

L-Tryptophan formic acid solvate at  
183 K

Christian B. Hübschle, Birger Dittrich and Peter Luger\*

Institut für Chemie, Kristallographie, Freie Universität Berlin, Takustraße 6,  
14195 Berlin, Germany

Correspondence e-mail: luger@chemie.fu-berlin.de

Received 18 June 2002

Accepted 11 July 2002

Online 10 August 2002

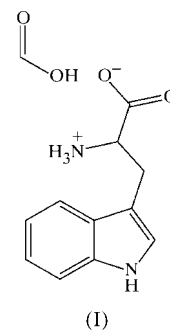
In the title compound,  $C_{11}H_{12}N_2O_2 \cdot CH_2O_2$ , at 183 K. L-tryptophan appears in the zwitterionic form, while the formic acid molecule is neutral. The formic acid molecule is the donor in a strong  $O-H \cdots O$  hydrogen bond to the carboxylate group of the tryptophan molecule, with a short  $O \cdots O$  contact of 2.487 (2) Å.

## Comment

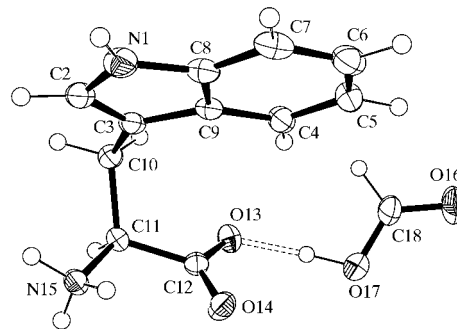
In conjunction with our ongoing work of comparative charge-density studies on the 20 naturally occurring amino acids, we have directed our interest to tryptophan, Trp, one of the essential  $\alpha$ -amino acids for humans. Compared with other amino acids, there are only a few tryptophan structures listed in the Cambridge Structural Database (Version 5.23, April 2002; Allen & Kennard, 1993), due to the difficulty of obtaining good quality crystals. The only solvent-free non-substituted tryptophan structure is that of DL-tryptophane (Bakke & Mostad, 1980), where the authors reported that 'crystals were not as good as could be wished'. While our attempts to obtain good quality crystals suitable for charge-density experiments also failed for pure tryptophan (for both the DL- and L-forms), we obtained nicely diffracting crystals of L-tryptophan formic acid solvate, (I), the X-ray structure of which was still unknown. Only the crystal structure of DL-tryptophan formate has been reported previously (Bye *et al.*, 1973), where it was shown that the tryptophan molecule was protonated and the formate anion had a deprotonated carboxyl group.

In contrast with these results, we found from our X-ray analysis at 183 K that, in the present case, the L-tryptophan is zwitterionic, as are most amino acids in the crystal form, and that the formic acid is neutral (Fig. 1). There is a strong  $O17-H17 \cdots O13$  hydrogen bond from the formic acid to the carboxylate group of L-tryptophan, with a remarkably short  $O \cdots O$  contact of 2.487 (2) Å (Table 2). In DL-tryptophan formate, a hydrogen bond with almost the same donor-acceptor  $O \cdots O$  distance was observed (2.492 Å). However, as already mentioned, the role of donor and acceptor was inverted in that case.

The opposite donor-acceptor situation is reflected in the C—O bond lengths at these sites. Normally, a charged carboxylate group has two more or less equal C—O bonds of  $\sim 1.23$ – $1.25$  Å, while in a neutral COOH group, the two C—O bond lengths differ by  $\sim 0.1$  Å. Due to the strong hydrogen bonds in both the L- and the DL-Trp formic acid derivatives, the C—O bond lengths in the COOH groups are less different [1.281 (3) and 1.213 (3) Å in the formic acid of the L-Trp derivative, and 1.295 and 1.214 Å in the carboxyl group of the DL-Trp structure]. On the other hand, the accepting C—O bonds in the carboxylate groups are lengthened [1.273 (2) versus 1.237 (2) Å for L-Trp, and 1.255 versus 1.232 Å in the  $COO^-$  group of the formate anion of the DL-tryptophan derivative]. The other bond lengths in the zwitterionic and cationic forms of L-Trp and DL-Trp are not affected and need no further discussion.



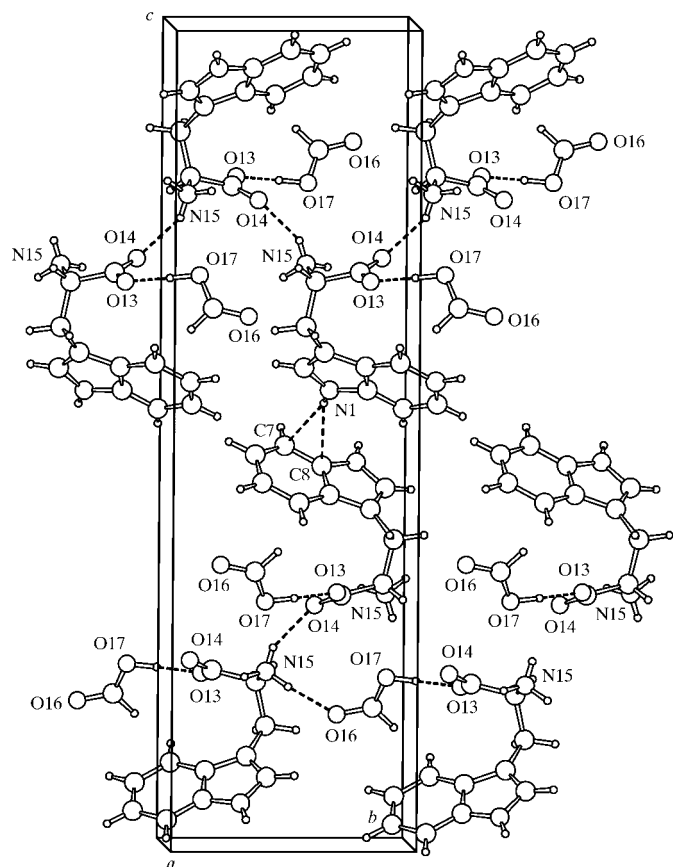
The overall molecular conformation, *i.e.* the relative orientation of the side chain with respect to the indole ring system, can best be described by the torsion angles along the C3—C10 and C10—C11 bonds (Table 3). While  $\chi^{2.1}$  [C2—C3—C10—C11 =  $-104.4$  (2) $^\circ$ ] has a similar value to that of L-tryptophan in the DL-tryptophan structure, the torsion angles  $\chi^1$  and  $\chi^{1.2}$  [for definition of nomenclature, see IUPAC-IUB (1970)] along C10—C11, indicating (+)- and (–)*gauche* arrangements, are different from both the free tryptophan molecular structure and the tryptophan cation in DL-tryptophan formate. Bakke & Mostad (1980) have summarized torsion angles for tryptophan derivatives and found that, in



**Figure 1**  
The molecular structure and atom-numbering scheme of (I). Displacement ellipsoids are plotted at the 50% probability level and H atoms are shown as small spheres of arbitrary radii.

most cases, C12 is *trans* to C3. The molecular conformation of (I) is roughly similar to that of the hydrochloride of DL-tryptophan ethyl ester (Vijayalakshmi & Srinivasan, 1975). Due to the *cis* arrangement of the carboxylate group relative to the ring system and the strong hydrogen bonding to the formic acid molecule, a U-shaped arrangement of the entire donor-acceptor complex is formed.

In addition to the above-mentioned strong O17—H17···O13 hydrogen bond from the formic acid to the tryptophan zwitterion, three further hydrogen bonds exist, involving the H atoms of the amino group as donors. Two link tryptophan molecules and the third is a further weaker tryptophan—formic acid hydrogen bond with the donor on the tryptophan side. These hydrogen bonds establish a network mainly in the *x* and *y* directions close to  $z = \frac{1}{4}$ , while double layers of the indole residues assemble close to  $z = 0, z = \frac{1}{2}, \dots$  (Fig. 2). Neighbouring indole layers have weak N1—H1···C<sup>i</sup> linkages as short contacts, with  $H1 \cdots C7^i = 2.50$  (3) Å [ $N1 \cdots C7^i = 3.314$  (2) Å] and  $H1 \cdots C8^i = 2.56$  (3) Å [ $N1 \cdots C8^i = 3.351$  (2) Å; symmetry code: (i)  $\frac{1}{2} + x, \frac{1}{2} - y, 2 - z$ ]. An extensive overview of  $X-H \cdots \pi$  interactions has been given by Desiraju & Steiner (1999), where  $N-H \cdots \pi$  interactions are also discussed. No further close contacts of interest were found in (I).



**Figure 2**  
Packing diagram for (I) viewed along *a*.

tophan zwitterion, three further hydrogen bonds exist, involving the H atoms of the amino group as donors. Two link tryptophan molecules and the third is a further weaker tryptophan—formic acid hydrogen bond with the donor on the tryptophan side. These hydrogen bonds establish a network mainly in the *x* and *y* directions close to  $z = \frac{1}{4}$ , while double layers of the indole residues assemble close to  $z = 0, z = \frac{1}{2}, \dots$  (Fig. 2). Neighbouring indole layers have weak N1—H1···C<sup>i</sup> linkages as short contacts, with  $H1 \cdots C7^i = 2.50$  (3) Å [ $N1 \cdots C7^i = 3.314$  (2) Å] and  $H1 \cdots C8^i = 2.56$  (3) Å [ $N1 \cdots C8^i = 3.351$  (2) Å; symmetry code: (i)  $\frac{1}{2} + x, \frac{1}{2} - y, 2 - z$ ]. An extensive overview of  $X-H \cdots \pi$  interactions has been given by Desiraju & Steiner (1999), where  $N-H \cdots \pi$  interactions are also discussed. No further close contacts of interest were found in (I).

## Experimental

Crystals of (I) were grown by cooling a hot solution of L-tryptophan (purchased from Sigma) in a mixture of propan-2-ol and a small amount of formic acid (purchased from Merck AG).

### Crystal data

$C_{11}H_{12}N_2O_2 \cdot CH_2O_2$   
 $M_r = 250.25$   
Orthorhombic,  $P2_12_12_1$   
 $a = 5.3163$  (3) Å  
 $b = 8.1348$  (4) Å  
 $c = 27.259$  (2) Å  
 $V = 1178.87$  (12) Å<sup>3</sup>  
 $Z = 4$   
 $D_x = 1.410$  Mg m<sup>-3</sup>

Mo  $K\alpha$  radiation  
Cell parameters from 5433 reflections  
 $\theta = 2.1$ – $32.4^\circ$   
 $\mu = 0.11$  mm<sup>-1</sup>  
 $T = 183$  (2) K  
Needle, colourless  
 $0.2 \times 0.1 \times 0.1$  mm

### Data collection

Bruker SMART CCD area-detector diffractometer  
 $\omega$  and  $\varphi$  scans  
Absorption correction: empirical (SADABS; Blessing, 1995; Siemens, 1996)  
 $T_{\min} = 0.761, T_{\max} = 1.000$   
14 392 measured reflections

2492 independent reflections  
2104 reflections with  $I > 2\sigma(I)$   
 $R_{\text{int}} = 0.064$   
 $\theta_{\max} = 33.1^\circ$   
 $h = -8 \rightarrow 8$   
 $k = -11 \rightarrow 12$   
 $l = -41 \rightarrow 41$

### Refinement

Refinement on  $F^2$   
 $R[F^2 > 2\sigma(F^2)] = 0.051$   
 $wR(F^2) = 0.126$   
 $S = 1.06$   
2492 reflections  
219 parameters  
All H-atom parameters refined

$w = 1/[\sigma^2(F_o^2) + (0.0702P)^2 + 0.0077P]$   
where  $P = (F_o^2 + 2F_c^2)/3$   
 $(\Delta/\sigma)_{\max} < 0.001$   
 $\Delta\rho_{\max} = 0.37$  e Å<sup>-3</sup>  
 $\Delta\rho_{\min} = -0.24$  e Å<sup>-3</sup>

All H atoms were found in difference Fourier maps and were refined freely; C—H distances were in the range 0.91 (3)–1.01 (3) Å.

**Table 1**

Selected geometric parameters (Å, °).

N1—C8	1.366 (3)	C10—C11	1.540 (3)
N1—C2	1.371 (3)	C11—N15	1.491 (3)
C2—C3	1.368 (3)	C11—C12	1.524 (3)
C3—C9	1.446 (3)	C12—O14	1.237 (2)
C3—C10	1.496 (3)	C12—O13	1.273 (2)
C5—C6	1.405 (3)	O16—C18	1.213 (3)
C6—C7	1.382 (3)	O17—C18	1.281 (3)
C8—N1—C2	109.31 (17)	C7—C8—C9	122.2 (2)
C3—C2—N1	110.36 (19)	N15—C11—C12	109.39 (16)
N1—C8—C7	130.23 (19)	O14—C12—O13	126.46 (19)
N1—C8—C9	107.57 (18)	O16—C18—O17	124.4 (2)

**Table 2**

Hydrogen-bonding geometry (Å, °).

$D-H \cdots A$	$D-H$	$H \cdots A$	$D \cdots A$	$D-H \cdots A$
N15—H15A···O13 <sup>i</sup>	0.89 (3)	2.12 (3)	2.938 (2)	153 (2)
N15—H15A···O14	0.89 (3)	2.25 (3)	2.641 (2)	106 (2)
N15—H15B···O16 <sup>ii</sup>	0.95 (3)	1.85 (3)	2.786 (2)	167 (3)
N15—H15C···O14 <sup>iii</sup>	0.82 (3)	2.01 (3)	2.750 (2)	151 (2)
O17—H17···O13	0.98 (3)	1.51 (3)	2.487 (2)	171 (4)

Symmetry codes: (i)  $1 + x, y, z$ ; (ii)  $1 + x, y - 1, z$ ; (iii)  $1 - x, y - \frac{1}{2}, \frac{3}{2} - z$ .

**Table 3**

Selected torsion angles ( $^{\circ}$ ) for L-tryptophan formic acid solvate, (I), DL-tryptophan formate, (II), and DL-tryptophane, (III). All angles refer to the *S* configuration at C11, *i.e.* the L-form.

Torsion	Nomenclature <sup>†</sup>	(I) <sup>‡</sup>	(II) <sup>§</sup>	(III) <sup>¶</sup>
C2—C3—C10—C11	$\chi^{2,1}$	-104.4 (2)	105.1	-106.6
C3—C10—C11—C12	$\chi^{1,2}$	-69.8 (2)	-174.6	68.6
C3—C10—C11—N15	$\chi^1$	53.5 (2)	-53.7	-168.6
N15—C11—C12—O13	$\psi^2$	172.45 (15)	-175.3	-19.5
N15—C11—C12—O14	$\psi^1$	-9.1 (2)	-0.5	156.0
C10—C11—C12—O13		-63.2 (2)	-52.6	101.9

<sup>†</sup> For definition of nomenclature, see IUPAC-IUB (1970). <sup>‡</sup> This work. <sup>§</sup> Bye *et al.* (1973). <sup>¶</sup> Bakke & Mostad (1980).

Data collection: *SMART* (Siemens, 1996); cell refinement: *SMART*; data reduction: *SAINT* (Siemens, 1996; program(s) used to solve structure: *SHELXS97* (Sheldrick, 1997); program(s) used to refine structure: *SHELXL97* (Sheldrick, 1997); molecular graphics: *ORTEPIII* (Burnett & Johnson, 1996) and *SCHAKAL99* (Keller & Pierrard, 1999); software used to prepare material for publication: *PLATON* (Spek, 1990).

Supplementary data for this paper are available from the IUCr electronic archives (Reference: SK1566). Services for accessing these data are described at the back of the journal.

## References

- Allen, F. H. & Kennard, O. (1993). *Chem. Des. Autom. News*, **8**, 1, 31–37.
- Bakke, O. & Mostad, A. (1980). *Acta Chem. Scand. Ser. B*, **34**, 559–570.
- Blessing, R. H. (1995). *Acta Cryst. A* **51**, 33–38.
- Burnett, M. N. & Johnson, C. K. (1996). *ORTEPIII*. Report ORNL-6895. Oak Ridge National Laboratory, Tennessee, USA.
- Bye, E., Mostad, A. & Rømming, C. (1973). *Acta Chem. Scand.* **27**, 471–484.
- Desiraju, G. R. & Steiner, T. (1999). *The Weak Hydrogen Bond in Structural Chemistry and Biology*. New York: Oxford University Press Inc.
- IUPAC-IUB Commission on Biochemical Nomenclature (1970). *J. Mol. Biol.* **52**, 1–17.
- Keller, E. & Pierrard, J.-S. (1999). *SCHAKAL99*. University of Freiburg, Germany.
- Sheldrick, G. M. (1997). *SHELXS97* and *SHELXL97*. University of Göttingen, Germany.
- Siemens (1996). *SMART* (Version 5.054), *SAINT* (Version 6.28A) and *SADABS* (Version 2.03). Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.
- Spek, A. L. (1990). *Acta Cryst. A* **46**, C-34.
- Vijayalakshmi, B. K. & Srinivasan, R. (1975). *Acta Cryst. B* **31**, 999–1003.