Available online at www.sciencedirect.com
ScienceDirect
Mendeleev Commun., 2021, **31**, 81–83

Mendeleev Communications

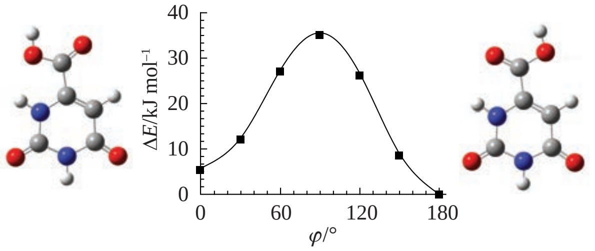
Equilibrium molecular structure of orotic acid from gas-phase electron diffraction data and quantum-chemical calculations

Ekaterina P. Altova,*[a,b] Anatolii N. Rykov,[a] Natalja Vogt*[a,b] and Igor F. Shishkov[a]

[a] *Department of Chemistry, M. V. Lomonosov Moscow State University, 119991 Moscow, Russian Federation. E-mail: altovaek@gmail.com*
[b] *Section of Chemical Informational Systems, University of Ulm, D-89081 Ulm, Germany. E-mail: natalja.vogt@uni-ulm.de*

DOI: 10.1016/j.mencom.2021.01.025

The accurate equilibrium structures of planar *anti* and *syn* orotic acid conformers with different positions of the carboxyl group with respect to the C=C bond were determined from gas-phase electron diffraction data using the dynamic model and taking into account anharmonic vibrational corrections calculated with *ab initio* force field. The high accuracy of coupled-cluster computations was exploited in the analysis of structural effects due to electron-acceptor substituents.



Keywords: equilibrium molecular structures, *syn* and *anti* conformers, pseudo-conformer model, orotic acid, gas-phase electron diffraction, coupled-cluster computations.

Orotic acid (6-carboxyuracil, vitamin B13) is a key precursor in the biosynthesis of pyrimidine nucleobases.[1] However, its molecular structure has not been accurately determined experimentally.[2] Two orotic acid conformers with a *synperiplanar* conformation of the carboxyl group and the *anti* and *syn* positions of a carbonyl group with respect to the C=C bond of the uracil ring were detected in the matrix isolation IR spectra. Their amounts were estimated at 85 and 15%, respectively, from the experimental intensities of bands due to ν(O–H) and ν(N–H).[3] The presence of a negligibly small amount of the further *anti* conformer with the *anticlinal* conformation of the carboxyl group was predicted by quantum-chemical calculations (for instance, 1.8% at 480 K, M06/aug-cc-pVTZ).[3] The aim of this work was to determine the accurate equilibrium structures of the lowest energy orotic acid conformers by gas-phase electron diffraction (GED) and high-level *ab initio* computations.

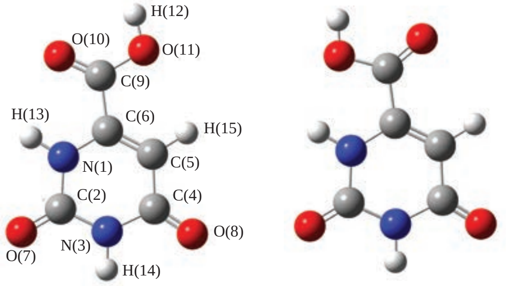
The GED experiment was carried out on an EG-100M apparatus with an accelerating voltage of ~60 kV for the long (LD = 362.28 mm) and short (SD = 193.94 mm) nozzle-to-film distances at a nozzle temperature of 520(3) K.[†] A commercial sample (purity, 98%) was used without further purification. The total electron scattering intensities $I(s)$[†] were obtained by the initial processing of electron diffraction patterns using a standard algorithm.[4]

The quantum-chemical calculations (MP2[5] and B3LYP[6,7]) were carried out using the Gaussian program package,[8] except for the CCSD(T)[9,10] structure optimization performed with the Molpro program.[11,12] According to the MP2 and B3LYP calculations (with the cc-pVTZ basis set[13]), the lowest energy *anti* and *syn* conformers of orotic acid are characterized by the s-*trans* and s-*cis* conformations of the O(10)=C(9)–C(6)=C(5) fragment, respectively; they have the *synperiplanar* carboxyl group and the planar overall structures of C_s symmetry (Figure 1). These conformers are separated by a very high barrier to internal rotation of the carboxyl group around the C(6)–C(9) bond (~ 35.6 kJ mol⁻¹, MP2/cc-pVTZ).[†] The planar structures were also optimized at the CCSD(T) level of theory with all electrons being correlated (ae) in conjunction with the cc-pwCVTZ basis set.[14] The final structure of the CCSD(T)_ae/cc-pwCVQZ quality (Table 1) was estimated using the additivity approximation

$$r_e^{BO} = r_e(\text{CCSD(T)\_ae/cc-pwCVTZ}) + \Delta r_{T\to Q}, \qquad (1)$$

where structural changes due to the enlargement of the basis set from triple- to quadruple-ζ size $\Delta r_{T\to Q} = r(\text{cc-pwCVQZ}) - r(\text{cc-pwCVTZ})$ were estimated at the MP2_ae level.

The harmonic and anharmonic (cubic) force fields, which are required for the calculations of vibrational corrections to the experimental internuclear distances $r_{a,ij}$ in order to derive the semiexperimental equilibrium structure, r_e^{se}, were computed at the MP2/cc-pVTZ level. The total corrections $\Delta r_{ij} = r_{a,ij} - r_{e,ij}^{se}$ including harmonic shrinkages, kinematic and dynamic anharmonicity effects and centrifugal distortions[†] were calculated according to Sipachev.[15–17] An attempt to describe the experimental data with a mixture of two conformers did not lead to satisfactory



anti　　*syn*

Figure 1 Molecular models of two lowest-energy conformers of orotic acid.

[†] See details in Online Supplementary Materials.

Table 1 Equilibrium structures of the *anti* and *syn* conformers of orotic acid from the GED data and *ab initio* calculations (bond lengths in Å and angles in °).

Distances and angles	r_eGED ^{a,b}	r_aGED ^{a}	$r_e^{BO\ c}$	r_eGED ^{a,b}	r_aGED ^{a}	$r_e^{BO\ c}$
		anti			*syn*	
C(2)–N(1)	1.377(5)^3	1.389(5)	1.377	1.379(5)^3	1.391(5)	1.379
N(3)–C(2)	1.380(5)^3	1.390(5)	1.381	1.378(5)^3	1.388(5)	1.378
C(4)–N(3)	1.394(5)^3	1.404(5)	1.394	1.395(5)^3	1.406(5)	1.395
C(5)–C(4)	1.463(3)^4	1.473(3)	1.463	1.462(3)^4	1.471(3)	1.461
C(6)–N(1)	1.371(5)^3	1.382(5)	1.372	1.375(5)^3	1.386(5)	1.375
C(6)=C(5)	1.348(13)^6	1.355(13)	1.347	1.348(13)^6	1.355(13)	1.346
O(7)=C(2)	1.209(1)^5	1.213(1)	1.208	1.209(1)^5	1.213(1)	1.209
O(8)=C(4)	1.212(1)^5	1.215(1)	1.211	1.211(1)^5	1.215(1)	1.211
C(9)–C(6)	1.490(3)^4	1.500(3)	1.490	1.493(3)^4	1.503(3)	1.492
O(10)=C(9)	1.205(1)^5	1.209(1)	1.204	1.199(1)^5	1.203(1)	1.198
O(11)–C(9)	1.337(13)^6	1.347(13)	1.335	1.349(13)^6	1.358(13)	1.347
H(12)–O(11)	0.971(9)^7	0.986(9)	0.966	0.971(9)^7	0.986(9)	0.966
H(13)–N(1)	1.013(9)^7	1.029(9)	1.008	1.011(9)^7	1.027(9)	1.006
H(14)–N(3)	1.014(9)^7	1.030(9)	1.009	1.014(9)^7	1.030(9)	1.009
H(15)–C(5)	1.081(9)^7	1.096(9)	1.076	1.082(9)^7	1.097(9)	1.076
N(3)–C(2)–N(1)	114.0(8)^8		113.4	114.2(8)^8		113.6
C(4)–N(3)–C(2)	127.7(4)^9		128.5	127.7(4)^9		128.4
C(5)–C(4)–N(3)	113.7(6)^{10}		113.3	113.6(6)^{10}		113.3
O(7)=C(2)–N(1)	123.4(11)^{11}		123.3	123.0(11)^{11}		123.0
O(8)=C(4)–C(5)	125.4(11)^{11}		125.4	125.6(11)^{11}		125.6
C(9)–C(6)–N(1)	110.8(7)^{12}		113.1	115.1(7)^{12}		117.4
O(10)=C(9)–C(6)	122.5(4)^{13}		122.4	124.5(4)^{13}		124.5
O(11)–C(9)–C(6)	113.1(4)^{13}		113.0	111.4(4)^{13}		111.4
H(12)–O(11)–C(9)	106.7^d		106.7	106.9^d		106.9
H(13)–N(1)–C(2)	118.1^d		118.1	116.7^d		116.7
H(14)–N(3)–C(2)	115.1^d		115.1	115.2^d		115.2
H(15)–C(5)–C(4)	119.0^d		119.0	119.6^d		119.6

^{a} Values in parentheses are estimated total errors (2.5σ). ^{b} Parameters with equal superscripts were refined in one group; differences between parameters in each group were assumed at the values of the r_e^{BO} structure. ^{c} Computed structure of CCSD(T)_ae/cc-pwCVQZ quality [see equation (1)]. ^{d} Assumed at the computed value.

results (R_f = 4.2%). Alternatively, the model of pseudo-conformers was applied in the GED analysis. The lowest frequency vibration [~67 cm^{-1} (*anti*) and ~57 cm^{-1} (*syn*), MP2/cc-pVTZ] is a pure mode describing the carboxyl group rotation around the C(6)–C(9) bond (a potential energy distribution for torsion in this mode is about 90%). Therefore, it could be separated from other vibrations. The weighting of pseudo-conformers was carried out according to a potential energy function from MP2/cc-pVTZ calculations.† Table 1 summarizes the structural analysis data and the determined structures of *anti* and *syn* conformers. The application of the model of pseudo-conformers reduced the disagreement factor R_f by 0.9% (to 3.3%) due to a strong anharmonicity of the dynamic mode.[18] Figure 2 shows the radial distribution functions obtained by Fourier transformation of the final molecular electron scattering intensities $sM(s)$.†

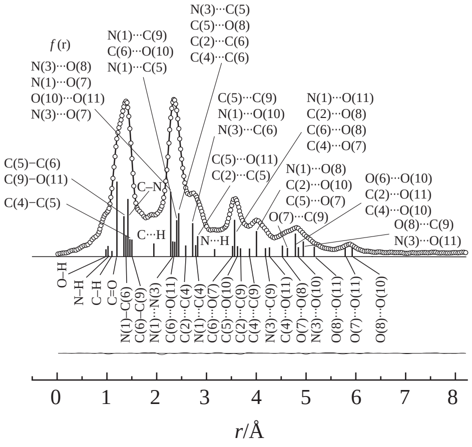


Figure 2 Experimental radial distribution curve $f(r)_{exp}$ (open circles) derived with a damping factor of exp(0.002 s^2) and its theoretical counterpart $f(r)_{theor}$ (solid line) for the dynamic model of orotic acid with vertical bars of molecular terms. Difference curve, $\Delta f(r) = f(r)_{exp} - f(r)_{theor}$.

For both conformers (see Table 1), the computed structures of CCSD(T)_ae/cc-pwCVQZ quality, r_e^{BO}, are consistent with the r_e^{se} ones to within a few thousandths of Å and a few tenths of degree for the bond lengths and angles, respectively. The high accuracy of this quality r_e^{BO} structure was confirmed in many studies.[19–22] The high-accuracy computed structure was exploited in the analysis of fine structural effects due to electron-acceptor substituents. The bond angle in the ring at the carboxyl group $\angle_e^{BO}$ [C(5)=C(6)–N(1)] = 122.6° (*anti* conformer) increased by 0.9° in comparison to this angle in uracil ($\angle_e^{BO}$ [C(5)=C(6)–N(1)] = = 121.7°).[20] Note that the above effect was observed in 5-fluorouracil due to the fluorine substituent. The angle $\angle_e^{BO}$ [C(4)–C(5)–C(6)] in 5-fluorouracil[21] also increased (by 1.5°) with respect to this angle in uracil.[20]

This work was supported by the Dr. B. Mez-Starck Foundation (Germany) and the Russian Foundation for Basic Research (grant no. 20-03-00747).

Online Supplementary Materials

Supplementary data associated with this article can be found in the online version at doi: 10.1016/j.mencom.2021.01.025.

References

1 J. Victor, L. B. Greenberg and D. L. Sloan, *J. Biol. Chem.*, 1979, **254**, 2647.
2 N. Vogt and J. Vogt, *Structure Data of Free Polyatomic Molecules*, Springer, 2019.

3 R. Wysokińnski, K. Helios, L. Lapinski, M. J. Nowak and D. Michalska, *Vib. Spectrosc.*, 2013, **64**, 108.

4 Y. V. Vishnevskiy, *J. Mol. Struct.*, 2007, **833**, 30.

5 C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618.

6 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.

7 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.

8 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09*, Gaussian, Wallingford, CT, 2010.

9 G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.*, 1982, **76**, 1910.

10 K. Raghavachari, G. W. Trucks, J. A. Pople and M. Head-Gordon, *Chem. Phys. Lett.*, 1989, **157**, 479.

11 H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby and M. Schütz, *WIREs Comput. Mol. Sci.*, 2012, **2**, 242.

12 H.-J. Werner, P. J. Knowles, F. R. Manby, M. Schütz, P. Celani, G. Knizia, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklaß, P. Palmieri, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni and T. Thorsteinsson, *MOLPRO*, 2009.

13 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007.

14 K. A. Peterson and T. H. Dunning, *J. Chem. Phys.*, 2002, **117**, 10548.

15 V. A. Sipachev, *J. Mol. Struct.: THEOCHEM*, 1985, **121**, 143.

16 V. A. Sipachev, *Struct. Chem.*, 2000, **11**, 167.

17 V. A. Sipachev, *J. Mol. Struct.*, 2001, **567**–**568**, 67.

18 I. V. Kochikov, Yu. I. Tarasov, N. Vogt and V. P. Spiridonov, *J. Mol. Struct.*, 2002, **607**, 163.

19 N. Vogt, J. Demaison, D. N. Ksenafontov and H. D. Rudolph, *J. Mol. Struct.*, 2014, **1076**, 483.

20 N. Vogt, L. S. Khaikin, O. E. Grikina and A. N. Rykov, *J. Mol. Struct.*, 2013, **1050**, 114.

21 L. S. Khaikin, N. Vogt, A. N. Rykov, O. E. Grikina, J. Vogt, I. V. Kochikov, E. S. Ageeva and I. F. Shishkov, *Mendeleev Commun.*, 2018, **28**, 236.

22 N. Vogt, D. N. Ksenafontov, D. Savelev and A. N. Rykov, *Mendeleev Commun.*, 2020, **30**, 660.

Received: 21st October 2020; Com. 20/6344