The Crystal Polymorphism of Tetrolic Acid ($CH_3C{\equiv}CCOOH$): A Molecular Dynamics Study of Precursors in Solution, and a Crystal Structure Generation

A. Gavezzotti,* G. Filippini, J. Kroon, B. P. van Eijck and P. Klewinghaus

Abstract: The possible configurations of two molecules of tetrolic acid in a solvent box containing 226 carbon tetrachloride molecules are studied by molecular dynamics with the GROMOS package and force field over periods of up to 2000 picoseconds. The cyclic hydrogen-bonded dimer was the most persistent configuration, but events leading to the cleavage of one hydrogen bond and thus the formation of a precursor to the crystal catemer motif were found to occupy up to 10% of the simulation times. The experimental bond-breaking enthalpy was correctly reproduced. Two different crystal structure generation procedures were employed to reproduce the two observed polymorphic crystal structures and to predict other possible polymorphs; in all cases, some unobserved structures had more cohesive packing energies than the observed ones. The possible application of molecular dynamics in the study of the preliminary steps in crystal nucleation is discussed.

Keywords
crystal structure predictions · molecular dynamics · polymorphism · tetrolic acid

Introduction

The prediction of crystal structure from molecular structure is a thriving field of research.[1-3] In this respect, polymorphism is a crucial issue, being at the same time a source of concern and a relevant phenomenon that helps to throw some light on the process of crystal construction.[4] Computational approaches to crystal structure prediction usually go through the following steps: 1) assumption or optimization of the gas-phase molecular structure; 2) construction of a geometrical model of the crystal; 3) enumeration of possible crystal structures with their packing energies, and choice of the best structure on an energy basis. Aside from the problems inherent in the choice of the force fields, and the possible effects of temperature, such an approach ignores the molecular recognition in bi- or multimolecular clusters that must occur in solution, and that dictates the kinetics of the very first nucleation stages.

This paper tackles the problem of the "social background" of crystallizing molecules in solution vis-à-vis the final crystal structure. A combination of molecular dynamics and crystal structure generation methods is used. Computational chemistry in solution is now a mature discipline,[5-8] but has so far been used mostly to study pure liquids or the intramolecular structure of solute molecules in supramolecular chemistry or in systems of biological interest. We aim explicitly at the modelling and, as far as possible, understanding of multimolecular crystal precursors in solution for small organic molecules. Tetrolic acid is an ideal candidate for a case study, since two crystal phases with two different and clearly characterized hydrogen bonding schemes have been determined by X-ray single-crystal diffraction.[9] Its small size and relatively easy parameterization help to reduce the computational effort. Correspondences and differences between molecular recognition patterns across the hydrogen bonds, as well as the basic recognition motifs in the crystal structures, will be discussed.

Crystal structures of tetrolic acid: An increasing number of molecules exhibiting polymorphism in the crystalline state are appearing in the literature. The choice of tetrolic acid for this case study depended on the presence of two crystal phases with two different and clearly characterized hydrogen bonding schemes determined by X-ray single-crystal diffraction. Moreover, the computational effort is reduced owing to the small size of the acid and its relatively easy parameterization.

Tetrolic acid crystallizes in two modifications, an α and a β form (Table 1); both phases were grown from solution in an apolar solvent (*n*-pentane).[9] Crystals of the α form are triclinic, space group $P\bar{1}$; the structure consists of hydrogen-bonded centrosymmetric dimers over the carboxylic group, arranged in

[*] Prof. A. Gavezzotti
Dipartimento di Chimica Strutturale e Stereochimica Inorganica
Università di Milano, via Venezian 21, I-20133 Milano (Italy)
Fax: Int. code +(2)7063-5288
e-mail: gave@stinch12.csmtbo.mi.cnr.it
Dr. G. Filippini
Centro CNR per lo Studio delle Relazioni tra Struttura e Reattività Chimica
c/o Dipartimento di Chimica Fisica ed Elettrochimica
Università di Milano, via Golgi 19, I-20133 Milano (Italy)
Prof. J. Kroon, Dr. B. P. van Eijck, P. Klewinghaus
Department of Crystal and Structural Chemistry
Bijvoet Center for Biomolecular Research, Utrecht University
NL-3584 CH Utrecht (The Netherlands)

  0947-6539/97/0306-0893 $ 17.50+.50/0

Table 1. Crystal data (Å, °) for tetrolic acid polymorphs (ref. [9]).

	a	b	c	α	β	γ
$P\bar{1}$ structure						
experimental	7.320	5.099	7.226	83.97	117.46	112.0
reduced [a]	5.099	7.185	7.226	113.28	96.03	109.15
$P2_1$ structure						
experimental	7.887	7.121	3.937	–	100.18	–

[a] Niggli cell reduction procedure.

planar layers (Figure 1 a). Crystals of the β form are monoclinic, space group $P2_1$; molecules are arranged in approximately planar chains, propagating along the twofold screw axis through a single O–H···O hydrogen bond between pairs of molecules (the so-called catemer motif; Figure 1 b). This motif also appears in the crystal structure of formic acid,[10] acetic acid,[11] and in the α form of oxalic acid.[12] The structural motifs occurring in organic carboxylic acids have been analyzed in several review papers.[13–15]

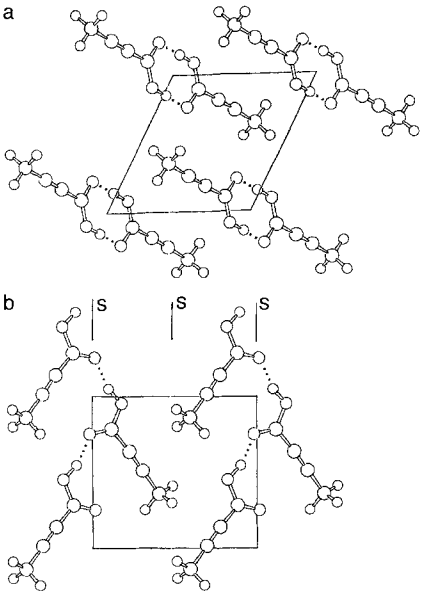

Figure 1. Crystal structures of tetrolic acid (ref. [9]): a) $P\bar{1}$, b) $P2_1$. See Table 1 for cell parameters. The positions of the carboxylic acid hydrogen atoms in the $P\bar{1}$ structure are disordered. Dotted lines: hydrogen bonds; S denotes a screw axis.

Computational Details

Energy minimization (EM) and molecular dynamics (MD) calculations were carried out with the GROMOS package[16] on a Silicon Graphics INDY workstation. A centrosymmetric dimer of tetrolic acid, with starting molecular dimensions taken from the crystal structures, was inserted into a previously equilibrated rectangular box containing 226 CCl_4 molecules (starting box dimensions 29.0, 31.4, 41.0 Å). Periodic boundary conditions were applied. The solvent was modelled in the united atom approximation, and two Lennard-Jones parameter sets were tried.[17, 18] The tetrolic acid molecules were modelled with a united atom for the methyl group in the charge group approximation. The full force field (Table 2) is the standard GROMOS one, supplemented by a few estimated force constants for the acetylenic part. Stretch and bend equilibrium parameters were taken from the experimental X-ray values.

Table 2. Force-field parameters.

charge parameters (electrons)					
hydroxyl oxygen	−0.40		hydroxyl hydrogen		0.40
carbonyl oxygen	−0.38		carbonyl carbon		0.38
bending	k, kcal mol^{-1} rad^{-2}		α^0, °		
O=C–O	120		123		
O=C–C	120		124		
O–C–C	120		113		
C≡C–C	200		180		
H–O–C	95		113		
dihedral (kcal mol^{-1})					
H–O–C=O, H–O–C–C			2 [1 − cos(2ϕ)]		
improper dihedral (energy in kcal mol^{-1}, ϕ in rad)					
C=O···O–C			20ϕ		
bond stretching, k (kcal mol^{-1} Å^{-2}), d (Å)					
O–H	750	1.000	C–O	900	1.310
C=O	1200	1.204	C–C	900	1.441
C≡C	1500	1.178	C–CH	900	1.455
Lennard-Jones parameters for CCl_4 [a]					
	C_{12} (kcal mol^{-1} Å^{12})		C_6 (kcal mol^{-1} Å^6)		
set A	1.544×10^9		7.755×10^4 (ref. [17])		
set B	1.361×10^9		6.350×10^4 (ref. [18])		

[a] Lennard-Jones parameters for solute: see ref. [16].

In the MD simulations, after an initial energy minimization the system was kept in contact with temperature ($T = 298$) and pressure ($p = 1$ bar) baths, with velocity rescaling (relaxation times 0.20 ps for solvent and 0.05 ps for solute molecules; 5 ps for isotropic pressure rescaling). The time step was 0.001 ps in all cases (see ref. [8] for details). Several runs were conducted (Table 3) with different parameter sets and cutoff distances.

Table 3. Overview of MD runs.

Parameter set	Cutoff, Å	Time, ps	Label
A	14	1750	PARA
A	13	850	PARA1
B	14	2000	PARB

The molecular recognition pattern between the two tetrolic acid molecules was analyzed by calculating average geometrical parameters (typical sampling times were 0.05 ps for coordinates and 0.1 ps for energies and box dimensions). Distances and angles used for this purpose are defined in Table 4 and Figure 2; as a reference for recognition patterns, several idealized models can be proposed, mostly suggested by obvious chemical concepts as well as by inspection of the animations of molecular motion during the MD runs. The crystal structures of the two polymorphs roughly correspond to DIM (dimeric α form) and *syn-cis*-CAT (catemeric β form); therefore, the geometrical parameters for these two configurations were taken from the crystal structures, and are collected in Table 4.

Table 4. Geometrical parameters for idealized bimolecular configurations, from the crystal structures (distances R in Å, angles α in °). For definitions and numbering, see Figure 2.

Parameter		DIM (α phase)	*syn-cis*-CAT (β phase)
$R1$	2–11	1.656	5.741
$R2$	4–9	1.656	1.683
$R3$	2–14	7.450	3.916
$R4$	7–9	7.450	7.422
α1	1-2-10	119	97
α2	1-3-9	116	121
α3	8-10-2	116	8
α4	8-9-3	119	137
α5	1-8-12	180	77
α6	8-1-5	180	168

0947-6539/97/0306-0894 $ 17.50+.50/0 Chem. Eur. J. 1997, 3, No. 6

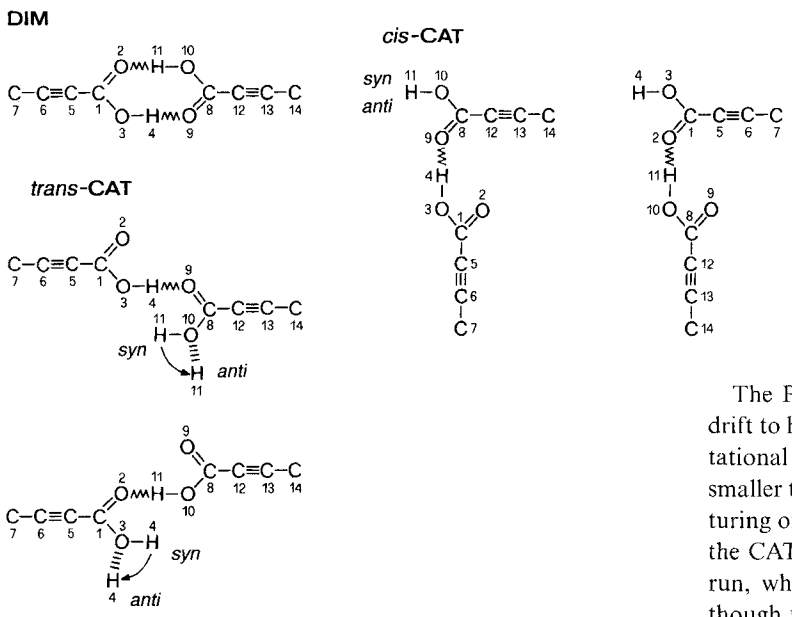

Figure 2. Definition of possible configurations for a bimolecular complex of tetrolic acid, with atom numbering.

Packing analysis of the existing crystal structures, and crystal structure generation for several phases of tetrolic acid, were carried out by means of the OPEC-PROMET package[19] with the accompanying crystalline force field[15] and, in parallel, by the Utrecht crystal packing program UPACK[2] with the same force field as used in the MD simulations.

Results

Molecular dynamics in solution: The starting configuration of the two solute molecules in all MD runs, imposed as almost perfect *cis*-DIM, was preserved during the largest fraction of time as expected, since in a nonpolar solvent the carboxylic acid groups cannot be favourably solvated and preferentially adopt a shielded cluster arrangement against each other. Table 5 shows average geometries of such DIM phases; aside from the effects of molecular vibration, the dimer preserves its typical structure and H-bonding geometry. When a simulation was run in water, disruption of the molecular complex occurred almost instantly.

Table 5. Average geometrical parameters with rmsd's for DIM configurations during MD runs. rmsd = $[\sum(R - \bar{R})^2/N_{obs}]^{1/2}$. See also Tables 3 and 4.

	PARA (1150 ps)	PARB (550 ps)
$R1$	1.76(16)	1.78(18)
$R2$	1.76(16)	1.77(18)
$R3$	7.19(13)	7.19(15)
$R4$	7.19(13)	7.19(14)
$\alpha1$	134(8)	133(9)
$\alpha2$	100(4)	100(5)
$\alpha3$	100(4)	100(5)
$\alpha4$	134(8)	133(9)
$\alpha5$	165(9)	164(9)
$\alpha6$	165(9)	164(10)

During all MD runs, the DIM pattern is occasionally disrupted by cleavage of one hydrogen bond. These occurrences, called CAT events, are easily spotted by a sudden change in one O–H···O distance; time intervals in which either $R1$ or $R2$ (see Table 4 and Figure 2 for definitions) grew greater than 3.7 Å were considered as real CAT events, rather than fluctuations. This apparently large threshold was chosen because the main aim was to spot the *syn-cis*-CAT geometry as found in the β crystal phase, where this distance is as large as 5.7 Å. However, one hydrogen bond was preserved in all CAT events, so that the two molecules never drifted completely apart during our simulations.

The PARA and, to a lesser extent, PARA1 runs showed a drift to higher than normal density; this is most likely a computational artifact, since one box dimension tended to become smaller than twice the cutoff radius, introducing a partial structuring of the solvent to an almost crystalline lattice. Results for the CAT events will therefore be discussed only for the PARB run, where the density and total energy are equilibrated, although the average density (1.31 $g\,cm^{-3}$) is lower than that of pure CCl_4 (1.59 $g\,cm^{-3}$). We emphasize, however, that CAT events of approximately the same significance were observed in all MD runs, irrespective of density and of convergence problems.

Figure 3 shows plots of $R1$, $R2$, $R3$, $R4$, $\alpha5$ and $\alpha6$ during the CAT2 and CAT3 events in the PARB simulation. Events occurring during the first 1300 ps are analyzed in detail, while the

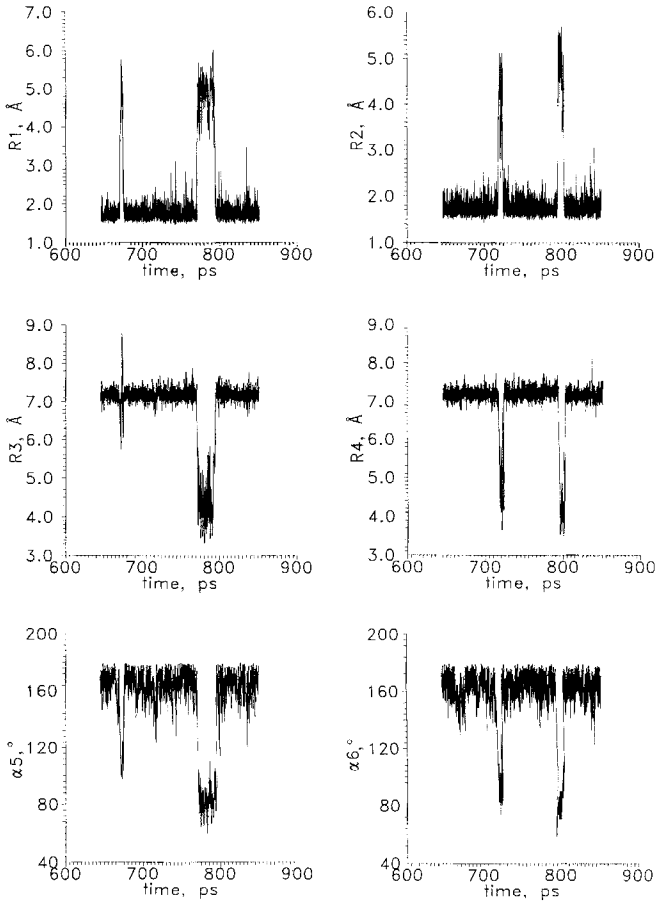

Figure 3. Evolution over time of relevant geometrical parameters (see Table 4 and Figure 2 for definitions) during the CAT1 (670 ps), CAT2 (720 ps) and CAT3 (780–820 ps) events.

simulation was carried over to 2000 ps just to check the overall stability of the system. These results are typical and illustrate all the observed possibilities. Table 6 shows the corresponding averaged geometrical parameters.

Table 6. Average geometrical parameters with rmsd's for CAT events during MD runs (see also captions to Tables 3 and 4).

	CAT2	CAT3A	CAT3B	CAT4	CAT5
$R1$	1.79(14)	4.82(49)	1.83(24)	4.07(100)	1.78(18)
$R2$	4.08(66)	1.75(14)	4.79(57)	1.77(15)	4.41(91)
$R3$	7.25(14)	4.54(58)	7.26(21)	6.71(69)	7.30(18)
$R4$	4.95(73)	7.28(15)	4.35(53)	7.28(17)	6.62(117)
$\alpha1$	124(13)	109(10)	122(14)	106(14)	125(16)
$\alpha2$	50(19)	107(6)	28(18)	108(9)	72(20)
$\alpha3$	105(6)	26(15)	107(8)	72(18)	110(8)
$\alpha4$	114(12)	129(13)	104(15)	128(18)	101(17)
$\alpha5$	162(10)	83(9)	157(15)	129(14)	159(11)
$\alpha6$	97(13)	163(10)	83(11)	162(11)	132(23)

The CAT2 event is a real *cis*-CAT one; $R1$ and $R3$ are as they should be for the preserved hydrogen bond, while $R4$ (4.95 Å) is somewhat longer than the contact in the crystalline catemer (3.92 Å). $\alpha5$ and $\alpha6$ show the expected 170–80° pattern (see Table 4).

The CAT3 event is also a true *cis*-CAT one, even closer to the ideal crystal geometry ($R3$ 4.54 Å). Interestingly, after 23 ps (CAT3A), an intermediate flipover transition (about 1 ps) switches the *cis*-CAT complex over to its permuted one, by cleavage of the 4–9 bond and reformation of the 2–11 one (CAT3B). The higher rmsd's of the averaged $R1$ and $R3$ ($R2$ and $R4$) distances demonstrate, however, that the *cis*-CAT configurations are rather flexible in solution, as could have been expected; crystalline configurations are perforce frozen ones. In fact, visual inspection of the animations during the CAT events shows that the simulated CAT complex is not always coplanar, as it is in the crystal. Clearly, planarity is enforced only at the stage of string and layer formation, and even more as layers are superimposed to form a three-dimensional structure.

The *syn–anti* equilibrium fluctuates in all CAT events. It appears that the position of the carboxylic hydrogen atom is nearly immaterial once the hydrogen bond has been cleaved, at least to the approximation inherent in the present force field.

The CAT4 and CAT5 events are much less easy to analyze, let alone classify. After cleavage of one hydrogen bond, the structure can be broadly assimilated to a *trans*-CAT one, although with large fluctuations, in particular away from coplanarity. The large rmsd's of averaged $R1$ and $R3$ ($R2$ and $R4$) distances confirm this.

Figure 4 shows the electrostatic contribution to the interaction energy between the two tetrolic acid molecules during the CAT events. The bond-breaking energy, within the GROMOS force field, is of an essentially electrostatic nature, since the corresponding Lennard-Jones energy profiles only show marginal fluctuations. A rough estimate of the hydrogen bond breaking energy is 8–11 kcal mol^{-1}, as reckoned from manual averaging of the energy profiles. This value compares reasonably well with the experimental dimerization energy of the first four aliphatic carboxylic acids,[20] 14–16 kcal mol^{-1} or 7–8 kcal mol^{-1} per hydrogen bond.

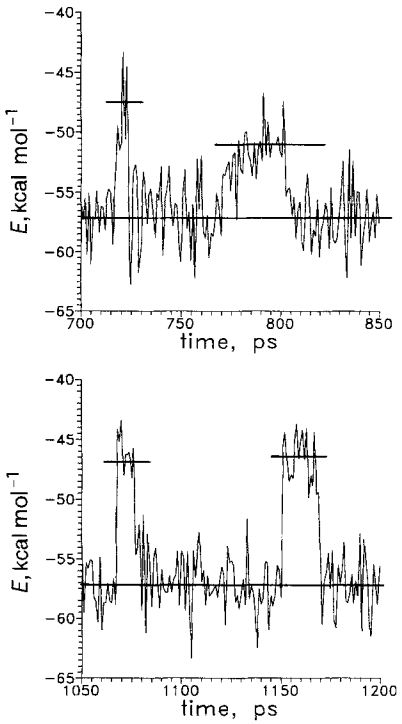

Figure 4. Solute–solute electrostatic energy (kcal mol^{-1}) during the CAT events. The horizontal bars are manual estimates of the averages, from which the hydrogen bond breaking energy is estimated. Left to right, top to bottom: CAT2, CAT3, CAT4, CAT5.

Crystal structure simulations: The PROMET procedure[19] was used to generate crystal structures for tetrolic acid in the two observed space groups. Briefly, the procedure consists of building preliminary cohesive structural units (in the present case, centrosymmetric dimers or molecular ribbons along a screw axis) and then expanding these in three or two directions, respectively, to build complete three-dimensional crystal structures. The cohesive energies of preliminary structural units and of full crystal structures are calculated by means of standard intermolecular potentials.[15]

The centrosymmetric dimer for the search in space group $P\bar{1}$ was easily constructed, with a cohesive energy of 11.2 kcal mol^{-1}. The search for stable molecular ribbons along the screw axis posed a much more challenging problem, since there are many possible arrangements for a complex formed by just one O–H···O hydrogen-bond linkage. Screw ribbons with a wide range of cohesive energies were considered as promising and carried over to the translational search.

The basic structural motif shown in Figure 5a propagates into a corrugated ribbon with a translational distance of 4 Å (screw pitch 2 Å) and the highest cohesive energy (18.8 kcal mol^{-1}). The ribbon then forms a very stable $P2_1$ crystal structure (see Table 7, entry "CORRUG"). On the other hand, the basic motif that led to the computational reproduction of the observed structure is shown in Figure 5b; its cohesive energy is only 10.8 kcal mol^{-1}, but it does show the expected *cis*-CAT arrangement, and, unlike the motif in Figure 5a, it is essentially planar. The packing energy of the full crystal struc-

0947-6539/97/0306-0896 $ 17.50+.50/0 Chem. Eur. J. 1997, 3, No. 6

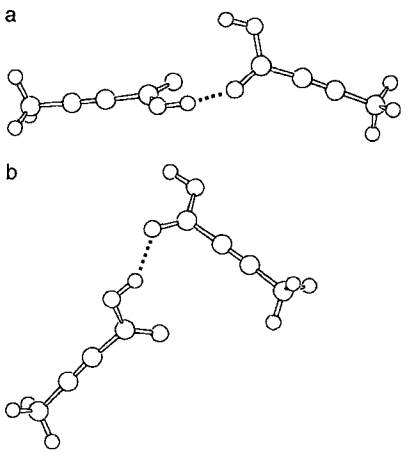

Figure 5. a) The basic structural motif along a screw axis for the corrugated CAT crystal structure (see Table 7), as determined during the preliminary search with PROMET. b) As for a) for the generation of the *cis*-CAT crystal structure, as experimental.

Table 7. Cell dimensions (Å and °) and packing energies (kcal mol^{-1}) for crystal structures generated by the PROMET and UPACK procedures.

	a	b	c	α	β	γ	−PE	V_{cell}
$P\bar{1}$ structures								
exptl	7.320	5.099	7.226	83.97	117.46	112.0	–	–
reduced [a]	5.10	7.18	7.23	113.3	96.0	109.2	17.3	221.1
opt [b]	5.22	6.56	6.94	113.1	96.1	109.2	18.3	198.5
PROMET [c]	5.15	7.00	7.16	119.3	94.2	112.6	18.6	196.7
UPACK	6.89	5.13	6.62	87.1	113.8	116.8	–	188.5
$P2_1$ structures								
exptl	7.887	7.121	3.937	–	100.18	–	17.5	217.6
opt [b]	7.73	6.75	3.91	–	102.1	–	18.4	199.6
PROMET [c]								
cis-CAT	7.63	6.72	3.99	–	104.4	–	18.7	198.0
CORRUG	8.25	4.31	5.71	–	94	–	18.5	202.5
trans-CAT	6.59	5.15	6.38	–	95	–	16.6	214.9
UPACK								
cis-CAT	8.31	6.24	4.00	–	111.3	–	–	193.3

[a] After cell reduction according to Niggli. [b] After structure relaxation by program PCK 83 [27] and the same force field as used in PROMET calculations. [c] See text.

ture (Table 7) does not differ appreciably from that of the structure obtained from the previous motif.

The cohesive energy of the ribbon corresponding to the pure *trans*-CAT configuration is 13.4 kcal mol^{-1}. The crystal structure obtained from this completely flat ribbon consists of layers, as shown in Figure 6; although it is a minimum in the crystal potential energy hypersurface, its packing energy (16.6 kcal mol^{-1}) is significantly lower than that of the other two structures. Thus, corrugation of the layer (Figure 5a) allows a translational periodicity of only 4 Å and a better packing.

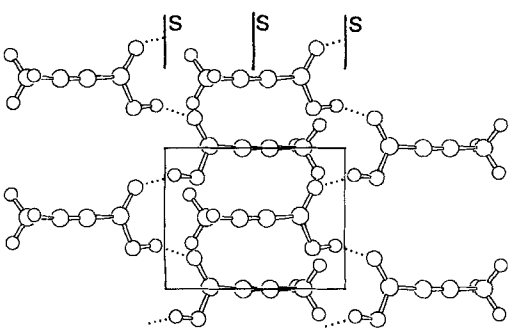

Figure 6. Hydrogen-bonding motif in the layers of the planar *trans*-CAT crystal structure (see Table 7). S denotes a screw axis.

Figure 7 shows an overview of the crystal structures generated by PROMET in the two space groups. As usual,[2, 21] the number of generated structures is very large, and increases at lower packing energies.

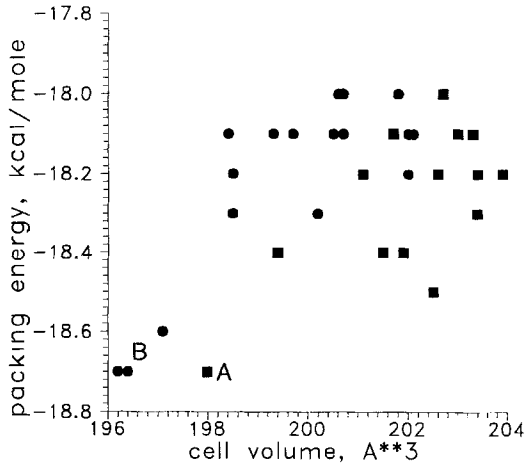

Figure 7. Scatterplot of packing energies for $P\bar{1}$ (●) and $P2_1$ (■) crystal structures of tetrolic acid, generated by PROMET, after screening for identical structures and cell reduction. **A** labels the observed $P2_1$ structure; the **B** cluster contains several $P\bar{1}$ structures, essentially identical in spite of screening.

For comparison, we also employed UPACK,[2] which uses a different crystal structure prediction method, based on a systematic search of the parameter space, and has now been extended to monoclinic and triclinic space groups, still assuming one molecule in the asymmetric unit (details will be published elsewhere). Briefly, rigid molecules are subjected systematically to translations and rotations in crystal cells whose dimensions are also variable, retaining an approximately constant density. The GROMOS force field (Table 2) was used, and a cutoff radius of 9 Å was used in the search procedure. Structures with energy lower than a given threshold were clustered, subjected to further rigid body energy minimization, and clustered again. The resulting independent structures were optimized under relaxation of all structural parameters (cutoff 12 Å). Only space groups $P\bar{1}$ and $P2_1$ were considered (11 and 8 variable parameters, respectively, at constant density). The lowest energy structure is in $P\bar{1}$, and structures within 5 kcal mol^{-1} of that are shown in Table 8. At less cohesive energies, structures without hydrogen bonds or that are otherwise improbable begin to appear.

The experimental structures are not the ones with lowest energies, but differences are hardly significant in view of the simplicity of the force field. There is considerable room for improvement, as can be seen from a comparison of calculated and observed cell parameters (Table 7). The corrugated structure found by PROMET (Table 7) was also found, with a slightly lower energy than the experimental one. However, in the

Table 8. Energy ordering for crystal structures generated by the UPACK procedure.

Space group $P\bar{1}$		Space group $P2_1$	
relative energy [a]	remarks [b]	relative energy [a]	remarks [b]
0	D	0.3	C
0	D	0.4	C "CORRUG"
0.1	D *C2/m*	0.5	C Expt
0.3	D	0.8	C
0.8	D	1.6	C
0.8	D	2.0	C
0.9	D Expt	2.9	C
1.0	D	3.3	C
1.1	D	3.7	C
1.7	D	4.0	C
4.1	C	4.2	C $P2_1/m$
4.2	C $P2_1/m$	4.3	C
		4.5	C

[a] In kcal mol^{-1}. [b] D: dimer; C: catemer; Expt: experimental X-ray structure. When a space group is given, the structure corresponds (within a small tolerance) to a special position in that space group; the two $P2_1/m$ structures are identical.

GROMOS force field the *trans*-CAT structure found by PROMET changed into the corrugated structure upon energy minimization. Catemer structures are very unlikely in $P\bar{1}$, while dimer structures cannot occur in $P2_1$ with one molecule in the asymmetric unit.

In a few cases (marked in Table 8) the structures correspond to special positions in space groups with higher symmetry. No doubt a search in other space groups would produce many more potential crystal structures.

Discussion and Conclusions

There is overwhelming experimental evidence for the formation of "precursors" in solution, either as aggregates or clusters,[22] and it is also possible that small clusters may assume different structures when passing from solution to mature crystals.[23–25] Keeping that in mind, the following considerations are the reasonable outcome of the different computational approaches:

1) Our calculations quite reasonably predict that in an apolar solvent the hydrogen-bonded cyclic dimer is the most persistent configuration, and yield the correct hydrogen bond breaking enthalpy; we conclude that the force field is adequate and that CAT events are not artifacts caused by bad parameterization.
2) CAT events are ubiquitous in all simulations with different solvent parameters; their frequency is about 4–5 events every 1000 ps, and the time spent in the CAT configuration is 10% or less.
3) The MD simulation is an effective tool for probing solution configurations, and demonstrates that both motifs found in the crystal structures are already present in solution. Although the DIM configuration largely predominates, the presence of *syn-cis*-CAT configurations, some of which are very similar to what is found in the crystal, proves that even short-lived species can nucleate and survive into crystal structures. However, the *trans*-CAT configuration is found in solution but does not grow into a crystal structure; appearance in solution is thus a necessary but not sufficient condition for nucleation and growth.
4) *anti* conformations are observed only in the free OH group at CAT events, but never for the OH engaged in the hydrogen bond. The discussion of this equilibrium is, however, unreliable based on just two molecules in the solute model.
5) In solution, short but significant fractions of time are spent in the planar *cis*-CAT configuration, although, as expected, CAT configurations are very flexible. Planarity of *cis*-CAT configurations is presumably completely enforced only at later stages of crystal organization or growth.
6) The computer generation of crystal structures leads to reconstruction of the two observed ones, thus confirming the effectiveness of both the PROMET and UPACK procedures. Besides these two, in all cases a number of other stable and close-packed structures is found: observed structures are recognized when cell parameters and space group are known, their packing energy being invariably among the highest but not always the highest one. In this case, the disorder in the carboxylic group can lower the free energy of the α phase by entropic effects. While such calculations may be useful as an aid in the interpretation of X-ray data,[26] predicting the crystal structure in an a priori manner—the fully ab initio crystal structure prediction—remains a major challenge for theoretical chemistry.
7) Combined results from MD and crystal structure generation demonstrate that although the planar *trans*-CAT configuration occurs in solution, it cannot grow into a full crystal because the three-dimensional packing of such a motif is less favourable (16.6 against 18.5 kcal mol^{-1} packing energy in PROMET; collapse to the "corrugated" structure in UPACK); this conclusion agrees with previous analyses.[13] The corrugated, nonplanar ribbon structure is calculated from two different force fields and computational procedures, and could therefore be a third, as yet unobserved, polymorph of tetrolic acid. However, subtle directional effects or cooperative stabilizing effects in planar continuous chains, which are beyond the scope of pairwise additive potentials used here, cannot be ruled out. These results are a substantial contribution to the understanding of the crystal packing of tetrolic acid.

Finally, we emphasize again that the simulation of precursors in solution is a promising complement to crystal structure generation procedures towards the prediction of crystal structures, offering at least a preliminary view of the nucleation kinetics. Clearly, however, more consistent modelling of the highly cooperative nucleation phenomenon should comprise a much larger number of solute molecules.

Acknowledgments: A European Community Human Capital and Mobility grant (No. ERB-CHRX-CT94-0469) is acknowledged for Milan – Utrecht exchanges and the stay of P. K. in Milan. A. G. acknowledges partial funding from Ministero dell'Università e della Ricerca Scientifica e Tecnologica. Crystal structure diagrams were drawn with the SCHAKAL program. We are grateful to Prof. W. F. van Gunsteren for his permission to use the GROMOS program.

Supplementary material available from the authors: Lennard-Jones parameters used in the molecular dynamics simulations; cell dimensions and fractional atomic coordinates for the most significant crystal structures generated by PROMET and UPACK.

Received: November 29, 1996 [F 531]

[1] A. Gavezzotti, *Acc. Chem. Res.* **1994**, *27*, 309–314.
[2] B. P. van Eijck, W. T. M. Mooij, J. Kroon, *Acta Crystallogr.* **1995**, *B51*, 99–103.
[3] *Theoretical Aspects and Computer Modeling of the Molecular Solid State* (Ed.: A. Gavezzotti), Wiley, Chichester, **1997**.
[4] A. Gavezzotti, G. Filippini, *J. Am. Chem. Soc.* **1995**, *117*, 12299–12305.
[5] W. L. Jorgensen, *Chemtracts Org. Chem.* **1991**, *4*, 91–119.
[6] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
[7] A. D. MacKerell, Jr., J. Wiockiewicz-Kuczera, M. Karplus, *J. Am. Chem. Soc.* **1995**, *117*, 11946–11975.
[8] W. F. van Gunsteren, H. J. C. Berendsen, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 992–1023.
[9] V. Benghiat, L. Leiserowitz, *J. Chem. Soc.* **1972**, 1763–1768.
[10] F. Holtzberg, B. Post, I. Fankuchen, *Acta Crystallogr.* **1953**, *6*, 127–130.
[11] P.-G. Jönsson, *Acta Crystallogr.* **1971**, *B27*, 893–898.
[12] J. L. Derissen, P. H. Smit, *Acta Crystallogr.* **1974**, *B30*, 2240–2242.
[13] L. Leiserowitz, *Acta Crystallogr.* **1976**, *B32*, 775–802.
[14] Z. Berkovitch-Yellin, L. Leiserowitz, *J. Am. Chem. Soc.* **1982**, *104*, 4052–4064.
[15] A. Gavezzotti, G. Filippini, *J. Phys. Chem.* **1994**, *98*, 4831–4837.
[16] W. F. van Gunsteren, *GROMOS, Groningen Molecular Simulation Program Package*, University of Groningen (The Netherlands), **1987**.
[17] J. E. Coon, S. Gupta, E. McLaughlin, *Chem. Phys.* **1987**, *113*, 43–52.
[18] D. W. Rebertus, B. J. Berne, *J. Chem. Phys.* **1979**, *70*, 3395–3400.
[19] A. Gavezzotti, PROMET 5, *A Program for the Generation of Possible Crystal Structures from the Molecular Structure of Organic Compounds*, University of Milan (Italy), **1995**.
[20] a) S. Murata, M. Sakiyama, S. Seki, *J. Chem. Thermodyn.* **1982**, *14*, 723–731; b) C. H. D. Calis-Van Ginkel, G. H. M. Calis, C. W. M. Timmermans, C. G. de Kruif, H. A. J. Oonk, *ibid.* **1978**, *10*, 1083–1088.
[21] A. Gavezzotti, *Acta Crystallogr.* **1996**, *B52*, 201–208.
[22] M. A. Larson, in *Advances in Industrial Crystallization* (Eds.: J. Garside, R. J. Davey, A. G. Jones), Butterworth/Heinemann, Oxford, **1991**.
[23] S. P. Weinbach, K. Kjaer, W. B. Bouwman, G. Gruebel, J-F. Legrand, J. Als-Nielsen, M. Lahav, L. Leiserowitz, *Science* **1994**, *264*, 1566–1570.
[24] I. Weissbuch, L. Leiserowitz, M. Lahav, *Adv. Mater.* **1994**, *6*, 952–956.
[25] G. Corongiu, E. Clementi, in *Computational Approaches in Supramolecular Chemistry* (Ed.: G. Wipff), NATO ASI Series, Kluwer, Dordrecht, **1994**.
[26] A. Gavezzotti, G. Filippini, *J. Am. Chem. Soc.* **1996**, *118*, 7153–7157.
[27] D. E. Williams, *PCK83: QCPE Program 548*, Quantum Chemistry Program Exchange, Chemistry Department, Indiana University, Bloomington, Indiana, **1983**.
[28] E. Keller, *SCHAKAL92, A Program for the Graphic Representation of Molecular and Crystallographic Models*, University of Freiburg (Germany), **1993**.