Theor Chim Acta (1991) 80: 51–62

Theoretica
Chimica Acta

A mathematical model of the logical structure of chemistry. A bridge between theoretical and experimental chemistry and a general tool for computer-assisted molecular design

II. Conformational and reaction changes

Jaroslav Koča

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

Received March 19, 1990; received in revised form October 30, 1990/Accepted December 17, 1990

Summary. Proposals of heuristic convertibility functions and general drivers for conformational and reaction potential energy (hyper)surfaces study are suggested. The functions and drivers are formulated under the mathematical model of the logical structure of chemistry which has been published in the previous paper. A part of the described theory has been realized as the computer program system DAISY.

Key words: Conformational analysis – Chemical reaction – Potential energy hypersurface – Mathematical model

1. Introduction

In the previous paper [1], a mathematical model of the logical structure of chemistry has been introduced. The model consists of two main parts. The first of them is based on the phenomenon of chemical convertibility, and models the logic of chemical changes utilizing potential energy hypersurface, PES. The second part of the model is directed to the heuristic prediction of chemical changes. It is based on PES path generation and analysis. From the formal point of view, there are two essential notions serving as a basis, a heuristic convertibility function Γ' and a general driver DRV. The function Γ' predicts interconversions between various nuclear configurations. These interconversions are verified by the general driver DRV. Proposals of concrete functions Γ' and drivers DRV for conformational and reaction changes elucidation are subjected in the presented paper.

2. Heuristic convertibility functions Γ'

We will turn our attention to the problem of the formulation of heuristic convertibility functions for conformational and reaction changes. Configurational changes are omitted. Note that group theoretical formalism of Ugi et al.

[2] may be used as a good starting point for the formulation of the heuristic convertibility functions for configurational problems elucidation.

2.1. *Heuristic convertibility functions for conformational changes study*

Two functions of this kind are discussed. The former is the general conformational convertibility function Γ^{c}_{P}. The latter is the function Γ^{c}_{1} which may be used for the elucidation of conformational changes of open chain molecules.

Let us suppose that studying conformational changes only one molecule composed of atoms of a set A is being of interest. Let card A be N. Any nuclear configuration C of the molecule can be described as a set of $3N-6$ internal coordinates which can be split into two subsets. Let it be a set E of M explicitly followed and a set I of implicitly followed internal coordinates. Formally:

$$C = \{c_1, c_2, \ldots, c_{3N-6}\} = \{i_1, i_2, \ldots, i_{3N-6-M}\} \cup \{e_1, e_2, \ldots, e_M\} = I \cup E. \quad (1)$$

Since changes of dihedral angles of the so-called free rotatable bonds (i.e. single bonds) usually induce much less energy changes than those of the other internal coordinates [3], the set E will usually be a subset of the set of dihedral angles expressing a rotation around single bonds.

Example 1. Cyclohexane (a) and 3-methylene-1,4-pentadiene (b).

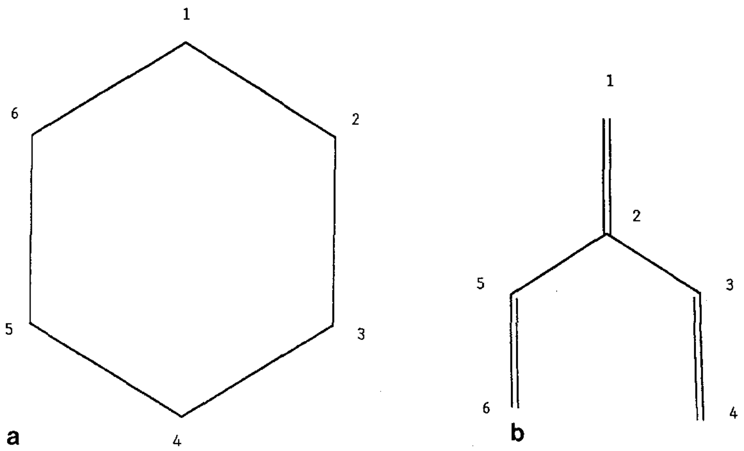

Fig. 1.

Using typical notation *a-b-c-d* for the description of a dihedral angle we can write $E = \{$1-2-3-4, 2-3-4-5, 3-4-5-6, 4-5-6-1, 5-6-1-2, 6-1-2-3$\}$ for molecule (a), and $E = \{$1-2-3-4, 1-2-5-6$\}$ for molecule (b).

Describing the functions Γ^{c}_{P} and Γ^{c}_{1} we will use a parameter $D_{\min}$. Let $D_{\min}$ denote the smallest value of the difference between the same dihedral angles of two nuclear configurations which will be taken into account.

2.1.1. The function Γ^{c}_{P}. The function Γ^{c}_{P} $(P \leqslant M)$ is a general convertibility function which can be used for an arbitrary conformational change study. It is defined in the combinatorial frame and based on the so-called set of "intermediate conformations", S, which is derived by the following construction.

Let v_i, v_j be two nuclear configurations of the molecule under study. Using notation of Ref. [1], we may write $v_i, v_j \in {}^0V''$. From Eq. (1), we have:

$$v_i = \{i_1^i, \ldots, i_{3N-6-M}^i\} \cup \{e_1^i, \ldots, e_M^i\}, \tag{2a}$$

$$v_j = \{i_1^j, \ldots, i_{3N-6-M}^j\} \cup \{e_1^j, \ldots, e_M^j\}. \tag{2b}$$

Let the index set T be defined as:

$$T = \{t \mid D(e_t^i, e_t^j) = 1\}, \tag{3}$$

where the function $D(e_t^i, e_t^j)$ is set to unity if the difference of the dihedral angles e_t^i, e_t^j is greater than $D_{\min}$. In all other cases, we set $D(e_t^i, e_t^j)$ equal to zero. Let T be ordered in an arbitrary but fixed way. Accordingly, T is the index set containing indices of explicitly followed coordinates in which v_i and v_j differ by the difference greater than $D_{\min}$ value. Let us take a nuclear configuration v_k:

$$v_k = \{i_1^k, \ldots, i_{3N-6-M}^k\} \cup \{e_1^k, \ldots, e_M^k\}. \tag{4}$$

We say that v_k is an "intermediate nuclear configuration" of v_i, v_j if:

$$D(e_i^i, e_i^k) = 0 \quad \text{and} \quad D(e_i^j, e_i^k) = 0 \quad \text{for } i \notin T, \tag{5a}$$

$$\text{there exists at least one } i \in T \text{ such that } D(e_i^i, e_i^k) = 1, \tag{5b}$$

$$\text{there exists at least one } i \in T \text{ such that } D(e_i^j, e_i^k) = 1. \tag{5c}$$

Let us denote the set of all the "intermediate nuclear configurations" as S_{ij}.

Let us define a function O for each dihedral angle e_i for $i \in T$. The function O will inform about the direction of the rotation of the dihedral angle e_i^i towards e_i^k under the condition that e_i^j is not met. We define the function O as unity if the above rotation is counterclockwise, and as -1 in the opposite case. Formally:

$$O(e_i^i, e_i^k, e_i^j) = \begin{cases} 1, & \text{if } x_2 - \text{mod}_3(x_1) = 1, \\ 2, & \text{if } D(e_i^i, e_i^k) = 0 \text{ or } D(e_i^k, e_i^j) = 0 \\ -1, & \text{otherwise} \end{cases} \tag{6}$$

The values of x_1 and x_2 are obtained by the following construction:

(a) Let the triple e_i^i, e_i^k, e_i^j of dihedral angles be expressed in the bounds of 0–360 degrees.

(b) Let the above triple be ordered as an increasing series y_1, y_2, y_3.

(c)
$$\begin{aligned} x_1 &= 1, \quad \text{if } y_1 = e_i^i, & x_2 &= 1, \quad \text{if } y_1 = e_i^k, \\ x_1 &= 2, \quad \text{if } y_2 = e_i^i, & x_2 &= 2, \quad \text{if } y_2 = e_i^k, \\ x_1 &= 3, \quad \text{if } y_3 = e_i^i, & x_2 &= 3, \quad \text{if } y_3 = e_i^k. \end{aligned}$$

Define the function Γ_P^c as follows:

$$\Gamma_P^c(v_i, v_j) = \begin{cases} 0, & \text{if } T = \varnothing, \\ 2, & \text{if card } T > P \text{ or if there exists a subset} \\ & S'_{ij} \text{ of the set } S_{ij} \subset {}^0V'' \text{ such that} \\ & \text{(a) card } S'_{ij} = 2^{\text{card } T}, \\ & \text{(b) } O(e_i^i, e_i^k, e_i^j) \neq 2 \text{ for each } v_k \in S'_{ij} \text{ and } t \in T, \\ & \text{(c) } (O(e_t^i, e_t^k, e_t^j)) \neq (O(e_t^i, e_t^l, e_t^j) \text{ for each} \\ & \quad v_k, v_l \in S'_{ij} \text{ and } t \in T. \text{ This relation} \\ & \quad \text{is understood as the inequality of vectors.} \\ 1, & \text{otherwise.} \end{cases} \tag{7}$$

Following the above definition one can say that the nuclear configuration v_i is convertible to v_j in the sense of the function Γ^c_P if v_i and v_j do not differ in more than P dihedral angles and, at the same time, if there exists such a combination of rotations of dihedral angles which transfers v_i to v_j in such a way that no other nuclear configuration from the basic set $^0V''$ is met. The basic relation between two general convertibility functions Γ^c_P and $\Gamma^c_{P'}$ is formulated by the following theorem.

Theorem 1. Let Γ^c_P and $\Gamma^c_{P'}$ be defined for the same value of $D_{\min}$. If $P' \leqslant P$ then $\Gamma^c_{P'} = 1$ implies $\Gamma^c_P = 1$ for each $v_i, v_j \in {}^0V''$.

Proof. Because of Eq. (3), the set T is the same for both the Γ^c_P and $\Gamma^c_{P'}$. The same can be said about the set S_{ij} which is defined by Eq. (5a–c). Let $\Gamma^c_{P'}(v_i, v_j) = 1$. It implies, because of Eq. (7), that there does not exist a set S'_{ij} which fulfills the feature (b) in Eq. (7) and card $T \leqslant P'$. Since $P' \leqslant P$, we have $T \leqslant P' \leqslant P$, i.e. $T \leqslant P$ and $\Gamma^c_P(v_i, v_j) = 1$.

2.1.2. The function Γ^c_1. The function Γ^c_1 is a special case of Γ^c_P for $P = 1$. In a very similar form, it has been used for conformational changes study of open chain molecules in the program system DAISY [4]. Two conformations v_i, v_j are seen to be interconvertible within the set $^0V''$ in the sense of Γ^c_1 if there exists only one k, $1 \leqslant k \leqslant M$, such that the difference between dihedral angles e^i_k and e^j_k is greater than $D_{\min}$, and if there exists such a rotation of this free rotatable bond which transfer v_i to v_j and no other nuclear configuration from the set $^0V''$ is met during this rotation.

Example 2. Let us consider the nuclear configurations $v_1, \ldots, v_7$ of the molecule

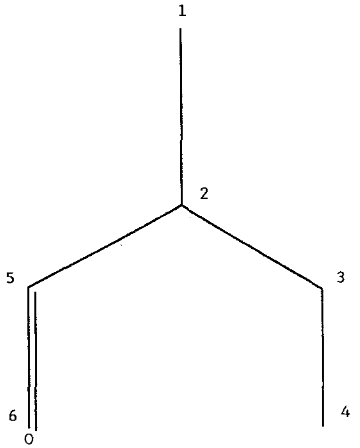

Fig. 2.

described by $E = \{1\text{-}2\text{-}3\text{-}4,\ 1\text{-}2\text{-}5\text{-}6\}$ as:

$$v_1 = \{0, 0\}, \qquad v_2 = \{0, 180\}, \qquad v_3 = \{180, 180\},$$

$$v_4 = \{120, 90\}, \qquad v_5 = \{120, -90\}, \qquad v_6 = \{-120, 90\},$$

$$v_7 = \{-120, -90\}.$$

For an arbitrary $D_{\min}$ less than 60 degrees the convertibility predicted by Γ_1^c (dashed lines) and Γ_2^c (dashed and all lines) is visualized by the following graph.

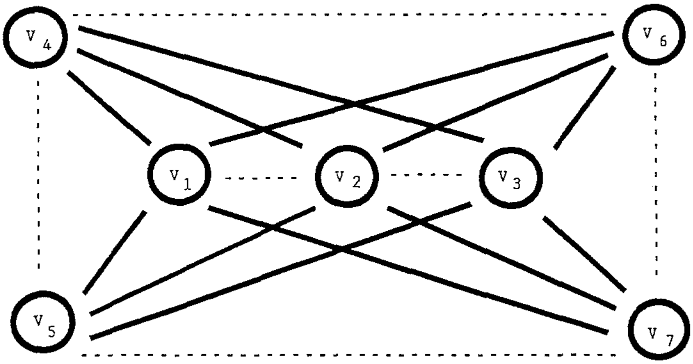

Fig. 3.

The above example illustrates that the number of predicted conversions increases rapidly by increasing of P in Γ_P^c.

2.2. *Convertibility functions for reaction* PES *study*

The problems connected to the formulation of the general convertibility function for reaction PES study are closely related to the problems solved in the frame of computer-assisted synthesis design. The generation of the starting set ${}^0V''$ is the first of them. In general, the set ${}^0V''$ is a subset of Family of Isomeric Ensembles of Molecules over a set of atoms A, FIEM(A) [5], or a subset of Family of Isomeric Synthons, FIS(A) [6–8]. However, the choice of ${}^0V''$ is dependent on the domain of PES which we are interested in. If only one reaction will be of interest to us, then ${}^0V''$ may be composed of only two nuclear configurations, reactant and product of the reaction. If the synthesis of a structure is in the center of our attention, then ${}^0V''$ should be composed of the structure and its possible precursors. In the first approximation, we use the notion of a conversion neighbor, CN, for the definition of reaction convertibility function Γ^R. We say that a chemical species with the nuclear configuration v_j is a CN of a chemical species with the nuclear configuration v_i if the minima corresponding to v_j and v_i along PES are neighbors, i.e. they are separated by a critical point of the PES of index 1. Note that this definition is a special case of a more general definition of "neighbor" and "strong neighbor" relations introduced by Mezey [9]. The set of all CN of v_i will be denoted as $\mathscr{S}(v_i)$. Formally, the general reaction convertibility function Γ^R is defined as follows:

$$\Gamma^R(v_i, v_j) = \begin{cases} 0, & \text{if } D_c(v_i, v_j) = 0, \\ 1, & \text{if } v_j \in \mathscr{S}(v_i), \\ 2, & \text{otherwise.} \end{cases} \tag{8}$$

The function D_c is a distance function which identifies in a heuristic manner whether two nuclear configurations are in the same catchment region of a critical point of PES. This function will not be further specified here. Note that the mechanism introduced below can work without this function.

In order to incorporate a heuristic concept we have to substitute the exact definition of CN by a heuristic definition. One of the possibilities is the synthetic precursor/successor (SPS) introduced in the frame of the so-called synthon model of organic chemistry [6–8, 10]. CN may be understood as SPS without virtual atoms.

However, the starting set ${}^0V''$ may be "non-dense" in the sense that there does not exist such a couple v_i, v_j in ${}^0V''$ that those are CN. It is, for example, the situation when ${}^0V''$ has only two elements, reactant and product of a many-step synthesis. In order to find the convertibility on such a "non-dense" set we have to complete this set by new nuclear configurations, and to create a new starting set ${}^0V''$. For this reason, the convertibility network algorithm, CNA, has been developed.

The convertibility network algorithm – algorithm 1

The algorithm is based on the so-called multilateral generation which is a generalization of the bilateral algorithm used in the program RAIN for reaction network production [11], and in the algorithm for reaction distance evaluation [12, 8]. The general distance function $D_G(v_i, v_j)$, which measures a distance of two nuclear configuration v_i, v_j, is used by the algorithm. In real situations, this function is substituted, for example, by chemical distance CD (in the RAIN program), by reaction distance RD [13], [7, 8, 12] or by a geometry based distance.

Let us suppose that there exists $v \in \mathscr{S}(v_i)$ such that $D_G(v_i, v_j) > D_G(v, v_j)$ for arbitrary $v_i, v_j \in {}^0V''$. Let us consider the following denotation: ${}^0V'' = \{v_1, \dots, v_k\}$, let $G'_H = (V'_H, E'_H)$ and $G_H = (V_H, E_H)$ be an auxiliary graph with vertex set V_H and V'_H and edge set E_H and E'_H, respectively. Let $I, F_i, C_i, Q_i^m, R_i^m$ be auxiliary sets. Further, let us suppose that if $v' \in \mathscr{S}(v)$ then $v \in \mathscr{S}(v')$ for any v.

```
 0. Initialization: $V_H := {}^0V'', E_H := \emptyset, I := \{1, 2, \dots, k\},$
                    $V'_H := \emptyset, E'_H := \emptyset, F_i := I - \{i\},$
                    $R_i^0 := Q_i^0 := \emptyset, R_i^m := m,$ for $i \in I$ and $m > 0.$
 1. $m := -1$
 2. $m := m + 1$
 3. Cycle for $i := 1, \dots, k$
 4. For each $v \in R_i^m$ do step 5
 5. If there exists $q \in F_i$ and $v' \in \mathscr{S}(v), v' \notin Q_i^m$ such that $D_G(v, v_q) > D_G(v', v_q)$
    then for each $q$ and $v'$ of that property put:
        $R_i^{m+1} := R_i^{m+1} \cup \{v'\}, \quad V_H := V_H \cup \{v'\}, \quad E_H := E_H \cup \{v, v'\}.$
    If there do not exist such $q$ and $v'$ put $R_i^{m+1} := \emptyset.$
 6. Put $Q_i^{m+1} := Q_i^m \cup R_i^{m+1}.$
 7. End of cycle $i$.
 8. Cycle for $i := 1, \dots, k - 1$
 9. Cycle for $j := i + 1, \dots, k$
10. If $j \notin C_i$ then
        if $Q_i^{m+1} \cap Q_j^{m+1} \neq \emptyset$ then
```

```
    $C_i := C_i \cup \{j\}$, $C_j := C_j \cup \{i\}$,
    $E'_H := E'_H \cup \{$all the edges of which all the paths directing from $v_i$ to $v_j$ in $G_H$ via
        any vertex $v$ from $Q_i^{m+1} \cap Q_j^{m+1}$ are composed$\}$,
    $V'_H := V'_H \cup \{v, v' \mid \{v, v'\} \in E'_H\}$.
11. End of cycle $j$.
12. End of cycle $i$.
13. $C_i := C_i \cup C_j, j \in C_i$, for $i := 1, \ldots, k$.
14. $F_i := F_i - C_i$, for $i := 1, \ldots, k$.
15. If $C_1 \neq I$ then go back to step 2.
16. Put ${}^0V'' := V'_H$.
17. The end of the algorithm.
```

It is easy to see that, in the sense of Eq. (8), $\Gamma^R(v, v') = 1$ iff $\{v, v'\}$ is an edge of the graph G'_H, i.e. $\{v, v'\} \in E'_H$. We call the graph G'_H a **convertibility network graph, CNG**. Note that this graph is heuristic estimation of the convertibility graph ${}^kG^c_{\mathrm{PES}(A)}$ defined in Ref. [1]. Of course, if the set ${}^0V''$ contains only two elements then algorithm 1 is bilateral. Although the general multilateral algorithm may be reformulated as a sequence of bilateral algorithms by the following algorithm 2, this reformulated algorithm is not equivalent to the original one.

The algorithm 2

Let us suppose the same denotation as used in the algorithm 1.

Let I' be a set of all $\binom{k}{2}$ couples which can be constructed from the elements of the set I, i.e. $I' = \{\{i, j\} \mid i, j, \in I, i \neq j\}$. Let $G'' = (V''_H, E''_H)$ be a graph, C'_i be auxiliary sets for $i \in I$.

```
0. Initialization: $V'' := {}^0V''$, $E''_H := \emptyset$, $C'_i := \{i\}$, for $i \in I$.
1. $d := \min\{D_G(v_i, v_j))\} \mid \{i, j\} \in I'\}$.
2. Take an arbitrary couple $\{i, j\} \in I'$ such that $D_G(v_i, v_j) = d$.
3. ${}^0V'' := \{v_i, v_j\}$.
4. Do algorithm 1.
5. $V''_H := V''_H \cup V'_H$, $E''_H := E''_H \cup E'_H$, $C'_i := C'_i \cup C'_j \cup \{1 \mid v_1 \in V'_H\}$, $C'_j := C'_i$.
6. $I' := I' \cup \{\{k, l\} \mid v_k \in V'_H, v_l \in V''_H - V'_H\} - \{\{j, t\} \mid j, t \in C'_i\}$.
7. If $I' \neq \emptyset$ then go back to step 1.
8. The end of the algorithm.
```

Comparing the above algorithms one can find out that algorithm 1 is more "memory consuming" but it is faster because generally more bilateral steps are realized in one multilateral step. In spite of that, some "local parts" of the entire convertibility network may be omitted by algorithm 2. It is schematically visualized in Fig. 4.

If the distances of v_1, v_2; v_2, v_3; v_1, v_3 are similar then all the local parts C, . . . , F of the network may be closed in the same step of algorithm 1. The whole network for this case is schematically visualized in Fig. 4a. This situation cannot appear using algorithm 2 for which the final network is schematically visualized in Fig. 4b. It can be seen the local parts E and F are omitted.

Intuitively, it seems that the resulting graph of algorithm 2, G'', is a subgraph of the resulting graph of algorithm 1, G'_H, but this inclusion is not generally satisfied, and can be proved only for special types of the distance D_G. Further,

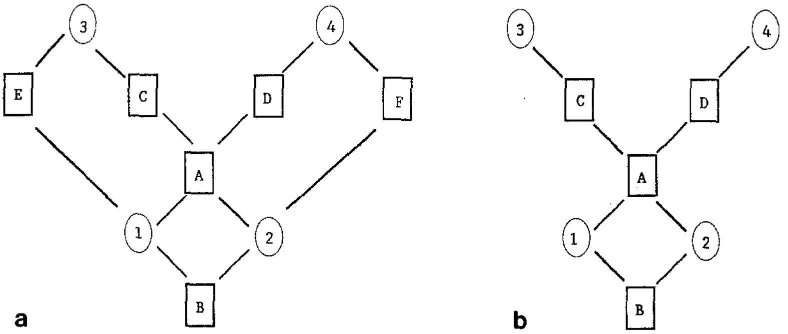

Fig. 4. The schematic picture of possible results of algorithms 1(a) and 2(b) on the basic set of four vertices

algorithm 2 may be modified to such a sequence of the bilateral algorithms which are realized only in between elements of the basic set ${}^0V''$. In this case, the final CNG, i.e. G'' is a subgraph of G'_H.

The computer implementation of CNA based on the above mentioned synthon model of organic chemistry is in progress [14]. In this implementation, the role of CN is played by the synthon precursor/successor, and a combination of chemical and reaction distance is used as the function D_G.

The starting set of nuclear configurations has been completed by use of the CNA. Now, further levels of the heuristic convertibility graph $G''_{\mathrm{PES}(A)}$, defined in [1], may be constructed. The algorithm DAISY [1] will be used by this construction. The algorithm requires a driver as an important part. The drivers for conformational and reaction changes study are discussed in the next section.

3. The drivers DRV

The general driver DRV role in the model has been described in the first part [1]. Now, we turn our attention to the formulation of conformational and reaction drivers. Note that the configurational driver can be developed in a similar way, and the group theory approach to theoretical stereochemistry of Ugi et al. [2] can serve as a good inspiration and basis.

3.1. The conformational driver DRV^c

Two principally different functors DRV^c can be proposed for the general conformational convertibility function Γ^c_P. It is driver DRV^c_I for independent driving of each explicitly followed internal coordinate (cf. Eq. (1)) and that for simultaneous driving of a subset of n internal coordinates ($n \leqslant P$) called DRV$^c_{Sn}$. We will turn our attention to the DRV^c_I which will be constructed in the following manner. Let us suppose the general form DRV(v_i, v', v_j, P, S, Γ', Θ, U, C) used in Ref. [1]. Let v_i and v_j be expressed in the same way as in Eqs. (2a) and (2b), respectively. Let S_{ij} be a set of "intermediate" nuclear configurations defined in Sect. 2.1.1, in our case for the set C. Now, we should separate two different cases:

(a) $v' = v_i$, i.e. DRV should generate the first step on the path $v_i \to v_j$.

(b) $v' \neq v_i$, i.e. v' is an "internal" point of the path $v_i \to v_j$.

(a) Let T be the set defined by Eq. (3), $n = \text{card}\, T$, i.e. the conformations v_i and v_j differ in n dihedral angles. Since each dihedral angle may be driven in $+$ or $-$ direction then there exist 2^n possibilities of how to perform driving of only one dihedral angle. Formally, one can write that a set W is created by all the possibilities.

$$W = \{(w_1, \ldots, W_n) \mid w_i = 1, -1, \text{ for } i = 1, \ldots, n\}. \tag{9}$$

Let us define a subset W' of W as follows:

$$W' = \{(w_1, \ldots, w_n) \mid \text{there exists } v_k \in S_{ij} \text{ such that } O(e^i_{ti}, e^k_{ti}, e^j_{ti}) = w_i, \quad \text{for } i = 1, \ldots, n\} \tag{10}$$

where the function O is defined by Eq. (6).

The set W' is an expression of all the possible rotations of dihedral angles of the nuclear configuration v_i which would be pointed to an "intermediate" nuclear configuration between v_i and v_j, i.e. to a nuclear configuration from the set S_{ij}. Let us define the set W'' as follows:

$$W'' = W - W'. \tag{11}$$

Let us define an auxiliary vector X over the set W'' in the following manner:

$$X = (x_{t1}, \ldots, x_{tn}), \tag{12}$$

such that $x_{ti} = 1 \Leftrightarrow w''_{ti} = 1, \quad \text{for } i = 1, \ldots, n,$

$x_{ti} = -1 \Leftrightarrow w''_{ti} = -1, \quad \text{for } i = 1, \ldots, n,$

$x_{ti} = 2, \quad$ otherwise.

The driver DRV^c_I performs a separate rotation of each dihedral angle ti with the step S counterclockwise if $x_{ti} = 1$, in the opposite direction if $x_{ti} = -1$, or sequentially in both directions if $x_{ti} = 2$. The ordering of the dihedral angle rotation may be different in various computer implementation. To start rotation with such an angle from the set T in which the nuclear configurations differ maximally is one of the possibilities. However, the driver needs a memory in order to eliminate repetition of rotation of the same dihedral angle in the same direction verifying the same conversion $v_i \to v_j$.

(b) In this case, the situation is simpler because it is known which dihedral angle is driven as well as the direction of the driving. Therefore, DRV starts from the value of this dihedral angle of the nuclear configuration v', and adds the step S.

The driver DRV^c_I is combinatorial. Thus, it is important, mainly for molecules with more degrees of conformational freedom, to combine the driver with some heuristic. The driver which uses the convertibility function Γ^c_1 has been tested as a part of the program system DAISY [4], versions 1.1 and 1.2 for conformational analysis of open chain molecules. In this case, the combinatorial explosion is fully eliminated. For systems with cyclic moieties, it is possible to include this driver using the convertibility function Γ^c_P where P is the cycle largeness. In general, the problem is more complicated because some of the internal coordinates are dependent. One of the possibilities to overcome this fact seems to be utilizing the ring-puckering theory [15].

3.2. The reaction driver DRV^R

The driver DRV^R verifies conversions predicted by the function Γ^R which is defined by Eq. (8), i.e. DRV^R generates reaction paths along PES between two CN. The approach is based on the idea that any reaction can be understood as a composition of elementary reactions. From the PES point of view, the elementary reaction is a process of changing a starting nuclear configuration and corresponding electron redistribution to the "transition state" or "unstable intermediate", or changing "transition state" or "unstable intermediate" to the final state, i.e. product. However, the notion of "unstable intermediate" has to be understood relatively because some "unstable intermediate" can appear as products which may be isolated by changing conditions. The abstract notion of elementary reaction can be seen as mechanisms connected to elementary processes of reorganization of valence electrons, ESRE [16, 17], i.e., heterolysis and homolysis of a bond, donation and colligation of radicals and redox processes. ESRE have been used in a similar meaning recently [6, 7, 8, 18, 19]. The starting point of the work of the driver DRV^R is the construction of the so-called elementary reaction network which is creacted by the elementary reaction network algorithm, ERNA. From the general point of view, it is similar to algorithm 1 for a bilateral case.

The elementary reactions network algorithm – algorithm 3

We use the following denotation. Let R be a set of all the elementary reactions to be considered, $R = \{r_1, \ldots, r_n\}$. Let v_1, v_2 be starting structures. Define auxiliary graphs $G_H = (V_H, E_H)$ and $G'_H = (V'_H, E'_H)$. A distance function D_R is used as a tool to measure the distance between single nuclear configurations. This function is dependent on the elementary reactions used. The above mentioned reaction distance RD may serve as an example.

```
 0. Initialization: $V'_H := \emptyset$, $E_H := \emptyset$, $V_H := \{v_1, v_2\}$,
                    $Y_1^0 := \{v_1\}$, $Y_2^0 := \{v_2\}$,
                    $Q_1^0 := Y_1^0$, $Q_2^0 := Y_2^0$,
                    $Y_1^m := Y_2^m := Q_1^m := Q_2^m := \emptyset$,   for $m > 0$.
 1. $m := -1$
 2. $m := m + 1$
 3. Cycle for $t := 1, 2$
 4. Do 5 for each $v \in Y_t^m$
 5. Do 6 and 7 for each $r_i \in R$
 6. $v' := v \oplus r_i$, where the operation $\oplus$ denotes a formal application of the
    elementary reaction $r_i$ to the nuclear configuration $v$.
 7. If $D_R(v, v_t) > D_R(v', v_t)$ then $Y_t^{m+1} := Y_t^{m+1} \cup \{v'\}$, $V_H := V_H \cup \{v'\}$, $E_H :=$
    $E_H \cup \{v, v'\}$.
 8. $Q_t^{m+1} := Q_t^m \cup Y_t^{m+1}$
 9. If $t = 1$ and $Q_1^{m+1} \cap Q_2^m \neq \emptyset$ then $Q := Q_1^{m+1} \cap Q_2^m$ and to to step 12.
10. If $t = 2$ and $Q_1^{m+1} \cap Q_2^{m+1} \neq \emptyset$ then $Q := Q_1^{m+1} \cap Q_2^{m+1}$ and go to step 12.
11. The end of the cycle $t$.
12. $E'_H := E'_H \cup$ {all the edges of which all the paths going from $v_1$ to $v_2$ in $G_H$
```

via any vertex v from the set Q are composed},

$$V'_H := V'_H \cup \{v, v' \mid \{v, v'\} \in E'_H\}.$$

13. The end of the algorithm.

The final graph G'_H contains all the paths which have to be verified. In other words, G'_H should contain all the reaction mechanisms of interest studying the change of the chemical species with the nuclear configuration v_1 to the chemical species with the nuclear configuration v_2.

The construction of the elementary reaction network utilizing ERNA is the first step of the elaborating of a conversion predicted by the function Γ^R by the general driver DRV^R. This step has always to be performed when DRV$^R(v_i, v', v_j, P, S, \Gamma', \Theta, U, C)$ is called and $v_i = v'$. The network generated and saved in the auxiliary graph G'_H is then elaborated by the verification of each elementary reaction. Each "internal" vertex of the network is seen as a "transition state" or a "unstable intermediate". The number of paths between v_1 and v_2 in the graph G'_H is returned by DRV^R in the variable P. A part of the elementary reaction network for the conversion $(CH_3)_2HCI + OH^- \rightarrow (CH_3)_2HCOH + I^-$ is visualized in the following scheme:

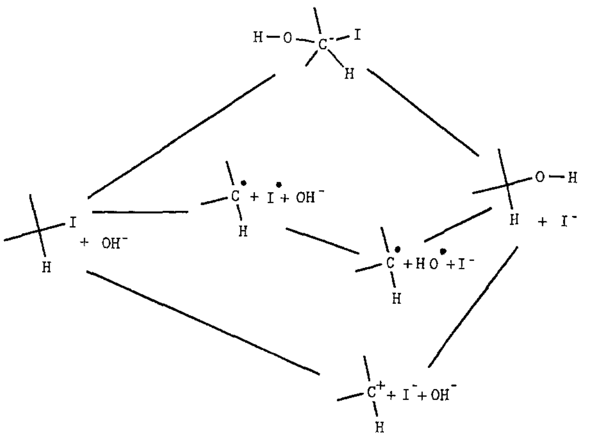

Fig. 5.

It is supposed in the above example that the elementary reactions are identical to the ESRE used in the synthon model of organic chemistry [7, 8]. In such a case, the final graph G'_H of algorithm 3 is a subgraph of the graph of Family of Isomeric Synthons, $G_{\mathrm{FIS}(A)}$. Computer implementation of the ERNA algorithm for this type of elementary reactions is in progress [14]. Note that the set of elementary reactions can contain "parts" of processes taking place simultaneously, for example, in the S_N2 mechanism. The above ideas can be combined with the reaction drivers currently used, cf. for instance Ref. [20].

Acknowledgements. The author wishes to express his appreciation for stimulating discussions with Prof. V. Kvasnička and Dr. J. Pospíchal (Slovak Technical University, Bratislava), Prof. M. Kratochvíl and Dr. V. Potůček (Research Institute of Pure Chemicals, Lachema, Brno), Dr. E. Hladká and Dr. L. Matyska (Masaryk University, Brno). I would like to thank Dr. B. Herzlík, Prof. J. Jonas and Dr. L. Matyska for language corrections.

References

1. Koča J Theor Chim Acta 80
2. Ugi I, Dugundji J, Kopp R, Marquarding D (1984) Perspectives in theoretical stereochemistry. Lecture notes in chemistry, vol 36, Springer, Berlin Heidelberg
3. Wilson S (1986) Chemistry by computers. Plenum Press, New York
4. Koča J, Carlsen PHJ (submitted) J Comput Chem
5. Dugundji J, Ugi I (1973) Top Curr Chem 39:19
6. Koča J (1988) Coll Czech Chem Commun 53:1007
7. Koča J (1989) J Math Chem 3:73; 3:91
8. Koča J, Kratochvíl M, Kvasnička V, Matyska L, Pospíchal J (1989) A synthon model of organic chemistry and synthesis design. Lecture notes in chemistry, vol 51. Springer-Verlag, Berlin Heidelberg
9. Mezey PG (1982) Theor Chim Acta 60:409
10. Kvasnička V, Pospíchal J (1990) Int J Quant Chem 38:253
11. Fontain E (1987) Die bilaterale Generierung von Reaktinsnetzwerken, dissertation, Technische Universität, München; Fontain E, Bauer J, Ugi I (1987) Chem Let 16:37
12. Kvasnička V, Pospíchal J (1989) J Math Chem 3:161
13. Koča J (1988) Coll Czech Chem Commun 53:3119
14. Hladká E (personal communication)
15. (a) Cremer D, Pople JA (1975) J Amer Chem Soc 97:1354
    (b) Diez E, Esteban AL, Guilleme J, Bermejo F (1987) J Mol Struct 70:61
    (c) Liwo JA, Ciarkowski J (1989) J Mol Struct (Theochem) 183:331
16. Kratochvíl M (1983) Chem Listy 77:225
17. Kratochvíl M, Koča J, Kvasnička V (1984) Chem Listy 78:1
18. Koča J, Kratochvíl M, Kunz M, Kvasnička V (1984) Coll Czech Chem Commun 49:1247
19. Schubert W (1979) MATCH 6:213
20. Halgren TA, Lipscomb WN (1977) Chem Phys Let 49:225; Dewar MJS, Healy EF, Stewart JJP (1984) J Chem Soc Faraday Trans 80:227