Reaction mechanism studies made simple using simulated annealing. Potential energy surface exploration

Didier Siri[a,*], Anouk Gaudel-Siri[a,b], Jean-Marc Pons[b], Daniel Liotard[c], Michel Rajzmann[b]

[a]*Laboratoire de Chimie Théorique et Modélisation Moléculaire, UMR-CNRS 6517, Université de Provence, boîte D42, 13397 Marseille Cedex 20, France*
[b]*Laboratoire de Réactivité en Synthèse Organique (ReSo), UMR-CNRS 6516, Université d'Aix-Marseille, boîte D12, 13397 Marseille Cedex 20, France*
[c]*Laboratoire de Physico-Chimie Théorique, UMR-CNRS 5803, Université de Bordeaux I, 33405 Talence, France*

Abstract

A new and easy way to study reaction mechanisms by theoretical means is proposed. Simulated annealing is used to explore the potential energy surface of three different systems. This approach enables a facile finding of most stationary points involved in the various possible reaction paths arising from a given molecular system and thus allows one to determine a whole set of competing reactions. Therefore, by a comparison of the activation barriers, we can qualitatively estimate whether a reaction should be stereoselective or if side-products may arise. This paper deals with the $[2 + 2]$ cycloaddition leading to β-lactones and the thermolysis of thiirane and thiiranium cation. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Simulated annealing; Competing reactions; Energy; Gradient norm; Semiempirical method

1. Introduction

Starting from given reactants, the obtention of several different products arising from various competing reactions is a common problem to most experimental chemists. Moreover, the same product can sometimes be accounted for through different reaction pathways. Therefore, a knowledge of the mechanism of these reactions and even of the energies involved (activation energies and heats of reaction) is of great help to try to minimize the formation of

undesirable products. Simulated annealing appears as a valuable tool which can be used towards this goal.

Simulated annealing is a general method of function optimization which has been proposed by Kirkpatrick et al. [1]. In its original form, simulated annealing is based on the Metropolis algorithm and leads only to the global minimum. A Markov chain is generated, resulting from a random walk on the potential energy surface. If random moves lead downhill in energy, the new geometry is accepted. Moves uphill are also accepted but only with a *P* probability in accordance with the Boltzmann's law. The annealing process starts at a sufficiently high temperature so that the *P* probability to pass from one local minimum to another is large. Then the temperature is lowered and a new Markov chain is generated at each temperature. The

* Corresponding author. Tel.: +33-4912-884-14; fax: +33-4919-885-12.
  E-mail address: siri@dexter.u-3mrs.fr (D. Siri).

decrease in temperature continues until no significant changes in geometry are observed: the global minimum is probably reached. To the best of our knowledge, simulated annealing has only been used in chemistry and biochemistry for the search, by energy minimization, of the global minimum of a system and, in some cases, other stable conformations (energy was calculated by molecular mechanics [2–9], Lennard-Jones type potential [10,11], EHT [12], DFT [13–16] and mixed (classical and DFT [17,18]) methods applied to clusters, surfaces [19,20] and various small molecules [21–23]).

Nearly 10 years ago, Bockisch et al. proposed a modification of the method that allows to keep a trace of the exploration of the potential energy surface [24]. Two functions can be optimized: energy and energy gradient norm and, therefore, both minima and transition states of a given system could be found. In this paper, we provide the first applications of this concept. What are the advantages of such an approach? In traditional methods, one starts from the reactants and/or the products and the *expected* reactions and reaction paths are investigated sequentially. This requires in each case to *postulate* a structure for the transition state and to locate it by a trial and error method. This approach is time-consuming and only allows finding transition states which are searched. On the contrary, through this new method based on simulated annealing, several transition states can be found simultaneously, even those that were *not anticipated* by the chemist. Some transition states will be irrelevant but this procedure leads to a significant gain of time, requires less human insight and guidance and can provide interesting new ideas.

In order to illustrate the potentiality of this new approach of reaction mechanism study, we have chosen to apply it to the study of two simple known systems:

The [2 + 2] cycloaddition reaction between formaldehyde and ketene and the further decarboxylation of the so-formed β-lactone. Theoretical studies on both reactions, cycloaddition [25–28] and decarboxylation [29–31] were reported in the literature.

The thermolysis of thiirane. This reaction also was studied both on experimental and theoretical grounds [32].

Finally, to tentatively predict the outcome of a system, we have studied the possible reactions of thiirane in the presence of a proton (acid catalysis).

2. Computational methods

Simulated annealing calculations were performed with the semiempirical AMPAC 6.55 package [33]. Among the semiempirical methods, we chose to use AM1 [34] and SAM1 [35–38] which are known to give reliable structures [35,39,40]. The SAM1 method allows one to take into account the sulfur d orbitals, the absence of which led to unreliable results with AM1. Energies were in some occasions calculated with CI. Actually, an efficient method for semiempirical calculations with configuration interaction is implemented in AMPAC. This strategy has been successfully tested on many molecules, deliberately including transition states and molecules with open shells in their ground states [41].

Structures resulting from semiempirical calculations can then be used as starting points for further ab initio calculations. Indeed, in our own experience, AM1 results were very close to ab initio (HF/6-31G*) ones [28].

All simulated annealing calculations were carried out with a starting temperature giving an *RT* value of 300 or 400 kcal mol^{-1}. Penalty functions were added in order to discard irrelevant configurations before any energy calculation. These penalty functions deal with the moment of inertia of the system and configurations where atoms are within a 0.6 Å distance. In a first step, we performed a simulated annealing calculation based on the energy function starting from a stable structure of the studied reaction. Various minima belonging or not to the reaction path under study were found. In the second step, we performed a simulated annealing calculation based on gradient norm minimization from the zero index structures previously obtained even if they did not belong to the studied reaction, provided they were chemically relevant. Minima, transition states and higher order critical points were obtained. Critical points were all confirmed by calculation of the vibrational frequencies. High order critical points could usually be brought down to structures of index one, provided they had no more than two or three negative
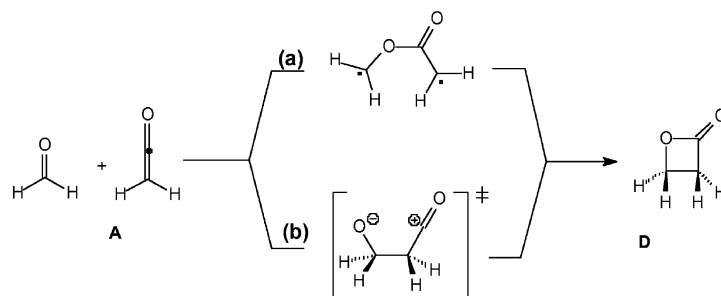
Scheme 1. Biradical (**a**) and zwitterionic (**b**) [2 + 2] cycloadditions leading to β-lactone.

eigenvalues. In the third step, IRC calculations were performed on both sides of each transition state in order to obtain the full reaction paths. If new stable intermediates were found, we restarted the search at step 2; otherwise the procedure was stopped. IRC calculations were run with the EPC algorithm [42] which was modified in 1997 [43] to allow to follow both sides of a reaction path from the normal vibration mode associated to a selected imaginary frequency. This modification generally allows the elimination of spurious eigenvalues when starting from a high index critical point. When success was not met with this strategy, the CHAIN algorithm was used. Both the energy and gradient norm local methods of minimization rely on a trust region strategy [44]. Since the original version [45], the simulated annealing method was improved in the following directions:

random shifts exhalted according to soft modes, penalty function improvements, and quenching performed by trust region minimisers.

The efficiency of these minimisers allows now to perform a quenching of all geometrical parameters of a molecule, starting from simulated annealing (focused on a small number of geometrical degrees) generated conformations.

3. Applications

3.1. [2 + 2] Cycloaddition reaction between ketene and formaldehyde and decarboxylation of 2-oxetanone

In our previous semiempirical and ab initio studies [25,26], we identified two competing reaction pathways:

The first one (mechanism (**a**) Scheme 1) involves the nucleophilic attack of a lone pair of the oxygen atom of the formaldehyde on the central carbon atom of the ketene and the further electrocyclization of the
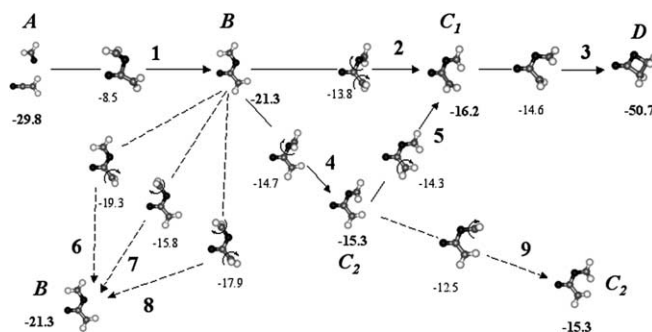

Fig. 1. [2 + 2] Cycloaddition reaction between ketene and formaldehyde. Mechanism (**a**). Numbers refer to reactions and letters to stable structures. Energies are given in kcal mol^{-1} and are written in bold for stable structures and in plain for transition states.
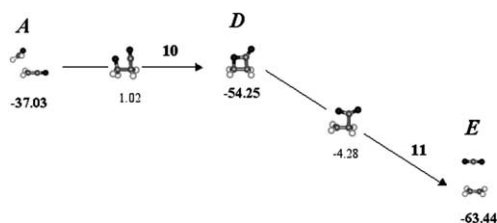
Fig. 2. [2 + 2] Cycloaddition reaction between ketene and formaldehyde (mechanism (**b**)) and decarboxylation of 2-oxetanone. Numbers refer to reactions and letters to stable structures. Energies are given in kcal mol^{-1} and are written in bold for stable structures and in plain for transition states.

adduct. It is a stepwise process with a marked biradical character.

The second mechanism (mechanism (**b**) Scheme 1), also studied by Cossio at an ab initio level [27,28], proceeds through the nucleophilic attack of the terminal carbon atom of the ketene on the carbon atom of the formaldehyde and the further cyclization of the adduct. It is an asynchronous concerted mechanism with a high zwitterionic character.

Mechanism (**a**). Since mechanism (**a**) has a biradical character, the Hartee–Fock determinants involved in the CI were built from the two highest single occupied SCF molecular orbitals. Starting from β-lactone **D**, a simulated annealing based on the energy function led to biradical intermediate **C$_1$** and products (ethylene and CO$_2$) resulting from the decarboxylation of β-lactone **D** (vide infra). From stable configuration **C$_1$**, a simulated annealing based on the gradient norm was performed and provided us transition states of reactions **2**, **3**, **4**, **6**, **8**, **9** and intermediate **B**. Transition state of reaction **7** was obtained in addition to the ones already obtained from **C$_1$** when calculations were performed from **B**. Then, starting from the transition structure of reaction **2**, a last simulated annealing based on the gradient norm led to transition state of reaction **5**. The transition state of reaction **1** was obtained from a configuration close to **B**. Two reaction paths could be identified from all these points; they are reported in Fig. 1: reaction path **I**, involving steps **1**, **2** and **3** and reaction path **II**, involving steps **1**, **4**, **5** and **3**. Both pathways are consistent with mechanism (**a**) (Scheme 1). Based on the activation energies, step **1** is in both cases the rate determining step. The other steps involved in reaction paths **I** and **II** are very

close in energy (less than 1 kcal mol^{-1} difference between step **2** and steps **4** and **5**). One can note that in step **2** the two methylene groups rotate simultaneously while in reaction path **II** these rotations occur one after the other (steps **4** and **5**) and that the second option is slightly favored. This energy difference would probably be enhanced if the hydrogen atoms were substituted by more bulky groups. Aside from these three pathways, four other transition states involved in rotations around single bonds were characterized (steps **6**, **7**, **8** and **9**). Since we carried out this study on the parent system, these steps did not lead to new conformations and the activation energies associated to these steps are rather low. With substituted molecules, mechanism (**a**) would probably lead to a mixture of diastereomers since rotations around three single bonds may occur.

Mechanism (**b**). Since mechanism (**b**) is a closed-shell process, Hartee–Fock determinants of the CI were built from closed-shell HOMO and LUMO. Results are reported in Fig. 2.[1] It appears that the only reaction found is one associated to mechanism (**b**). It is a concerted but asynchronous process involving a synperiplanar approach of the reactants (step **10**). Since no rotation can occur in that process, the reaction is more likely to be stereoselective when starting from substituted substrates. As established in our previous study [25,26], the reaction paths associated with mechanism (**a**) are favored over the single reaction path associated with mechanism (**b**). Indeed, the lowest activation energy is of 21.3 kcal mol^{-1} in the first case whereas it is of 38.0 kcal mol^{-1} in the second case. However, we have shown that this preference is reversed when a Lewis acid is introduced [26]. Moreover, such a result is in agreement with experimental observation [46].

Decarboxylation. Along with the reaction path associated to mechanism (**b**) we also found the transition state and products (**E**: ethylene and carbon dioxide) related to the decarboxylation of β-lactone **D** (Fig. 2 step **11** and Scheme 2). The value of the activation energy is 50 kcal mol^{-1}, identical to the one reported in 1989 in a previous AM1 study [29–31], and is very close to the experimental value

---

[1] Since the zero order CI functions are different (CI being partial) energy differences occur between the two mechanisms for common compounds, reactants **A** and product **D**.
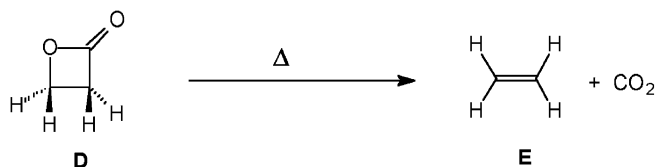
Scheme 2. Decarboxylation of 2-oxetanone.

(43.11 kcal mol^{-1} [47]) and to ab initio ones [31] (HF/6-31G*: 51.00 kcal mol^{-1}; MP2/6-31 + G*: 41.67 kcal mol^{-1}). The concerted mechanism we found is also in agreement with the ones reported previously (including those obtained with a difluoro β-lactone [30]) for an uncatalyzed process. The characterization of the transition state associated to decarboxylation (step **11**) is a good illustration of the potentiality of the simulated annealing approach. Indeed, we were able to find this reaction path even though we were not looking for it.
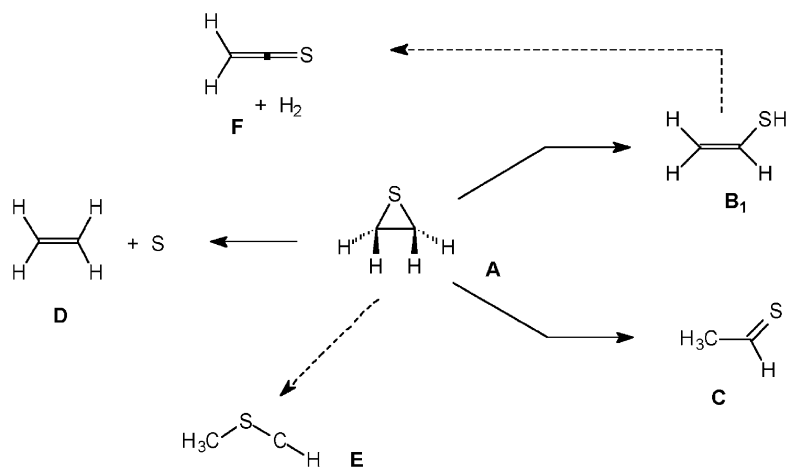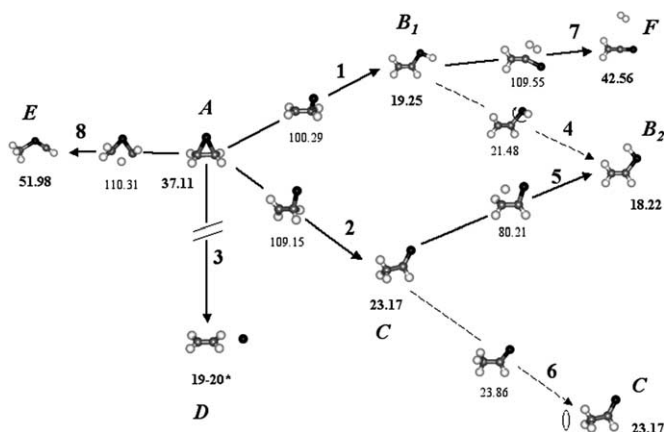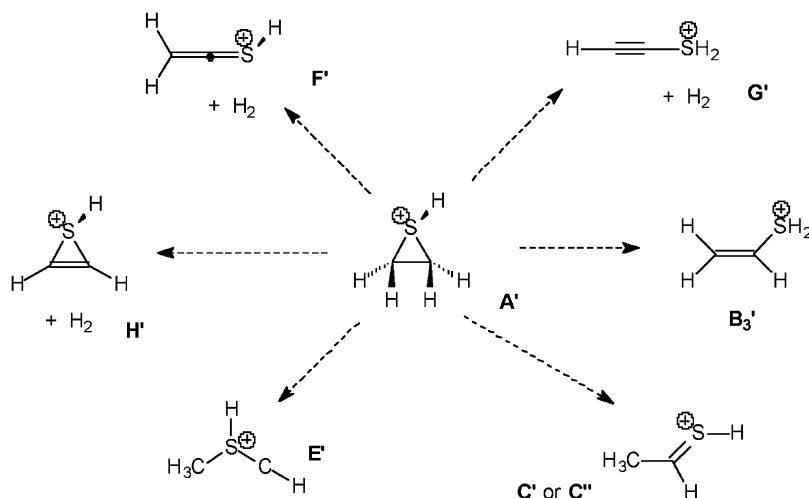

Scheme 3. Thiirane thermolysis.

Fig. 3. Thermolysis of thiirane. Numbers refer to reactions and letters to stable structures. Energies are given in kcal mol^{-1} and are written in bold for stable structures and in plain for transition states.

Scheme 4. Thermolysis of thiiranium ion.

3.2. Thiirane thermolysis

We then applied this simulated annealing approach to the study of a reaction on which both experimental data and quantum calculations were available [32], the thermolysis of thiirane (Scheme 3).

Starting from **A**, the procedure described in Section 3.1 was applied and we obtained all the reactions reported in Fig. 3. The three reactions reported in the experimental study [32] are found by calculations: reaction paths **1**, **2** and **3**. Activation energies for **1** and

2 are rather high (60–80 kcal mol^{-1}) as can be expected for thermolysis processes. Heats of formation of the products resulting from **3** were greatly overestimated by the SAM1 calculation (80.3 kcal mol^{-1} using UHF method ($\langle S^2 \rangle = 2.007$) to account for the S atom in the triplet state) because of the presence of atomic sulfur. Indeed the most stable form of sulfur is the S_8 form. Since its SAM1 heat of formation is of 20.3 kcal mol^{-1} and the SAM1 heat of formation of ethylene is 15.9 kcal mol^{-1}, the energy of ethylene and sulfur is more likely to be
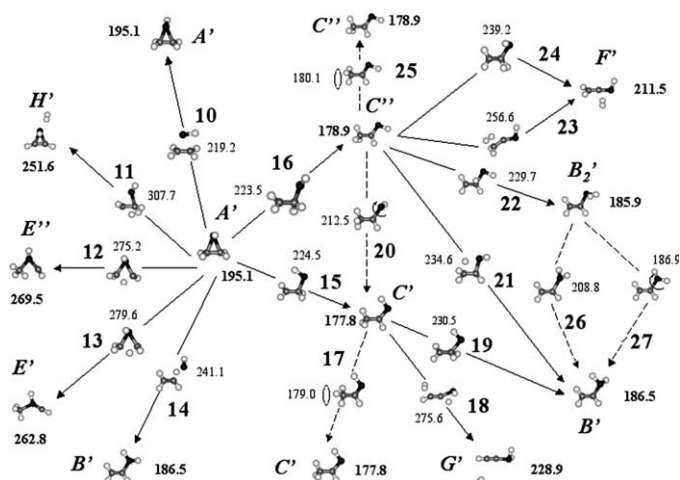

Fig. 4. Thermolysis of thiiranium ion. Numbers refer to reactions and letters to stable structures. Energies are given in kcal mol^{-1} and are written in bold for stable structures and in plain for transition states.

Table 1
Activation energies (kcal mol^{-1}) of selected reactions described both from thiirane and thiiranium ion

Thiirane		Thiiranium ion	
Reaction	E_a (kcal mol^{-1})	Reaction	E_a (kcal mol^{-1})
1	63.2	14	46.0
2	72.0	15	29.4
5	57.0	21	55.7
8	73.2	12	80.1

around 18–19 kcal mol^{-1}. Using SAM1 theory (even including a configuration interaction involving eight frontier molecular orbitals), we were not able to characterize a transition state for reaction **3**. Reactions **7** and **8** leading to the thioketene **F** and carbene **E** require comparable activation energies but they are highly endothermic. It is therefore not surprising not to found these products in the experiment.

With minimum calculations, simulated annealing allowed us to find reaction paths for all the transformations experimentally observed. We then decided to explore the influence of acid catalysis (H$^+$) on the same reaction.

3.3. Thiiranium thermolysis

Calculations were also carried out with SAM1 method. Starting from **A**$'$, seven stable configurations were obtained. From each minimum, a simulated annealing based on the gradient norm was performed and led to 20 transition states. The remaining configurations (minima and transition states) were obtained by the same procedure. Main reaction paths are reported in Scheme 4 and Fig. 4. Other transformations such as permutation of two hydrogen atoms or even three were also found. Since they have less chemical significance, we have not included them.

From a general point of view, an acid catalysis should induce a significant reduction of the activation energies involved in the various possible reactions. Indeed, the addition of a proton results in a global reduction of activation energies (Table 1). Thus, the activation energy of reaction **15** is only of 29.4 kcal mol^{-1} while the uncatalyzed process (reaction **2**) required 72.0 kcal mol^{-1}. However, the reduction can be less spectacular and sometimes nearly disappears. For example, reaction **5** (uncatalyzed process)

requires 57.0 kcal mol^{-1} and reaction **21** (related catalyzed reaction) requires 55.7 kcal mol^{-1}.

From **B**$'_2$ to **B**$'$, rotation of the $-SH_2^+$ group (reaction **27**) is favored over inversion of the sulfur atom (reaction **28**) by over 20 kcal mol^{-1}. This gap would probably be reduced with the introduction of more bulky substituents on the sulfur atom.

Reaction **18** leads to acetylene derivative **G**$'$; however the activation energy nearly reaches 100 kcal mol^{-1}, a value comparable to the one involved in reaction **12** (leading to protonated carbene **E**$''$). The highest transition state of the whole exploration is the one involved in reaction **11** leading to thiirene **H**$'$.

The reactions, analogous to the one experimentally observed with thiirane in the absence of catalyst are all found. The presence of the additional proton doubles the number of possible reactions from the thiiranium cation. Thus, interconversion reactions between different forms appear. Some reactions, such as **11** and **19**, seem to be specific to the catalyzed process since they were not detected in the absence of catalyst.

4. Concluding remarks: strategy and future developments

This study shows that the simulated annealing method applied to the gradient norm and coupled to the IRC method is a very promising approach for the study of competing mechanisms. The continuously growing power of computers should allow the exploration of larger systems in the near future. Performances could also be enhanced by the substitution of the gradient norm minimizer, which is a non-selective algorithm, by one specially dedicated to transition state search. IRC calculations could then be run automatically.

Indeed, the most commonly used algorithms for transition state search are of 'eigenvector following' type [48–52]. However, these methods turn out to be totally unstable when starting from conformations generated by a simulated annealing based on gradient norm minimization. Decisive progress will only arise with new generation methods, able to move a reaction path and not a single point on the PES [53–55]. An example of this kind of algorithm has recently been

implemented in AMPAC [56] (keyword: FULLCHN). The method is robust, less time-consuming than classical ones and provides a reaction path equivalent to what would have been found by the IRC method. When simulated annealing will become compatible with these new methods, it should lead to non-local strategies for the determination of transition states. These new techniques should give reasonable answers to the huge challenge of the theoretical study of competing chemical reactions.

References

[1] S. Kirkpatrick, C.D. Gellat Jr., M.P. Vecchi, Science 220 (1983) 671.

[2] S.R. Wilson, W. Cui, J.W. Moskowitz, K.E. Schmidt, Tetrahedron Lett. 29 (1988) 4373.

[3] F. Guarnieri, S.R. Wilson, Tetrahedron 48 (1992) 4271.

[4] L. Dosen-Micovic, Tetrahedron 51 (1995) 6789.

[5] S.R. Wilson, W. Cui, J.W. Moskowitz, K.E. Schmidt, J. Comput. Chem. 12 (1991) 342.

[6] A. Roitberg, R. Elberg, J. Chem. Phys. 95 (1991) 9277.

[7] V.I. Abkevich, A.M. Gutin, E.I. Shakhnovich, J. Chem. Phys. 101 (1994) 6052.

[8] E. Fadrna, J. Koca, J. Phys. Chem. B 101 (1997) 7863.

[9] I. Yarovsky, M.T.W. Hearn, M.I. Aguilar, J. Phys. Chem. B 101 (1997) 10962.

[10] R.E. Kunz, P. Blaudeck, K.H. Hoffman, R.S. Berry, J. Chem. Phys. 108 (1998) 2576.

[11] V.K.W. Cheng, J. Chem. Phys. 108 (1998) 8456.

[12] Y. Wong, B. Schubert, R. Hoffman, J. Am. Chem. Soc. 114 (1992) 2367.

[13] R.O. Jones, J. Chem. Phys. 99 (1993) 1194.

[14] R.O. Jones, G. Gantefoer, S. Hunsicker, P. Pieperhoff, J. Chem. Phys. 103 (1995) 9549.

[15] S. Hunsicker, R.O. Jones, J. Chem. Phys. 105 (1996) 5048.

[16] D. Wei, D.R. Salahub, J. Chem. Phys. 106 (1997) 6086.

[17] R.L. Asher, D.A. Micha, P.J. Brucat, J. Chem. Phys. 96 (1992) 7683.

[18] C.H. Xu, C.Z. Wang, C.T. Chan, K.M. Ho, Phys. Rev. B 47 (1993) 9878.

[19] G. Toth, J. Phys. Chem. 101 (1997) 8871.

[20] K. Ng, D. Vanderbilt, Phys. Rev. B 59 (1999) 10132.

[21] R.O. Jones, D. Hohl, J. Am. Chem. Soc. 112 (1990) 2590.

[22] U. Röthlisberger, M.L. Klein, J. Am. Chem. Soc. 117 (1995) 42.

[23] S. Miura, M.E. Tuckerman, M.L. Klein, J. Chem. Phys. 109 (1998) 5290.

[24] F. Bockish, D. Liotard, J.-C. Rayez, B. Duguay, Int. J. Quantum Chem. 44 (1992) 619.

[25] J.-M. Pons, A. Pommier, M. Rajzmann, D. Liotard, J. Mol. Struct. (Theochem) 313 (1994) 361.

[26] J.-M. Pons, M. Oblin, A. Pommier, M. Rajzmann, D. Liotard, J. Am. Chem. Soc. 119 (1997) 3333.

[27] B. Lecea, A. Arieta, G. Roa, J.M. Ugalde, F.P. Cossio, J. Am. Chem. Soc. 116 (1994) 9613.

[28] B. Lecea, A. Arrieta, X. Lopez, J.M. Ugalde, F.P. Cossio, J. Am. Chem. Soc. 117 (1995) 12314.

[29] A. Moyano, M.A. Pericas, E. Valenti, J. Org. Chem. 54 (1989) 573.

[30] R. Ocampo, W.R. Dolbier, M.D. Bartberger, R. Paredes, J. Org. Chem. 62 (1997) 109.

[31] I. Morao, B. Lecea, A. Arrieta, F.P. Cossio, J. Am. Chem. Soc. 119 (1997) 816.

[32] W.S. Chin, B.W. Ek, C.Y. Mok, H.H. Huang, J. Chem. Soc., Perkin Trans. 1 (1994) 883.

[33] AMPAC 6.55, Semichem, PO Box 1649, Shawnee Mission, KS 66222, 1998.

[34] M.J.S. Dewar, E.V. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902.

[35] M.J.S. Dewar, C. Jie, G. Yu, Tetrahedron 49 (1993) 5003.

[36] A.J. Holder, R.D. Dennington, C. Jie, Tetrahedron 50 (1994) 627.

[37] A.J. Holder, E.M. Evleth, in: D.A. Smith (Ed.), Modeling the Hydrogen Bond, American Chemical Society, Washington, DC, 1994, p. 113.

[38] A.J. Holder, C. Jie, D. Liotard, in preparation.

[39] T. Clark, J. Mol. Struct. (Theochem) 530 (2000) 1.

[40] A.J. Holder, R.D. Dennington II, J. Mol. Struct. (Theochem) 401 (1997) 207.

[41] D. Liotard, A.J. Holder, J. Chem. Inf. Comput. Sci. 39 (1999) 587.

[42] D. Liotard, Int. J. Quantum Chem. 44 (1992) 723.

[43] AMPAC 6.0, Semichem, 7204 Mullen, Shawnee, KS 66126, 1997.

[44] D. Liotard, A.J. Holder, to be published.

[45] AMPAC 5.0, Semichem, 7204 Mullen, Shawnee, KS 66126, 1994.

[46] P.J. Kocienski, B. Pelotier, J.-M. Pons, H. Prideaux, J. Chem. Soc., Perkin Trans. 1 (1998) 1373.

[47] H.M. Frey, I.M. Pidgeon, J. Chem. Soc., Faraday Trans. 1 81 (1985) 1087.

[48] H.B. Schlegel, J. Comput. Chem. 3 (1982) 214.

[49] J. Simons, P. Jorgensen, H. Taylor, J. Ozment, J. Phys. Chem. 87 (1983) 2745.

[50] J. Baker, J. Comput. Chem. 7 (1986) 385.

[51] P. Culot, G. Dive, V.H. Nguyen, J.M. Chuysen, Theor. Chim. Acta 82 (1991) 189.

[52] F. Jensen, J. Chem. Phys. 102 (1995) 6706.

[53] L.L. Stacho, M.I. Ban, J. Math. Chem. 11 (1992) 405.

[54] L.L. Stacho, M.I. Ban, Theor. Chim. Acta 83 (1992) 433.

[55] L.L. Stacho, M.I. Ban, Theor. Chim. Acta 83 (1993) 535.

[56] D. Liotard, et al., to be published. The method is implemented in AMPAC 70, 2001.