Calculated heats of formation of sterol diene isomers compared with synthetic yields of isomerisation reactions of $\Delta^{5,7}$ sterols

D. Roeland Boer,[*a] **Huub Kooijman,**[a] **Jaap van der Louw,**[c] **Marinus Groen,**[c] **Jan Kelder**[b] **and Jan Kroon**[a]

[a] *Department of Crystal & Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, The Netherlands*
[b] *Department of Molecular Design and Informatics, N.V. Organon, Oss, The Netherlands*
[c] *Department of Medicinal Chemistry, N.V. Organon, Oss, The Netherlands*

Heats of formation of five series of diene sterol isomers were calculated and compared with synthetic yields of acid-catalysed isomerisation reactions starting from $\Delta^{5,7}$ isomers. Calculations were based on molecular mechanics, using the MM3 program package. For each of the five $\Delta^{5,7}$ starting compounds, three possible reaction paths were considered, in which heats of formation were calculated for theoretically possible intermediate double bond isomers. Similar results are found for all five series. The starting compounds are found to have the unfavourable heats of formation compared to all other isomers considered within one series. In general, isomerisation reactions of diene sterols ultimately yield spiro compounds when allowed to proceed for a sufficient amount of time. These compounds are found to have the lowest heats of formation in each series. However, they were not formed in the reactions considered in this paper, because the reactions were stopped after the desired isomer was formed in excess, before the spiro compounds could occur. Most compounds identified as products in the syntheses have a favourable heat of formation compared to the isomers preceding the (most stable) spiro compounds. However, when the reactions were carried out at low temperature, isomers with less favourable heats of formation could be trapped. We show that calculated heats of formation correspond well with synthetic yields and we suggest they can be a useful tool in planning syntheses.

Introduction

Acid-catalysed isomerisation of $\Delta^{5,7}$ sterols to, among others, $\Delta^{8,14}$ or $\Delta^{7,14}$ compounds, occurs through a pathway involving a sequence of protonation and deprotonation steps. During the reaction, other isomers are formed as well and the crude product usually is a mixture, which is undesirable. It would be convenient to be able to predict which isomers will most likely be formed and to control the relative yields of the isomers by changing reaction conditions using knowledge about the stability of the different isomers. In this paper, yields from syntheses of sterol double bond isomers, taken from the literature and obtained by us, are compared with calculated heats of formation of a range of possible isomers. Heats of formation were calculated with MM3[1,2] using the standard force field parameters. The force field is well parametrised for hydrocarbon compounds and has been shown to calculate heats of formation for a number of hydrocarbons with a standard deviation of 0.41 kcal mol^{-1} compared to experimental values.[2]

Since the reactions are reversible, the isomers that are formed during the reaction are assumed to be in equilibrium. As a consequence, the heat of formation differences between double bond isomers should be indicative of the isomer stability, in the same way as has been reported for global steric-energy minima of steroids and their occurrence in geological sediments as molecular fossils.[3] In principle, it is possible to include kinetic effects by calculating formation enthalpies for carbocations.[4] These are not considered in this paper.

Five compound series are considered here. Scheme 1 is a key to the structural formulas of compounds of those series (**1**, **2**, **3**, **4** and **5**). The first series (**1**) consists of structures derived from 3-benzoyl-$\Delta^{5,7}$-cholesterol. The second series is analogous to series **1** with two extra methyl groups attached to C-4 of ring A. Series **3** consists of 3-hydroxy analogs of series **2**. The cholestane side chain attached to C-17 of the compounds in series **1** and **2** is replaced by an ergostane side chain to give the com-

pounds for series **4** and **5** respectively. Synthetic data on a subset of structures in series **1**,[5,6] **4**[5,7] and **5**[7] were reported by other authors. The syntheses of compounds **2f**, **3f** and **2h**, **3h** (the 5α-$\Delta^{6,8(14)}$ and 5α-$\Delta^{8,14}$ isomers respectively) are reported here. The latter are of significant biological and pharmaceutical interest, as they are related to FF-MAS, a signalling molecule which induces resumption of meiosis of the human oocyte (immature egg cell).[8]

Three reaction paths were considered for each series (**A**, **B** and **C** in Scheme 1). Path **A** leads to isomers containing (conjugated) double bonds in rings B and C, path **B** to 5α isomers with conjugated double bonds in rings B, C and D and path **C** to the corresponding 5β isomers. Depending on the particular reaction conditions within each series, isomers **f**, **g**, **h**, and **p** were found as end products. In general, isomerisation reactions yield spiro compounds like **i** and **q** when allowed to progress for a considerable amount of time.[3]

Experimental

Syntheses

The synthetic route used to prepare the 5α-$\Delta^{6,8(14)}$ and 5α-$\Delta^{8,14}$ isomers of the 3-OH and benzoylated sterol cholestane series (**2f,h** and **3f,h**) is shown in Scheme 2.

(3β,5α,20R)-4,4-Dimethylcholesta-8,14-dien-3-ol (3h) and (3β,5α,20R)-4,4-dimethylcholesta-6,8(14)-dien-3-ol (3f). A mixture of (3β,20R)-4,4-dimethylcholesta-5,7-dien-3-ol (**3a**, 2.32 g, 5.63 mmol), ethanol (96%, 42 ml), toluene (6 ml), and concentrated hydrochloric acid (6 ml) was heated under reflux for 2 h. After cooling, the mixture was poured into a saturated aqueous solution of NaHCO$_3$. The product was extracted into diethyl ether; the combined organic phases were washed with brine, dried over Na$_2$SO$_4$ and concentrated under reduced pressure. The crude product (3.69 g) was purified by column chromato-
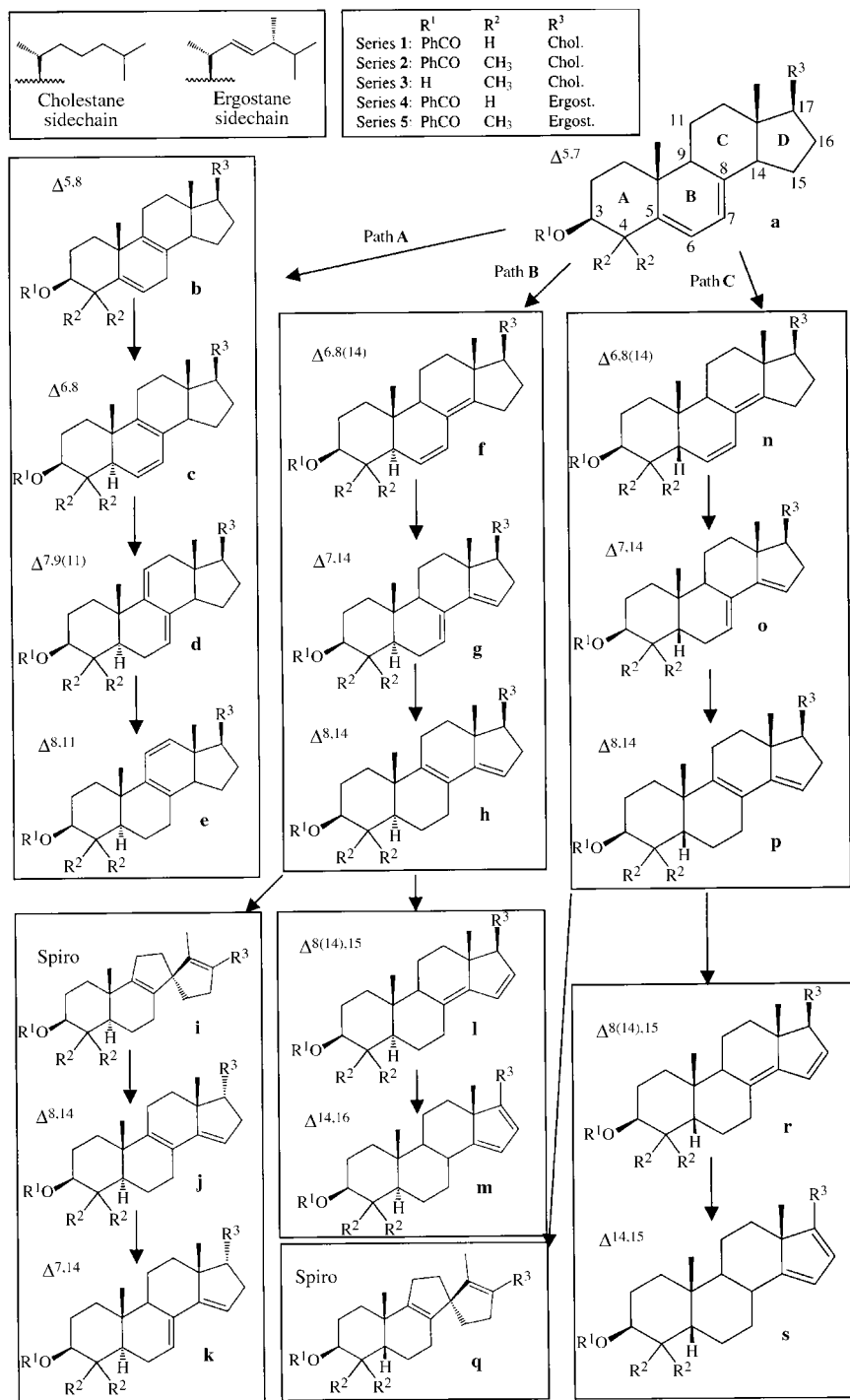
Scheme 1 Molecular structures of all isomers in the five series. Three possible reaction paths (**A**, **B** and **C**) are included.

	R¹	R²	R³
Series 1:	PhCO	H	Chol.
Series 2:	PhCO	CH₃	Chol.
Series 3:	H	CH₃	Chol.
Series 4:	PhCO	H	Ergost.
Series 5:	PhCO	CH₃	Ergost.

graphy to yield a mixture of **3h** and **3f** (2.83 g, ratio 6:1). **3h**: ¹H NMR (CDCl₃): δ 5.35 (m, 1H), 3.24 (m, 1H), 2.43–0.80 (m), 1.04 (s, 3H), 1.02 (s, 3H), 0.93 (d, J = 6.2 Hz, 3H), 0.88 (s, 3H), 0.85 (s, 3H), 0.83 (s, 3H), 0.81 (s, 3H); **3f**: ¹H NMR (CDCl₃): δ 6.23 (dd, J = 10.0 and 3.0 Hz, 1H), 5.61 (dm, J = 10.0 Hz, 1H), 3.30 (m, 1H), 2.34 (m, 2H), 2.00 (m, 1H), 1.90 (m, 2H), 1.79–0.97 (m), 1.08 (s, 3H), 0.94 (d, J = 8.6 Hz, 3H), 0.89 (s, 3H), 0.88 (s, 3H), 0.86 (s, 3H), 0.83 (s, 3H), 0.68 (s, 3H).

(3β,5α,20R)-4,4-Dimethylcholesta-8,14-dien-3-ol benzoate (2h) and (3β,5α,20R)-4,4-dimethylcholesta-6,8(14)-dien-3-ol benzoate (2f).

A mixture of (3β,20R)-4,4-dimethylcholesta-5,7-dien-3-ol benzoate (**2a**, 2.58 g, 5.0 mmol), ethanol (96%, 70 ml), toluene (20 ml), and concentrated hydrochloric acid (10 ml) was heated under reflux for 1 h. The reaction mixture was worked up as described above to produce a mixture of **2h** and **2f** (2.61 g, ratio 1:5). Prolonged reaction results in a shift in ratio between **2h** and **2f** via 3:2 (5 h heating) to 5:1 (21 h heating). **2h**: ¹H NMR (CDCl₃): δ 8.06 (m, 2H), 7.61–7.40 (m, 3H), 5.38 (m, 1H), 4.75 (m, 1H), 2.48–0.75 (m); **2f**: ¹H NMR (CDCl₃): δ 8.06 (m, 2H), 7.61–7.40 (m, 3H), 6.27 (dd, J = 10.4 and 3.0 Hz, 1H), 5.62 (dd, J = 10.4 and 2.0 Hz, 1H), 4.82 (dd, J = 10.8 and 5.0 Hz, 1H), 2.48–0.75 (m).

Heat of formation calculations

MM3 calculates heats of formation based on a bond-energy increment method. For each bond appearing in the structure, a fixed contribution is added by the program, which depends on the atom type and substitution number of the atoms that form the bond. However, the formation enthalpy also depends on

Table 1 Heats of formation (ΔH_f in kcal mol^{-1}) and relative yields (in % where possible) of the compounds (denoted by letter according to Scheme 1) of the five isomer series

	Series 1			Series 2		Series 3		Series 4			Series 5	
	ΔH_f	Yield[a]	Yield[b]	ΔH_f	Yield[c]	ΔH_f	Yield	ΔH_f	Yield[a]	Yield[d]	ΔH_f	Yield[d]
a	−113.0			−120.3		−101.7		−92.9			−100.8	
Path A												
b	−115.7			−122.4		−103.6		−96.0			−103.1	
c	−113.1			−120.3		−101.3		−93.6			−101.0	
d	−118.2			−124.2		−105.1		−98.6			−104.5	
e	−113.9			−119.8		−100.7		−94.1			−100.1	
Path B												
f	−119.7		3	−127.5	20	−107.6	15	−100.2			−106.6	
g	−119.2	Trace	48	−124.3		−105.7		−99.5	4	75	−104.7	75
h	−121.8	Major	27	−127.6	80	−108.7	85	−102.4	72	25	−107.5	25
i	−128.5			−135.2		−116.3		−108.0			−114.1	
j	−114.7			−120.9		−103.1		−96.6			−102.3	
k	−117.8			−123.6		−105.9		−99.5			−105.2	
l	−117.9			−124.7		−105.4		−98.2			−104.0	
m	−124.7			−133.0		−113.3		−108.0			−113.3	
Path C												
n	−120.4			−128.3		−108.8		−98.4			−109.4	
o	−117.5			−122.9		−103.9		−94.9			−102.5	
p	−122.9	Minor	22	−127.4		−107.9		−102.5	23		−107.7	
q	−130.2			−135.0		−115.4		−108.7			−113.8	
r	−117.7			−121.5		−101.3		−97.4			−101.7	
s	−125.7			−124.5		−109.5		−103.3			−109.6	

[a] Reported by Dolle *et al.*[5] [b] Reported by Wilson and Schroepfer.[6] [c] Synthesis according to Scheme 2. [d] Reported by Dolle and Kruse.[7]
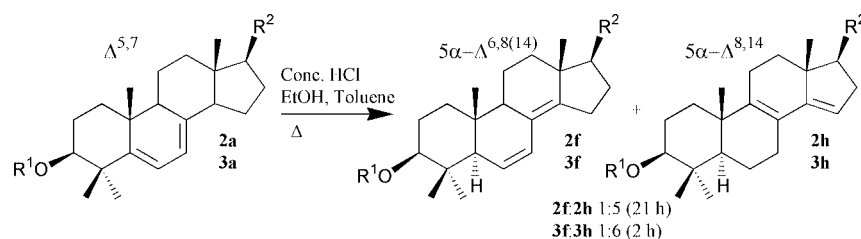
Scheme 2 Synthesis and relative yields of compounds **2f**, **2h** (R^1 = PhCO), **3f** and **3h** (R^1 = H). R^2 corresponds to the cholestane sidechain, shown in Scheme 1.

conformation and electronic configuration. These are taken into account as much as possible using formula (1).

$$\Delta H_f = \Sigma E_{bonds} + E_{st} + E_{SCF} + E_{T/R} + E_{TORS} + E_{POP} \quad (1)$$

The ΣE_{bonds} term represents the bond and structural increments of the molecular structure. E_{st} represents the steric energy of the molecule. If a conjugated system is present in the structure, the Self Consistent Field energy (E_{SCF}) for that system has to be included as well. Translation and rotational energy are included with the $E_{T/R}$ term which is taken to be $4RT$. The E_{TORS} term represents the corrections needed in case wide torsional motions are possible in the molecule. When this occurs, the harmonic approximation for the potential energy with respect to the internal coordinate is no longer valid. Therefore, the differences in energy levels in the potential well will be smaller than those calculated using the harmonic approximation. The MM3 program does not calculate the torsional correction terms. However, Wertz and Allinger[9] proposed a value of 0.42 kcal mol^{-1} for the E_{TORS} term, based on theoretical considerations by Pitzer and Gwinn.[10] In the calculations described here, the heats of formation were corrected by adding the TORS term multiplied by the number of rotations for which the harmonic approximation was not valid (five times the above mentioned correction term for series **1**, **2**, **4** and **5**, three times for series **3**). E_{POP} represents the energy contribution originating

from the fact that higher energy conformations are present. It is calculated using a Boltzmann weighted summation of the heats of formation for an ensemble of conformers generated at 298 K. The ensemble was generated by randomly changing the atom positions in the molecule 200 times using the stochastic search algorithm in MM3. The conformations found were minimised and only unique ones were retained.

Results and discussion

Calculated heats of formation for compound series **1** to **5** are given in Table 1, along with the available relative synthetic yields obtained by us and taken from the literature. For series **1**, no explicit yields are given in the article by Dolle *et al.*[5] However, it was stated that similar results were obtained as for the synthesis of the 17-ergostane-3-benzoylated series (series **4**), reported in the same article.

The heat of formation calculations were consistent for all five series with respect to the relative stabilities of the different isomers within a series. The starting compounds (**1a**–**5a**) were found to have the most unfavourable heat of formation compared to almost all other isomers within a series. Compounds in path **A** (Scheme 1) generally had a more unfavourable heat of formation than those of paths **B** and **C** for each series. Given the finding that the isomers (**b**–**e**) in path **A** are not formed in appreciable yields in any of the syntheses, the calculation results

are in agreement with the synthetic data. For each series, spiro compounds **1i–5i** and **1q–5q** were found to have the lowest heats of formation in paths **B** and **C**, respectively. It is known that isomerisation reactions, such as those described in this paper, yield spiro compounds like **1i–5i** and **1q–5q** when the reactions are allowed to proceed for longer time spans.[3] However, the reactions described in this paper and those reported in the literature[5–7] were stopped after the desired isomer was formed in excess over the other isomers. In all syntheses, the isomer of interest occurred 'up-stream' from the spiro compounds in the reaction paths **B** and **C** (Scheme 1). Therefore, the spiro compounds and the 5α- and 5β-$\Delta^{14,16}$ isomers **m** and **s**, which in most cases also have favourable heats of formation, are not formed.

In the reactions reported by Dolle *et al.*[5] (series **1** and **4**), the major products are the 5α- and 5β-$\Delta^{8,14}$ isomers (**1h,4h** and **1p,4p** respectively). These isomers are more stable than any of the other isomers within path **B** and **C** respectively, except for the 5α- and 5β-$\Delta^{14,16}$ isomers **m** and **s** and spiro compounds **i** and **q** (see discussion above). The syntheses reported in the present paper (series **2** and **3**) yield the 5α-$\Delta^{6,8(14)}$ and 5α-$\Delta^{8,14}$ isomers **f** and **h** as major products. These isomers have lowest heats of formation in path **B**, with the exception of isomer **m** and spiro compound **i** (see above). So, for the syntheses shown in Scheme 2 and those reported by Dolle *et al.*,[5] the major products are the more stable compounds (in terms of heat of formation) occurring in the first part of the formation paths **B** and **C**, which lead to the 5α- and 5β-$\Delta^{8,14}$ isomers **h** and **p**. It is interesting to note that compounds **2f** and **2h** have quasi identical calculated heats of formation, yet their yield ratio changes from 5:1 after one hour to 1:5 after 21 hours. This phenomenon is not caused by selective decomposition of compound **2f**, since no appreciable loss of total product mass was observed.[11] It follows that compound **2f** is converted into compound **2h** and that the latter compound actually has a lower heat of formation. The relatively slow conversion of **2f** into **2h** is probably caused by a kinetic barrier between the two compounds which has to be overcome to reach the lower heat of formation well of the latter compound.

The synthesis described by Wilson and Schroepfer[6] (series **1**) yields the 5α-$\Delta^{7,14}$ isomer (**1g**) as the major product. As was the case in the syntheses described by Dolle *et al.*,[5] the 5α-$\Delta^{8,14}$ and 5β-$\Delta^{8,14}$ isomers (**1h** and **1p**) were also formed in considerable amounts. However, the 5α-$\Delta^{7,14}$ isomer (**1g**) has an unfavourable heat of formation compared to both **1h** and **1p**, but is isolated in highest yields. Similar results are found for the syntheses reported by Dolle and Kruse[7] (series **4** and **5**). The 5α-$\Delta^{7,14}$ and 5α-$\Delta^{8,14}$ isomers are found in a 3:1 yield ratio (**4g**:**4h** and **5g**:**5h**). When comparing reaction temperatures of the syntheses that yield the **g** isomer (Dolle and Kruse[7] and Wilson and Schroepfer[6]) with those that yield the **h** isomer in highest yields (Dolle *et al.*[5] and the syntheses reported in this paper), we found that the former are considerably lower (−60 to −40 °C) than the latter (25–80 °C). We therefore hypothesise that at low temperature, the isomers formed sequentially in the reaction path **B** are not in equilibrium during their formation. Thus, at low temperature, the reaction is slowed down considerably and the isomers are trapped. Dolle *et al.*[5] describe the temperature-dependent formation of compounds **4g** and **4h**, and report that at −30 °C, isomer **g** was predominant, whereas the $\Delta^{8,14}$ isomer (**4h**) was formed in excess when raising the reaction temperature to 25 °C. In addition, Wilson and Schroepfer[6] report increased formation of **1h** at the expense of **1g** after raising the reaction temperature from −55 to 15 °C. These findings support our hypothesis that at low temperature, equilibrium is not reached and the isomers preceding the stable 5α-$\Delta^{8,14}$ isomer (**h**) occurring in path B can be trapped.

In some cases, the 5β-isomers of path **C** are more stable in terms of calculated heats of formation than the corresponding 5α-isomers of path **B**, whereas the latter are formed in higher yields. This holds for isomer pairs **1h** and **1p** (5α-$\Delta^{8,14}$ and 5β-$\Delta^{8,14}$), **2f** and **2n** (5α-$\Delta^{6,8(14)}$ and 5β-$\Delta^{6,8(14)}$), **3f** and **3n** (5α-$\Delta^{6,8(14)}$ and 5β-$\Delta^{6,8(14)}$). It is well known that double bond sterols and steroids are more susceptible to reactions on the α-side than on the β-side. The C-19 methyl group and the methyl group attached to C-4 projecting towards the β-side of the molecule shield the Δ^5 double bond at this side of the molecule, while the α-side is far less sterically hindered. Therefore, protonation of C-5 will occur in preference at the α-side of the molecule, even if the heats of formation of the 5β-isomers are more favourable.

According to the reaction mechanism, the 17α-isomer **j** in path **B** could be formed from isomer **i**. In the article by Dolle *et al.*,[5] an unknown compound was in the first instance thought to be the 5α-17-$\Delta^{8,14}$ isomer **4j**. However, crystal structure analysis showed that this isomer was in fact the 5β-17β-$\Delta^{8,14}$ isomer (**4p**), which is more obvious when comparing the heats of formation of these isomers. Knowledge of the relative stabilities of the different isomers that can occur during synthesis may therefore help to identify products with unknown structure.

Conclusion

Calculated heats of formation of the sterol isomer compounds considered are consistent with the formation of these compounds in syntheses. Firstly, the $\Delta^{5,7}$ (**1a–5a**) starting compounds have most unfavourable heats of formation compared to all other compounds within a series. Secondly, compounds in path **A**, which were found to have unfavourable heats of formation compared to compounds in paths **B** and **C**, are not formed. Most favourable heats of formation are found for spiro compounds **i** and **q**, which are well known end products of sterol isomerisation reactions. These are not formed if the reactions are stopped at an isomer occurring up-stream in the formation pathways **B** and **C**. The same holds for the $\Delta^{14,16}$ isomers **m** and **s**, which are also found to have favourable heats of formation.

Syntheses described in this paper (series **2** and **3**) and those reported by Dolle *et al.*[5] (series **1** and **4**) yield isomers **f**, **h** and **f**, **h**, **p**, respectively, as major products. The heats of formation of these compounds are the most favourable in the reaction routes leading towards spiro compounds **i** and **q**. When the temperature during the reaction is lowered, it is possible to trap the more unstable 5α-$\Delta^{7,14}$ isomer (**g**), which was the case in the syntheses reported by Wilson and Schroepfer[6] (series **1**) and those reported by Dolle and Kruse[7] (series **4** and **5**).

Calculation of heats of formation for isomeric compounds correlate well with synthetic yields of isomerisation reactions. When used for predictions, they may lead to new insights when planning a synthesis. Thus, it can be a useful tool for synthetic chemists designing a new experiment. In addition, the calculated stabilities of the different isomers can help to identify unknown products.

References

1 N. L. Allinger, Y. H. Yuh and J.-H. Lii, *J. Am. Chem. Soc.*, 1989, **111**, 8551.
2 Y.-H. Hii and N. L. Allinger, *J. Am. Chem. Soc.*, 1989, **111**, 8566.
3 R.-M. Liu, X. F. D. Chillier, P. Kamalaprija, U. Burger and F. O. Gülaçar, *Helv. Chim. Acta*, 1996, **79**, 989.
4 A. van Duin, PhD thesis, Delft University of Technology, The Netherlands, 1998.
5 R. E. Dolle, S. J. Schmidt, D. Eggleston and L. I. Kruse, *J. Org. Chem.*, 1988, **53**, 1563.
6 W. K. Wilson and G. J. Schroepfer, Jr., *J. Org. Chem.*, 1988, **53**, 1713.
7 R. E. Dolle and L. I. Kruse, *J. Org. Chem.*, 1986, **51**, 4047.
8 A. G. Byskov, C. Y. Andersen, L. Hordhom, H. Thogersen, X. Guoliang, O. Wassmann, J. V. Andersen, E. Guddal and T. Roed, *Nature*, 1995, **374**, 559.
9 D. H. Wertz and N. L. Allinger, *Tetrahedron*, 1979, **35**, 3.
10 K. S. Pitzer and W. D. Gwinn, *J. Chem. Phys.*, 1942, **10**, 428.
11 Data not shown.