedral angle values when a standard MOPAC grid (A) and the method of independent one-dimensional pathways (B) is applied. The values are for pyrolytic elimination of *t*-butyl iodide (see Figs 1 and 2). Each new starting geometry is written in italics. The calculation proceeds in the sequence *column a* → *column b* → *column c*

A						B					
<i>a</i>		<i>b</i>		<i>c</i>		<i>a</i>		<i>b</i>		<i>c</i>	
H-I	C-I	H-I	C-I	H-I	C-I	H-I	C-I	H-I	C-I	H-I	C-I
3.0	2.1	1.4	2.2	3.0	2.3	3.0	2.1	3.0	2.2	3.0	2.3
2.9	2.1	1.5	2.2	2.9	2.3	2.9	2.1	2.9	2.2	2.9	2.3
2.8	2.1	1.6	2.2	2.8	2.3	2.8	2.1	2.8	2.2	2.8	2.3
2.7	2.1	1.7	2.2	2.7	2.3	2.7	2.1	2.7	2.2	2.7	2.3
2.6	2.1	1.8	2.2	2.6	2.3	2.6	2.1	2.6	2.2	2.6	2.3
2.5	2.1	1.9	2.2	2.5	2.3	2.5	2.1	2.5	2.2	2.5	2.3
2.4	2.1	2.0	2.2	2.4	2.3	2.4	2.1	2.4	2.2	2.4	2.3
2.3	2.1	2.1	2.2	2.3	2.3	2.3	2.1	2.3	2.2	2.3	2.3
2.2	2.1	2.2	2.2	2.2	2.3	2.2	2.1	2.2	2.2	2.2	2.3
2.1	2.1	2.3	2.2	2.1	2.3	2.1	2.1	2.1	2.2	2.1	2.3
2.0	2.1	2.4	2.2	2.0	2.3	2.0	2.1	2.0	2.2	2.0	2.3
1.9	2.1	2.5	2.2	1.9	2.3	1.9	2.1	1.9	2.2	1.9	2.3
1.8	2.1	2.6	2.2	1.8	2.3	1.8	2.1	1.8	2.2	1.8	2.3
1.7	2.1	2.7	2.2	1.7	2.3	1.7	2.1	1.7	2.2	1.7	2.3
1.6	2.1	2.8	2.2	1.6	2.3	1.6	2.1	1.6	2.2	1.6	2.3
1.5	2.1	2.9	2.2	1.5	2.3	1.5	2.1	1.5	2.2	1.5	2.3
1.4	2.1	3.0	2.2	1.4	2.3	1.4	2.1	1.4	2.2	1.4	2.3

driven parameter is bond length (*B*), angle (*A*) or dihedral angle (*D*). The value

$$(\text{limit} - \text{start})/\text{step}$$

must be an integer. It expresses the number of steps, and it has to be the same for each line in the *name.drv* file, otherwise the program will not run. If one of the three parameters *start*, *limit*, *step* is replaced by a star, it is automatically calculated from the two remaining ones. If all the three parameters are replaced by stars, they are substituted by the parameters from the first line. An example of the input file for the Diels–Alder addition of ethene and butadiene is shown in the Appendix together with the corresponding driver control file (*name.drv*).

Results of the calculation are saved in a standard MOPAC output file (unit 6). Additionally, driver generates files *stepXXX* and *geomXXX* (*XXX* is number of the step). The first is the MOPAC *arc* file (unit 12). In the second one (unit 19), the input geometry for the next step is saved. It can be used for restarting in case of any failure.

4. APPLICATIONS

4.1. Two-dimensional cut of PES—a method of independent one-dimensional pathways

The first application for which we have utilized the driver is the so-called *method of independent one-dimensional pathways*, which suppresses some disadvantages of the standard *grid* method implemented in MOPAC.

The method separates the energy surface into a certain number of one-dimensional pathways. It is performed in the same way as the grid method but these pathways are *independent* of each other. It means that the starting geometry for a new pathway is independent of the final geometry of the pathway calculated before. It is shown in Table 1. Each

pathway is calculated by the above described DRIVER method. The method is illustrated for the same reaction as the above discussed GRID. Figure 2 shows the results.

It is seen that the application of the *independent one-dimensional pathways* method generates a smooth energy surface giving the information about the real reaction pathway and transition state. The example demonstrates the successful application of the DRIVER method in the case when the GRID method failed.

4.2. Diels–Alder addition of ethene and butadiene

The Diels–Alder reaction of ethene and butadiene has widely been studied in the literature (see, for example, Bernardi *et al.*, 1988; Craig & Stone, 1994; Jorgensen *et al.*, 1989; Kelley & Williams, 1986). We have used the symmetrical driver on two coordinates, bond lengths ${}^1\text{C}-{}^{11}\text{C}$ and ${}^4\text{C}-{}^{12}\text{C}$. The results are pictured in Fig. 3, which shows numbering of atoms, the contour diagram of the two-dimensional cut of PES where the path produced is superimposed by a heavy line, and the energy profile obtained. Table 2 reveals that driver results are in good qualitative agreement with those obtained by other methods.

Table 2. Energy barriers calculated and measured for the Diels–Alder reaction of ethene and butadiene by various methods. The results presented here are just illustrative and do not cover all the published data

Method	Barrier (kcal/mol)
This study	41.0
4-31 G ^a	32.0
RHF/6-31 G//3-21 G ^b	45.9
MP2/6-31 G//3-21 G ^b	15.9
Experimental ^c	27.5

^aBernardi *et al.*, 1988.

^bJorgensen *et al.*, 1989.

^cRowley & Steiner, 1951.

4.3. Multidimensional driving on *t*-butyl halide

We have selected bond breaking on *t*-butyl fluoride, which leads to the *t*-butyl cation. However, in the gas phase this reaction requires a very high energy. When driving just the ¹C–F bond [see Fig. 1(a)] then the system prefers an elimination reaction leading to H–F rather than to overcome the high energy barrier leading to the carbonium ion. So, to mimic this reaction one needs a special tool.

The situation can be treated by application of the MOPAC keyword SYMMETRY, especially in this case of a high-symmetry system. We have applied the DRIVER method. Simultaneous driving of four internal coordinates has been performed, angles ⁴C–¹C–F, ³C–¹C–F, ²C–¹C–F, and the distance ¹C–F. Both AM1-UHF and AM1-RHF calculations were run. The results are pictured in Fig. 4.

The results show that the homolytic change will preferably take place in this case (UHF energy is substantially lower). It is seen from the figure that both SYMMETRY and DRIVER produce very similar results. However, as the SYMMETRY can only be used for symmetrical systems, the DRIVER is a general method.

5. CONCLUSIONS

A method has been developed which is able to control multidimensional linear driving along PES. All the internal coordinates, bond lengths, angles, and dihedral angles may be driven. The method provides a new tool to estimate transition states. It is substantially faster than the MOPAC *grid* option, and more simple to use compared to the SADDLE option. The method is written as an integral part of the MOPAC semiempirical quantum chemistry

package, and it requires only small changes within the MOPAC source code itself. It is activated by the key word DRIVER. The source code is available upon request on the e-mail address: jkoca@chemi.muni.cz

Acknowledgements—The authors would like to thank the Czech Academic Supercomputer Centre in Brno for providing them with access to their computer facilities. Our thanks go also to Mr Laurence Benjamin, B.A., for language corrections. The research has partially been supported by the Grant Agency of the Czech Republic, Grant no. 203/94/0522. This financial support is gratefully acknowledged.

REFERENCES

- Bernardi F., Bottoni A., Field M. J., Guest M. F., Hillier I. H., Robb M. A. & Venturini A. (1988) *J. Am. Chem. Soc.* **10**, 3050.
 Craig S. L. & Stone A. J. (1994) *J. Org. Chem.* **90**, 1663.
 Dewar M. J. S., Healy E. F. & Stewart J. J. P. (1984) *J. Chem. Soc., Faraday Trans.* **80**, 227.
 Eyring H. & Polanyi M. (1931) *Z. Phys. Chem.* **B12**, 279.
 Fukui K. (1970) *J. Phys. Chem.* **74**, 4161.
 Jorgensen L. W., Houk K. N., Loncharich R. J. & Blake J. F. (1989) *J. Am. Chem. Soc.* **111**, 9172.
 Kelley C. & Williams T. (1986) *Gnuplot. An Interactive Plotting Program*.
 Koča J. (1991a) *Theor. Chim. Acta* **80**, 29.
 Koča J. (1991b) *Theor. Chim. Acta* **80**, 51.
 Koča J. (1994) *J. Mol. Struct. (Theochem.)* **308**, 13.
 Koča J. & Carlsen P. H. J. (1992) *J. Mol. Struct. (Theochem.)* **257**, 105.
 Kolossvary I. & Guida W. C. (1993) *J. Am. Chem. Soc.* **115**, 2107.
 Mezey P. G. (1987) *Potential Energy Hypersurfaces*. Elsevier, Amsterdam.
 Muller K. (1980) *Angew. Chem. Int. Ed. Eng.* **19**, 1.
 Rowley D. & Steiner H. (1951) *Discuss. Faraday Soc.* **10**, 198.

APPENDIX

An example of the input file for the Diels–Alder addition of the ethene and butadiene (*example.dat*) and the corresponding driver control file (*example.drv*).

example.dat:

AM1 PRECISE DEBUG GEO-OK DRIVER
 Diels–Alder reaction

C	0.000000	0	0.000000	0	0.000000	0	0	0	0
C	1.3336000	1	0.000000	0	0.000000	0	1	0	0
C	1.3335600	1	124.881600	1	-125.128890	1	3	2	1
H	1.1053900	1	114.795030	1	179.108210	1	2	3	1
H	1.0975000	1	122.938370	1	-0.283010	1	1	2	3
H	1.0976500	1	121.903570	1	179.736280	1	1	2	3
H	1.1054200	1	114.793760	1	54.092810	1	3	2	1
H	1.0976300	1	122.870450	1	-0.129260	1	4	3	2
H	1.0975200	1	121.991820	1	179.824990	1	4	3	2
C	5.0000000	0	99.697390	1	68.570980	1	1	2	3
C	5.0000000	0	99.592270	1	68.003300	1	4	3	2
H	1.0982800	1	122.710270	1	88.311830	1	11	12	4
H	1.0982900	1	122.718500	1	-91.595210	1	11	12	4
H	1.0982800	1	122.709850	1	87.008910	1	12	11	1
H	1.0982900	1	122.719010	1	-92.897140	1	12	11	1

example.drv:

11 5.0 1.4 -0.1 B
 12 5.0 1.4 -0.1 B

The file *example.drv* must be present in the working directory.