Tetrahedron Vol. 34, pp. 2939–2945
© Pergamon Press Ltd., 1978. Printed in Great Britain

0040-4020/78/1001-2939/$02.00/0

# AUTOMATIC ESTIMATION OF RING STRAIN ENERGIES

JOHANN GASTEIGER[1a*] and OLIVER DAMMER[1b]

Institut für Organische Chemie der Technischen Universität München, D-8046 Garching, West Germany

and

Fachbereich Informatik, Fachhochschule München, D-8000 München, West Germany

(*Received in UK 2 December 1977; Accepted for publication 25 April 1978*)

**Abstract**—A simple model has been developed which allows to estimate ring strain energies of monocyclic and bicyclic compounds. It can also be applied to polycyclic systems with reasonable accuracy. A computer program based on this model has been implemented which relies on a topological representation of molecular structures and performs the estimation of the effects of ring strain both on molecular energies and on reaction enthalpies.

The quantitative treatment of equilibria and reaction rates in organic chemistry requires information on the energies of molecules. Experimental thermochemical data can be gathered only with large expense in time and effort and are therefore available for a rather small collection of compounds.[2,3] Thus recourse has to be made to other approaches to molecular energies. In principle, heats of formation can be obtained through quantum mechanical calculations. At present, such an approach is feasible only for quite small species. Even there, the computations are rather time-consuming and expensive. The values obtained quite often lack the accuracy and reliability required for the prediction of chemical phenomena.

With the theoretical approach not yet being of general utility various models for estimating heats of formation have been designed. A large amount of work has been dedicated to the development of empirical force field methods.[4] Only a limited number of classes of compounds has been treated; 3-membered rings have yet escaped parametrization. A further drawback is that molecular mechanics calculations have to go through many iterations until the energy minimum is reached, requiring a fair amount of computer time.

Another approach to the estimation of heats of formation of molecules is based on additivity schemes.[2] Here, the heat of formation of a particular molecule is calculated from additive structural contributions. Various methods have been suggested, their accuracy depending predominantly on the number of parameters which is taken into account. These schemes are attractive because of their simplicity and yet, their accuracy and reliability is high.

In developing a computer program for generating reactions and planning organic syntheses (EROS)[5] we were faced with the problem of calculating reaction enthalpies. As a fast, accurate, and reliable method was required we developed a computer program[6] which is based on an additivity scheme for the estimation of heats of formation. From the various additivity schemes proposed, the one introduced by Allen[7] was chosen because it allowed the implementation of a generally applicable computer program.

## *Strain energy*

Because unbranched hydrocarbons and their derivatives fit additivity laws for heats of formation so well they are often taken as standards. Deviations from these additivity schemes are attributed to strain and the differences between observed heats of formation and the values estimated from additivity relations are called strain energies. According to this definition strain energies can have negative or positive values due to stabilization or destabilization of the molecule.

Large amounts of strain are often experienced with cyclic structures. In this paper we will be dealing with destabilizing effects through the presence of rings in a molecule. Stabilizing factors, e.g. those responsible for —negative—resonance energies, will be dealt with in a forthcoming paper.

The ring strain energy of a cyclic structure can only be assigned relative to a reference state.[8] In our system this standard is given by the value estimated from the bond and interaction parameters of the Allen scheme. As our parameters for the alkanes were obtained from a least-squares analysis of 49 alkanes,[9] in effect, this group of compounds constitutes the reference state. These alkanes comprise unbranched and branched compounds. Gauche 1,4-interactions are absorbed into the parameters. As only two of the compounds possess 1,5-interactions they are not taken into account and therefore contribute to the strain of a molecule. The Allen scheme is numerically equivalent to the group additivity scheme of Benson.[11] Both methods therefore lead to the same values for the ring strain energy if consistent parameters are used. Table 1 gives our parameters for the estimation for the heats of formation of alkanes.

With these values the estimated heat of formation of cyclopropane is −14.67 kcal/mole. As the experimental

Table 1. Structural units, notations, and values (in kcal/mole) for the estimation of heats of formation of alkanes according to the Allen scheme

| C - C | C - H | C - C - C | C - C(C)C |
|---|---|---|---|
| B(CC) | B(CH) | G(CCC) | D(CCC) |
| + 4.49 | - 4.14 | - 1.10 | + 0.06 |

heat of formation of cyclopropane is +12.73 kcal/mole, its strain energy amounts to 27.4 kcal/mole.

Rather than to treat the various mechanisms leading to ring strain separately—and to be faced with the difficult task of assigning values to the individual contributions—we decided to deal with ring strain energies as a whole.

*The model*

Our objective was to develop a model for calculating ring strain energies which should be easily amenable to algorithmic representation and should result in a very fast program, at least a hundred times faster than molecular mechanics calculations. We were prepared to sacrifice some accuracy for the sake of speed.

We have developed a simple model for estimating ring strain energies. As a first approximation the strain energy of a polycyclic compound was considered as being additive in the strain energies of the constituting rings. The strain energy of a ring with a given size was taken as having a constant value, independent of the specific molecule at hand and its substituent pattern. To improve on that approximation, correction parameters for combinations of two rings were introduced.

The constituting rings of a molecular structure were taken from the smallest set of smallest rings (SSSR). We have developed a computer program which determines the SSSR from the list of bonds of a molecule.[12] Aside from its general importance[13,14] the SSSR is particularly appropriate for the purpose of estimating ring strain. It is with the small rings where most of the strain energy resides. Further, it would be erroneous to assign strain energies to rings which are envelopes of smaller ones. For example, the strain energy which has to be assigned to cyclodecane is no longer present in decaline although the two 6-membered rings of decaline have a 10-membered ring as envelope. In decaline the nonbonded interactions largely responsible for the strain energy of cyclodecane do not exist any more.

For the individual rings from the SSSR the strain energy of the monocyclic carbocyclic compound is taken. Heterocyclic compounds are not yet treated explicitly as long as not sufficient experimental data are available. From the presently known data it appears that the error introduced by substituting the value of the carbocycle for a heterocyclic compound is small. For example, the strain energies of cyclopropane, oxirane, and aziridine are 27.5, 27.2 and 27.1 kcal/mole, respectively. Similar trends hold for the larger rings.[2] Only for sulfur compounds strain energies are usually smaller; for thiirane the value is 19.9 kcal/mole. The reasons lie presumably in the longer C–S bond lengths and the smaller C–S–C bending forces. In the following we will confine the discussion to carbocyclic compounds.

For the monocycles values are assigned according to (1) ring size, (2) number of double bonds and (3) number of allene or triple bonds. Table 2 lists values for 3- to 8-membered rings.

With our reference state for defining ring strain energies[9] for the compounds examined quite often nearly the same ring strain energies were obtained as given by Cox and Pilcher[2] or Benson.[11] In situations where more than one double bond occurs in one ring no differentiation has been made as to the location of these double bonds—except of allene type bonds—because only scant experimental data are available. With these values good estimates of ring strain energies can be made for substituted monocyclic compounds. This demonstrates the constancy of these contributions for a certain type of ring, the values being more or less independent of individual substituent patterns.

Table 2. Ring strain energies of monocyclic hydrocarbons (a) Ref. 2; (b) Ref. 15; (c) Estimate, this work; (d) Ref. 11; (e) Ref. 11, average of the values of two isomers; (f) Ref. 15, average value of several isomers; (g) Ref. 16

| ring size | saturated | 1 double bond | 2 double bonds | allene bond | triple bond |
|---|---|---|---|---|---|
| 3 | 27.5[a] | 53.8[a] | - | - | - |
| 4 | 26.5[a] | 30.0[a] | 50.0[c] | - | - |
| 5 | 6.2[a] | 5.7[a] | 6.0[d] | 30.0[c] | 30.0[c] |
| 6 | 0.0[a] | 1.3[a] | 2.6[e] | 20.0[c] | 20.0[c] |
| 7 | 6.2[a] | 4.7[b] | 4.3[f] | 15.0[c] | 15.0[c] |
| 8 | 9.3[b] | 4.9[b] | 6.1[f] | 10.0[c] | 10.0[g] |

For bicyclic compounds generally reasonable values for strain energies can be obtained by adding the strain energies of the individual rings. Appreciable deviations occur for systems where the two rings have more than two atoms in common (e.g. norbornane) or when one of the rings is a 3-membered one. To correct for all those deviations, even if they are only minor ones, we extended our scheme to also take into account the additional strain introduced—or relieved—by joining two rings.

Individual 2-ring combinations take into account: (1) the size of the two rings, (2) the number of atoms common to both rings (bridge atoms), and (3) the number of bridge atoms participating in double bonds. Tables 3–5 list the additional ring strain energies introduced when joining 3- to 8-membered rings through one, two, or three atoms, respectively.

Where possible, experimental data were employed but quite often recourse had to be taken to molecular mechanics calculations.

In each case the values were calculated consistently. This means that from the total ring strain energy of a bicyclic structure the strain energies of the constituent rings listed in the same reference were subtracted. This is necessary as the reference states of thermochemical and molecular mechanics methods differ somewhat. For many cases, neither experimental values nor molecular mechanics calculations are available. Then, rough estimates were made by us. These estimates are open to

Table 3. Extra ring strain energy for combinations of two rings with *one* joint atom (spiro-compounds); values our estimates unless otherwise noted (a) Ref. 2

| | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|
| 3 | 9.0[a] | | | | | |
| 4 | 5.0 | 3.0 | | | | |
| 5 | 2.0 | 0 | 0 | | | |
| 6 | 0 | 0 | 0 | 0 | | |
| 7 | 0 | 0 | 0 | 0 | 0 | |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 |

Table 4. Extra ring strain energy for combinations of two rings with *two* joint atoms; values our estimates unless otherwise noted. When different stereochemical arrangements at the ring junction are possible the value for the one with lower energy is listed (a) Ref. 2; (b) Ref. 11; (c) Ref. 4a; (d) Ref. 17

| | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|
| 3 | 10.2[a] | | | | | |
| 4 | 1.5[b] | -1.9[c] | | | | |
| 5 | -1.4[a] | -3.1[c] | -2.1[c] | | | |
| 6 | 0.9[a] | 0[a] | 0.1[c] | -0.9[c] | | |
| 7 | -4.5[a] | -2.5 | -0.8[d] | -0.5 | 0 | |
| 8 | -6.5[a] | -3.0 | -1.0 | 0 | 0 | 0 |

Table 5. Extra ring strain energy of combinations of two rings with *three* joint atoms; values our estimates unless otherwise noted (a) Ref. 17; (b) Ref. 18; (c) Ref. 4a

| | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|
| 4 | 38.4[a]<br>11.0[b] | | | | |
| 5 | 7.6[c] | 2.4[c] | | | |
| 6 | 8.1[c] | 3.4[c] | 6.7[c] | | |
| 7 | 7.5 | 3.0 | 4.0 | 3.0 | |
| 8 | 7.0 | 3.0 | 4.0 | 3.0 | 2.0 |

criticism; we hope that future work will provide us with data for these compounds.

The values of Tables 3–5 directly show the error introduced by an approach based entirely on additivity of the ring strain energy of individual rings. As can be seen, this approximation has quite a wide range of validity. But the error for the extremes of ring combinations was considered too high and forced us to adopt the present model which corrects for these errors. Thus, the ring strain energy of bicyclic compounds is reproduced exactly. The individual values of Tables 3–5 give a quantitative measure of the additional effects (bond angle deformations, eclipsed and gauche interactions, other nonbonded interactions) introduced when joining two rings through one, two or three atoms.

Values for additional systems (e.g. bicyclo[2.2.2]octane) were accessible and have been included. They are too scant as to warrant tabular listing. Molecular mechanics calculations on bridgehead olefins have appeared.[20] We used them to obtain values for the ring strain parameters for these 2-ring combinations.

*Application of the model*

The present model was designed to account for ring strain energies in molecules with monocyclic or bicyclic structures. It is an additivity scheme employing constant values for a given monocycle and additional increments for the junction of two monocycles to a bicyclic system. All mono- and bicyclic structures are considered explicitly. Therefore, their ring strain energy is reproduced exactly. For systems which differ in stereochemistry at the ring junction only a single value is used. Presently, the model has been developed for handling all combinations of 3- to 8-membered rings. Additivity of strain energies of rings separated through one or more bonds is well established. Therefore, molecules which contain several such ring systems are treated correctly, too. Only additional steric or polar interactions of substituents on the ring systems are not accounted for.

Our model is of an empirical nature; the values are obtained from thermochemical data or from molecular mechanics calculations which are an empirical method, too, parameterized, for its part, on thermochemical data. The accuracy of the model for mono- and bicyclic compounds is limited only by the availability and accuracy of thermochemical data.

(A) *Polycyclic structures*

Naturally, it is tempting to test the performance of this model with polycyclic systems. As only sporadic experimental data on polycyclic compounds are available values from molecular mechanics calculations[4a] were used as a test sample.

From the compounds listed in Ref. 4*a* all systems consisting of 4-, 5- or 6-membered rings were taken (Table 6). The biadamantanes, diadamantanes and triadamantanes were disregarded but otherwise no selection was performed. In column 3 of Table 6 the values for the strain energy calculated by the molecular mechanics method of Schleyer[4a] are contained.

Three approaches to the estimation of the ring strain energy of these polycyclic structures were taken.

First, the strain energies were calculated by adding the strain energies of the monocycles of the SSSR (Table 6, column 4). As the molecular mechanics calculations[4a] and the bond and group additivity scheme used by us[9] have slightly different reference states for unstrained molecules, the strain energies of the monocycles listed in Ref. 4*a* were taken to make the values consistent. The estimates thus obtained are often quite close to the values from the molecular mechanics calculations. But the many cases with sizeable deviations make this approach unacceptable as a general method.

Secondly, the model developed here, i.e. considering monocycles and 2-ring combinations explicitly, was applied with values for the bicyclic correction parameters taken from Tables 3–5 (see also Table 7, set b, column 2); the results are in column 5 of Table 6. The procedure shall be illustrated with compound 3 (Fig. 1). As the correction parameters were obtained from bicyclic compounds their application in the estimation of ring strain energies of polyclic compounds is questionable. But on the whole, remarkable improvements over the values from additivity of strain energies of monocycles is obtained. But certain trends can be seen which prompted us to make modifications leading to our third method.

Again, the model of working with parameters for monocycles and 2-ring combinations was used. But we took into account certain deviations in the ring strain energy of combinations of two rings when they are incorporated into polycyclic structures. Without going into the details of our reasoning of the individual changes we just list the new parameters in Table 7, set c, column 3. It is our hope that future experimental data will provide a basis for determining these values through least squares methods.

These minor changes lead to sizeable improvements in the estimated ring strain energies (Table 6, column 6). The compounds studied cover a wide range of structural diversity, from tri- to ennea-cyclic systems (dodeca-

Table 6. Estimations of ring strain energies of polycyclic compounds (a) molecular mechanics calculation, Ref. 4a; (b) sum of strain energies of monocycles from the SSSR; (c) as in (b), but with corrections for two-ring combinations with values from Table 7, column 2; (d) as in (b), but with corrections for two-ring combinations with values from Table 7, column 3; (e) difference between values of columns (a) and (d)

| | Cycloalkane | Ring Strain Energy (kcal/mole) | | | | |
|---|---|---|---|---|---|---|
| | | a | b | c | d | e |
| 1 | anti-Tricyclo[4.2.0.$0^{2,5}$]octane | 74.55 | 78.90 | 75.10 | 78.90 | -4.35 |
| 2 | syn- Tricyclo[4.2.0.$0^{2,5}$]octane | 79.45 | 78.90 | 75.10 | 78.90 | 0.55 |
| 3 | Tricyclo[3.2.1.$0^{3,6}$]octane | 41.46 | 40.86 | 37.06 | 41.26 | 0.20 |
| 4 | Tricyclo[3.3.0.$0^{2,7}$]octane | 48.29 | 40.86 | 45.36 | 47.46 | 0.83 |
| 5 | Tricyclo[3.3.0.$0^{3,7}$]octane | 47.15 | 21.84 | 24.54 | 25.14 | 22.01 |
| 6 | exo- Tricyclo[4.2.1.$0^{2,5}$]nonane | 42.51 | 40.86 | 40.16 | 42.26 | 0.25 |
| 7 | endo-Tricyclo[4.2.1.$0^{2,5}$]nonane | 47.33 | 40.86 | 40.16 | 42.26 | 5.07 |
| 8 | Tricyclo[4.2.1.$0^{3,7}$]nonane | 22.57 | 21.84 | 20.04 | 21.24 | 1.33 |
| 9 | Tricyclo[4.3.o.$o^{3,7}$]nonane | 25.47 | 21.84 | 24.54 | 25.14 | 0.33 |
| 10 | Tricyclo[3.3.1.$0^{3,7}$]nonane | 20.07 | 15.99 | 20.69 | 21.29 | -1.22 |
| 11 | Tricyclo[4.3.0.$0^{3,8}$]nonane | 34.17 | 15.99 | 21.89 | 22.99 | 11.18 |
| 12 | Tricyclo[4.2.1.$0^{4,9}$]nonane | 36.28 | 40.86 | 32.56 | 37.36 | -1.08 |
| 13 | Tricyclo[5.3.0.$0^{4,10}$]decane | 20.74 | 15.99 | 17.29 | 17.89 | 2.85 |
| 14 | exo- Tricyclo[5.2.1.$0^{2,6}$]decane | 22.65 | 21.84 | 22.14 | 22.74 | -0.09 |
| 15 | endo-Tricyclo[5.2.1.$0^{2,6}$]decane | 27.11 | 21.84 | 22.14 | 22.74 | 4.37 |
| 16 | trans-Tricyclo[5.2.1.$0^{2,6}$]decane | 41.65 | 21.84 | 22.14 | 28.94 | 12.71 |
| 17 | exo- Tricyclo[5.2.1.$0^{1,5}$]decane | 23.86 | 21.84 | 22.14 | 22.74 | 0.94 |
| 18 | endo-Tricyclo[5.2.1.$0^{1,5}$]decane | 31.15 | 21.84 | 22.14 | 28.94 | 2.21 |
| 19 | Tricyclo[5.2.1.$0^{3,8}$]decane | 19.26 | 15.99 | 18.59 | 20.79 | -1.53 |
| 20 | Tricyclo[5.2.1.$0^{4,10}$]decane | 15.68 | 21.84 | 15.54 | 17.34 | -2.66 |
| 21 | Tricyclo[4.4.0.$0^{3,8}$]decane | 26.12 | 4.29 | 20.19 | 19.19 | 6.93 |
| 22 | Tricyclo[4.3.1.$0^{3,7}$]decane | 20.77 | 8.71 | 19.01 | 21.21 | -0.44 |
| 23 | Tricyclo[5.2.1.$0^{4,8}$]decane | 21.48 | 15.99 | 17.39 | 19.09 | 2.39 |
| 24 | Tricyclo[4.4.0.$0^{3,9}$]decane | 30.90 | 10.14 | 16.04 | 18.74 | 12.16 |
| 25 | Tricyclo[4.4.0.$0^{3,7}$]decane | 23.59 | 15.99 | 21.89 | 22.99 | 0.60 |
| 26 | Tricyclo[4.2.2.$0^{1,5}$]decane | 29.65 | 21.84 | 20.04 | 27.44 | 2.21 |
| 27 | Tricyclo[4.3.1.$0^{3,8}$]decane | 18.29 | 10.14 | 20.34 | 17.74 | 0.55 |
| 28 | syn- Tricyclo[4.2.1.$1^{2,5}$]decane | 40.38 | 15.99 | 22.79 | 22.79 | 17.59 |
| 29 | anti-Tricyclo[4.2.1.$1^{2,5}$]decane | 30.66 | 15.99 | 22.79 | 22.79 | 7.87 |
| 30 | Adamantane | 6.87 | 4.29 | 24.39 | 13.29 | -6.51 |
| 31 | [2.2.2]Propellane | 73.28 | 78.90 | 73.20 | 78.90 | -5.62 |
| 32 | [3.3.3]Propellane | 16.49 | 21.84 | 15.54 | 17.34 | -0.85 |
| 33 | [4.4.4]Propellane | 14.23 | 4.29 | 1.59 | 1.59 | 12.64 |
| 34 | Tetracyclo[6.3.1.$0^{2,6}$.$0^{6,10}$]dodecane | 18.22 | 11.57 | 28.57 | 21.17 | -2.95 |
| 35 | Tetracyclo[5.3.1.$1^{2,6}$.$0^{4,9}$]dodecane | 25.02 | 5.72 | 31.62 | 16.82 | 8.20 |
| 36 | Tetracyclo[6.4.0.$0^{2,10}$.$0^{5,9}$]dodecane | 39.44 | 17.42 | 27.92 | 32.32 | 7.12 |
| 37 | Pentacyclo[6.4.0.$0^{2,6}$.$0^{4,12}$.$0^{7,14}$]dodecane | 44.49 | 36.40 | 35.20 | 37.60 | 6.89 |
| 38 | exo,endo-Tetracyclo[6.2.1.$1^{3,6}$.$0^{2,7}$]-dodecane | 40.84 | 29.12 | 31.82 | 32.42 | 8.42 |
| 39 | exo,exo-Tetracyclo[6.2.1.$1^{3,6}$.$0^{2,7}$]-dodecane | 39.88 | 29.12 | 31.82 | 32.42 | 7.46 |

Table 6(*Cont.*)

| | Cycloalkane | Ring Strain Energy (kcal/mole) | | | | |
|---|---|---|---|---|---|---|
| | | a | b | c | d | e |
| 40 | endo,endo-Tetracyclo[6.2.1.1$^{3,6}$.0$^{2,7}$]-dodecane | 49.89 | 29.12 | 31.82 | 32.42 | 17.47 |
| 41 | Cubane | 165.87 | 131.5 | 116.3 | 131.5 | 34.37 |
| 42 | Tetracyclo[4.2.0.0$^{2,5}$.0$^{3,8}$]octane | 111.98 | 105.2 | 95.7 | 105.2 | 6.78 |
| 43 | Pentacyclo[4.3.0.0$^{2,5}$.0$^{3,9}$.0$^{4,7}$]nonane | 118.13 | 112.48 | 93.68 | 109.48 | 8.65 |
| 44 | Pentacyclo[4.4.0.0$^{2,5}$.0$^{3,10}$.0$^{4,7}$]decane | 112.75 | 106.63 | 97.13 | 106.63 | 6.12 |
| 45 | Hexacyclo[5.4.1.0$^{2,6}$.0$^{3,10}$.0$^{5,9}$.0$^{8,11}$]-dodecane | 57.53 | 62.7 | 47.7 | 57.0 | 0.53 |
| 46 | Hexacyclo[5.3.0.0$^{2,6}$.0$^{3,10}$.0$^{4,9}$.0$^{5,8}$]-decane | 135.72 | 138.78 | 113.78 | 133.78 | 1.94 |
| 47 | Heptacyclo[6.4.0.0$^{2,7}$.0$^{3,12}$.0$^{4,11}$.0$^{5,10}$.0$^{6,9}$]dodecane | 155.19 | 159.23 | 147.83 | 159.23 | -4.04 |
| 48 | Heptacyclo[12.4.0.0$^{2,7}$.0$^{4,17}$.0$^{5,10}$.0$^{8,13}$.0$^{11,16}$]octadecane | 115.40 | 8.58 | 43.38 | 21.18 | 94.22 |
| 49 | Hexacyclo[7.5.1.0$^{3,13}$.0$^{5,12}$.0$^{7,11}$.0$^{10,14}$]pentadecane | 38.61 | 43.68 | 22.68 | 28.68 | 9.93 |
| 50 | Dodecahedrane | 42.98 | 80.08 | 27.58 | 42.58 | 0.40 |

Table 7. Extra ring strain energy (E.R.S.E.) for combinations of two rings; values in kcal/mole (a) for identification numbers see note 21; (b) see Tables 3–5; (c) modified values; see text

| Identification [a] | | E.R.S.E. | |
|---|---|---|---|
| | | b | c |
| 4420 | | -1.9 | 0 |
| 4520 | | -3.1 | -1.0 |
| 4620 | | 0 | 0 |
| 5510 | | 0 | 0 |
| 5520 | cis | -2.1 | -1.5 |
| | trans | -2.1 | 4.7 |
| 5530 | | 2.4 | 2.4 |
| 5620 | cis | 0.1 | 1.2 |
| | trans | 0.1 | 0.1 |
| 5630 | | 3.4 | 3.4 |
| 6620 | cis | -0.9 | -0.9 |
| | trans | -0.9 | 1.8 |
| 6630 | | 6.7 | 3.0 |
| 6640 | | 10.1 | 10.1 |

hedrane). The compounds have not been selected to particularly suit our model. It was only that for these molecules data from molecular mechanics calculations[4e] were available for comparison.

Appreciable deviations between the values calculated by molecular mechanics and our estimates (Table 6, column 7) occur in the following cases:

1. *For compounds* **5**, **21** *and* **41**. Here the selection of the SSSR is largely responsible for the deviations. An additional ring of the same size as the ones from the SSSR is contained in the polycyclic structure. Including this additional ring in the estimation procedure leads to improved values for **5** (35.72 kcal/mole) and **41** (157.8 kcal/mole). For **21** a value of 35.52 kcal/mole is obtained. This value is too high as the disappearance of nonbonded interactions is not taken into account. Because of these improvements when including the extra ring we are working on an extension of our ring finding algorithm. In the calculations of the influence of ring strain energies on heats of reaction the problems with the SSSR are already taken care of (see below). Thus these deviations can already be eliminated.

2. *For compounds* **7**, **11**, **15**, **16**, **24**, **28**, **29**, **33**, **35–40**, **48**, **49**. Here, the rigidity of the polycyclic structure results in appreciable H..H nonbonded interactions. This can be seen from molecular models. Only for these compounds appreciable interactions of that sort are to be expected. The differences between the values from molecular mechanics calculations and our values constitute an estimate of the magnitude of these nonbonded interactions—if there are no additional strain mechanisms as in **5** and **33**. This allows a simple access to the value of these interactions which is otherwise hard to come by. The estimate for the strain energy of adamantane is too high, as in comparison to bicyclo[3.3.1]nonane H..H nonbonded interactions disappear.

3. *For compounds* **5**, **16**, **33**, **41–44**. In these cases part of the deviation is the result of bond angle strain.

When the fusion of more than two rings to a polycyclic

structure does not introduce either one of these effects—additional H..H nonbonded interactions, or additional bond angle strain—our additivity scheme works astonishingly well. Although initially developed for mono- and bicyclic systems, where exact account of ring strain energies is given, these effects in polycyclic compounds are reproduced well, too. Larger deviations can be accounted for by additional strain mechanisms, and we can gain a direct estimate of the magnitude of these effects.

The scheme is extremely simple; by inspection of a constitutional formula the monocycles and the number and types of two-ring combinations are deduced. The appropriate values are taken from a table. Adding them up leads to the estimate of the ring strain energy (Fig. 1).

3

| | |
|---|---|
| monocycles | |
| (4) | 26.30 |
| (5) | 7.28 |
| (5) | 7 28 |
| bicyclic combinations[a] | |
| (4520) | -3.1 |
| (4520) | -3.1 |
| (5530) | 2.4 |
| | 37.06 kcal/mole |

Fig. 1. Estimation of ring strain energies of polycyclic structures (a) for identification numbers see note 21.

(B) *Automatization*

This procedure, outlined above, was used as a basis for a computer program for the automatic estimation of ring strain energies. From a topological representation of molecules[5] the smallest set of smallest rings is searched for;[12] the SSSR is transferred in a bit matrix. The strain energies of the members of the SSSR are taken from lists incorporated in the program. Due consideration is given to double, triple, and allene type bonds. Then, all possible combinations of two rings from the SSSR are generated. In this process the number of joint atoms (bridge atoms) and the number of bridge atoms participating in double bonds are considered. The identification numbers thus obtained for the 2-ring combinations allow to find the appropriate values for the additional ring strain energies in another list. This program gives an estimate for the ring strain energy of a molecule using only its topological representation as input. No information on its stereochemistry, in particular, no 3-dimensional coordinates are necessary.

Our model can be used to estimate not only the ring strain energies of molecules but also the effect of ring strain energies on heats of reaction. Two approaches have been taken.

First, the change in the ring strain energy during a reaction is calculated as the difference in the strain energies of starting materials and products. Here, the ring strain energies of individual molecules are estimated as above. This version of the program is of advantage when the starting materials have many ways to react as then the strain energy of the starting materials has to be evaluated only once.

In the second approach, only those rings are considered which contain atoms that participate in the reaction. Thus, the strain energy at the reaction site is estimated directly. We have the advantage that only part of the molecule has to be scanned making the program faster. On the other hand, as each reaction might proceed at different places of a molecule, the starting materials have to be evaluated for each reaction.

Two examples (Figs. 2 and 3) shall illustrate our two approaches. (The atoms marked in the figures indicate the reaction site.)

Fig. 2. Reaction A.

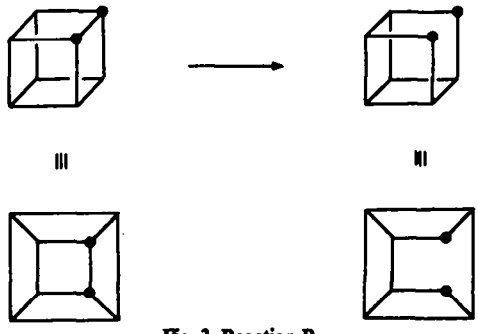

Fig. 3. Reaction B.

*Reaction A* (*Fig.* 2)

1. According to our first approach the estimates for the strain energies are: (In the following, the notation (n) represents the strain energy of an n-membered ring, the notation (ijk0) the additional strain resulting when a ring of size i is fused to a ring of size j through k atoms.)

| | |
|---|---|
| starting material | (6) + (3) + (4) + (3620) + (4620) |
| product | (3) + (4) |
| difference | (6) + (3620) + (4620) |

2. In the second approach, the respective strain energies are given by:

| | |
|---|---|
| starting material | (6) + (3620) + (4620) |
| product | — |
| difference | (6) + (3620) + (4620) |

As can be seen, in this example both approaches give the same results. Differences can occur in the few cases where through the selection of the SSSR an additional ring of the same size as some rings of the SSSR has been discarded. As already mentioned, this is the case with cubane (Fig. 3). The SSSR consists of five 4-membered rings, the additional 4-membered ring is linearly dependent on these five rings.

*Reaction B* (*Fig.* 3)

1. With the SSSR the first approach leads to:

| | |
|---|---|
| starting material | 5×(4)+8×(4420) |
| product | 4×(4)+5×(4420) |
| difference | (4)+3×(4420) |

2. With the second approach:

| | |
|---|---|
| starting material | 4×(4)+8×(4420) |
| product | 2×(4)+4×(4420) |
| difference | 2×(4)+4×(4420). |

Thus, the second approach duly accounts for the ring strain energy of two cyclobutane rings when one bond of cubane is broken.[23] The estimate for the release of strain energy by our model is 53.0 kcal/mole; molecular mechanics calculations[24] give 53.9 kcal/mole. This shows that our second method for estimating the effect of strain energies on heats of reaction gives the correct value and does not suffer from errors through the selection of the SSSR.

Simple search methods and addition of a few values are the only operations being performed. The program is therefore very fast, at least $10^3$–$10^4$ times faster than molecular mechanics calculations for a medium-sized molecule. In addition, our second approach considers only the immediate vicinity of the reaction site. Thus, with this method, computation times are rather independent of the size of the entire molecule. This is a big advantage over molecular mechanics or quantum mechanical calculations. There, computation times increase rapidly with the size of the molecules.

The program has been written in PL/1 and is routinely used in EROS,[5] our program system for generating sequences of reactions and for synthesis design. In this system reactions are generated through a mathematical formalism. The estimation of heats of reaction, and in this context of ring strain energies, is one of the evaluations performed on these reactions.

*Summary.* The model developed here allows to exactly reproduce ring strain energies of mono- and bicyclic structures. No explicit treatment of stereochemistry is necessary. The model is an additivity scheme; ring strain energies are estimated by adding up numbers, a procedure simple enough to be performed by inspection. The accuracy of the method is only limited by the availability of data. Deviations in the strain energies are only to be expected for substituted compounds with appreciable nonbonded or polar interactions.

As far as data on heterocyclic compounds are known they demonstrate that the ring strain energy estimates of carbocyclic compounds are good approximations for the strain energies of O- and N-heterocycles. With more data on heterocyclic compounds becoming available specific values for these systems could be assigned and incorporated into our model in a straightforward manner. Until now systems with 3- to 8-membered rings have been analyzed. It should be noted that molecules containing cyclopropane rings which have yet eluded treatment by molecular mechanics calculations are handled without difficulties.

With surprising success our model can be applied to polycyclic structures. Ring strain energies of a large variety of compounds are predicted correctly. Appreciable deviations occur only in polycyclic systems when additional sizeable bond angle deformations or nonbonded interactions are introduced. When data for comparison are available the magnitude of these effects can be estimated.

The model serves as a basis for a computer program for the automatic estimation of ring strain energies. As input only the constitution of a molecule is needed. The program is flexible as to allow the estimation of either the ring strain energy of a molecule or the change in the ring strain energies in the course of a reaction. This program is an integral part of our program system for generating and evaluating reactions (EROS).

It is our hope that more thermochemical data become available to be able to further improve on our values and to test the accuracy and limits of our model.

*Acknowledgement*—Financial support of our work by Deutsche Forschungsgemeinschaft is gratefully acknowledged.

REFERENCES

[1a]Institut für Organische Chemie, Technische Universität München, Lichtenberg-str. 4, D-8046 Garching, West-Germany; [b]Fachbereich Informatik, Fachhochschule München, D-8000 München 2, West-Germany.

[2]J. D. Cox and G. Pilcher, *Thermochemistry of Organic and Organometallic Chemistry*. Academic Press, New York (1970).

[3]D. R. Stull, E. F. Westrum and G. C. Sinke, *The Chemical Thermodynamics of Organic Compounds*. Wiley, New York (1969).

[4]For leading references: [a]E. M. Engler, J. D. Andose and P. v. R. Schleyer, *J. Am. Chem. Soc.* **95**, 8005 (1973); [b]N. L. Allinger, *Adv. Phys. Org. Chem.* **13**, 1 (1976).

[5]J. Gasteiger and C. Jochum, *Topics Curr. Chem.* **74**, 93 (1978).

[6]J. Gasteiger, *Comp. & Chem.* in press.

[7]T. L. Allen, *J. Chem. Phys.* **31**, 1039 (1959); A. J. Kalb, A. L. H. Chung and T. L. Allen, *J. Am. Chem. Soc.* **88**, 1938 (1966).

[8]For a comprehensive review: J. F. Liebman and A. Greenberg, *Chem. Rev.* **76**, 311 (1976).

[9]J. Gasteiger, P. Jacob and U. Strauss, manuscript submitted, data taken from Table I of Ref. 10.

[10]J. W. H. Kao and A. Chung-Phillips, *J. Chem. Phys.* **63**, 4143 (1975).

[11]S. W. Benson, *Thermochemical Kinetics* (2nd Edn.) Wiley, New York (1976).

[12]J. Gasteiger and C. Jochum, manuscript submitted.

[13]M. Plotkin, *J. Chem. Soc.* **11**, 60 (1971).

[14]A. Zamora, *J. Chem. Inf. Comput. Sci.* **16**, 40 (1976).

[15]R. B. Turner, B. J. Mallon, M. Tichy, W. v. E. Doering, W. R. Roth and G. Schröder, *J. Am. Chem. Soc.* **95**, 8605 (1973).

[16]R. B. Turner, A. D. Jarreff, P. Goebel and B. J. Mallon, *Ibid.* **95**, 790 (1973).

[17]P. v. R. Schleyer, J. E. Williams and K. R. Blanchard, *Ibid.* **92**, 2377 (1970).

[18]N. L. Allinger, M. T. Tribble, M. A. Miller and D. H. Wertz, *Ibid.* **93**, 1637 (1971).

[19]M. D. Newton and J. M. Schulman, *Ibid.* **94**, 773 (1972).

[20]U. Burkert, *Chem. Ber.* **110**, 773 (1977); O. Ermer, *Z. Naturforsch.* **32b**, 837 (1977).

[21]The identification numbers for the combination of two rings are constructed in the following manner: The first two digits give the sizes of the two rings, the third digit the number of joint atoms, and the last digit the number of bridge atoms participating in double bonds.

[22]J. Blair, J. Gasteiger, C. Gillespie, P. Gillespie and I. Ugi, *Tetrahedron* **30**, 1845 (1974).

[23]When one of the atoms taking part in a reaction is chosen as starting point for our ring finding algorithm all small rings containing these atoms are found even if the SSSR has several choices.

[24]R. Stober and H. Musso, *Angew. Chem.* **89**, 430 (1977).