Pergamon

Tetrahedron Vol. 51, No. 2, pp. 449-472, 1995

0040-4020/95 $9.50+0.00

0040-4020(94)00908-2

Investigations into the Stereochemistry of Cyclophane-Steroid Complexes via Monte Carlo Simulations

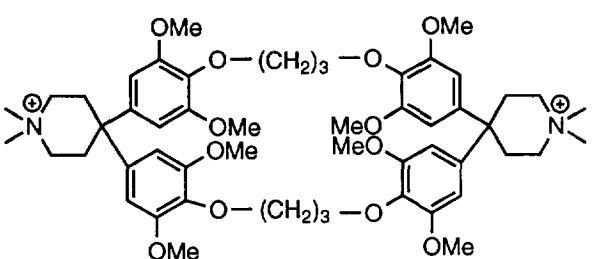
Heather A. Carlson and William L. Jorgensen*

Department of Chemistry, Yale University, New Haven, CT 06511-8118

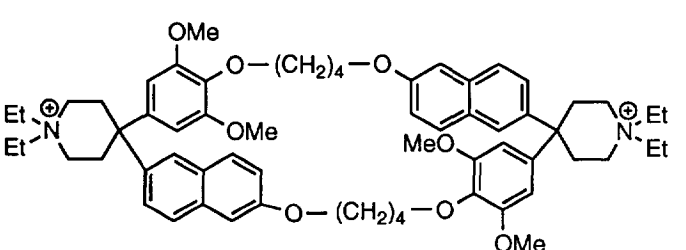
Abstract: *Monte Carlo simulations have been used to determine the $\Delta\Delta G_b$ for steroid-cyclophane complexes in four different conformations. Statistical perturbation theory provided changes of free energy in good agreement with experimental data. Hydrogen bonding and energy component analyses yielded insights into the interaction of the steroid with the cyclophane and with the mixed water/methanol solvent; this information should prove useful to steroid host design.*

INTRODUCTION

Molecular recognition is central to understanding biological activity and enzymatic systems. As models, many organic host-guest systems have been studied experimentally[1-5] and computationally.[6-10] The interest in cyclophanes and cyclodextrins as hosts is particularly strong because these systems are generally water soluble and yet contain hydrophobic pockets, so they share the enzyme characteristic of maintaining small controlled environments within aqueous solution.



1



2

Previously, we performed computer simulations[10] to help elucidate complex structures and the origins of variations in binding affinities for Diederich's octamethoxy tetraoxaparacyclophane, **1**, with substituted benzenes.[11] Carcanague and Diederich have also developed a larger cyclophane with a central cavity approximately 11 x 13 Å wide, **2**.[12] This cyclophane is of interest as it binds steroids including bile acids and corticoids. Efforts were also made to complex cholesterol; however, the complex was insoluble and a binding constant could not be obtained.

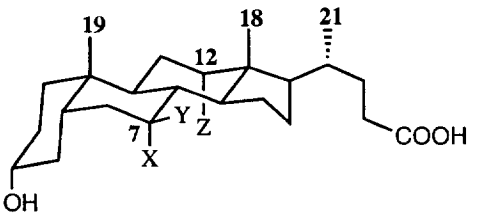
2 was found to bind cholic acid and four derivatives: deoxycholic acid, chenodeoxycholic acid, ursodeoxycholic acid, and lithocholic acid. The NMR data on the binding of these molecules are reproduced in Table 1.[12] As would be expected, binding strength increases with the hydrophobicity of the acids. An interesting note about the upfield shifts of the methyl groups is that cholic acid and deoxycholic acid have their strongest shifts at methyl 19, whereas the three more strongly binding acids have their weakest shifts at methyl 19 and their strongest shifts at methyl 18. This is consistent with a preference for cholic and deoxycholic acid to bind between rings B and C and the others to bind between rings C and D.[12]

No crystal structure for **2** is known; however, the crystal structure of **1** reveals point of inversion symmetry.[13] In fact, Diederich has developed more elaborate hosts, based on benzene crystal packing,

Table 1. Carcanague and Diederich's Association Constants, K_a, and Free Energies of Binding, $-\Delta G^\circ$, for the Steroid Complexes with **2** and the Upfield Complexation Shifts, $\Delta\delta$, Calculated for Saturation Binding, of the Proton NMR Resonances of the Steroid Methyl Groups in D_2O/CD_3OD (1:1 v/v) at T = 293 °K.[12,a]

Steroid Acid	X	Y	Z	K_a (L/mol)	$-\Delta G^\circ$ (kcal/mol)	$\Delta\delta$ of CH_3 C19	C18	C21
Cholic	OH	H	OH	145	2.90	0.73	0.56	0.25
Deoxycholic	H	H	OH	250	3.21	0.76	0.66	0.39
Chenodeoxycholic	OH	H	H	810	3.91	0.47	1.40	0.89
Ursodeoxycholic	H	OH	H	1750	4.35	0.30	1.49	1.23
Lithocholic	H	H	H	7075	5.18	0.55	1.49	0.90

[a] Accuracy of the K_a values is ± 10%. The solutions contained 0.01 M Na_2CO_3.



which all have a point of inversion.[14] This suggests that **2** may also have C_i symmetry; however, lacking a crystal structure, one should not overlook the possibility that the host can be C_2 symmetric with respect to a central axis perpendicular to the macrocycle plane.

Furthermore, because steroids are chiral, there is a difference in orienting the guests with the carboxylic acid side chain "up" versus "down" as in Figure 1. This results in 4 conformations for the host-guest complex which have been designated here as C_2 (c2), C_2 opposite orientation (c2o), point of inversion (i), and point of inversion opposite orientation (io). In solution, the forms may interconvert and the NMR results could correspond to an averaged structure. The present computational study was aimed at better understanding the structures and relative energetics for the stereoisomeric complexes along with reproducing the observed trends in binding affinities. The work is technically challenging in view of the size and flexibility of the systems. Mimicking the experimental mixed methanol/water solvent system also required significant additions to our modeling software.

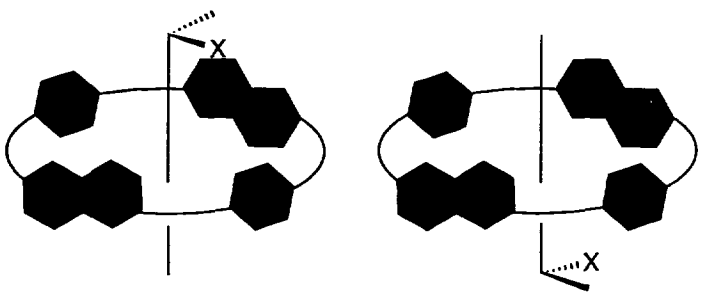


Fig. 1. Schematic display of the difference between two possible orientations of a chiral guest within the C_2 symmetric host.

COMPUTATIONAL PROCEDURE

All gas phase and solution calculations were conducted with the 3.4 version of the BOSS program[15] using the AMBER/OPLS force field to determine the potential energy of the system.[16-18] This program allows sampling of specified bond lengths, bond angles, and torsional angles of the solutes. The potential energy between two molecules, ΔE_{ab}, where a and b can be either solute or solvent, is determined by Coulomb and Lennard-Jones interactions between atoms *i* on a and atoms *j* on b, separated by a distance r_{ij} (eqn. 1). Geometric combining rules are applied for the Lennard-Jones parameters (eqns. 2).

$$\Delta E_{ab} = \sum_{ij} \{ q_i q_j e^2 / r_{ij} + 4\varepsilon_{ij} [(\sigma_{ij} / r_{ij})^{12} - (\sigma_{ij} / r_{ij})^6] \} \quad (1)$$

$$\varepsilon_{ij} = (\varepsilon_i \varepsilon_j)^{1/2} \qquad \sigma_{ij} = (\sigma_i \sigma_j)^{1/2} \quad (2)$$

Intramolecular nonbonded interactions are also treated by eqn. 1 for atoms separated by more than three bonds. The torsional energy, $V(\Phi)$, is given by a Fourier series for each dihedral angle, Φ, with coefficients obtained by fitting experimental data or reliable gas phase calculations (eqn. 3).

$$V(\Phi) = (V_1 / 2)[1 + \cos \Phi] + (V_2 / 2)[1 - \cos(2\Phi + \psi)] + (V_3 / 2)[1 + \cos 3\Phi] \quad (3)$$

All host and guest parameters used in this study appear in Table 2 including the AMBER atom types,[16] partial charges, Lennard-Jones parameters, and Fourier coefficients. A united-atom representation is only used for saturated CH_n groups; all other atoms are explicit. Harmonic bond stretching and angle bending force constants were obtained from the AMBER force field using the AMBER atom types.[16]

Studies began with determining possible structures for the host. The initial structure was obtained from an MM2 energy minimization using the Macromodel program;[19] further refinements were made using Fletcher-Powell (**FP**) gas-phase energy minimizations[20] in BOSS to find the lowest energy C_i and C_2 symmetric structures. By allowing no bond stretching, angle bending, or torsional variation within specific units, the host optimizations were restricted to impose rigid, planar aromatic units and the dimethyl piperidinium fragments found in the crystal structure of **1**.[13,21] The energy minimizations were performed while varying the torsional angles for the methoxy groups and the remaining bond lengths, angles, and

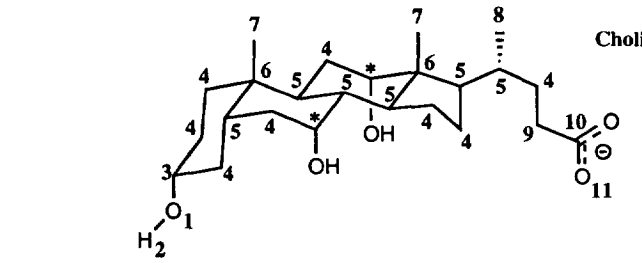
Table 2. OPLS Parameters for Cholic Acid and the Cyclophane[a]

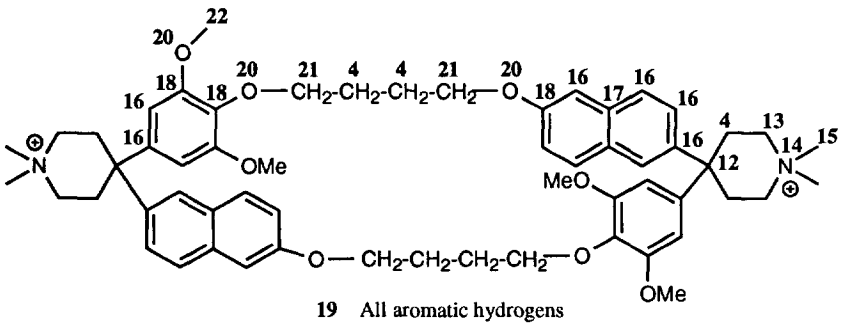
Type (OPLS)	Type (AMBER)	Nonbonded parameters		
		q, e	σ, Å	ε, kcal/mol
1	OH	-0.700	3.070	0.170
2	HO	0.435	0.000	0.000
3	CH	0.265	3.850	0.080
4	C2	0.000	3.905	0.118
5	CH	0.000	3.850	0.080
6	CT	0.000	3.800	0.050
7	C3	0.000	3.960	0.145
8	C3	0.000	3.910	0.160
9	C2	-0.100	3.905	0.118
10	C	0.700	3.750	0.105
11	O2	-0.800	2.960	0.210

12	CT	0.230	3.800	0.050
13	C2	0.250	3.905	0.118
14	N3	0.000	3.250	0.170
15	C3	0.250	3.960	0.145
16	CA	-0.115	3.550	0.070
17	CA	0.000	3.550	0.070
18	CA	0.135	3.550	0.070
19	HC	0.115	2.420	0.030
20	OS	-0.385	3.000	0.170
21	C2	0.250	3.800	0.118
22	C3	0.250	3.800	0.170

[a] Structures below indicate the OPLS atom-type designation. During a mutation of a hydroxyl group, the atoms noted with a * are linearly scaled from type 3 to type 4, while the oxygen and hydrogens are linearly scaled from type 1 and 2, respectively, to dummy atoms for which all parameters are 0.000.

Dihedral Angle (Φ) Type		Fourier Coefficients, kcal/mol V_1	V_2	V_3	ψ, rad
C-C-C-C	alkanes	1.411	-0.271	3.145	0
C-C-C=C	ethyl benzene	0.000	1.500	0.000	π
C=C-O-C	aryl ether	0.000	2.200	0.000	0
C-O-C-C	aryl ether	2.138	-1.060	2.814	0
C-C-O-H	alcohols	0.834	-0.116	0.747	0
C-C-C-C(=O)	carboxylate	2.290	-1.030	3.630	0
C-C-C-O(C)	ethers	0.650	-0.320	3.190	0
C-C-C(-O)2⁻	deprotonated acid	0.000	1.156	0.000	0





19 All aromatic hydrogens

torsional angles except within the aromatic rings and the piperidinium units.

Despite the conformational difference, the resultant energies are surprisingly close with the C_2 structure minimizing to -56.88 kcal/mol and the C_i form minimizing to -56.73 kcal/mol, as illustrated in Figure 2. The C_2 form has a saddle shape with the aromatic units pointing slightly out to form a bowl. However, the C_i host's macrocycle is rather flat, and its aromatic rings are parallel to the corresponding partner across the macrocycle. This parallel structure might be suitable for the observed binding of some smaller cyclophanes by **2** through π stacking.[12]

In the gas phase calculations for the steroid complexes, the acids were kept in their protonated form, otherwise deformation would occur to decrease the distance between the carboxylate group and the positively charged piperidinium units of the host. The structure for lithocholic acid was originally obtained with an MM2 energy minimization using the Macromodel program.[19] This structure was then used as the basis for the FP energy minimizations for all 5 bile acids. All cholic acid derivatives were completely flexible with all bonds, angles, and torsions varied in the minimizations. There was little difference in the structures of the steroids' cyclic frameworks, and a standard bile acid backbone was obtained by averaging all the acids' A, B, and C rings as well as the common oxygens and methyl groups in axial positions off those rings.

Gas phase FP energy minimizations of the host-guest complexes used this standard rigid A, B, and C ring backbone for the acids. The hydroxyl hydrogens, ring D, and the side chain atoms were allowed torsional freedom. For the cyclophane, the same bond angles and torsional angles were varied in the complex as in the host minimizations; however, no bond lengths were sampled except the implicitly varied macrocyclic ring closure bond. These constraints were consistent with those for the subsequent Monte Carlo (**MC**) simulations.

An exhaustive series of energy minimizations was conducted on the host-guest complexes. Starting positions for the guest were tried over approximately 9 Å of translation along an axis perpendicular to the plane of the macrocycle (the C_2 axis for the C_2 symmetric host) and over 360° of rotation about the same axis in 30° increments. This series yielded the minimum energy gas phase structures which were then used for the solution phase work.

The solutes were represented in Z-matrix form, and the MC sampling involved making changes, $\Delta\Theta$, and $\Delta\Phi$, to the designated internal coordinates. The same specifications from the gas phase complex studies were maintained in the solution simulations except that ring D was held rigid as its variations in the gas phase optimizations were negligible, and the carboxylic acid was deprotonated as one would expect under the experimental conditions.[12]

Statistical perturbation theory was applied to compute the relative free energies of binding ($\Delta\Delta G_b$) according to the cycle below where Host(guest) indicates a bound complex.

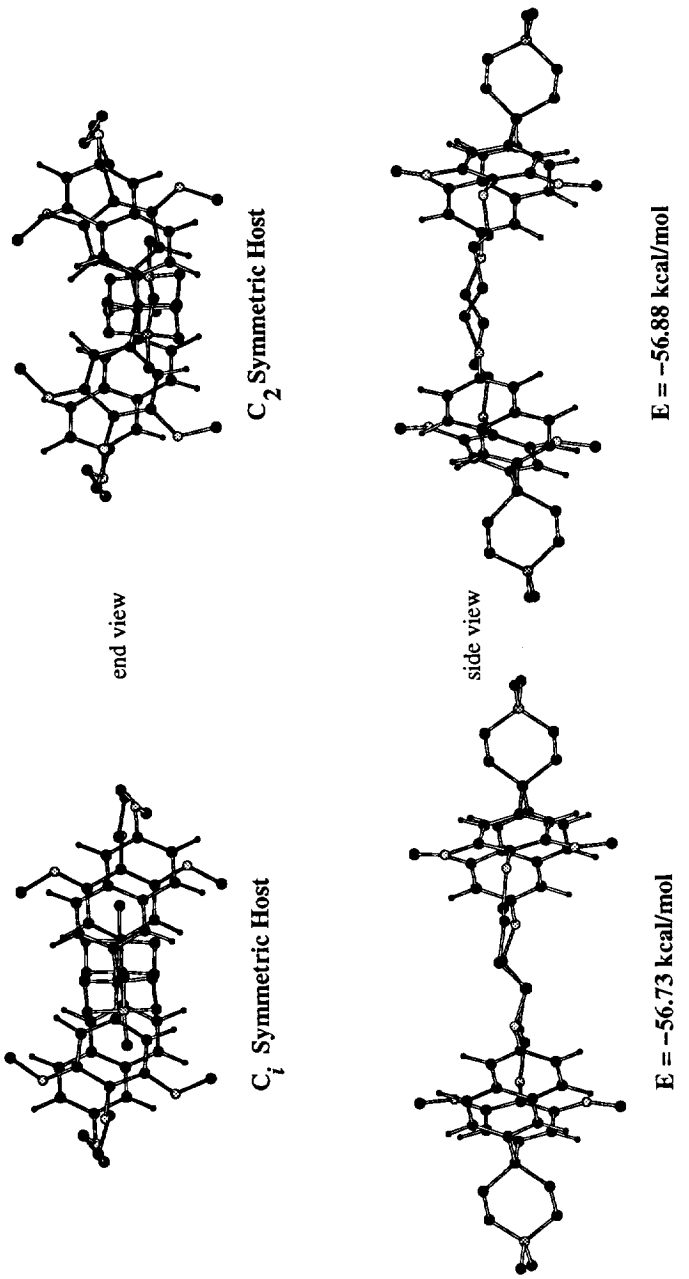
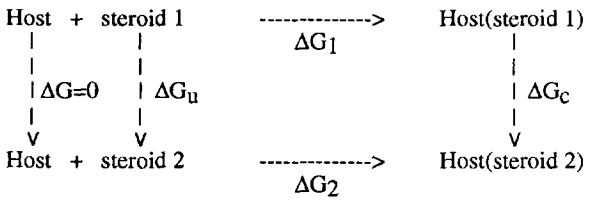


Fig. 2. Minimum energy structures for the cyclophane host.



$$\Delta\Delta G_b = \Delta G_c - \Delta G_u = \Delta G_2 - \Delta G_1$$

This cycle equates the experimental free energy change for binding, $\Delta G_2 - \Delta G_1$, with $\Delta G_c - \Delta G_u$. Rather than calculating a potential of mean force to obtain ΔG_1 or ΔG_2, ΔG_c and ΔG_u are computationally more accessible and require interconversions of the solvated guests bound and unbound to the cyclophane, respectively.[6] The perturbations are performed in small increments; the free energy for the step from i to j is computed using eqn. 4, where the average corresponds to sampling configurations for the reference state i, k_B is the Boltzmann constant, T is the temperature, and E_i and E_j are the total energies for the system in states i and j, respectively.[22]

$$\Delta G_{ij} = G_j - G_i = -k_B T \ln \langle \exp[-(E_j - E_i) / k_B T] \rangle_i \quad (4)$$

Since the only differences in states i and j are within the guest, all non-guest energetic terms of $(E_j - E_i)$, like solvent-solvent and host-solvent energies, cancel leaving only the contributions of intraguest, host-guest, and solvent-guest energies. The total change in free energy for the mutation is the sum of all ΔG_{ij} for each window.

The mutation of cholic acid to lithocholic acid was chosen for study and was conducted in two steps: cholic acid to chenodeoxycholic acid and chenodeoxycholic acid to lithocholic acid. Both mutations remove a single hydroxyl group, which required the charge and Lennard-Jones parameters of the oxygen and hydrogen atoms to be linearly scaled to zero, and at the same time, the attached carbon was linearly scaled from a polarized CH (atom type 3) to a nonpolar CH_2 (atom type 4). Simultaneously, the OH bond length was linearly contracted to 0.1 Å and the CO bond length was reduced to 0.3 Å.

The solution phase MC calculations utilized deprotonated cholic acid, chenodeoxycholic acid, and lithocholic acid in a mixed solvent. A 31 x 34 x 39 Å cell with periodic boundary conditions solvated the complexes with 680 TIP4P water[23] and 294 united atom methanol[24] molecules which corresponds to Carcanague and Diederich's 1:1 (v/v) mixture of D_2O/ CD_3OD.[12] A smaller cell of 26 x 23 x 35 Å with periodic boundary conditions solvated the uncomplexed cholic acid derivatives with 332 TIP4P water and 148 CH_3OH molecules. These cells provided at least 9Å of solvent around all solute atoms. The solvent-solvent and solute-solvent interactions were spherically truncated at a distance of 8.5 Å with quadratic

feathering to zero over the last 0.5 Å. All simulations were conducted at constant temperature, 298 K, and constant pressure, 1 atmosphere. It may be noted that this is the first application of the mixed solvent capabilities, which were implemented in BOSS, version 3.4.[15]

Metropolis and preferential sampling were used, and attempted moves for the solutes were made every 30 configurations. Though internal sampling was performed for the cyclophane, it was not translated or rotated as its size yielded little acceptance even for very small movements. Overall, moves of the cyclophane had ~30% acceptance, and the ranges for the motions of the acids were adjusted to give 15-25% acceptance. For the calculations without hosts, the acids had the same ranges for sampling internal coordinates and translations as were specified for the complexes, but rigid body rotations were not performed to increase the acceptance rates to ca. 26%. Mutations, with and without the host, were performed in 7-8 simulations with double-wide sampling to yield 14-16 incremental free energy changes for each hydroxyl group removal.

The mutations of the unbound steroids were carried out in the smaller cell as mentioned above. This cell was created by forming a TIP4P water box around cholic acid, substituting the appropriate number of methanol molecules for TIP4P waters, and equilibrating. The equilibrations covered 1M iterations of NVT MC and 2M iterations of NPT MC without sampling the solute and then a final 2M iterations of NPT MC allowing the solute to be sampled. All mutations of the unbound steroids were straightforward after this set-up of the solvent system. ΔG_{ij} in each subsequent window was obtained through 1M configurations of equilibration and 2M configurations of averaging. The solvent cell for the complex was obtained in the same fashion using the c2 conformer of cholic acid. 1M iterations of NVT and 3M iterations of NPT MC without solute sampling and then 3M NPT configurations with solute sampling were executed to equilibrate the solvent for the c2 complex. In order to use this solvent box for the other complexes, the alternate solute conformer was substituted for the complex in the original solvated system and additional equilibrations were conducted: 1M NVT and 1M NPT configurations without solute sampling and then 2M NPT configurations with solute moves. This method was used for the equilibration of the first simulation of the cholic acid to chenodeoxycholic acid mutation in each conformer. Subsequent windows started from the last configuration of the previous window and required only an additional 1M NPT configurations with solute sampling to equilibrate. All simulations of the complexes were averaged for 4M configurations in order to obtain the ΔG_{ij} in each window.

THERMODYNAMIC RESULTS

The energy minimizations for the complexes were conducted for all five cholic acid derivatives in all four conformers. The goal was to possibly exclude some of the conformers based on poor energetics or improper binding patterns. Table 3 shows that no conclusion could be reached from the gas phase energies alone since the different minima for each complex have similar energies. For illustration, the four

structures for the complex with cholic acid are shown in Figure 3. Complexes of cholic acid with the enantiomer of the C_2 symmetric host were also considered; however, they yielded energies 4-5 kcal/mol higher and were not pursued.

Thermodynamic results from the solution phase simulations appear in Table 4. Standard deviations ($\pm$ 1σ) were calculated from the fluctuations in the separate averages over each window. The largest uncertainty occurs in the mutations without the host. However, the results appear accurate as a second calculation for the C7 hydroxyl mutation (cheno to litho) with modified torsional sampling of the hydrogen resulted in a change of free energy of 12.62 $\pm$ 0.12 kcal/mol. The two results are within 1σ error of each other despite the fact that this second calculation had almost twice the number of windows in the simulation. It should be noted that in the solution-phase calculations, the conformers for the complexes did not interconvert, but stayed relatively close to their initial geometries.

The changes in free energy for the two mutations without the host are different. Inspection of the energetic contributions and hydrogen bonding of the systems reveals the origin of the 1.7 kcal/mol difference. It is interesting to note the contrast in the energetic contributions over the simulations in Figure 4. In cholic acid, the C7αOH and C12αOH hydroxyl oxygens are 4.4 Å apart, which allows favorable interactions between one's hydroxyl hydrogen and the other's oxygen as the hydrogen torsions are sampled. This interaction is apparent from the differences between the two mutations non-bonded energy contributions. The C12 hydroxyl removal (cholic to chenodeoxycholic acid) has a non-bonded energy contribution over 6 kcal/mol higher than the C7 hydroxyl removal. Simulations of the mutations in the gas phase, where the only energetic contribution is the intramolecular non-bonded energy, support this observation from the solution phase simulations. In the gas phase, the change in free energy for the protonated cholic to chenodeoxycholic acid mutation is 12.31 $\pm$ 0.02 kcal/mol, whereas the protonated chenodeoxycholic to lithocholic acid mutation results in a free energy change of 6.61 $\pm$ 0.02 kcal/mol.

Table 3. Results of the Energy Minimizations(kcal/mol) of Each Steroid in the Four Conformations[a]

Steroid	i	io	c2	c2o
Cholic Acid	**-63.67**	**-64.64**	**-65.60**[b]	**-65.68**
Deoxycholic Acid	-56.27	-58.20	**-57.48**	**-58.39**
Chenodeoxycholic Acid	**-59.43**	**-58.14**	-59.66	**-58.32**
Ursodeoxycholic Acid	**-59.47**	-59.33	-58.52	-59.64
Lithocholic Acid	**-55.34**	**-50.87**	**-50.28**	-50.33

[a] Numbers in bold face indicate structures which exhibited a similar binding pattern as those interpreted from NMR data: cholic and deoxycholic acids bound between C19 and C18 and the other three between C18 and C21.

[b] A second c2 cholic acid conformer (c2A) where the C7αOH is more solvent exposed minimized to **-61.55** kcal/mol.

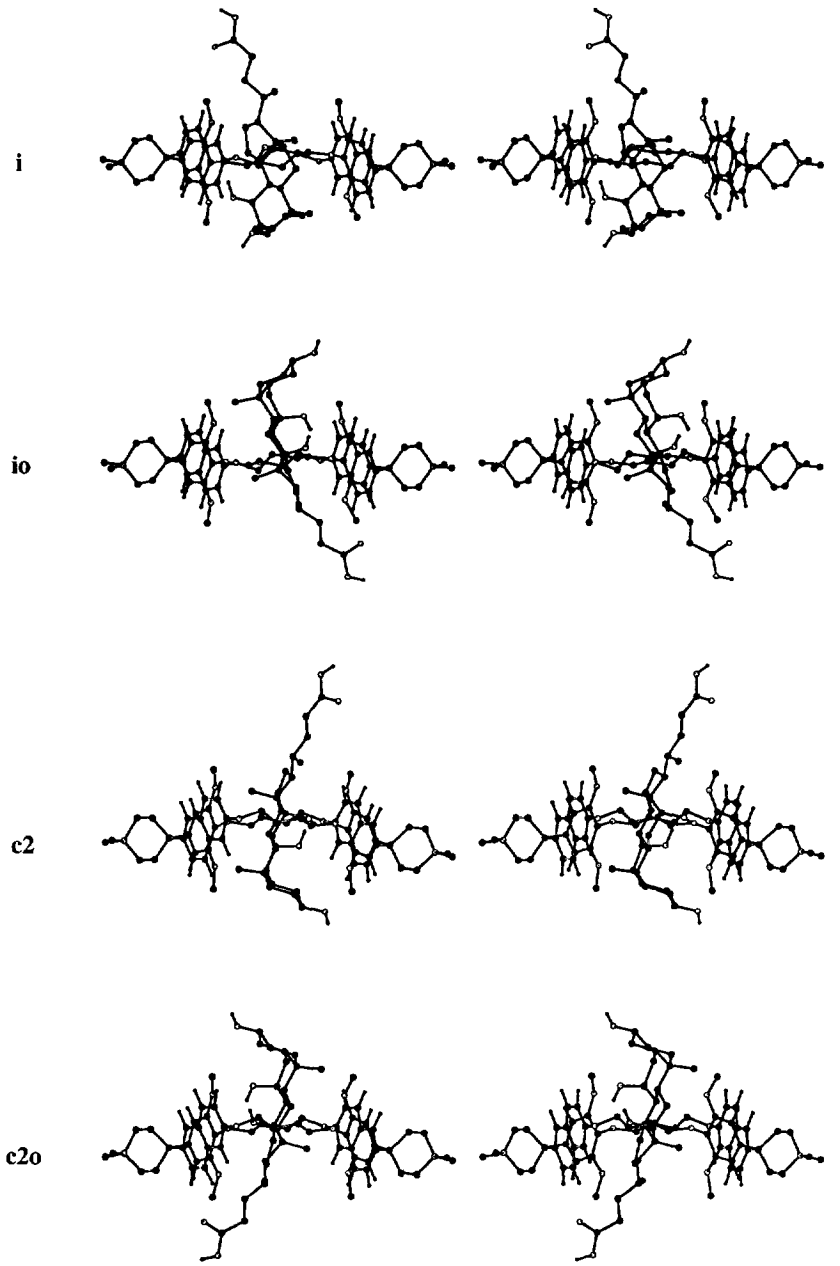


Fig. 3. Stereo views of the four conformers for the complex of protonated cholic acid obtained by gas–phase energy minimizations.

Table 4. Calculated and Experimental Free Energy Changes, in kcal/mol

Perturbation	Conformer	ΔG_c	ΔG_u	$\Delta\Delta G_b$[a]	Experimental[b]
Cholic to Cheno	i	13.20 ± 0.06	14.58 ± 0.13	-1.38 ± 0.14	-1.01 ± 0.08
	io	13.37 ± 0.04		-1.21 ± 0.14	
	c2	13.65 ± 0.05		-0.93 ± 0.14	
	c2o	13.50 ± 0.05		-1.08 ± 0.14	
	c2A[c]	12.03 ± 0.06		-2.55 ± 0.14	
Cheno to Litho	i	10.58 ± 0.09	12.86 ± 0.16	-2.28 ± 0.18	-1.27 ± 0.08
	io	11.75 ± 0.12		-1.11 ± 0.20	
	c2	8.32 ± 0.07		-4.54 ± 0.17	
	c2o	8.72 ± 0.06		-4.14 ± 0.17	
	c2A[c]	10.33 ± 0.11		-2.53 ± 0.19	
Cholic to Litho	i			-3.66 ± 0.23	-2.28 ± 0.08
	io			-2.32 ± 0.24	
	c2			-5.47 ± 0.22	
	c2o			-5.22 ± 0.22	
	c2A[c]			-5.08 ± 0.24	

[a] 298 °K.
[b] 293 °K.[12]
[c] Second c2 conformer where the C7αOH is more solvent exposed.

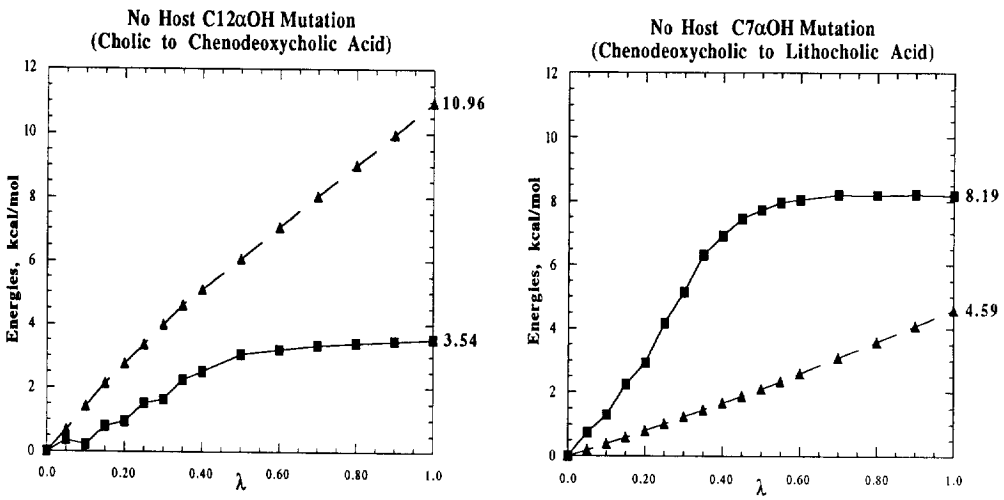


Fig. 4. Energetic contributions to ΔG_{ij} over the mutations calculated in the absence of the host. Solid squares (■) represent the steroid-solvent energy changes and the solid triangles (▲) represent the intraguest non-bonded energy changes.

It was observed during the simulations that the two closely spaced hydroxyl groups on C7 and C12 prevented complete solvation of both hydroxyl hydrogens. The definition of a hydrogen bond between the solute and a solvent molecule requires an O---H distance less than 3.0 Å and an attractive interaction of least -2.25 kcal/mol; results of hydrogen bonding analyses for the three acids to the solvent appear in Table 5. The non-integral numbers of hydrogen bonds arise from averaging the results over configurations saved every ca. 25,000 steps during the simulations. An isolated alcohol molecule in water normally forms two to three hydrogen bonds, while alcohols in neat liquids form two hydrogen bonds.[18,24] The reduced hydrogen bonding in the C12αOH mutation contributes to the 4.65 kcal/mol smaller loss in the steroid-solvent energy contributions in Figure 4. Hydrogen bonding analyses reveal that the increase in steroid-solvent energy contribution for the C7αOH mutation is not only from this hydroxyl group forming slightly more hydrogen bonds once the C12αOH is mutated (Table 5), but also arises from stronger hydrogen bonding to the solvent once that secondary interaction is removed.

It should be pointed out that these trends are not inherent to the C12 and C7 hydroxyls, but rather are inherent to which hydroxyl is mutated first. In other words, the first hydroxyl mutated exhibits larger non-bonded energy contributions. If the C7αOH had been mutated first, the trends of the first mutation should be qualitatively the same even though it is not the same hydroxyl group. For example, the subsequent C12αOH removal would show stronger solute-solvent interactions during the second mutation. This needs to be clarified before discussing related trends in the complexes which *are* inherent due to the C7 and C12 hydroxyls being in different environments.

When the steroids bind in the cyclophane, the C12αOH is well within the hydrophobic macrocycle, regardless of the conformer, and is never solvent exposed. However, the C7αOH is partially solvent exposed and forms one hydrogen bond with the solvent in each conformer. In Figure 5, the cumulative free energy changes for each conformer are displayed over the course of the two mutations, where the hydroxyl group is removed as λ goes from 0.0 to 1.0. From the graphs it is apparent that the results for the solvent exposed hydroxyl group are much more sensitive to conformation. Whereas the smaller

Table 5. Numbers of Hydrogen Bonds Between the Steroids and the Mixed Solvent During Simulations Without the Cyclophane Host

Steroid	C3αOH		C7αOH		C12αOH	
	Oxygen	Hydrogen	Oxygen	Hydrogen	Oxygen	Hydrogen
First Mutation						
Cholic to	1.25	0.86	1.18	0.96	1.01	0.59
Chenodeoxycholic	1.15	1.03	1.34	1.30		
Second Mutation						
Chenodeoxycholic	0.96	1.09	1.06	1.23		
to Lithocholic	1.08	1.03				

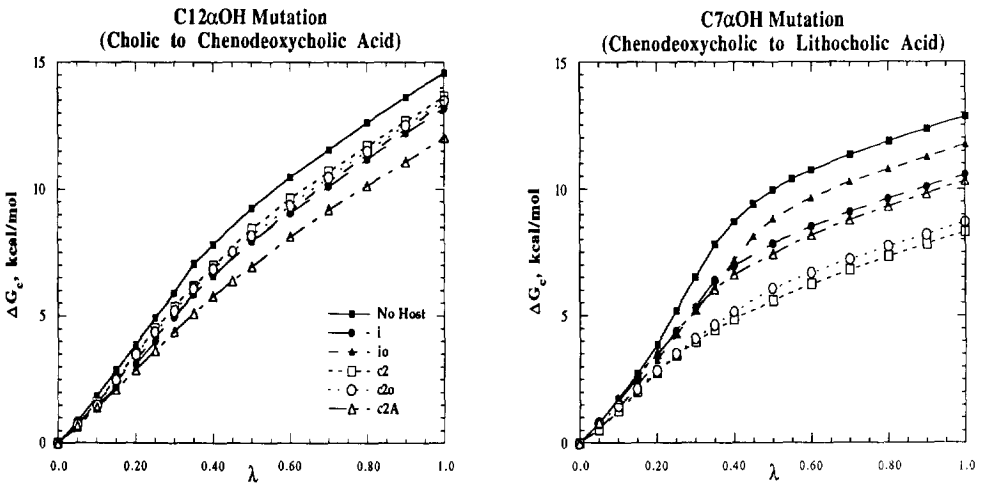


Fig. 5. Change in free energy, ΔG_C, over the course of the mutations for the complexed steroids. The results for the uncomplexed steroids, ΔG_U, are also shown for comparison.

differences between the conformers in the C12αOH mutations arise primarily from the non-bonded contributions (see Figure 6), the differences for the C7αOH case are dominated by the steroid-solvent energy contributions (see Figure 7).

It should be noted that there is relatively little variance in the contributions from the host-guest energies to the free energy changes. Nevertheless, the choice of host conformer does affect how the steroid interacts with the solvent and itself. The hydrogen bonding for the complexes was also analyzed, as summarized in Table 6. The solvent-solute hydrogen bonding definitions are the same as previously stated, but the definition of intersolute hydrogen bonding is based solely on a cutoff distance of 3.0 Å between appropriate atoms of opposite partial charge. This provides only a rough estimate for the possible hydrogen bonding, but there is no clear-cut way of partitioning the interaction energy between the two solutes to determine an energy cutoff. The possible host-guest interactions considered as hydrogen bonding involve the hydroxyl and carboxylate groups of the guest with the ether oxygens or aromatic carbons and hydrogens. Whether a hydroxyl hydrogen is within 3 Å of one aromatic carbon or four, it is counted as one hydrogen bond between the guest and the host's π system, but obviously, these two hydrogen bonding situations do not provide the same energetic change as the hydroxyl group is removed. Therefore, it is possible that the same number of intersolute hydrogen bonds could result in different energetic contributions. The results show little intersolute hydrogen bonding; the host-guest interactions involve primarily the Lennard-Jones terms.

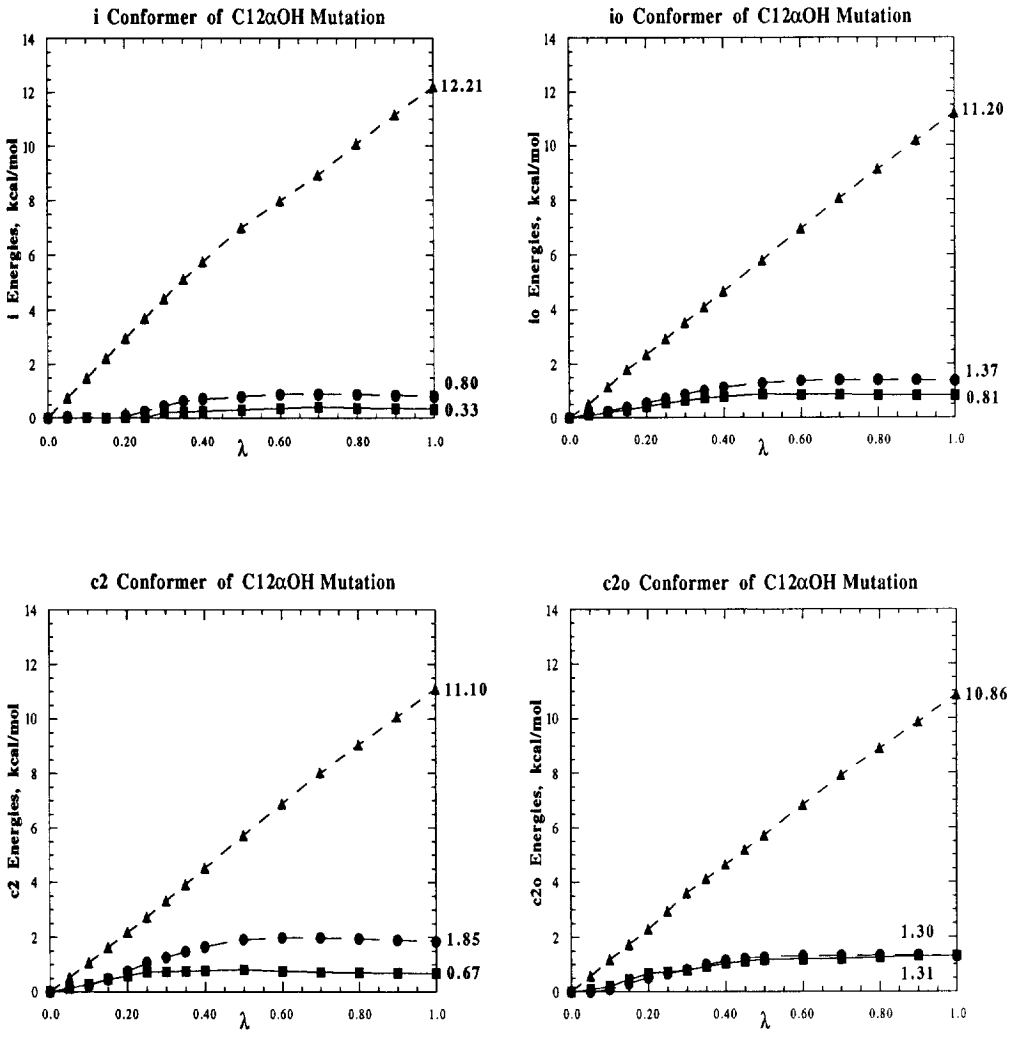


Fig. 6. Energetic contributions to ΔG_{ij} over the C12αOH mutations calculated for the guest complexed in the various conformations. Solid squares (■) represent the steroid-solvent energy changes, solid triangles (▲) represent the intraguest non-bonded energy changes, and solid circles (●) represent the host-guest energy changes.

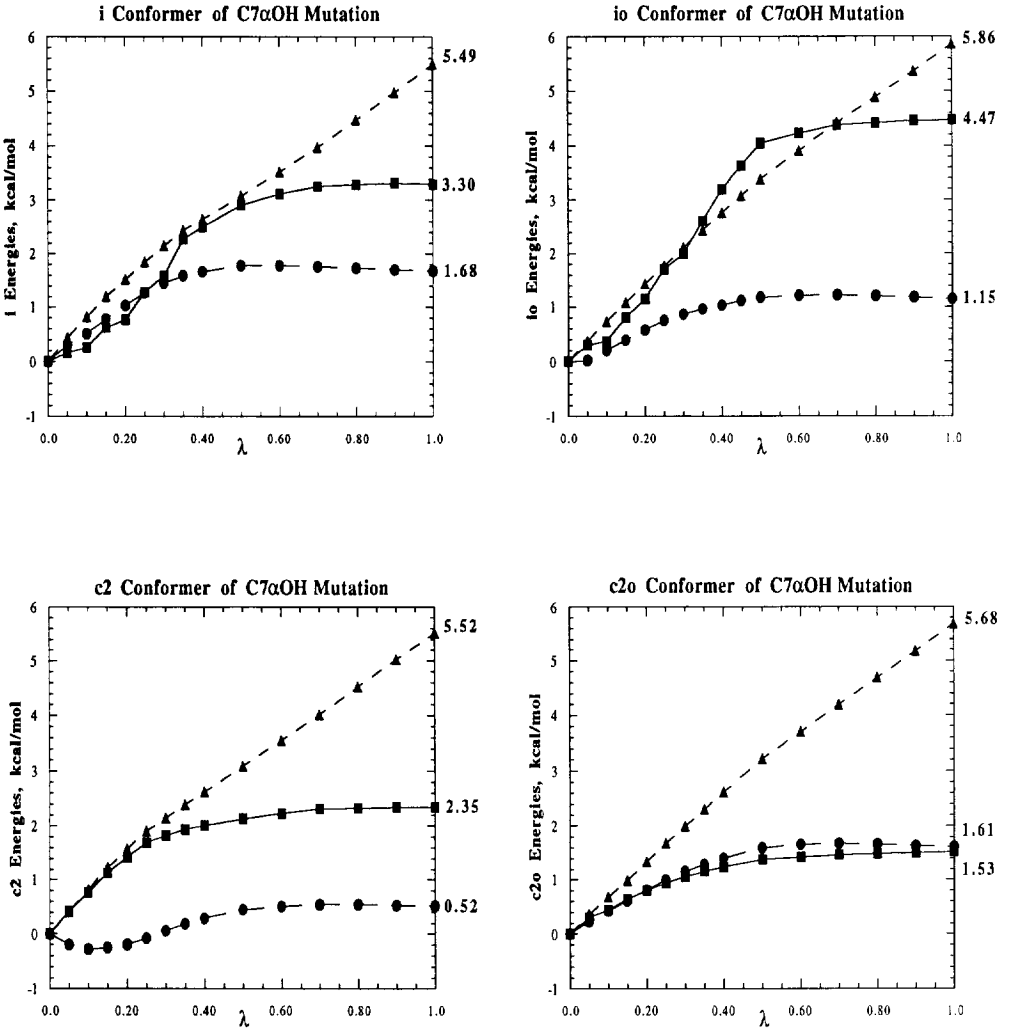


Fig. 7. Energetic contributions to ΔG_{ij} over the C7αOH mutations calculated for the guest complexed in the various conformations. Solid squares (■) represent the steroid-solvent energy changes, solid triangles (▲) represent the intraguest non-bonded energy changes, and solid circles (●) represent the host-guest energy changes.

Table 6. Numbers of Hydrogen Bonds Between the Steroids and the Host[a] and Mixed Solvent

Steroid Conformer	Carboxylate Oxygens		C3αOH Oxygen		C3αOH Hydrogen		C7αOH Oxygen		C7αOH Hydrogen		C12αOH Oxygen		C12αOH Hydrogen	
	Solvent	Host	Solvent	Host	Solvent	Host	Solvent	Host	Solvent	Host	Solvent	Host	Solvent	Host
Cholic Acid														
No Host	7.66		1.25		0.86		1.18		0.96		1.01		0.59	
i	6.99	0	1.03	0.75	0.91	0	1.00	0	0	0.96	0	0	0	0.17
io	5.49	0	0.86	0	1.06	0	0	0	0	0.91	0	0	0	0.24[a]
c2	7.55	0	1.04	0.09	0	0.36	0	0	0	0	0	0	0	0.99
c2o	5.64	0.51	1.04	0	1.03	0	0.94	0	0	0.56	0	0	0	0.54
c2A[b]	7.22	0	0.99	0	0.96	0	0	0	0.94	0	0	0	0	0
Chenodeoxycholic Acid (End of C12αOH Mutation)														
No Host	7.04		1.15		1.03		1.34		1.30					
i	6.81	0	1.11	0.68	1.01	0	1.00	0	0	0.92				
io	6.28	0	1.11	0	1.08	0	0.16	0	0	0.91				
c2	6.36	0	1.03	0.18	1.26	0	0.98	0	0	0				
c2o	6.34	0	1.04	0	0.98	0	0.64	0	0	0.69				
c2A[b]	7.45	0	0.94	0	0.98	0	0	0	0.98	0				
Chenodeoxycholic Acid (Start of C7αOH Mutation)														
No Host	7.45		0.96		1.09		1.06		1.23					
i	7.10	0	1.16	0.75	1.06	0	1.00	0	0	0.64				
io	5.74	0	1.18	0	1.03	0	0.73	0	0	0.98				
c2	7.16	0	1.02	0.46	0.79	0	0.98	0	0	0				
c2o	6.67	0	2.08	0	0.93	0	0	0	0	0.61				
c2A[b]	6.95	0	0.98	0	0.99	0	0	0	0.95	0				
Lithocholic Acid														
No Host	7.19		1.08		1.03									
i	6.47	0	1.05	0.47	1.03	0								
io	6.19	0	1.59	0	1.01	0								
c2	6.98	0	0.96	0.71	1.00	0								
c2o	6.30	0	1.18	0	1.00	0								
c2A[b]	6.94	0	0.97	0.44	0.98	0								

[a] Hydrogen bonding of the carboxylate and hydroxyl oxygens were to the hydrogens of the host's aromatic units. The hydroxyl hydrogens interacted with the host π systems in all but the one noted case where the hydrogen bond was to one of the methoxy oxygens.

[b] Denotes the second c2 conformer where the C7αOH is more solvent exposed.

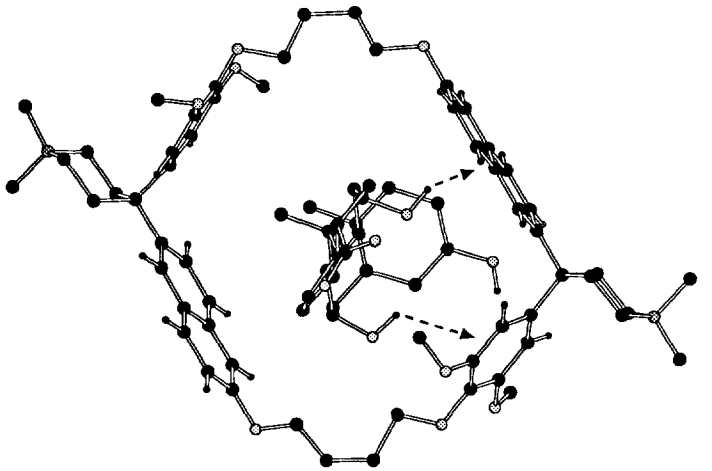
DISCUSSION

There are several characteristics of binding apparent in the energy breakdowns and hydrogen bonding data. The C12αOH is deep within the macrocycle and exhibits no hydrogen bonding to solvent; therefore, it is reasonable that the smallest contribution to its mutation is the guest-solvent energy (Figure 6). This mutated hydroxyl's primary interactions are with the steroid itself, and the secondary interaction is with the cyclophane. The host-guest energetic contributions are roughly correlated to the limited hydrogen bonding between the C12 hydroxyl hydrogen and the π systems of the host (Figure 6 and Table 6). Though there are variations in the individual energy components for the four conformers, they are not large and tend to compensate to make the ΔG_C's fall in a narrow range (Table 4).

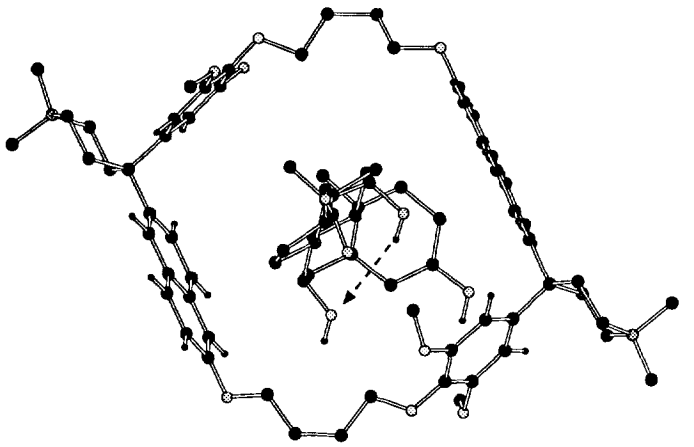
However, the behavior in the C7αOH mutation is different. The nonbonded energetic contribution is about half the value for the C12αOH mutation, which is also the case in the solution phase calculations of the guest alone (Figure 4). The variations in steroid-solvent interactions control the binding variability (Figure 7). The nonbonded energy change is also practically the same for each conformer since the C12 hydroxyl is already removed. Although the differences in the ΔG_C are directly related to the guest-solvent energy contributions, there is no relationship observed between the degree of hydrogen bonding to the solvent and the free energy changes (Table 4). In this case, the different host conformers are indirectly influencing the steroid-solvent energetic contributions in more ways than just limiting the hydrogen bonding to the solvent. The differences in the steroid-solvent interactions are assumed to be in part due to long distance interactions with the solvent which has different spatial relationships with the steroids in each conformer.

In order to study this solvent effect, a second conformation of c2 (c2A) was obtained by minimizing a structure of the steroid twisted 30° in the cavity so that the C7αOH was unable to hydrogen bond to a π system (Figure 8). The energetic contributions for the perturbations of this conformer are in Figure 9. Again the nonbonded contributions in the C7αOH mutation are about half those of the C12αOH mutation. For the removal of the C12αOH, the $\Delta\Delta G_b$ (Table 4) is too favorable suggesting that this hydroxyl group does not have enough interaction with the host; there is no hydrogen bonding between the C12αOH and the host in this conformer. Indeed, the C12 hydroxyl hydrogen is observed in the simulations to interact continually with the C7 hydroxyl oxygen in a pseudo-hydrogen bond. From these observations, it appears that the C12αOH should be hydrogen bonded to the macrocycle of the host in solution in order to obtain the experimental free energy changes.[12]

The main purpose of the c2A conformer was to inspect the solvent's effect on ΔG_C when the C7 hydroxyl group is more solvent exposed. The 30° twist oriented the hydroxyl group away from the π systems. Though the hydroxyl group still formed only one hydrogen bond to the solvent, the secondary interactions were increased as the hydroxyl group had less interference from the cyclophane. This increased the steroid-solvent energetic contribution by about 1.5 kcal/mol over the other c2 conformer



c2 Conformer



Second c2 Conformer with More Solvent Exposed C7 Hydroxyl Group

Fig. 8. The two c2 conformations where dashed arrows show the primary interactions of the hydroxyl groups within the complexes.

(2.35 versus 3.78 kcal/mol in Figures 7 and 9, respectively). It may be a general conclusion that a large part of the difference between the conformers' steroid-solvent energetic contributions to ΔG_C arise from longer-range solvent interactions. It also seems likely that the poor $\Delta\Delta G_b$ results for the c2 and c2o conformers in the C7αOH mutations are due to the fact that the C7αOH must not be solvent exposed enough, which would account for a too favorable $\Delta\Delta G_b$.

Except for the second c2 conformer, the $\Delta\Delta G_b$'s for the cholic to chenodeoxycholic acid mutations are in good agreement with the experimental values, with the c2o and the first c2 conformers providing the closest results. Considering the sizes of ΔG_C and ΔG_u, it is notable that the $\Delta\Delta G_b$ results are within the experimental uncertainty reported by Carcanague and Diederich. Even the results for the i and io conformers are well within an acceptable range for such calculations. The uniformity of these results does not rule out the representation of any of these host-guest conformations in solution. In the case of the C7αOH mutations, only the io conformer was able to reproduce the experimental $\Delta\Delta G_b$ within calculated error. The i and c2A conformers provided $\Delta\Delta G_b$'s about 1 kcal/mol too favorable, which may be due to the fact that the hydroxyl group is underexposed to solvent. The significantly poorer $\Delta\Delta G_b$ for the c2o and first c2 conformations suggests that these structures may not be well populated in solution.

Carcanague and Diederich's NMR data show a shift in binding patterns. However, a substantial shift was not observed over the course of the calculations. This may be problematic for the C7αOH mutation, where chenodeoxycholic and lithocholic acid are expected to have a different binding pattern

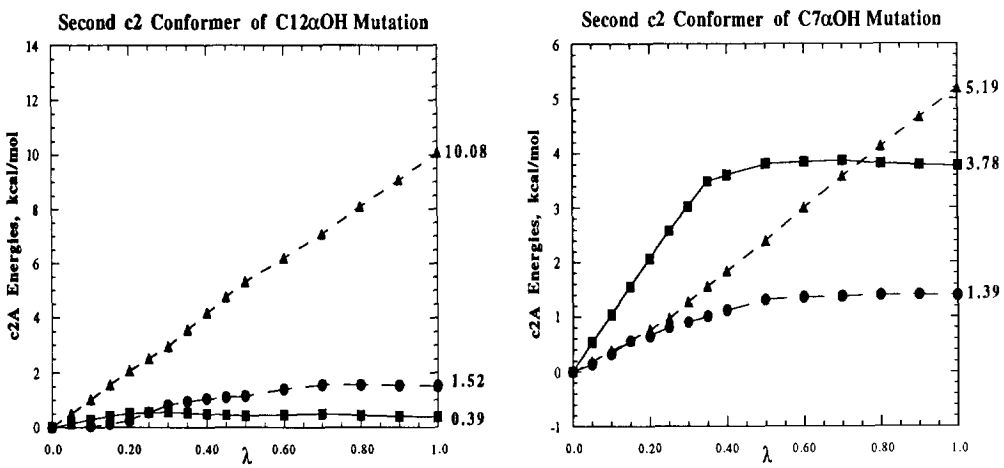


Fig. 9. Energetic contributions to ΔG_{ij} over the two hydroxyl mutations calculated for the guest complexed in the c2A conformation; symbols are the same as in Figures 6 and 7.

than cholic acid. It is possible that even with 4M configurations of averaging, there was inadequate opportunity for the steroid to shift completely. The relatively small errors for the i and second c2 conformers may reflect more complete shifts of the steroids. However, the approximately 3 kcal/mol discrepancy for the c2o and first c2 conformer could imply either a large shift should have occurred in these cases to expose the C7 hydroxyl more or these conformers are inappropriate representations of the cyclophane-steroid complexes in solution.

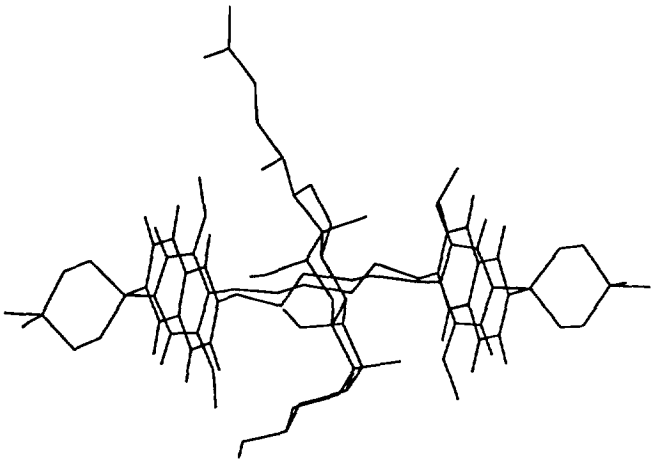
The io conformer for lithocholic acid shows the greatest similarity to its gas phase structure (Table 7). Possibly, this is the proper solution phase conformation since it has given the free energy change in best agreement with the experimental value. Furthermore, the RMS fitting between the gas phase and solution phase structures shows that the cholic acid structures exhibit about the same deformation upon solvation as they do after two mutations.

The behavior of the io conformer over the course of the two mutations suggests that the expected shifts are more subtle than just a translation of the guest. Observations of the complexes over the course of the simulations show that the motion of the guest is primarily roll and pitch with respect to the cyclophane macrocycle instead of translation and yaw. It is possible that the shifts observed by Carcanague and Diederich are due to the bound steroid pitching instead of purely translating (Figure 10). Translation from the lithocholic acid complex in Figure 10 to give binding between the B and C rings for the cholic acid complexes appears impossible owing to development of a steric clash between the A ring and cyclophane. The representations of the steroid in Figure 10 are based on the common configurational sampling of the steroids in the simulations and the fact that three of the original four conformers minimized to orient their hydroxyl groups toward the benzene and naphthalene rings bridged by a piperidinium unit. As can be seen from the figure, the guest can rotate in the host and does not have to shift translationally through the macrocycle to give the change in binding pattern.

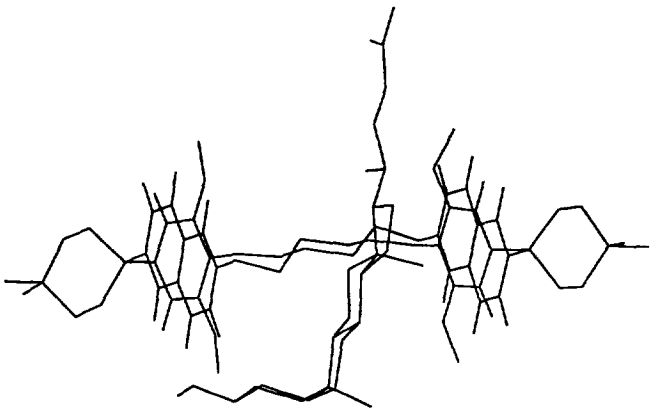
Table 7. RMS Fitting of the Common Atoms of Cholic and Lithocholic Acid Conformers Between Conformers and Phases, in Å

Conformer	Solvated Cholic Acid vs. Gas Phase Minimum Structure	Solvated Cholic Acid vs. Solvated Lithocholic Acid	Solvated Lithocholic Acid vs. Gas Phase Minimum Structure
i	0.49	0.49	2.26
io	0.44	0.48	0.41
c2	0.64	0.57	2.28
c2o	0.43	0.60	1.82
c2A[a]	0.37	0.39	2.15

[a] Second c2 conformer where the C7αOH is more solvent exposed.



Cholic Acid Binding Pattern



Lithocholic Acid Binding Pattern

Fig. 10. Example of the shift that can account for the change in binding pattern for the steroid guests.

For further studies involving steroid derivatives and this cyclophane host, the io orientation of the complex appears particularly promising for obtaining results to compare to experimental data. Studies could be expanded to calculate the $\Delta\Delta G_b$ for the other two cholic acid derivatives, and it would be interesting to conduct mutations between lithocholic acid and cholesterol to obtain the change in free energy of binding, which was not measurable in Carcanague and Diederich's study.

CONCLUSION

Excellent agreement with experiment for the C12αOH mutation was obtained for all of the conformers except the second c2 complex. In addition, because of the good performance of the io conformer in the C7αOH mutation, further studies for this host and steroids might best begin with this conformer. The energetic breakdowns of the ΔG_{ij} over the course of the calculations have proved to be a useful tool in sorting out the contributions to the variations in binding. This method and the hydrogen bonding analyses have also been useful for determining the differences in the environments for the two hydroxyl group mutations in the presence and absence of the host.[25]

REFERENCES AND NOTES

1. Diederich, F. *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 362-386.
2. Rebek, J., Jr. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 245-255.
3. Schneider, H.-J. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1417-1436.
4. *Supramolecular Chemistry*; Balzani, V.; De Cola, L. Eds.; NATO ASI Ser. C, *371*; Kluwer, Dodrecht: The Netherlands, **1992**.
5. Kearney, P. C.; Mizoue, L. S.; Kumpf, R. A.; Forman, J. E.; McCurdy, A.; Dougherty, D. A. *J. Am. Chem. Soc.* **1993**, *115*, 9907-9919.
6. Jorgensen, W. L. *Chemtracts - Org. Chem.* **1991**, *4*, 91-119.
7. Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 697-703.
8. Langley, D. R.; Golik, J.; Krishnan. B.; Doyle, T. W.; Beveridge, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 15-29.
9. Troxler, L.; Wipff, G. *J. Am. Chem. Soc.* **1994**, *116*, 1468-1480.
10. Jorgensen, W. L.; Nguyen, T. B. *Proc. Natl. Acad. Sci., USA* **1993**, *90*, 1194-1200.
11. Ferguson, S. B.; Seward, E. M.; Diederich, F.; Sanford, E. M.; Chou, A.; Inocencio-Szweda, P., Knobler, C. B. *J. Org. Chem.* **1988**, *53*, 5593-5595. Ferguson, S. B.; Sanford, E. M.; Seward, E. M.; Diederich, F. *J. Am. Chem. Soc.* **1991**, *113*, 5410-5419. Smithrud, D. B.; Wyman, T. B.; Diederich, F. *J. Am. Chem. Soc.* **1991**, *113*, 5420-5426.
12. Carcanague, D. R.; Diederich, F. *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 769-771.

13. Smithrud, D. B.; Sanford, E. M.; Chao, I.; Ferguson, S. B.; Carcanague, D. R.; Evanseck, J. D.; Houk, K. N.; Diederich, F. *Pure Appl. Chem.* **1990**, *62*, 2227-2236. Chao, I. *Ph.D. Thesis*, University of California, Los Angeles, 1992.
14. Klebe, G.; Diederich, F. *Phil. Trans. R. Soc. Lond. A* **1993**, *345*, 37-48.
15. Jorgensen, W. L. *BOSS* (Yale Univ., New Haven, CT), Version 3.4 **1993**.
16. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.
17. Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657-1666.
18. Jorgensen, W. L.; Nguyen, T. B. *J. Comp. Chem.* **1993**, *14*, 195-205.
19. Still, C. W. *Macromodel* (Columbia Univ., New York, NY), Version 3.5a **1989**
20. Schlegel, H. B. Optimization of Equilibrium Geometries and Transition Structures. In *Ab Initio Methods in Quantum Chemistry*; Lawley, K. P. Ed.; John Wiley and Sons, Ltd.: New York, 1987; pp. 249-286.
21. The ethyl groups on the piperidinium nitrogens of **2** were replaced by methyl groups in this study.
22. Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420-1426.
23. Jorgensen, W. L; Madura, J. D. *Mol. Phys.* **1985**, *56*, 1381-1392.
24. Jorgensen, W.L. *J. Phys. Chem.* **1986**, *90*, 1276-1284.
25. Gratitude is expressed to the National Institutes of Health and the National Science Foundation for support of this research.

(*Received* 8 *May* 1994)