A Molecular Mechanics Study of the Cholesteryl Acetate Crystal: Evaluation of Interconversion among r_g, r_z, and r_α Bond Lengths

Buyong Ma, Jenn-Huei Lii, Kuohsiang Chen, and Norman L. Allinger*

Contribution from the Computational Center for Molecular Structure and Design, Department of Chemistry, The University of Georgia, Athens, Georgia 30602

Received August 22, 1996. Revised Manuscript Received November 21, 1996[⊗]

Abstract: MM3 calculations on the cholesteryl acetate crystal were carried out to study the accuracy of the MM3 force field, and to evaluate previous approximations (Kuchitsu and Cyvin) in the relationships among r_g, r_z, and r_α bond lengths. It has been found that the previous approximations are good for treating the skeletons of relatively rigid molecules, for chemical bonds not involving hydrogen atoms. However, additional corrections have to be included to interconvert r_g, r_z, and r_α bond lengths for general purposes. New relationships among r_g, r_z, and r_α bond lengths have been derived. Using the improved approximations, our MM3 r_α structure for cholesteryl acetate in the crystal agrees well with experimental results (neutron diffraction, 20 K), including for the bonds involving hydrogen atoms.

Introduction

Information on molecular structure may suffer not only from computational or experimental errors, but also from the consequences of the different nature of the various physical phenomena that may be involved in its determination.[3] For example, a molecular structure may be determined by gas-phase electron diffraction, or by crystal diffraction (neutron diffraction). The molecular structures determined by these two kinds of experiments may differ not only as a result of the gas/crystal structural changes, but also as a result of the different definitions of molecular structure used in the experiments. Gas-phase electron diffraction experiments give the thermal average value of the internuclear distance (r_g), which is also our definition of "bond length" in MM3[4] and MM4[5] (and also MM1 and MM2).[6] However, neutron diffraction experiments give the average nuclear positions at thermal equilibrium, and from the average nuclear positions, the internuclear distance r_α is determined. The r_g and r_α bond lengths differ from the consequences of the perpendicular vibrations.[7] The relationship between r_g and r_α bond lengths has been derived by Kuchitsu and Cyvin,[1] and has been used to obtain experimental bond lengths for small relatively rigid molecules in the gas phase.[8] However, its validity for general purposes has not been previously examined.

In the present research, we studied the cholesteryl acetate crystal (Figure 1), including both the structure of the molecule in the crystal and the crystal cell constants, via MM3 calculations. Cholesterol and its derivatives have long been important for their biological significance. Computational studies of cholesterol derivatives were reported recently, including careful force field studies of the cholesteryl acetate crystal, by Hagler's group.[9] Our interest in the cholesteryl acetate crystal is in part due to the accurate experimental structure now available, which was determined at 20 K by neutron diffraction experiments. Neutron diffraction may determine the nuclear positions of the hydrogen atoms with a precision comparable to that for the non-hydrogen atoms. Therefore, the experimental structure of cholesteryl acetate provides a good opportunity for us to examine the accuracy of MM3 calculations, and the relationship between r_g and r_α bond lengths.

Computational and Theoretical Considerations

A. Relationships among r_g, r_e, r_z, and r_α Bond Lengths.[8,10]
1. r_g and r_e. The r_e bond length can be estimated from r_g, provided one knows the anharmonicity of the bond stretching, since the average displacement of the bond length at temperature T

$$\langle r \rangle_T = r_g - r_e \qquad (1)$$

$\langle \Delta r \rangle_T$ is estimated to be $\langle \Delta r \rangle_T = 1.5 C_{anh}\langle \Delta z^2 \rangle_T$, where C_{anh} is the anharmonicity constant. Therefore

$$r_e = r_g - 1.5 C_{anh}\langle \Delta z^2 \rangle_T \qquad (2)$$

⊗ Abstract published in *Advance ACS Abstracts,* January 15, 1997.

(1) (a) Kuchitsu, K.; Cyvin, S. J. In *Molecular Structures and Vibrations*; Cyvin, S. J., Ed.; Elsevier Publishing Co.: New York, 1972; p 196. (b) Kuchitsu, K. In *Accurate Molecular Structures*; Domenicano, A., Hargittai, I., Eds.; Oxford Science Publications: New York, 1992.

(2) Weber, H-P.; Craven, B.; Sawzik, P. *Acta. Crystallogr.* **1991**, *B47,* 116.

(3) Domenicano, A.; Hargittai, I. In *Accurate Molecular Structures*; Domenicano, A., Hargittai, I., Eds.; Oxford Science Publications: New York, 1992.

(4) Allinger, N. L. Yuh, Y. H.; Lii, J-H. Molecular Mechanics. The MM3 Force Field for Hydrocarbons I, II, and III. *J. Am. Chem. Soc.* **1989**, *111,* 8551−8582, and subsequent papers. The MM3 program is available to all users from Tripos Associates, 1699 South Hanley Rd., St. Louis, MO 63144, and to academic users only from the Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN 47405. The MM4 program will be available shortly from the same sources.

(5) Allinger, N. L.; Chen, K.; Lii, J-H. *J. Comput. Chem.* **1996**, *17,* 642 and following papers.

(6) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99,* 8127.

(7) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society Washington, DC, 1982.

(8) (a) Nakata, M.; Yamamoto, S.; Fukuyama, T.; Kuchitsu, K. *J. Mol. Struct.* **1983**, *100,* 143. (b) Egawa, T. Yamamoto, S.; Nakata, M.; Kuchitsu, K. *J. Mol. Struct.* **1987**, *156,* 213. (c) Nakata, M.; Kuchitsu, K. *J. Mol. Struct.* **1982**, *95,* 205. (d) Nakata, M.; Fukuyama, T.; Kuchitus, K. *J. Mol. Struct.* **1982**, *81,* 121.

(9) Liang, C.; Yan, L.; Hill, J.-R. Ewig, C. S.; Stouch, T. R. Hagler, A. T. *J. Comput. Chem.* **1995**, *16,* 883 and references cited therein.

(10) Hedberg, L.; Mills, I. M. *J. Mol. Spectrosc.* **1993**, *160,* 117.

(11) Ma, B.; Lii, J-H.; Schaefer, H. F.; Allinger, N. L. *J. Phys. Chem.* **1996**, *100,* 8763.
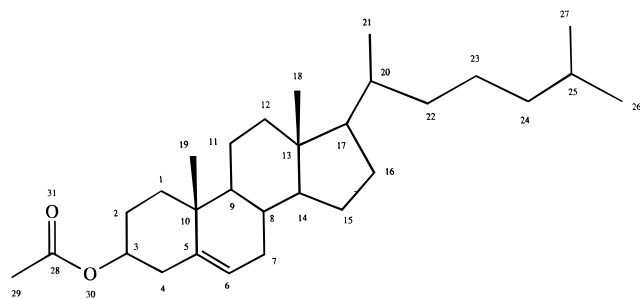
Figure 1. Cholesteryl acetate.

Equation 2 is the expression we used in MM3, and it has been proven to be a good approximation.[11]

2. r_e, r_z, and r_α. The vibrationally averaged bond length (r_v) may be calculated as[1]

$$r_v = [\langle\Delta x\rangle_v^2 + \langle\Delta y\rangle_v^2 + (r_e + \langle\Delta z\rangle_v)^2]^{1/2}$$

$$= r_e + \langle\Delta z\rangle_v + [\langle\Delta x\rangle_v^2 + \langle\Delta y\rangle_v^2]/2r_e \qquad (3)$$

Where local Cartesian coordinates are used, the z coordinate is defined along the equilibrium internuclear axis under consideration and the choice of the x and y axes is arbitrary.

Kuchitsu and Cyvin have argued that $\langle\Delta x\rangle_v$ and $\langle\Delta y\rangle_v$ are usually much smaller than 0.01 Å, and the term $[\langle\Delta x\rangle_v^2 + \langle\Delta y\rangle_v^2]/2r_e$ may be neglected.[1] Therefore, it is normally a good approximation to assume that

$$r_v = r_e + \langle\Delta z\rangle_v \qquad (4)$$

Usually, $\langle\Delta z\rangle_v$ may be expressed as[8a]

$$\langle\Delta z\rangle_v = 1.5 C_{anh}\langle\Delta z^2\rangle_v - K_v \qquad (5)$$

where $\langle\Delta z^2\rangle$ is the vibrational amplitude and K_v is the perpendicular vibrational amplitude

$$K_v = [\langle\Delta x^2\rangle_v + \langle\Delta y^2\rangle_v]/2r_e \qquad (6)$$

Therefore

$$r_v = r_e + 1.5 C_{anh}\langle\Delta z^2\rangle_v - K_v \qquad (7)$$

When we only consider the ground vibrational state, and include the small effect of centrifugal distortion ($\langle C\rangle_0$ or $\langle C\rangle_T$), we obtain the usual expressions

$$r_z = r_e + 1.5 C_{anh}\langle\Delta z^2\rangle_0 - K_0 - \langle C\rangle_0 \qquad (8)$$

$$r_\alpha = r_e + 1.5 C_{anh}\langle\Delta z^2\rangle_T - K_T - \langle C\rangle_T \qquad (9)$$

where r_z is the distance between average nuclear positions in the ground vibrational state at 0 K and r_α is the distance between average nuclear positions at T K. K_0 and K_T are the perpendicular vibrational amplitudes at 0 and T K, respectively. Equations 8 and 9 are the expressions used in MM3(94), and they are the same as the expression

$$r_\alpha = r_g - K_T - \langle C\rangle_T \qquad (10)$$

which is used in ASYM20.[10]

B. Examination of the Computational Accuracy of r_z and r_α. The molecular structures from MM3 and MM4 calculations may be directly compared to the r_g structures from electron diffraction. Recently, we demonstrated that since anharmonicity is included in the force field, the MM3 and MM4 programs

Table 1. Experimental and Optimized Unit Cell Constants for the Cholesteryl Acetate Crystal

	exptl (20 K)	exptl (123 K)	exptl (298 K)[a]	CVFF	CFF93	MM3
a (Å)	16.521	16.547	16.585	16.758	16.925	16.696
b (Å)	9.220	9.297	9.409	9.342	9.303	9.463
c (Å)	17.620	17.645	17.682	17.752	17.453	17.715
α (deg)	90.0	90.0		90.0	90.0	89.78
β (deg)	107.18	106.96		107.35	107.14	107.02
γ (deg)	90.0	90.0		90.0	90.0	89.82
volume (Å^3)	2564.2	2596.4	2643.4	2652.6	2626.0	2660.2
deviation (20 K)[b] (%)				+3.42	+2.41	+3.74
deviation (298 K)[b] (%)				+0.3[c]	−0.6[c]	+0.6

[a] Extrapolated from experimental data at 20 and 123 K. [b] Deviation of the computed volume from the experimental volume at 20 and 298 K. [c] Reference 15.

successfully convert r_g to r_e values.[11] However, occasionally for some molecules, we found that the calculated r_z and r_α bond lengths were shorter than r_e bond lengths. The nature of a Morse function is such that the anharmonicity always tends to lead to vibrationally averaged bond lengths being longer than equilibrium bond lengths. For diatomic molecules, and in general, one might anticipate that $r_\alpha > r_e$. Of course, $r_\alpha > r_e$ is not a golden rule.[12] For the molecules with a large amplitude bending vibration, the r_z value of the bond involved shrinks very much as a result of the bending. Since r_z and r_α are defined from the average nuclear positions at thermal equilibrium, it is physically possible in such a case for r_z and r_α to be shorter than the equilibrium bond length r_e. However, one should be suspicious when obtaining a result where r_z or r_α is shorter than r_e. One should also note that it is not possible to simultaneously measure r_e, r_z, and r_α at the current experimental level. In the gas phase one may measure either r_g or r_s structures, and then deduce r_e, r_z, and r_α from vibrational amplitude corrections. However, such deductions are strongly dependent on the accuracy of the physical model used, which is the main concern of the present work. An accurate crystalline r_α structure can be determined by low-temperature neutron diffraction experiments. We cannot determine a comparable experimental r_e structure. The gas-phase r_e structure is not comparable due to the large intermolecular interactions in the crystal. Hopefully in the near future, quantum mechanical calculations may provide accurate r_e structures in crystals. Nevertheless, we found that previously used relationships among r_g, r_e, r_z, and r_α failed badly in calculating the r_α structure of cholesteryl acetate (Table 1). Two main errors are responsible: (1) neglect of the term $[\langle\Delta x\rangle_v^2 + \langle\Delta y\rangle_v^2]/2r_e$ in eq 3, and (2) overestimation of the perpendicular vibrational amplitude K_T.

C. Corrections to the Expressions for r_z and r_α Bond Lengths. 1. The term $[\langle\Delta x\rangle_v^2 + \langle\Delta y\rangle_v^2]/2r_e$. This term may be neglected for bonds in relatively rigid molecules that do not involve hydrogen atoms, as shown by Kuchitsu and Cyvin. However, this neglect will cause problems when there are large amplitude motions, so it should be taken into account for molecules in general, as is shown below.

Similar to eq 5, we may have

$$\langle\Delta x\rangle_v \approx 1.5 C_{anh}\langle\Delta x^2\rangle_v \qquad (11)$$

$$\langle\Delta y\rangle_v \approx 1.5 C_{anh}\langle\Delta y^2\rangle_v \qquad (12)$$

Here, we assume that the anharmonicity constants are the same for all three directions of the vibration.

(12) Kuchitsu, K. Private communications.

From eq 6 we have

$$K_v^2 = [\langle\Delta x^2\rangle_v^2 + \langle\Delta y^2\rangle_v^2 + 2\langle\Delta x^2\rangle_v\langle\Delta y^2\rangle_v]/4r_e^2 \quad (13)$$

and for small values of $\langle\Delta x^2\rangle_v$ and $\langle\Delta y^2\rangle_v$, we may assume that

$$\langle\Delta x^2\rangle_v^2 + \langle\Delta y^2\rangle_v^2 \approx 2\langle\Delta x^2\rangle_v\langle\Delta y^2\rangle_v \quad (14)$$

Therefore

$$\langle\Delta x^2\rangle_v^2 + \langle\Delta y^2\rangle_v^2 = 2r_e^2 K_v^2 \quad (15)$$

Thus

$$[\langle\Delta x\rangle_v^2 + \langle\Delta y\rangle_v^2]/2r_e = 2.25C_{anh}^2(\langle\Delta x^2\rangle_v^2 + \langle\Delta y^2\rangle_v^2)/2r_e$$

$$= 2.25C_{anh}^2 r_e K_v^2 \quad (16)$$

Finally, we have

$$r_z = r_e + 1.5C_{anh}\langle\Delta z^2\rangle_0 - K_0 + 2.25C_{anh}^2 r_e K_0^2 - \langle C\rangle_0 \quad (17)$$

$$r_\alpha = r_e + 1.5C_{anh}\langle\Delta z^2\rangle_T - K_T + 2.25C_{anh}^2 r_e K_T^2 - \langle C\rangle_T \quad (18)$$

The validity of eq 14 depends on the degree of asymmetry of the perpendicular vibrational amplitudes.[12] For an ellipsoidal probability distribution of the amplitudes, one can estimate the validity by examining extreme cases. Equation 14 is exact when the distribution is cylindrical, i.e., $\langle\Delta x^2\rangle_v = \langle\Delta y^2\rangle_v$, and it breaks down if it is "polarized", e.g., $\langle\Delta x^2\rangle_v = 0$. Suppose $\langle\Delta x^2\rangle_v = 2\langle\Delta y^2\rangle_v$. It turns out that $0.8\{\langle\Delta x^2\rangle_v^2 + \langle\Delta y^2\rangle_v^2\} = 2\langle\Delta x^2\rangle_v\langle\Delta y^2\rangle_v$. So, it does not seem to be too bad to assume that they are equal, because in many cases this ellipsoid will be close to cyclindrical.[12]

The term $2.25C_{anh}^2 r_e K_0^2$ is normally negligible for non-hydrogen atoms in relatively rigid molecules, since K_T is 0.01 Å or less, as assumed by Kuchitsu and Cyvin. For bonds involving hydrogen atoms, or for torsional modes in saturated molecules (low frequencies), K_T may be on the order of 0.03. If $C_{anh} = 2.0$ and $r_e = 1.0$ Å (the C–H approximate bond length, for example), the value of the whole term is then on the order of 0.008 Å, a non-negligible amount. With eqs 17 and 18, the relationship $r_z > r_e$ is assured. However, we still found occasionally that $r_\alpha < r_e$, due to the overestimation of K_T.

2. Overestimation of K_T. K is calculated from the expression[10]

$$K_i = \sum_s 0.5\mathbf{L}_i^{ss}\langle Q_s^2\rangle$$

$$= \sum_s 0.5\mathbf{L}_i^{ss}A \coth(hc\nu/2kT) \quad (19)$$

where $\mathbf{L}_i^{ss}$ is a second-derivative $\mathbf{L}$ tensor. The function coth-$(hc\nu/2kT)$ in eq 19 is used to take account of the Boltzmann distribution. The function is 1.0 at 0 K. However, at T K ($T \neq 0$), when the molecule has a low vibrational frequency ν, the term $\coth(hc\nu/2kT)$ may blow up, and K_T will be overestimated.

Two approaches were tested to solve this problem. One was to set an upper limit for the value of the function $\coth(hc\nu/2kt)$. The other was to use a function $\exp[-a(1.0 - \nu/b)]$ to scale down the low-frequency (less than 500 cm^{-1}) contribution to K_T, where ν is the vibrational frequency and a and b are constants to be chosen. Both approaches succeeded. Upon fitting the experimental r_α structure of cholesteryl acetate, we set the upper limit value of $\coth(hc\nu/2kt) = 1.2$. Alternatively,

we used the scale function $\exp[-4.0(1.0-\nu/500)]$. The former method was chosen for incorporation into MM3(96).

There are two physical reasons to suppress K_T. First, low-frequency and large-amplitude motions existing in a molecule in the gas phase may be suppressed in the crystal phase due to intermolecular forces. There are two terms involving K_T in eq 18, with opposite effects on the r_α bond length. Note that the term $2.25C_{anh}^2 r_e K_T^2$ contains the anharmonicity constant C_{anh}, which is taken to have a fixed value for all vibrational modes. Generally speaking, C_{anh} should increase with decreasing vibrational frequency. Therefore, suppressing K_T will compensate in part for the errors introduced by holding C_{anh} fixed.

D. Computational Procedure. The crystal studies include two different, but related, aspects. One is the computation of the molecular structure in the crystal, and the other is the computation of crystal cell constants. Experimentally, the unit cell of the cholesteryl acetate crystal contains two nonequivalent molecules (for convenience, we will call them a "dimer"). First, we took the experimental dimer structure and generated the experimental crystal lattice cage with additional molecules by replication according to appropriate symmetry operations. The atoms which are more than 4.5 Å from the nearest atom of the central molecule (dimer) were trimmed off to make the problem more manageable. The energy of the dimer inside the cage was then minimized in the usual way, including all of the forces from the neighboring atoms (which were held immobile). After the dimer inside the cage had been optimized, it was replicated to yield a 3 × 3 × 3 block of unit cells, and the cell constants were then optimized. The calculation then alternated in successive iterations between the two procedures until the values for the unit cell constants and the structure converged. Thus, we obtained the MM3 structure (r_g) of the dimer in the crystal, together with the crystal cell constants.

The free dimer was also optimized in the gas phase to obtain the r_g structure and vibrational frequencies. The gas-phase r_g bond lengths were converted to r_z, r_α, and r_e bond lengths subsequently. The differences between the r_g and r_α structures in the gas phase were applied to the r_g structure in the crystal to obtain the r_α structure in the crystal.

Results and Discussion

We calculated the cell constants for the cholesteryl acetate crystal as indicated above. The computed cell constants are systematically a little larger than the experimental cell constants, as they should be. The computed cell volume is 3.74% larger than the experimental one at 20 K. The MM3 parametrization used is for room-temperature calculations, and the low-temperature unit cell would be expected to expand slightly when warmed to room temperature. The expected experimental volume at 298 K was extrapolated from the experimental volumes at 20 and 123 K, and our MM3 value is only 0.6% larger than the extrapolated value at 298 K (see Table 1). Hagler's group also studied the cell constants for the cholesteryl acetate crystal, using their CVFF and CFF93 force fields.[14] As indicated in ref 8, the CFF93 and CVFF force fields correspond to a temperature of 0 K, so their computed volumes should be smaller than the experimental volume at 20 K. However, Dr. Ewig[15] pointed out to us that the nonbonded terms in the CFF93 and CVFF force fields were actually optimized to reproduce the crystal structure at 298 K, not 0 K. Thus, all three force fields (MM3, CFF93, and CVFF) reproduced the experimental

(13) Ma, B.; Allinger, N. L. *J. Mol. Struct.,* in press.
(14) Hwang, M. J.; Stockfisch, T. P.; Hagler, A. T. *J. Am. Chem. Soc.* **1994,** *115,* 2515.
(15) Ewig, C. S. Private communications.

Table 2. Selected MM3 Calculated Bond Lengths Å and Comparison with Experiments[a]

bond	r_g	r_e	r_z old[b]	r_z new[c]	$r_\alpha{}^d$ old[b]	$r_\alpha{}^d$ new[c]	r_α (exptl)
$C_{25}-C_{24}$	1.5433	1.5334	1.5366	1.540	1.4740	1.540	1.524
$C_{26}-C_{25}$	1.5395	1.5297	1.5310	1.536	1.4529	1.536	1.534
$C_{27}-C_{25}$	1.5390	1.5292	1.5276	1.536	1.3798	1.536	1.529
$C_{29}-C_{28}$	1.4918	1.4829	1.4853	1.489	1.4565	1.489	1.505
$O_{30}-C_3$	1.4465	1.4385	1.4425	1.444	1.4278	1.444	1.463
$O_{30}-C_{28}$	1.3590	1.3515	1.3536	1.356	1.3265	1.356	1.344
$O_{31}-C_{28}$	1.2145	1.2087	1.2079	1.212	1.1781	1.212	1.208
$C_{25}-H_{68}$	1.1178	1.0934	1.0982	1.099	0.9943	1.099	1.096
$C_{26}-H_{70}$	1.1129	1.0893	1.0842	1.099	0.9645	1.099	1.104
$C_{27}-H_{73}$	1.1129	1.0893	1.0873	1.099	1.0343	1.099	1.100
$C_{29}-H_{77}$	1.1100	1.0866	1.0637	1.097	0.9039	1.097	1.081
$C_{29}-H_{76}$	1.1060	1.0829	1.0630	1.093	0.9232	1.093	1.083
$C_{29}-H_{75}$	1.1100	1.0866	1.0637	1.097	0.9081	1.097	1.089
RMS diff (r_α, mm3 − r_α exptl)					0.0551	0.007	
signed av diff (r_α, mm3 − r_α, exptl)					−0.034	−0.001	

[a] These bond lengths are those of the side chains. The rigidity of the ring systems reduces the vibrational motions, and that part of the molecule is fit with much less difficulty. [b] Old: calculated from eqs 8 and 9. [c] New: calculated from eqs 17 and 18, with the upper limit value of coth($hc\nu/2kt$) = 1.2 applied in the calculation of K_T. [d] The r_α values are the bond lengths in the crystal, in order to compare with experimental bond lengths. . The r_g, r_e, and r_z values are the bond lengths in the gas phase.

cell constants for the cholesteryl acetate crystal to approximately within experimental error.[16]

As may be seen in Table 2, using eq 18 and with the upper limit value of coth($hc\nu/2kt$) = 1.2, applied in the calculation of K_T, we now are able to ensure the relationships $r_z > r_e$ and $r_\alpha > r_e$. Our MM3 estimation of r_z bond lengths is known to agree with the experiments pretty well, with a RMS deviation of 0.0044 Å over the fourteen bonds for a total of six molecules surveyed in a separate study.[13]

Next we will focus on the comparison of our MM3 r_α structure of cholesteryl acetate with the experimental structure from neutron diffraction experiments at 20 K. We calculated the molecular structure of the cholesteryl acetate dimer by the procedure described above. Table 3 shows, as expected, that the MM3 r_g bond lengths are systematically longer than the experimental r_α bond lengths, especially for the C−H bonds. Additionally, the agreement of the MM3 r_α bond lengths with the experimental r_α bond lengths is satisfactory. An achievement here was the great improvement in the agreement of the MM3 calculation with experiments for the r_α bond lengths of the C−H bonds, especially for those of the side chains. The rigidity of the ring systems constrains the vibrational motions, and that part of the molecule is fit with much less difficulty. Using the approximations in the earlier literature for the interconversion of r_g and r_α, these RMS errors and the sums of the signed errors were 0.0551 and −0.0335 Å, respectively, for the side chains (Table 2). However, our new approach deals with the side chains with a precision equivalent to that obtained for rigid ring systems. The RMS deviation between the experimental and MM3 r_α bond lengths is now 0.0074 Å over all bonds in the cholesteryl acetate dimer, and 0.0067 Å over all bonds between heavy atoms (Tables 2 and 3).

Hagler's group also studied the molecular structure of cholesteryl acetate in the crystal, using their CVFF and CFF93

(16) Note that two different experimental determinations of the cholesterol crystal cell vectors gave *b* lengths differing by 1.5% (ref 14).

Table 3. RMS and Signed Average Deviations of the Calculated Structure of the Cholesteryl Acetate Dimer (in the Crystal Lattice) from the Experimental Crystal Structure[a]

			CVFF[b]	CFF93[b]	MM3 (optimized[c] cell unit) $r_{g,mm3} - r_{\alpha,exptl}$	MM3 (optimized[c] cell unit) $r_{\alpha,mm3} - r_{\alpha,exptl}$
position[d] (Å)		RMS	0.154	0.096	0.041	
bond	total	RMS	0.018	0.011	0.013	0.007
		av			0.009	−0.001
	X−Y	RMS	0.025	0.013	0.008	0.007
		av			0.004	−0.001
	C−H	RMS			0.016	0.008
		av			0.013	−0.001
angle	total	RMS	1.4	1.1	1.1	
		av			−0.2	
	non-H	RMS	1.3	1.0	0.6	
		av			−0.3	
torsion (deg)	total	RMS	3.6	2.3	1.7	
		av			−0.2	
	non-H	RMS	3.7	2.9	1.2	
		av			−0.2	

[a] Experiment: ref 2. [b] From ref 9. They reported the deviations for the two molecules (A and B) of the dimer separately, but the differences are quite small. The averaged values of molecules A and B are cited here. [c] The molecular structure was optimized with the restriction of both the experimental cell constants and the optimized cell constants. Only results from the optimized unit cell are reported here. The improvement from the use of the experimental unit cell is insignificant.

force fields.[14] Their results are listed in Table 3, together with the MM3 and experimental results. It may be seen that the MM3 results for the structure of the molecule are in better agreement with the experimental molecular structure than are the CVFF and CFF93 force fields. (Note that no force field parameters were adjusted in Hagler's work, or in the present study. The MM3(96) force field was used here.) If we take the RMS value of the discrepancies of the atomic positions in the cholesteryl acetate molecule in Table 3, for example, this is 0.154 Å with the CVFF force field, but is reduced to 0.096 Å with CFF93. The corresponding value with MM3 is only 0.041 Å. Note also that with MM3 the sums of the signed errors are only about 0.001 Å in bond lengths and 0.2° in angles (the corresponding data were not given for other force fields). That means the systematic errors in the MM3 calculations are really small, so that most of the error is random (and presumably experimental). Note that the standard literature methods for obtaining r_α from other kinds of bond lengths ("old method", Table 2) lead to disastrous errors in cases like this one, where bonds involve hydrogen atoms, or where there are low-frequency (torsional) modes between heavy atoms, either of which leads to vibrations with quite large perpendicular amplitudes.

Supporting Information Available: Four tables containing the Cartesian coordinates of the MM3 calculated cholesteryl acetate (r_α) crystal and the detailed comparisons between the MM3 and experimental structures (15 pages). See any current masthead page for ordering and Internet access instructions.

JA9629524