Relative stabilities of cholestadienes calculated by molecular mechanics and semi-empirical methods: application to the acid-catalyzed rearrangement reactions of cholesta-3,5-diene

Patrick Y. Schüpfner, Fazil O. Gülaçar *

Laboratoire de Spectrométrie de Masse, Université de Genève, 16 Boulevard d’Yvoy, 1211 Geneva 4, Switzerland

Abstract

The study of geochemical transformations undergone by ‘biological markers’ after their incorporation into sediments is an important field of organic geochemistry. Combined with laboratory simulation experiments, molecular mechanics calculations have been shown to be very useful to establish the reaction pathways, and to predict intermediate components and stable reaction end products, especially in the case of the acid-catalyzed isomerization reactions of steroid and terpenoid hydrocarbons. Many commercially available softwares are able to optimize (minimize) the geometries of molecules and compute some of their thermodynamical data with either molecular mechanics (MM) or semi-empirical methods of quantum chemistry. In order to verify the reliability of these methods, we have computed the relative thermodynamic stabilities of a large number of steradiene isomers with MM3 (Tripos Inc.), MM+ (HYPERCHEM™) and MM2 (Chem3D, CambridgeSoft Corp.) empirical force fields, and with AM1 and PM3 (HYPERCHEM™) semi-empirical methods. The calculation results of thermodynamic stabilities of steradiene isomers are used to explain the compounds produced by the rearrangement of cholesta-3,5-diene when treated with p-toluenesulfonic acid in acetic acid at 70°C. The end products, namely the spirosteradienes 7–8, obtained by this treatment are the most stable steradiene isomers according to all computational methods. The relative thermodynamic stabilities of cholestadienes are also consistent with the mechanism postulated for the spirosteradiene formation proceeding through a pathway including cholestadienes 2–6 as intermediates. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Cholestadiene isomerization; Acid catalyzed rearrangement; Thermodynamic stabilities; Molecular mechanics calculations; Semi-empirical calculations

1. Introduction

Sterols are well-known ‘biological markers’ due to their skeleton preservation in sedimentary environments for a long geological record and to their structural specificity or distribution reflecting their sources. Early diagogenesis of sterols has been the subject of several studies and it has been shown that sterenes and steradienes are widespread products generated in the water column and in shallow sediments (MacKenzie et al., 1982). During the burial process, they undergo further geological transformations, depending on their structure and on the nature of the sediment matrix, to thermodynamically more stable products such as steranes, diasteranes and/or aromatic steroids. The formation of these compounds and their identification in ancient sediments may be well assisted by laboratory experiments that simulate the diagensis of supposed precursor biomarkers. Thermodynamic data on the relative stabilities of the compounds involved may also be useful to predict or to interpret the transformation pathways during maturation. As experimentally determined values are often unavailable in literature, the thermodynamic properties are estimated by computations with molecular mechanics or semi-empirical methods. Thus, the use of MM2 force field (Allinger, 1977) has been of great help in the past to improve the understanding of
the results of laboratory simulation experiments. For example, the distribution at thermodynamic equilibrium of sterane isomers at C-5, C-14 and C-17 (van Graas et al., 1982) and of a range of cholestenes (de Leeuw et al., 1989) have been estimated by this approach and the results obtained were in agreement with the sedimentary distributions. The same method was also applied to reinvestigate the mixture at hypothetical equilibrium obtained after mild acid treatment of 4-methylcholest-4-ene (Rechka et al., 1992) and to calculate the thermodynamic stabilities for a range of possible rearrangement products of Δ4- and Δ5-sterenes including diasterenes and rearranged spirosterenes (de Leeuw et al., 1993). Based on these investigations an explanation of the sterene/sterane distribution patterns observed in Recent sediments has been proposed and it has been concluded that double-bond migration in steroidal monoenes occurs only via tertiary carbocations. Recently, a force field for tertiary carbocations has been developed in order to predict the isomerization rates of sterenes (van Duin et al., 1996).

With the increasing power of microcomputers in recent years, semi-empirical methods like MNDO, AM1, PM3, etc., also became accessible for use without the need of mini or mainframe computers. Thus, geometry optimizations and thermodynamical properties can now be computed by these methods on a desktop in less than 2 h for molecules with ca. 30 carbon atoms.

We report here a comparison between thermodynamic stabilities of a series of cholestadienes calculated by molecular mechanics and semi-empirical methods, and the rearrangement reactions of cholest-3,5-diene by acid-catalysis. A brief description of these two categories of computational methods is also presented.

1.1. Molecular mechanics — empirical force fields

In molecular mechanics (MM), a molecule is described as a collection of atoms regarded as spherical balls that are bonded together by elastic or harmonic forces and that interact through space by non-bonded forces, such as van der Waals attraction, steric repulsion, or electrostatic interactions (Allinger, 1976; Burkert and Allinger, 1982). These forces are represented by classical potential energy functions describing the structural characteristics of the compound (bond distances and angles, torsion angles, non-bonding interactions, etc.). The parameters of these potential energy functions must be optimized to obtain the best fit between calculated and experimental properties of the molecule. The collection of both potential energy functions and parameters, permitting the calculations of the absolute energy as a function of the molecular geometry, is called a force field. This absolute or “steric” energy (SE) is obtained from the sum of bonding energies like stretching, bending, or torsion energy, and of non-bonding energies like van der Waals, steric and electrostatic energies. Hence, this energy has no intrinsic physical meaning; it is only a measure of the intramolecular deformation from a hypothetical situation in which bonding interactions are at minimum values. Nevertheless, the SE values are useful for comparisons of stabilities for different conformations of a molecule since the number of chemical bonds of each type is conserved. The enthalpy of formation values are of course more reliable for comparing energies between molecules differing in their bonding arrangements or in molecular size.

The enthalpies of formation, Δ\(H_{\text{form}}^{\circ}\), reported in Table 1 have been calculated for the molecules in the gas phase at 298 K. In molecular mechanics the enthalpy of formation is build up from several contributions. In addition to the SE which reflects the strain effects, the energies of bond formation (BE) must evidently be considered in Δ\(H_{\text{form}}^{\circ}\) calculation, and can simply be obtained by the sum of appropriate bond or group increments. Nevertheless, this contribution is not sufficient, and vibrational, rotational and translational energy terms must be considered as well. The vibrational contribution at 298 K can be directly included in BE terms, although the two latter contributions take part in the Δ\(H_{\text{form}}^{\circ}\) calculation with a contribution of \(RT/2\) (where \(R = 1.987\) cal deg\(^{-1}\) mol\(^{-1}\) is the ideal gas constant and \(T\) the Kelvin temperature) for each rotational and translational degree of freedom in the molecule, and an additional \(RT\) to convert energy to enthalpy (i.e. \(6/2\ RT + RT = 2.4\) kcal/mol for a nonlinear molecule), leading to the following equation:

\[
\Delta H_{\text{form}}^{\circ} = \text{SE} + \text{BE} + 4RT
\]

However, this latter equation can be applied only for structures with localized bonds. In delocalized electronic systems, a variation of bond lengths can be observed (e.g. in naphthalene), and these differences in bond lengths and force constants are well determined by the bond orders.

In MM3, a calculation on π-MO (molecular orbitals) is performed with an iterative, planar VESCF (variable electronegativity self-consistent field) method to determine the bond orders from which necessary force constants are assigned and then used in the molecular mechanics force field. The enthalpy of formation is finally obtained by the equation \(\Delta H_{\text{form}}^{\circ} = \text{EC} + \text{BE} + \text{ESCF} + 4RT\), where EC is the corrected steric energy and ESCF is the energy from VESCF calculation. Energy differences between molecules differing only in the configuration of one or more asymmetric centers are the same for the enthalpy of formation (\(\Delta \Delta H_{\text{form}}^{\circ}\)) and for the steric energy. ESCF and BE terms do not change for different epimers as they possess the same bonds and group increments for the calculations.
1.2. Semi-empirical methods of quantum chemistry

In order to avoid solving the complete Schrödinger equation, the Born–Oppenheimer approximation allows to simplify the Schrödinger equation for a molecule by separating the treatment of the nuclei and the electrons, the nuclei being much heavier than the electrons and their motion being relatively slow. In molecular mechanics (MM), the Born–Oppenheimer approximation is also used since only the motions of the nuclei are considered. The electrons, which are simply supposed to be in an optimum distribution about the nuclei, are not studied explicitly. In semi empirical methods, on the other hand, it is the electronic structure of the molecule that is examined using fixed nuclear positions. But, even after Born–Oppenheimer approximation, the electronic Schrödinger equation for the electron distribution remains very complex to solve. Nevertheless, diverse approximations can be made. First it may be assumed that each electron is affected by an effective average field due to all other electrons and nuclei, simplifying the electronic Hamiltonian by a sum of one-electron Hamiltonians. The self-consistent field (SCF) or Hartree–Fock approximation uses a modified one-electron Hamiltonian, called Fock operator, which contains terms accounting for electronic interactions (Coulomb and exchange integrals). The electronic wave function may also be approximated by describing molecular orbitals as linear combination of atomic orbitals (LCAO).

The direct calculations of SCF approximation are referred to as ab initio or Gaussian methods and can accurately calculate the electronic properties of small molecules. In semi-empirical methods, the complex resolution of electron–electron and electron–nuclear interactions are approximated by the use of empirical formulae and parameters in order to reduce the number of integrals to calculate. The two methods employed in this paper, Austin Model 1 (AM1) and Parameterized Model 3 (PM3) (Stewart, 1990), are based on Neglect of Diatomic Differential Overlap (NDDO) approximation. Only valence atomic orbitals are considered in the calculations and the core electrons are accounted for the calculations in the effective nuclear charge, and several differential overlap terms are neglected. Thus, the more complex portions of the calculation are replaced in the semi-empirical methods with parameters derived from experimental data and the electronic properties of larger molecules can be calculated. The enthalpy of formation is finally obtained by the equation

$$\Delta H_{\text{form}}^A = E_{\text{Elect}} + E_{\text{nuc}} - \Sigma E_{\text{el}} + \Sigma \Delta H_{\text{form}}^A \quad \text{(Stewart, 1989)}$$

where $E_{\text{Elect}}$ is the total electronic energy, $E_{\text{nuc}}$ is the nuclear-nuclear repulsion energy, $E_{\text{el}}$ is the calculated energy of formation of the gaseous atom A from its ion and $\Delta H_{\text{form}}^A$ is the experimental enthalpy of formation for atom A.

1.3. Semi-empirical methods vs. molecular mechanics

For MM methods, the parameters used in the potential energy functions have been estimated for all possible bonds between atom types (e.g. sp-, sp$^2$-, sp$^3$-hybridized carbon, hydrogen...). Additional parameters have been included to improve the accuracy of the results for certain functional groups (e.g. in MM3 force field cyclopentane, cyclobutane, and cyclopropane rings have their own parameters, while higher cycloalkanes are treated as deformed acyclic alkanes). All these parameters are determined by experimental data or by ab initio calculations. The parameter set in molecular mechanics can contain a large number of values unlike in semi-empirical methods which use only single-atom parameters to compensate the approximations made to solve the Schrödinger equation.

Comparisons with experiment of MM3, AM1 and PM3 enthalpies of formation show clearly that the best results are obtained using MM3, then PM3 which is an optimization of AM1 parameters. However, the absolute values of enthalpy of formation should be considered with caution because of the heavy dependence on the experimentally determined parameters. Furthermore, calculated values can only properly be gauged when compared with experimental results unrelated to those used in determining the parameters, so calculations on new molecules should be considered with caution (especially those obtained by molecular mechanics). In the present paper, we are dealing with cholestanediene hydrocarbons which contain atoms and bonds that are fully covered by the parameters of all the computational methods used.

1.4. Acid-catalyzed isomerization of cholesta-3,5-diene

The clay minerals in a sediment matrix, like kaolinite or montmorillonite, show acidic properties which can catalyze the rearrangement reactions of organic compounds (Sieskind et al., 1979). The medium we used to simulate the diagenesis of cholesta-3,5-diene is an anhydrous mixture of p-toluene sulfonic acid (TsOH) in acetic acid (AcOH), a mixture widely used for the acid-catalyzed rearrangements of steroid hydrocarbons (Blunt et al., 1969; Kirk and Shaw, 1970, 1975; Wolff et al., 1986; Peakman and Maxwell, 1988a–c; Peakman et al., 1988, 1992; de Leeuw et al., 1989; Liu et al., 1996). The acetic acid is a poorly solvating solvent with a low dielectric constant (6.62 at 70°C) and, in anhydrous acetic acid, even strong electrolytes have small dissociation constants ($<10^{-5}$) (Kolthoff and Brown, 1956). Furthermore, it appears that in anhydrous TsOH/AcOH monounsaturated compounds can rearrange only via tertiary carbocations which seems to be also the case in sediments (de Leeuw et al., 1989, 1993). Thus, the anhydrous TsOH/AcOH mixture simulates quite well...
the mild acid conditions to which the sedimentary organic matter is subjected.

2. Results and discussion

2.1. Acid-catalyzed backbone rearrangement of cholesta-3,5-diene

The distribution of the rearrangement products of cholesta-3,5-diene 1 in the presence of p-TsOH in AcOH at 70 °C is monitored by GC and GC/MS (Fig. 1). While the starting material is still present, the first rearranged product 7a appears after 1 h and is identified as (20R)-5x-12(13→14)-abeo-cholesta-8,13(17)-diene by co-injection with the fully characterized product isolated during similar experiments on 5a(H)-cholesta-8,14-diene (Liu et al., 1996). With further reaction time, three other isomers of 7a appear and after 5 days, GC/MS analyses show that these four major products have virtually identical mass spectra, and correspond to (20R)-, (20S)-5β and (20R)-, (20S)-5x epimers of 12(13→14)-abeo-cholesta-8,13(17)-diene, respectively (Liu et al., 1996). The chromatograms taken after 8 and 12 days (not included in Fig. 1) show that the epimerization at C-20 reaches a constant ratio of (57/43) between 20R (7a,b) and 20S (8a,b). On prolonged reaction time, 5x-then 5β-spirosteradienes disappear sequentially and the monoaromatic compounds 9–11 become predominant products. These compounds have been conclusively identified (Fig. 2), but they are not the subject of this paper.

We can reasonably postulate that the isomerization of cholesta-3,5-diene 1 to spiro compounds 7–8 proceeds through a pathway including cholestadienes 2–6 as intermediates and via allylic cations by reversible protonation-deprotonation reactions, even though no intermediate showed up experimentally, except for cholesta-4,6-diene 2 that coelutes with cholesta-3,5-diene 1. To use thermodynamic stabilities for the determination of product distribution, we must make sure that we are dealing with an equilibrium mixture in which all components can revert to each other by reversible reactions. It is obvious that cholestadienes 1–8 do not reach a state of equilibrium and are only intermediates finally yielding to monoaromatic steroids 9–11. However, we can assume that the reversible isomerization is rapid compared to the formation of monoaromatic compounds arising from irreversible dehydrogenation and rearrangement processes. Indeed, chromatograms taken after 8 and 12 days (not included in Fig. 1) show that the epimerization of spirodienes at C-20 reaches a constant ratio (57/43) between 20R (7a,b) and 20S (8a,b). Thus, we expect the relative thermodynamic stabilities of cholestadienes involved to be in accord with the observed evolution.

Monoaromatic steroids 9–11 are not included in the calculations, because no equilibrium exists between them and the cholestadienes. Furthermore their relative amount seems to depend on kinetic rather than on thermodynamic considerations. More studies are clearly required to fully understand their formation.

2.2. Energy calculation results

Geometry optimizations of all possible cholestadiene isomers which can be produced by double bond migration via allylic and/or tertiary carbocations have been done using MM3 (Tripos Inc.), MM+ (HYPERCHEM™) and MM2 (Chem3D, CambridgeSoft Corp.) empirical force fields, as well as with AM1 and PM3 (HYPERCHEM™) semi-empirical methods. All optimizations have been done with the conformation of the molecules corresponding to the global minimum. Enthalpies of formation have been computed with MM3, AM1 and PM3 following geometry optimization using the same method. All energy values in the Table 1 are valid for molecules in the gas phase at 298 K, nevertheless they can be assumed to reflect the stabilities of these molecules in the sediment. Table 1 shows only the results obtained for the conjugated dienes which are the most stable ones and which can be produced by intermediacy of allylic carbocations. The geometry corresponding to the global minimum energy presents for all cholestadienes a side chain adopting an all-anti configuration. The methyl groups C-27 and C-20 point in the same direction as was observed with cholestane isomers (van Graas et al., 1982).

For an easier comparison of the values obtained with different methods, the energy differences are shown relative to the most stable isomer. Examination of the Table 1 shows that although there is a coarse parallelism in the order of stabilities, there are also several cases with significant differences. As a general rule, the homocyclic dienes are less stable than the heteroannular dienes. However, discrepancies with this rule are observed for 5β-cholesta-6,8-diene which has a higher stability than the heteroannular diene 1 by both AM1 and PM3 while it is more stable than 1 and 2 by MM3.

The enthalpies of formation calculated by MM3, AM1 and PM3 show that spirostadienes 7 and 8 are the most stable, in accord with the experimental results. However, the order of relative stabilities of spiro compounds 7, 8 (a,b) changes following the computational method (only MM+ and MM2 find the same sequence 7b < 8b < 7a < 8a). Considering the stereochemistry at C-20 for spirostadienes, we can observe that enthalpies of formation and steric energies calculated with MM3 and AM1 are lower for the (20S)-isomers, while the (20R)-isomers are more stable according to PM3, MM+ and MM2 methods. The latter results seem to be more consistent with the experiment since 20S/20R...
Fig. 1. Gas chromatograms (SE-54) showing the evolution of the products distribution on acid treatment (p-TsOH/AcOH) of cholesta-3,5-diene (1) (see Fig. 2 and Table 1 for the numbered compounds).
ratios for the two epimeric pairs increased to a maximum value of only 43/57 as already mentioned. However, whatever the computational method used, the (20R)- and (20S)-isomers of the spirosteradienes are very close in relative stability and at thermodynamic equilibrium the ratio 20R/20S should be close to 1.

All MM methods give the 5β(H)-epimers more stables than 5α(H)-epimers except for cholesta-7,14-diene. AM1 method produces rather the reverse results, whereas no general feature is observed with PM3. Based on MM methods, at equilibrium we would expect a predominance of spirosteradienes 7b + 8b over 7a + 8a.

Fig. 2. Compounds involved in the acid-catalyzed rearrangement of cholesta-3,5-diene (1) (see Table 1 for compound names). Products 3–6 are not experimentally observed but they are obligatory intermediates from 1 to 7 and 8.

Table 1

<table>
<thead>
<tr>
<th>Name</th>
<th>$\Delta H_f$ (298 K) (kcal/mol)</th>
<th>$\Delta$ Steric energy (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MM3</td>
<td>AM1</td>
</tr>
<tr>
<td>(20S)-5β(H)-spirocholesta-8,13(17)-diene$^{\text{a}}$ (8b)</td>
<td>0.00$^b$</td>
<td>0.16</td>
</tr>
<tr>
<td>(20R)-5β(H)-spirocholesta-8,13(17)-diene (7b)</td>
<td>0.02</td>
<td>0.21</td>
</tr>
<tr>
<td>(20S)-5α(H)-spirocholesta-8,13(17)-diene$^{\text{a}}$ (8a)</td>
<td>0.94</td>
<td>0.00$^c$</td>
</tr>
<tr>
<td>(20R)-5α(H)-spirocholesta-8,13(17)-diene$^{\text{a}}$ (7a)</td>
<td>0.96</td>
<td>0.03</td>
</tr>
<tr>
<td>(20R)-5β(H)-cholesta-8,14-diene (6b)</td>
<td>4.62</td>
<td>1.90</td>
</tr>
<tr>
<td>(20R)-5α(H)-cholesta-8,14-diene (6a)</td>
<td>5.32</td>
<td>1.23</td>
</tr>
<tr>
<td>(20R)-5β(H)-cholesta-6,8(14)-diene (4b)</td>
<td>5.93</td>
<td>0.86</td>
</tr>
<tr>
<td>(20R)-5α(H)-cholesta-6,8(14)-diene (4a)</td>
<td>7.26</td>
<td>0.94</td>
</tr>
<tr>
<td>(20R)-5β(H)-cholesta-7,14-diene (5a)</td>
<td>8.59</td>
<td>2.83</td>
</tr>
<tr>
<td>(20R)-5α(H)-cholesta-7,9(11)-diene</td>
<td>9.02</td>
<td>3.26</td>
</tr>
<tr>
<td>(20R)-5β(H)-cholesta-7,14-diene (5b)</td>
<td>9.63</td>
<td>3.74</td>
</tr>
<tr>
<td>(20R)-5α(H)-cholesta-6,8-diene</td>
<td>10.99</td>
<td>5.83</td>
</tr>
<tr>
<td>(20R)-cholesta-3,5-diene (1)</td>
<td>11.51</td>
<td>5.45</td>
</tr>
<tr>
<td>(20R)-cholesta-4,6-diene (2)</td>
<td>11.76</td>
<td>5.92</td>
</tr>
<tr>
<td>(20R)-5α(H)-cholesta-6,8-diene</td>
<td>13.80</td>
<td>5.76</td>
</tr>
<tr>
<td>(20R)-cholesta-2,4-diene</td>
<td>13.83</td>
<td>5.95</td>
</tr>
<tr>
<td>(20R)-cholesta-5,7-diene (3)</td>
<td>14.04</td>
<td>5.99</td>
</tr>
<tr>
<td>(20R)-5α(H)-cholesta-1,3-diene</td>
<td>17.27</td>
<td>11.06</td>
</tr>
</tbody>
</table>

$^a$ Spirocholesta-8,13(17)-diene corresponds to 12(1314)-abeo-cholesta-8,13(17)-diene.
$^b$ Enthalpy of formation = –70.31 (kcal/mol).
$^c$ Enthalpy of formation = –70.56 (kcal/mol).
$^d$ Enthalpy of formation = –72.98 (kcal/mol).
$^e$ Steric energy = 50.458 (kcal/mol).
$^f$ Steric energy = 38.469 (kcal/mol).
$^g$ Steric energy = 37.137 (kcal/mol).
while AM1 and PM3 both predict the contrary. Our experimental results cannot allow settling this point since it is obvious that the equilibrium state between the 5α- and 5β-spirosteradienes have never been attained during the experiment. 5β-spirosteradienes begin to predominate after ca. 10 days of rearrangement. However, we believe that this is due to a more rapid disappearance of 5α-spirosteradienes to generate mono-aromatic compounds.

3. Conclusion

Computed enthalpies of formation are consistent with the experimental results since all methods lead to spirosteradienes as the most stable products and the order of stabilities are almost the same. The enthalpies of formation computed by MM3 seem to correlate well with those obtained by PM3. However, the absolute values of the differences between the enthalpies of formation computed by different methods may be as important as several kcal/mol. MM3 calculation results appear to be the most consistent with observed general rules such as the highest stabilities of heteroannular dienes and of 5β(H)-spiroidienes. However, the lack of relevant geochemical data about the steradienes is a restrictive factor for this study.

Although the results reported here do not allow to compare the reliability of the different methods, it is clear that their use for the determination of the thermodynamic stabilities in the field of polyenic hydrocarbons should be considered with caution.

Acknowledgements

We thank A. van Duin and G. van Graas for their helpful comments and suggestions. Financial support by the Swiss National Foundation (Grant No. 20-52429.97) is gratefully acknowledged.

References


