Computational Investigations of Carbenium Ion Reactions Relevant to Sterol Biosynthesis

Corky Jenson and William L. Jorgensen*

Contribution from the Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107

Received May 5, 1997. Revised Manuscript Received August 24, 1997[⊗]

Abstract: Computational studies of carbenium ions relevant to sterol biosynthesis via lanosterol synthase were undertaken to determine fundamental energetics underlying cyclization steps. Ab initio B3LYP/6-31G*//B3LYP/6-31G* calculations were performed for the addition of 2-methyl-2-propyl cation and 2-methylpropene to represent a tertiary cation → tertiary cation cyclization. Solvent effects were included by Monte Carlo (MC) simulations in methylene chloride, methanol, and THF. The picture that emerges for a cation−olefin cyclization is one of barrierless collapse at short distance, while desolvation and conformational barriers are expected for initial separations beyond ca. 5 Å. Thus, the cyclization that forms the sterol B ring likely proceeds in barrierless concert with A-ring formation in the preorganized environment of a cyclase enzyme. However, C-ring formation appears to involve a tertiary → secondary cation rearrangement. This was modeled by ab initio and MC calculations for the interconversion of the C_{11} cations, **9** and **10**. The 12 kcal/mol higher energy for the secondary ion is only reduced to ca. 10 kcal/mol by solvation. Though this is consistent with initial formation of the tertiary ion, force-field calculations on the full protosteryl cations show that the equilibrium between the isomeric ions can be readily shifted by selective placement of nucleophilic groups from the protein backbone or side chains including the indole ring of tryptophans.

Introduction

The biosyntheses of sterols are rich in novel chemistry.[1] The end products are stereochemically complex precursors of molecules that play central roles in cellular structure and intercellular signaling. One extensively studied sequence is the conversion via lanosterol synthase of 2,3-oxidosqualene (**1**) to a protosteryl cation, **2**, which then undergoes a series of suprafacial 1,2-shifts to form lanosterol (**3**) (Figure 1). The overall process is striking for its stereocontrol in the generation of four new rings and seven asymmetric centers. A fundamental understanding of the enyzme mechanism would have benefits for the design of biomimetic catalysts and selective inhibitors. The related experimental work has featured studies of sterol cyclase activities on a variety of substrates[1d,2] and emulation of the chemical processes central to sterol cyclization.[1a−c,3] At the crux of understanding is the question of whether the enzymatic conversion of **1** to **2** occurs via a "nonstop" process, or whether a stepwise mechanism obtains.[1] Earlier work has supported the intermediacy of discrete carbocations for non-enyzmatic cation−olefin cascades,[4] and recently Corey et al. presented compelling evidence for the formation of carbocationic intermediates in the cyclization of 20-oxa-2,3-oxidosqualene by yeast lanosterol sythase.[2d] In the closure of ring C, two products were isolated that implicate the expansion of an initially formed cyclopentylcarbinyl cation, **4**, to the cyclohexyl isomer, **5**, before trapping (Figure 2).[2d] The result can be rationalized by noting that the tertiary cation is the product of Markovnikov addition, which is the kinetically preferred pathway in similar nonenzymatic reactions.[5] Details of the enzymatic cyclization of the other rings are unclear, though Corey et al. have now established that oxirane cleavage of **1** and A-ring formation are concerted.[2e]

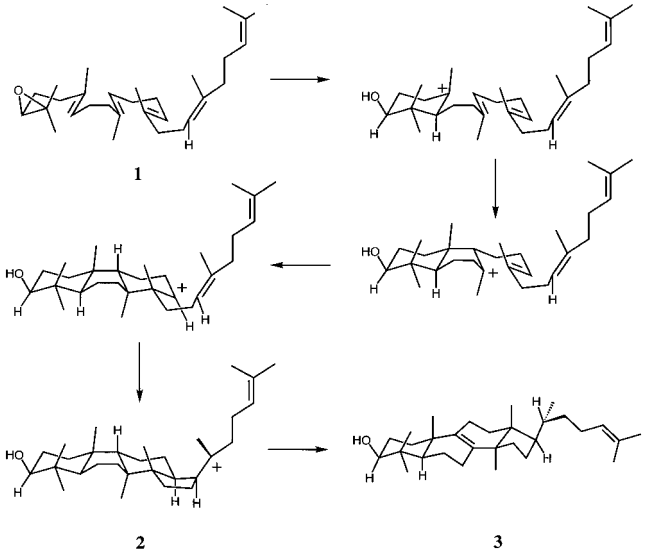


Figure 1. Hypothetical course for the conversion of 2,3-oxidosqualene (**1**) to lanosterol (**2**).

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1997.

(1) (a) Eschenmoser, A.; Ruzicka, L.; Jeger, O.; Arigoni, D. *Helv. Chim. Acta* **1955**, *38*, 1890. (b) Johnson, W. S. *Acc. Chem. Res.* **1968**, *1*, 1. (c) van Tamelen, E. E. *Acc. Chem. Res.* **1968**, *1*, 111. (d) Abe, I.; Rohmer, M.; Prestwich, G. D. *Chem. Rev.* **1993**, *93*, 2189.

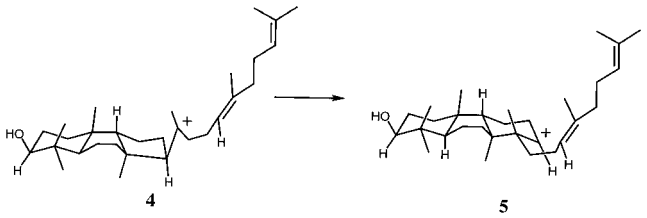
(2) (a) Corey, E. J.; Virgil, S. C.; Sarshar, S. *J. Am. Chem. Soc*. **1991**, *113*, 8171. (b) Krief, A.; Scauder, J.-R.; Guittet, E.; Herve du Penhoat, C.; Lallemand, J.-Y. *J. Am. Chem. Soc.* **1987**, *109*, 7910. (c) Corey, E. J.; Virgil, S. C. *J. Am. Chem. Soc.* **1991**, *113*, 4025. (d) Corey, E. J.; Virgil, S. C.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Singh, V.; Sarshar, S. *J. Am. Chem. Soc.* **1995**, *117*, 11819. (e) Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, *119*, 1277. (f) Corey, E. J.; Cheng, H..; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, *119*, 1289.

(3) (a) Nishizawa, M.; Takenaka, H.; Hayashi, Y. *J. Org. Chem.* **1986**, *51*, 806. (b) Pock, R.; Mayr, H.; Rubow, M.; Wilhelm, E. *J. Am. Chem. Soc.* **1986**, *108*, 7767. (c) Taylor, S. K. *Org. Prep. Proced. Int.* **1992**, *24*, 247. (d) Barrero, A. F.; Altarejos, J.; Alvarez-Manzaneda, E. J.; Ramos, J. M.; Salido, S. *J. Org. Chem.* **1996**, *61*, 2215. (e) Corey, E. J.; Wood, H. B., Jr. *J. Am. Chem. Soc.* **1996**, *118*, 11982. (f) Tietze, L. F. *Chem. Rev.* **1996**, *96*, 115. (g) Li, T.; Lerner, R. A.; Janda, K. D. *Acc. Chem. Res.* **1997**, *30*, 115.

(4) Nishizawa, M.; Takenaka, H.; Hayashi, Y. *J. Am. Chem. Soc.* **1985**, *107*, 522.

(5) (a) Sauers, C. K. Ph.D. Thesis University of Illinois, Urbana, IL, 1957. (b) Johnson, W. S.; Owyang, R. *J. Am. Chem. Soc*. **1964**, *86*, 5593.

S0002-7863(97)01424-8 CCC: $14.00

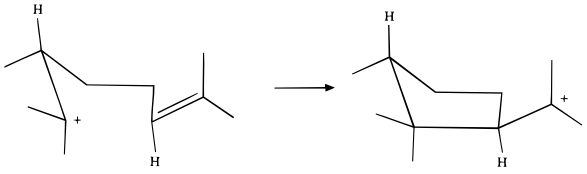


4 **5**

Figure 2. Expansion of the C-ring from the initially formed cyclopentylcarbinyl cation **4** to the cyclohexyl cation **5**.

Numerous basic issues about cation−olefin reactions and carbenium ion rearrangements have been raised that led to the present computational study. In particular, we set out to clarify the energetic profile of a fundamental cation−olefin addition in the gas phase and in solution. The central point is the existence or absence of a barrier to the reaction. The simplest model for a tertiary cation → tertiary cation reaction, as possibly in formation of sterol rings A and B, was pursued (Figure 3); the gas-phase reaction surface for the addition of 2-methyl-2-propyl cation (t-Bu$^+$), **6**, to 2-methylpropene (isobutylene), **7**, to form 2,2,4-trimethyl-4-pentyl cation (**8**) was determined using ab initio molecular orbital calculations, and the influence of solvation was taken into account through Monte Carlo (MC) statistical mechanics simulations in methylene chloride, tetrahydrofuran (THF), and methanol.

Delving further into the actual enzymatic transformation, Corey's work on the formation of the C ring has led us to clarify the relative energetics of models for the competing secondary and tertiary protosteryl cations, both in the gas phase and in solution. Ab initio calculations were performed on the C_{11} fragments corresponding to ring C and its substituents capped as methyl groups, i.e., 2-(2,2,3-trimethylcyclopentyl)-2-propyl cation (**9**) and 1,1,2,2,3-pentamethyl-6-cyclohexyl cation (**10**) (Figure 4). The energetics of the ring closure from 2,3,7-trimethyloct-6-en-2-yl cation (**11**) were also examined. The



11 **9**

influence of solvation on the **9** → **10** rearrangement was then considered through MC simulations in methylene chloride and THF. In the absence of detailed structural information for the cyclase, interactions with the enzyme that might favor the expansion of the initially formed cyclopentylcarbinyl cation to the cyclohexyl isomer are unknown. To gain insights on this issue, a molecular mechanics study was also performed for the full protosteryl cations **4** and **5** with probe molecules representative of nucleophilic subunits of proteins. This illuminates the optimal positioning and the possible extent of influence of single, advantageously placed side-chain or backbone groups.

Computational Details

Gas-Phase Cation−Olefin Addition. Full geometry optimizations were performed for **6** and **7** and the product of their addition, **8**, using restricted Hartree−Fock (RHF) calculations and the 3-21G and 6-31G* basis sets.[6] Electron correlation effects were included with hybrid Hartree−Fock density functional theory calculations, B3LYP/6-31G*//B3LYP/6-31G*. A reaction path was also generated by progressively lengthening the C2−C3 bond of the product and optimizing the remaining variables in C_s symmetry. Initially, the path was followed with geometries optimized at the RHF/3-21G level and with energy evaluations from RHF/6-31G* calculations (6-31G*//3-21G). Subsequently, the path was also determined at the B3LYP/6-31G*//B3LYP/6-31G* level.

Fluid Simulations for the Cation−Olefin Addition. Monte Carlo statistical mechanics simulations were then performed in order to determine the effects of solvation on the reaction profile. The simulations used methodology identical to that applied in recent studies of Diels−Alder reactions, and Claisen and Mislow-Evans rearrangements.[7] The optimizations with the 3-21G basis set provided a 53-frame "movie" of the reaction, which was initially solvated in boxes containing 390 molecules of methylene chloride, THF, or methanol, which had dimensions of approximately 30 × 30 × 45, 33 × 33 × 49, and 26 × 26 × 39 Å, respectively. The interactions between the reacting system and the solvent molecules are represented by a classical potential function where the intermolecular potential energy, ΔE_{ab}, consists of Coulombic and Lennard-Jones terms between the atoms i in molecule a and the atoms j in molecule b, which are separated by r_{ij}:

$$\Delta E_{ab} = \sum_i \sum_j \{q_i q_j e^2/r_{ij} + 4\epsilon_{ij}[(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^6]\}$$

The OPLS united-atom models for the solvents, which have been shown to reproduce well the thermodynamic properties of the pure solvents,[8] were adopted. For the 3-21G reaction path, the charges for the solute atoms came from the single-point RHF/6-31G* calculations with use of the CHELPG procedure for fitting the charges to the electrostatic potential surface.[6] The charges for the B3LYP/6-31G* reaction path came from B3LYP/6-31G* CHELPG calculations. Standard OPLS all-atom (AA) Lennard-Jones parameters (σ, ϵ) were adopted for the solutes;[9] specifically, σ (in Å) and ϵ (in kcal/mol) were 3.50 and 0.66 for an sp^3 carbon, 3.55 and 0.076 for an sp^2 carbon, 2.50 and 0.030 for hydrogen on an sp^3 carbon, and 2.42 and 0.030 for hydrogen on an sp^2 carbon. For atoms that changed type, σ and ϵ were scaled linearly along the reaction path. Also, geometric combining rules were used to obtain σ_{ij} and ϵ_{ij}. All solutes were represented in an all-atom format, i.e., one interaction site at each atomic nucleus.

The changes in free energy of solvation along the reaction path were calculated via Monte Carlo simulations with statistical perturbation theory.[10] This allows the computation of the changes in free energy for perturbing the system from one frame of the movie to the two frames on either side of it by double-wide sampling.[10b] For example, in order to span the 53 frames for the 3-21G pathway, 26 MC simulations in each solvent were carried out. Every simulation involved 2 × 10^6 configurations for equilibration, followed by 2 × 10^6 configurations of averaging in the isothermal, isobaric (NPT) ensemble at 25 °C and 1 atm. The simulations were performed with periodic boundary conditions, Metropolis sampling, and truncation of the intermolecular interaction energies at 12 Å, based roughly on the separation of the centers of mass. Internal degrees of freedom were not varied; only translations and rigid-body rotations for the solvent molecules were

(6) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian94*, Revision A; Gaussian, Inc.; Pittsburgh, PA, 1994.

(7) (a) Jorgensen, W. L.; Blake, J. F.; Lim D.; Severance, D. L. *Trans. Faraday Soc.* **1994**, *90*, 1727. (b) Jones-Hertzog, D. K.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1995**, *117*, 9077.

(8) (a) Jorgensen, W. L. *J. Phys. Chem.* **1986**, *90*, 1276. (b) Briggs, J. M.; Matsui, T.; Jorgensen, W. L. *J. Comput. Chem.* **1990**, *11*, 958. (c) Lim, D.; Hrovat, D. A.; Borden, W. T.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1994**, *116*, 3494.

(9) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(10) (a) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420. (b) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050.

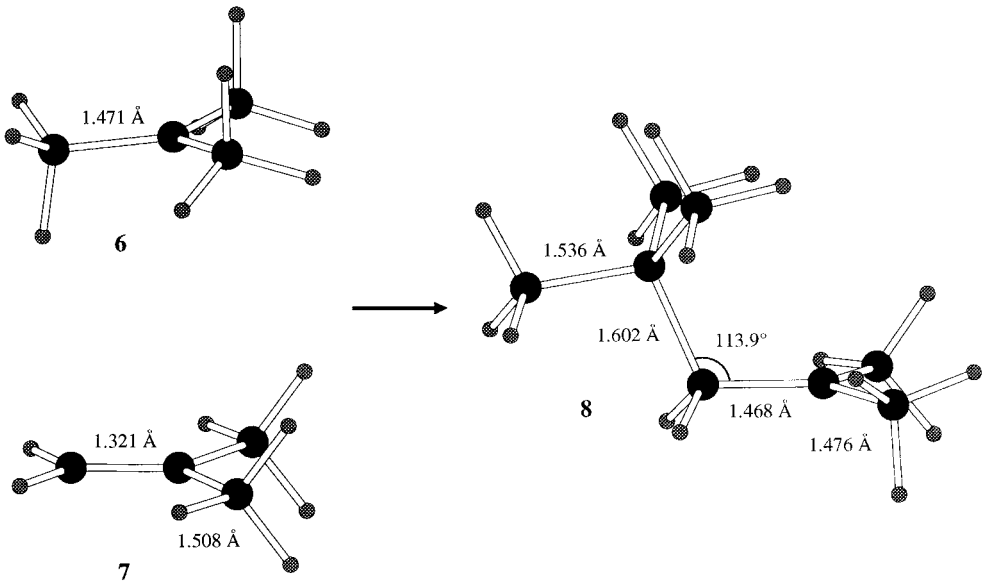


Figure 3. RHF/6-31G*-optimized structures of 2-methyl-2-propyl cation (**6**), 2-methylpropene (**7**), and the product of their reaction, 2,2,4-trimethyl-4-pentyl cation (**8**).

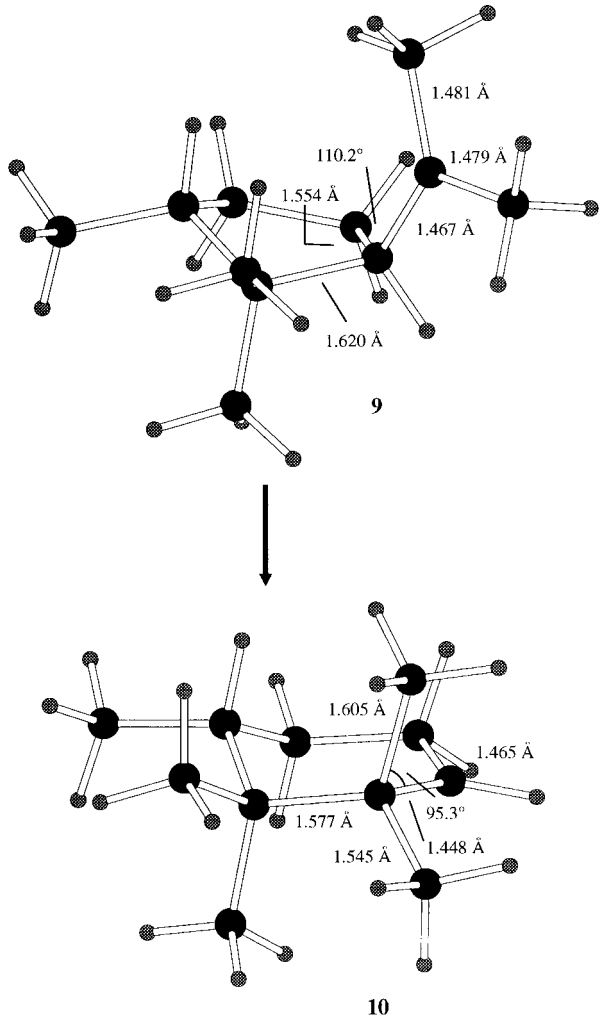


Figure 4. RHF/6-31G*-optimized structures for the C_{11} cations **9** and **10**.

sampled. In order to compute the total free energy change along the pathway in solution, the gas-phase relative energies were added to the computed free energies of solvation. The sensitivity of the results to the ab initio procedure was tested by performing analogous MC simulations for the B3LYP/6-31G* reaction path in methylene chloride. All Monte Carlo simulations were executed with the BOSS program.[11]

C-Ring Rearrangement and Formation. In order to gain insights into the formation of the sterol C ring, full geometry optimizations were performed for **9** and **10** with ab initio calculations at the RHF/3-21G//RHF/3-21G and RHF/6-31G*//RHF/6-31G* levels. In addition, the effects of electron correlation were incorporated by B3LYP/6-31G*//RHF/6-31G* calculations. The energy change for the ring closure to form **9** from **11** was also obtained from RHF/6-31G*//RHF/6-31G* calculations. A conformational search was performed with the BOSS program and OPLS-AA force field to find the lowest-energy structure for **11**, which was used in the ab initio calculations.

The influence of solvation on the energetics for the **9** → **10** rearrangement was explored with Monte Carlo simulations. The solute parameters consisted of CHELPG charges from the B3LYP/6-31G*//RHF/6-31G* calculations and the Lennard-Jones parameters, which were described above. Two solvents were treated using cubic cells of 260 solvent molecules with dimensions of ca. 30 × 30 × 30 Å for methylene chloride and 33 × 33 × 33 Å for THF. The details of the MC simulations were the same as above, including truncation of the intermolecular interactions at 12 Å. The structures were perturbed from the tertiary cation to the secondary isomer in a series of six steps or "windows" with 1×10^6 configurations of equilibration and 2×10^6 configurations of averaging for each window.

Molecular Mechanics Studies of Interactions with the Protosteryl Cations. Molecular mechanics calculations were performed with the BOSS program to elucidate interactions for the protosteryl cations **4** and **5** with individual solvating molecules. The geometries of the cations were determined by BFGS optimizations;[12] the lowest-energy conformation for the C20−C27 terminus was used along with a gauche conformation about the C15−C16 bond to provide a geometry positioned for D-ring formation from **5** and the observed four-membered ring formation from **4**.[2d] The nonbonded parameters for the cations were standard OPLS-AA values[9] merged with the B3LYP/6-31G*//RHF/6-31G* CHELPG charges from the C_{11} model systems, **9** and **10**. Some bond-stretching and angle-bending parameters involving the cationic carbons, which were not available in the OPLS-AA force field, were adopted from Fry et al.[13] In order to obtain correct conformational energetics for rotation about bonds attached to cationic carbons, a twofold torsion term with $V_2 = -1.0$ kcal/mol also needed to be added for C−C−C$^+$−C,H. The probe molecules were *N*-methylacetamide, dimethyl sulfide, imidazole, benzene, and indole; OPLS-AA parameters[9]

(11) Jorgensen, W. L. *BOSS, Version 3.6*; Yale University: New Haven, CT, 1995.

(12) Press, W. H.; Teukolsky, S. A.; Vettering, W. T.; Flannery, B. P. *Numerical Recipes in FORTRAN*, 2nd ed.; Cambridge Univ. Press: Cambridge, U.K., 1992.

(13) Fry, J. L.; Engler, E., M.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1972**, *94*, 4628.

and geometries from 6-31G*//6-31G* optimizations[6] were used in these cases. For the complexes between the protosteryl cations and the probe molecules, the structures of the individual molecules were fixed at their optimized geometries and conformational searching was performed for the six intermolecular degrees of freedom. The lowest-energy structures that were obtained from BFGS optimization of the six degrees of freedom are reported here.

Results and Discussion

Gas-Phase Reaction of *t*-Bu$^+$ and Isobutylene. The energy changes for conversion of the reactants, **6** and **7**, to the product, **8**, from the RHF/3-21G//RHF/3-21G, RHF/6-31G*//RHF/6-31G*, and B3LYP/6-31G*//B3LYP/6-31G* calculations are −25.4, −19.8, and −22.9 kcal/mol, respectively. These figures correspond to the gas-phase transformation at 0 K. In spite of the bond formation, the correlation energy is similar for the reactants and product. The optimized geometries of **6**, **7,** and **8** at the RHF/6-31G*//RHF/6-31G* level are shown in Figure 3. A notable point is the long C2−C3 bond in **8**, 1.602 and 1.665 Å from the 6-31G* and B3LYP optimizations, which reflects the hyperconjugation with the cationic 2p-orbital on C4. Vibrational frequencies were also computed at the 6-31G* level. In combination with the RHF/6-31G*//RHF/6-31G* energies, the resultant energy, enthalpy, entropy, and free energy for the transformation at 298 K are −17.6 kcal/mol, −18.2 kcal/mol, −47.6 cal/mol·K, and −4.0 kcal/mol, respectively.[14] The energy changes of ca. −20 kcal/mol reasonably reflect the replacement of a CC π-bond (65 kcal/mol) by a CC σ-bond (85 kcal/mol). For a further sense of the accuracy of calculations at this level, it may be noted that the $\Delta E°$ including the zero-point correction for *t*-Bu$^+$→ methyl-bridged 2-butyl cation is 14.6 kcal/mol from both RHF/6-31G*//RHF/6-31G* and B3LYP/6-31G*//B3LYP/6-31G* calculations, which compares favorably to a value of 13.8 kcal/mol from MP4/6-31G**//MP2/6-31G** calculations,[15a] a gas-phase experimental ΔH^{298} of 16.2 ± 1.6 kcal/mol,[16] and a ΔH of 14.5 ± 0.5 kcal/mol in SO_2ClF solution.[17]

The gas-phase energy profiles from both the RHF/6-31G*//RHF/3-21G and B3LYP/6-31G*//B3LYP/6-31G* calculations and several structures from the B3LYP calculations are shown in Figure 5. The addition occurs essentially without activation energy in the gas phase; however, the process has two stages. When the reactants are 5−6 Å apart, the *t*-Bu$^+$ cation approaches the olefin in a canted manner with hydrogens from two methyl groups directed toward the π-bond. A flat region is then found in the B3LYP/6-31G*//B3LYP/6-31G* profile, as well as with RHF/6-31G*//RHF/6-31G* (not shown), from 4 to 3.5 Å, while it is pushed into 2−3 Å in the RHF/6-31G*//RHF/3-21G profile. In this region, the *t*-Bu$^+$ reorients to allow formation of the C2−C3 bond, and the C2−C3−C4 angle widens to ca. 100°. Shallow minima were located in the RHF/6-31G*//RHF/6-31G* and B3LYP/6-31G*//B3LYP/6-31G* energy surfaces at C2−C3 distances of 3.6 and 4.0 Å with net interaction energies of −5.4 and −6.7 kcal/mol relative to the separated reactants; the barrier to collapse to the product is about 0.5 kcal/mol. These minima are tenuous; they might not survive with still higher level calculations and they should not be assigned significance as π-complexes. The reliable point is that there is a flat region

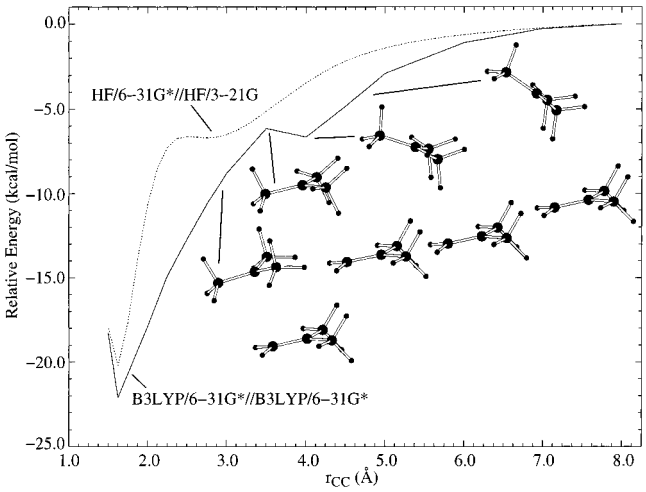


Figure 5. Potential energy changes along the gas-phase reaction paths for the addition of 2-methyl-2-propyl cation (**6**) to 2-methylpropene (**7**) from RHF/6-31G*//RHF/3-21G (dotted line) and B3LYP/6-31G*//B3LYP/6-31G* (solid line) calculations. Some B3LYP/6-31G*-optimized structures are also shown.

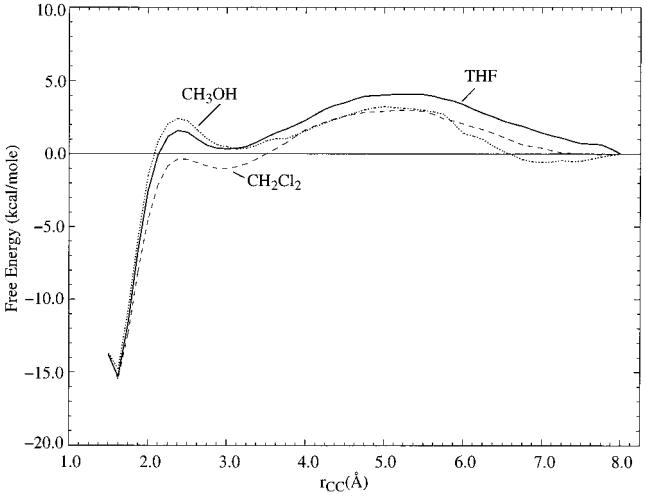


Figure 6. Free energy profiles from the Monte Carlo simulations in three solvents for the addition of 2-methyl-2-propyl cation (**6**) to 2-methylpropene (**7**) using the 6-31G*//3-21G reaction path.

in the gas-phase energy surfaces at a C2−C3 separation of 3−4 Å, and inside this region, the energy declines rapidly as C2 and C3 rehybridize, the bond forms between them, and charge is transferred from isobutylene to the electrophile. Experimentally, it may be noted that the present cation−olefin addition has been observed to occur at −115 °C on a cooled surface and the process can be reversed upon warming to −80 °C.[18]

Solvent Effects on the Cation−Olefin Addition. The free energy profiles for the 6-31G*//3-21G reaction path in the three solvents are shown in Figure 6. For the B3LYP/6-31G*//B3LYP/6-31G* path, the free energy profile in methylene chloride is shown in Figure 7 along with its components, the changes in the gas-phase potential energy and the free energy of solvation. The profiles for the 6-31G*//3-21G and B3LYP pathways in methylene chloride are very similar and support the qualitative correctness of the results. In solution, a 3−4 kcal/mol barrier is introduced with a maximum at 5−6 Å separation. Isolated *t*-Bu$^+$ has a solvent molecule well-positioned on both π-faces, with either a chlorine or an oxygen from the solvent molecules near the central carbon. This is illustrated by the stereoplot in Figure 8 of a representative

(14) Vibrational frequencies were scaled by 0.91, and scaled frequencies below 500 cm^{-1} were treated as rotations for computing ΔH^{298}.

(15) (a) Sieber, S.; Buzek, P.; Schleyer, P. v. R.; Koch, W.; Carniero, J. W. M. *J. Am. Chem. Soc.* **1993**, *115*, 259. (b) Ibrahim, M. R.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1989**, *111*, 819.

(16) Schultz, J. C.; Houle, F. A.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1984**, *106*, 3917.

(17) Bittner, E. W.; Arnett, E. M.; Saunders, M. *J. Am. Chem. Soc.* **1976**, *98*, 3734.

(18) Saunders, M.; Lloyd, J. R. *J. Am. Chem. Soc.* **1977**, *99*, 7090.

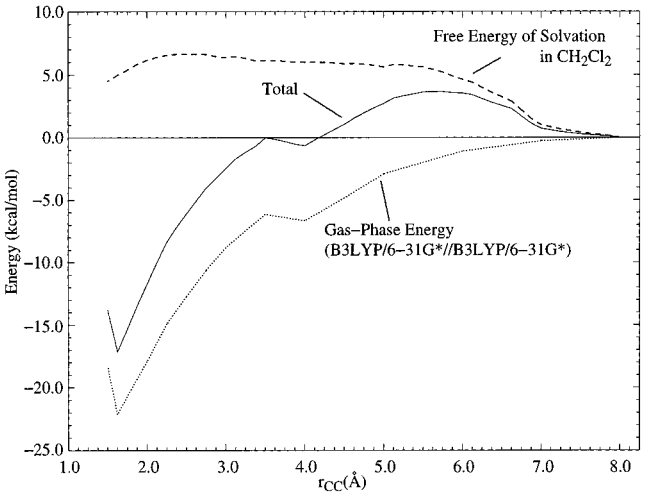


Figure 7. Computed free energy profile for the addition of **6** to **7** in methylene chloride along with its components, the changes in gas-phase potential energy and free energies of solvation, for the B3LYP/6-31G*//B3LYP/6-31G* reaction path.

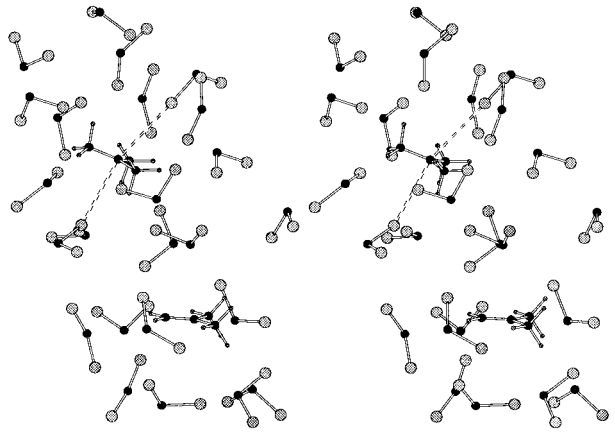


Figure 8. Stereoplot of a representative configuration from the MC simulation of **6** and **7** at 8 Å separation in methylene chloride. Only the nearest solvent molecules are shown. Hydrogens of methylene chloride are implicit. The highlighted C−Cl distances are 3.57 and 3.47 Å.

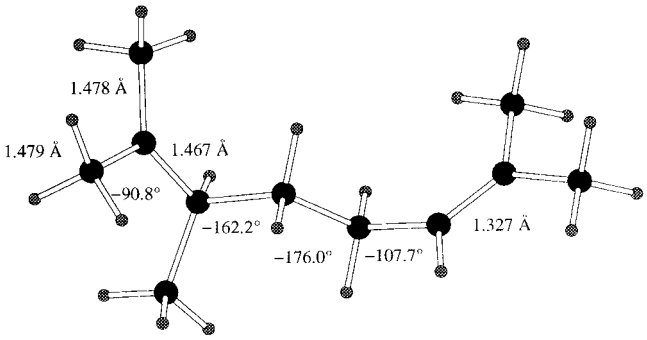
configuration with the reactants separated by 8 Å in methylene chloride. Ejection of the solvent between the reactants accounts for the barrier. Solvation also strengthens the minimum near 3 Å separation in Figure 6; the effect increases with increasing solvent polarity. This occurs because desolvation of the reactants continues up to about 2 Å separation, while the gas-phase energy profile is flat in the 2−3 Å region for the 6-31G*/3-21G path or the 3−4 Å region in the B3LYP path (Figure 5). The poorer solvation at shorter distances results from both the extrusion of solvent and the progressive charge delocalization, which is not reversed until the product is well formed.

It is clear that in the gas phase the reaction proceeds without an activation barrier and is exothermic by ca. 20 kcal/mol. It is also clear that transfer to solution introduces a net desolvation barrier of 3−4 kcal/mol and tends to reinforce the potential existence of an intermediate at 3−4 Å separation. This suggests that for intramolecular tertiary → tertiary cation−olefin additions barrierless collapse can be expected if the cationic center and the double bond are inside the maximum in the desolvation barrier, i.e., if they are separated by less than ca. 5 Å. This situation is likely in the preorganized state of a substrate such as **1** bound to a cyclase enzyme. At larger initial separations of cationic site and olefinic bond, activation barriers can be expected, which could include a conformational component as well as removal of a solvent molecule or other solvating group from the reacting face of the cationic site. As referenced above, Corey's recent results indicate that, after protonation of the oxirane in 2,3-oxidosqualene (**1**) by Asp 456 of lanosterol synthase, cleavage of the oxirane and formation of ring A are concerted.[2e] The degree of preorganization, which is reflected in the observed stereocontrol, coupled with the present results suggests that the formation of ring B would then proceed without activation energy.

Gas-Phase Studies of C-Ring Formation and Expansion. The RHF/6-31G* optimized structures for **9** and **10** are shown in Figure 4. Both ions show strong hyperconjugation with the C−C bond that best eclipses the cationic 2p-orbital. These bonds lengthen to 1.620 Å in **9** and 1.605 Å in **10**. For the secondary ion, the hyperconjugation is also reflected in the small bond angle of only 95.3° to the eclipsing methyl group. The ab initio calculations confirm the expected higher energy for the secondary ion. The gas-phase energy changes for the ring expansion are 10.8, 12.4, and 12.3 kcal/mol at the RHF/3-21G//RHF/3-21G, RHF/6-31G*//RHF/6-31G*, and B3LYP/6-31G*//RHF/6-31G* level, respectively. With vibrational frequencies from the RHF/6-31G* calculations and using the RHF/6-31G*//RHF/6-31G* energy difference,[14] the energy, enthalpy, entropy, and free energy change at 298 K are 12.4 kcal/mol, 11.9 kcal/mol, −5.1 cal/mol·K, and 13.4 kcal/mol, respectively.

The energy difference of 12 kcal/mol is somewhat less than the usual ca. 14 kcal/mol separation for simple, acyclic isomeric secondary and tertiary carbenium ions.[15] Clearly, for **9** and **10** there are differences in ring strain and steric interactions that affect this base value and that would be difficult to gauge reliably in the absence of the present calculations. It should be noted that RHF/6-31G*//RHF/6-31G* optimizations were also carried out for the alternative envelope structure of **9** with the methyl group on the methine carbon of the ring in a pseudoaxial orientation; this structure was 0.6 kcal/mol higher in energy than the conformer in Figure 4. RHF/6-31G*//RHF/6-31G* optimizations were also performed for C_8 analogs of **9** and **10** lacking the three methyl groups on the ring carbons in **9** and the corresponding methyl groups for **10**, i.e., dimethylcyclopentylcarbinyl cation and 2,2-dimethylcyclohexyl cation. The latter, secondary ion was found to be only 8.6 kcal/mol higher in energy than the tertiary ion. Reduction fom the standard 14 kcal/mol can be attributed mostly to the added ring strain for the five-membered ring of the tertiary ion. However, the increase in the tertiary to secondary difference from 8.6 kcal/mol in the unsubstituted case to 12.4 kcal/mol for **9** to **10** does suggest that the steric strain associated with the adjacent quaternary centers in **10** (and in **5**) hinders the ring expansion by several kcal/mol.

For the formation of **9** from **11**, the lowest-energy conformer for **11** that was found in the conformational search is illustrated in Figure 9. The energy change for conversion of this extended structure to **9** is −9.6 kcal/mol at the RHF/6-31G*//RHF/6-31G* level. The lessened favorability in comparison to the −19.8 kcal/mol for the formation of **8** from **6** + **7** can be attributed to the ring strain and steric interactions between the *syn* substituents in **9**. From RHF/6-31G*//RHF/6-31G* vibrational frequencies,[14] the computed enthalpy, entropy, and free energy changes for the conversion of **11** to **9** at 298 K are −8.1 kcal/mol, −15.4 cal/mol·K, and −3.6 kcal/mol. Thus, the enthalpy change also ends up 10 kcal/mol less favorable than for the cation−olefin addition to form **8**. For the enzymatic formation of the protosteryl cation **4**, preorganization can be expected to enhance the exothermicity of this step by several kilocalories per mole by raising the conformational energy of



11

Figure 9. RHF/6-31G*-optimized structure for cation **11**. Selected bond lengths and CCCC dihedral angles are shown.

the precursor. Thus, although formation of **4** is not as favorable as in an unconstrained tertiary → tertiary cation−olefin addition, it is still expected to be exothermic by ca. 10 kcal/mol.

Solvent Effects on the Ring Expansion. The gas-phase calculations have established an energy difference of 12 kcal/mol for the ring expansion of **9** to **10** as a model for the conversion of the protosteryl cation **4** to **5**. The next question is how the difference would be affected by transfer to a condensed-phase environment. The simple expectation is that the more charge-localized secondary ions should be better solvated. However, the distributions of charge for both secondary and tertiary carbenium ions are complex with substantial delocalization to both carbons and hydrogens.[19] In view of this and the uncertain steric influence of the substituents, the magnitude of the solvent effects for the present rearrangement was unclear. The Monte Carlo simulations for the 1,2-shift proceeded smoothly and with high precision; the resultant changes in free energies of solvation for the transformation of **9** to **10** are -1.0 ± 0.2 and -1.5 ± 0.2 kcal/mol in methylene chloride and THF, respectively. Thus, the presence of a low-dielectric medium including the interior of a protein is expected to have little influence on the equilibrium in the absence of additional specific interactions or steric requirements. In the only prior study of a carbocation rearrangement with similar methodology, the classical structure of the 2-norbornyl cation was found to be better solvated in water than the bridged, nonclassical form by only 0.7 kcal/mol.[20] Thus, except in extreme cases such as planar vs perpendicular allyl cations,[21] the computational results support the view that "solvent effects play a limited role on the relative energies of different equilibrium structures and transition states of carbocations".[22]

There is, however, some interesting structural information that can be garnered from the Monte Carlo simulations concerning the complexation of the carbenium centers in the two cations. As shown in Figure 10 for the THF solutions, two solvent molecules are complexing the carbenium center in **10**, but **9** is complexed by just one solvent molecule. One face of the carbenium center in **9** is blocked by an exocyclic methyl group, though for **10** the solvating THF molecules are bent away from the ring and the geminal methyl groups on the α-carbon. Although the absolute energetic effects of this differential complexation are small in the present solvents, it does illustrate the potential for specific interactions with the carbenium centers that could influence the equilibrium for the protosteryl cations in the active site of a cyclase enzyme. Nevertheless, the estimated energy difference of ca. 10 kcal/mol between the protosteryl cations **4** and **5** is consistent with Corey's results that **4** is formed first enzymatically and that its 20-oxa analog has sufficient lifetime that it can be trapped before the ring expansion.[2d]

Interactions of the Protosteryl Cations with Protein Fragments. A three-dimensional structure of lanosterol synthase, or any sterol cyclase, is not available owing to difficulties associated with the size and membrane-bound nature of the proteins. However, some insights on stabilizing interactions are emerging from the amino acid sequences that have been decoded from the genes,[23] from site-directed mutagenesis,[2e,f] and from studies of mechanism-based inhibitors.[2f,24] The reliance of older models on the placement of anionic sites to stabilize the cationic sites formed by the cyclizations[25] has been supplanted by a focus on potential cation−π interactions as key stabilizing elements.[26] In particular, studies of the known cyclase sequences reveal unusually high frequency and conservation of tyrosine and tryptophan residues across a number of species.[23]

The present probe molecules were chosen to represent nucleophilic groups of protein backbones and side chains: *N*-methylacetamide (NMA) for amide carbonyl groups of the backbone and side chains of asparagine and glutamine; dimethyl sulfide, imidazole, benzene, and indole for the side chains of methionine, histidine, phenylalanine, and tryptophan. Though the OPLS force field generally describes ion−molecule interactions well, the interaction energies for alkali cation−π complexes are known to be underestimated by a factor of about 2 using force fields that do not explicitly incorporate polarization.[27] However, the discrepancies are not expected to be as great for interactions with more delocalized cations such as secondary and tertiary carbenium ions. This was tested by optimizing structures for π-complexes of *tert*-butyl cation with benzene and indole both with the OPLS-AA force field and with ab initio RHF/6-31G* calculations. For the latter calculations, the RHF/6-31G*-optimized geometries for the separated species were kept fixed and the six intermolecular degrees of freedom were optimized. It should be noted that if the intramolecular constraints are relaxed the *tert*-butyl cation/benzene complex collapses to the σ-complex (Wheland intermediate) with an interaction energy of −14.9 kcal/mol at the MP2(fc)/6-31G*//RHF/6-31G* level.[28] The interaction energies were also computed at the B3LYP/6-31G* level using the present structures from the RHF/6-31G* calculations. The lowest-energy π-structures from the OPLS-AA optimizations are illustrated in Figure 11 along with the interaction energies from the different procedures. The OPLS-AA- and RHF/6-31G*-

(19) Wiberg, K. B.; Schleyer, P. v. R.; Streitwieser, A. *Can. J. Chem.* **1996**, *74*, 892.
(20) Schreiner, P. R.; Severance, D. L.; Jorgensen, W. L.; Schleyer, P. v. R.; Schaefer, H. F., III *J. Am. Chem. Soc.* **1995**, *117*, 2663.
(21) Mayr, H.; Förner, W.; Schleyer, P. v. R. . *J. Am. Chem. Soc.* **1979**, *101*, 6032. Cournoyer, M. E.; Jorgensen, W. L. . *J. Am. Chem. Soc.* **1984**, *106*, 5104.
(22) Saunders, M.; Jiménez-Vázquez, H. A. *Chem. Rev.* **1991**, *91*, 375.
(23) (a) Buntel, C. J.; Griffin, J. H. *J. Am. Chem. Soc.* **1992**, *114*, 9711. (b) Zheng, S.; Buntel, C. J.; Griffin, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7370. (c) Buntel, C. J.; Griffin, J. H. In *Isopentenoids and Other Natural Products, Evolution and Function*; Nes, W. D., Ed.; ACS Symposium Series 562; American Chemical Society; Washington, DC, 1994. (d) Baker, C. H.; Matsuda, S. P. T.; Liu, D. R.; Corey, E. J. *Biochem. Biophys. Res. Commun.* **1995**, *213*, 154.
(24) (a) Dodd, D. S.; Oehlschlager, A. C.; Georgopapadakou, N. H.; Polak, A.-M.; Hartman, P. G. *J. Org. Chem.* **1992**, *57*, 7226. (b) Abe, I.; Prestwich, G. D. *Lipids* **1995**, *30*, 231. (c) Zheng, Y. F; Dodd, D. S.; Oehlschlager, A. C. *Tetrahedron* **1995**, *51*, 5255. (d) Goldman, R. C.; Zakula, D.; Capobianco, J. O.; Sharpe, B. A.; Griffin, J. H. *Antimicrob. Agents Chemother.* **1996**, *40*, 1044.
(25) Johnson, W. S.; Lindell, S. D.; Steele, J. *J. Am. Chem. Soc.* **1987**, *109*, 5852.
(26) Dougherty, D. A. *Science* **1996**, *271*, 163.
(27) Kumpf, R. A.; Dougherty, D. A. *Science* **1993**, *261*, 1708. Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177.
(28) Schleyer, P. v. R.; Buzek, P.; Müller, T.; Apeloig, Y.; Siehl, H.-U. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1471.

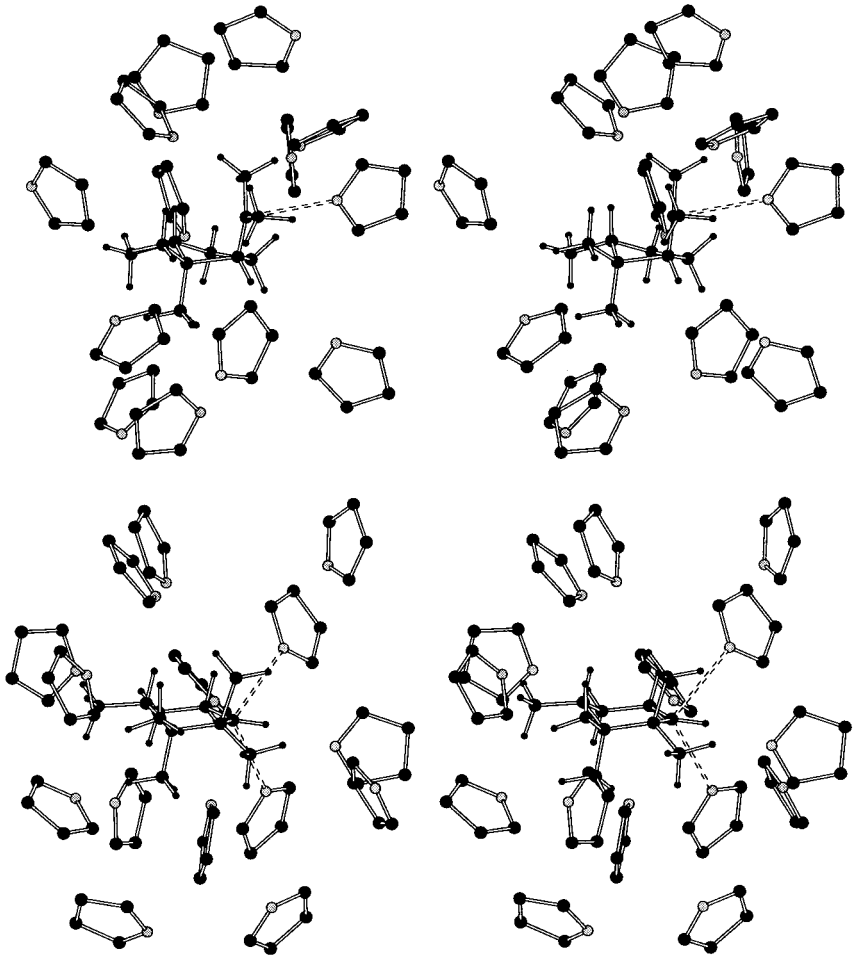


Figure 10. Stereoplots of representative configurations from the MC simulations of cations **9** (top) and **10** (bottom) in THF. Only the nearest solvent molecules are shown. The hydrogens of THF are implicit.

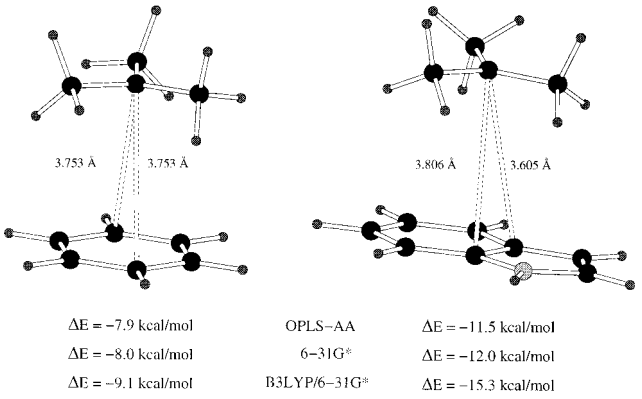


Figure 11. Structures optimized with the OPLS-AA force field for complexes of 2-methyl-2-propyl cation (**6**) with benzene and indole. The energy changes for formation of the complexes from the separated ion and molecule are given for the three computational procedures.

optimized structures are similar; e.g., the distances from the carbenium carbon to the ring center of benzene and C8 and C9 of indole are 3.48, 3.81, and 3.61 Å from the OPLS-AA optimizations and 3.75, 3.60, and 3.99 Å from the RHF/6-31G* calculations. As noted in Figure 11, the OPLS-AA and RHF/6-31G* interaction energies are within 0.5 kcal/mol, near 8 kcal/mol for benzene and 12 kcal/mol for indole. The B3LYP single-point calculations do predict increased attraction, though the discrepancies are small compared to the case of alkali cations.[27]

The search for low-energy structures showed that the two protosteryl cations preferred different complexation motifs. For **4,** both faces of the carbenium carbon are partially blocked, one by the B-ring, particularly, C7 and its α-hydrogen, and the other by the C21 methyl group. Consequently, the lowest-energy complexes are varied in Figure 12. The most attractive interaction (−15.5 kcal/mol) is with NMA, which has its oxygen placed on the front face (toward ring B) of the carbenium carbon. Relaxation of the intramolecular constraints would lead to some enhancement of the interaction energies and shorter distances to the carbenium carbons.[28] Dimethyl sulfide, imidazole, and indole all find favorable sites tucked under the incipient D-ring, while benzene has a similar interaction energy in that position and on the back face, as illustrated. For protosteryl cation **5**, the carbenium carbon is relatively unshielded and two low-energy structures were obtained in each case, one with the solvating molecule on either π-face. Both positions and the interaction energies are shown in Figure 13. In view of the reduced shielding and the greater electron demand of the secondary ion, the interactions are significantly more favorable for each probe molecule with **5** than with **4**. Furthermore, the most favorable position for the probe is on the steroidal α-face, under the incipient D-ring. Thus, the calculations suggest that a cyclase enzyme would be served well by positioning a nucleophilic group such as a tryptophan side chain in this position. Substantial additional stabilization can be achieved by placement of a second nucleophilic group on the β-face, e.g., by sandwiching the substrate between two aryl side chains, as illustrated in Figure 13. Furthermore, the individual interactions are sufficiently strong, 10−20 kcal/mol, that judicious placement of the nucleophilic groups in the enzyme could overcome the intrinsic preference for the tertiary cation **4** over the secondary isomer **5**. Thus, although the experimental evidence supports

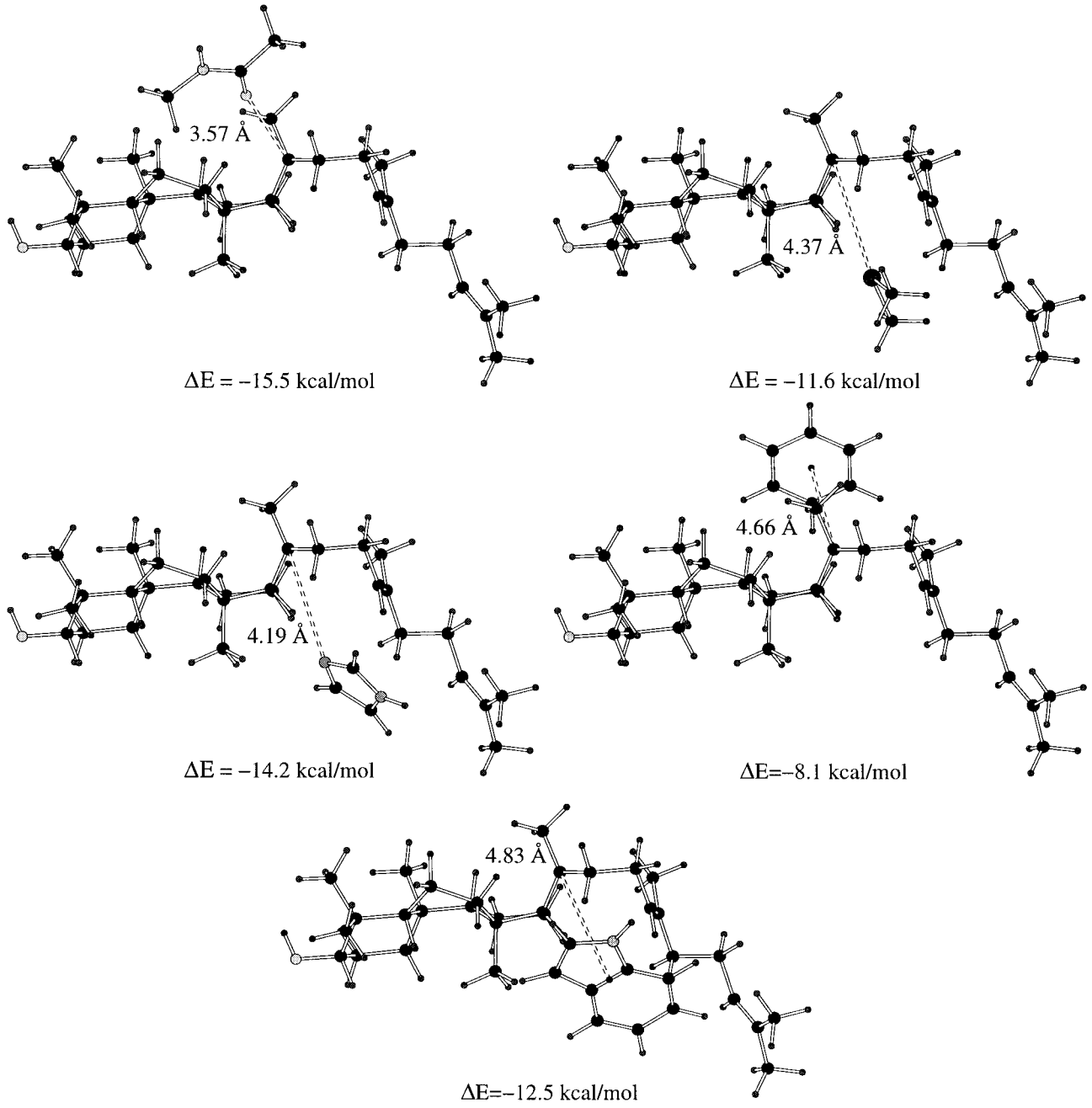


Figure 12. Optimized complexes for the protosteryl cation **4** and NMA, dimethyl sulfide, imidazole, benzene, and indole. Distances to the cationic center and complexation energies are given. For benzene and indole, the distances are from the ring center and center of the C8−C9 bond.

the intermediacy of **4** in the generation of ring C,[2d] a catalyst could be constructed to much enhance the direct formation of **5**.

Conclusion

The present study has addressed the energetics of cation−olefin additions, the ring expansion of a tertiary cyclopentylcarbinyl cation to a secondary cyclohexyl cation, and interactions between the protosteryl cations **4** and **5** and stabilizing groups. The principal findings were as follows: (1) The addition of *tert*-butyl cation to isobutylene is exothermic by about 20 kcal/mol in the gas phase and does not have an activation barrier. In solution, a 3−4 kcal/mol barrier for desolvation is introduced at a separation of 5−6 Å. (2) The intramolecular closure of the extended structure **11** to yield **9** in the gas phase is exothermic by only 8 kcal/mol owing to the ring strain and steric interactions in the cyclopentylcarbinyl cation. Formation of **9** from a more preorganized conformer of **11** is expected to be several kilocalories per mole more exothermic. (3) The rearrangement of the tertiary cation **9** to the secondary cyclohexyl cation **10** is endothermic by 12 kcal/mol in the gas phase. Better solvation of the secondary ion only reduces this difference by 1−2 kcal/mol. Shielding substituents keep both ions from being optimally solvated. (4) Molecular mechanics calculations revealed the most favorable interaction sites and energies for models of the protosteryl cations **4** and **5** with nucleophilic fragments of proteins. Approach to the carbenium carbon in **4** is sterically restricted, while **5** yields better bound complexes with the solvating molecule positioned on either face of the cationic center.

The results are consistent with the following ideas on sterol biosynthesis. It is anticipated that, when bound in the active site of lanosterol synthase, 2,3-oxidosqualene is in a preorganized state with a fold conducive to the incipient cyclization cascade. Under these circumstances, the tertiary cation → tertiary cation cyclization that forms ring B likely proceeds in

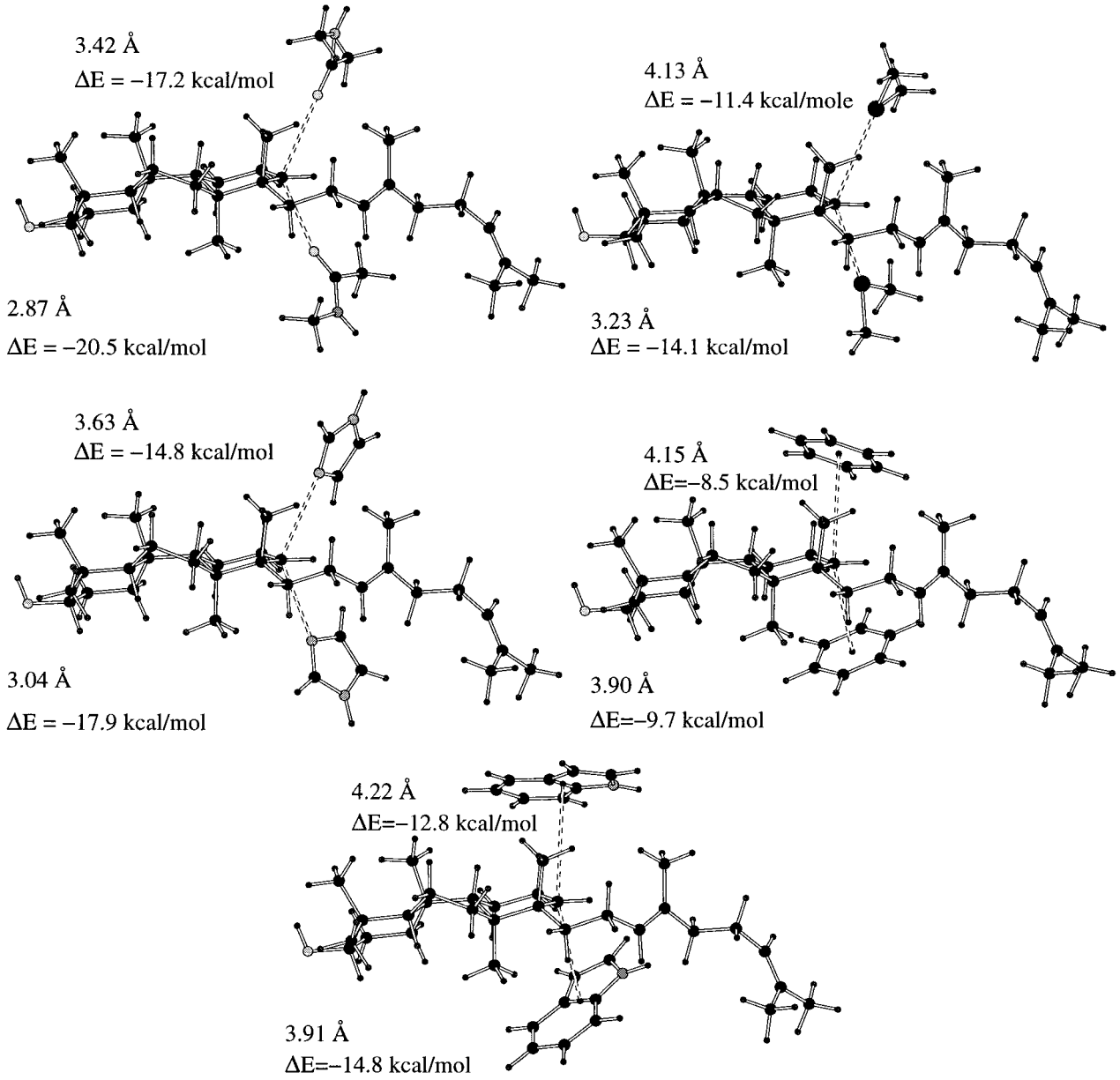


Figure 13. Optimized complexes for the protosteryl cation **5** and NMA, dimethyl sulfide, imidazole, benzene, and indole. Distances to the cationic center and complexation energies are given.

barrierless concert with the formation of ring A. The next addition yields **4**, which is the first intermediate with a significant lifetime. Though judicious placement of cation-stabilizing groups in the enzyme can do much to offset the intrinsic 12 kcal/mol higher energy for the secondary ion, there is evidence that **4** is a true intermediate.[2d] Stabilization of sites in the substrate that develop positive charge could occur through cation−π interactions with aryl groups of side chains or through Lewis acid−base interactions with lone-pair-containing heteroatoms. Ring C is completed upon overcoming the barrier to expansion to **5**. A particularly favorable position for placement of a cation-stabilizing group such as an aromatic side chain is on the α-face of the incipient D-ring. Ring D can be expected to form in barrierless concert with the ring expansion since generation of ring D should be exothermic by ca. −25 kcal/mol: −20 kcal/mol for the bond formation, −14 kal/mol for the secondary to tertiary carbenium ion difference, and roughly +9 kcal/mol for formation of the five-membered ring and interactions between *syn* substituents. While the present results have helped clarify important energetic and structural issues for carbenium ion chemistry with relevance to cyclization of 2,3-oxidosqualene, clearly, much detail on this fascinating enzymatic process remains to be discovered.

Acknowledgment. Gratitude is expressed to the National Science Foundation and to the National Institute of General Medical Sciences for support of this research. Prof. E. J. Corey is also thanked for discussions and preprints.

Supporting Information Available: RHF/6-31G* optimized structures in PDB format and nonbonded parameters for carbocations **9** and **10** (6 pages). See any current masthead for ordering and Internet access instructions.

Note Added in Proof. Crystal structures of three terpenoid cyclases including squalene → hopene cyclase have just been reported and further support the stabilization of carbocationic intermediates by cation−π and Lewis acid−base interactions as in Figures 12 and 13: Wendt, K. U.; Poralla, K.; Schulz, G. E. *Science* **1997**, *277*, 1811. Starks, C. M.; Back, K.; Chappell, J.; Noel, J. P. *Science* **1997**, *277*, 1815. Lesburg, C. A.; Zhai, G.; Cane, D. E.; Christianson, D. W. *Science* **1997**, *277*, 1820.

JA9714245