Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Atomic volumes and charges in a system with a strong hydrogen bond: L-tryptophan formic acid

S. Scheins,[a] B. Dittrich,[a] M. Messerschmidt,[a] C. Paulmann[b] and P. Luger[a]*

[a]Institut für Chemie/Kristallographie der Freien Universität Berlin, Takustrasse 6, Berlin D-14195, Germany, and [b]Mineralogisch-Petrologisches Insitut, Universität Hamburg, Grindelallee 48, Hamburg D-20146, Germany

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

Received 24 June 2003
Accepted 22 January 2004

A high-resolution X-ray diffraction data set was collected within 48 h at 100 K with synchrotron radiation and area detection. A full topological analysis was applied to the resulting electron-density model. This analysis was followed by a Bader partitioning making use of the zero-flux surfaces of the electron-density gradient vector field. The atomic and bonding properties obtained were compared with the results of previous experimental studies, and with theoretical calculations for the title complex and free tryptophan as reported in the literature. The agreement between experiment and theory is similar to the agreement between different theoretical calculations. There is no charge transfer *via* the strong hydrogen bond; however, its strength is indicated by the very small atomic volume of the H atom involved.

1. Introduction

Thanks to recent technical advances in X-ray crystallography, especially the introduction of CCD area detectors, the time needed to collect high-resolution diffraction data for experimental charge-density determination has been reduced to one or a few days rather than the weeks or even months required with point detectors. Such improvements made charge-density studies on larger molecules and comparative investigations of an entire class of compounds possible. The latter type of study has been carried out for the 20 naturally occurring amino acids, and charge-density and related topological data for 13 of the 20 compounds have been published. For the remaining compounds, it has not been possible to grow crystals of sufficient quality to allow a successful experimental charge-density determination.

Moderate-quality crystals of tryptophan have been obtained previously, the only solvent-free non-substituted tryptophan structure reported in the literature being that of DL-tryptophan (Bakke & Mostad, 1980). These authors mentioned poorly diffracting crystals, which resulted in a data set with more than 70% of the reflections being unobserved. Until now, our attempts to obtain crystals suitable for a charge-density data collection also failed for pure tryptophan (for both the DL- and the L-form); however, we have obtained nicely diffracting crystals of L-tryptophan formic acid solvate, (I) (Hübschle *et al.*, 2002). We therefore directed our interest to this system, which gave us, in addition, the opportunity to study topological details of a short O—H···O hydrogen bond. We compare the results of this charge-density study with experimental results for the amino acids that have been studied previously.

In a series of papers, Matta & Bader (2000, 2002, 2003), have published topological and atomic properties of all 20 genetically encoded amino acids, the data being based on theoretical calculations. This class of compounds is thus the first for which a complete set of theoretical charge-density data is available and for which the corresponding experimental studies are approaching completeness.

Structural information on an electronic scale obtained in this way is very useful, since such data can be used to model electron-density details of large systems of biological interest (Jelsch *et al.*, 1998, 2000; Pichon-Pesme & Lecomte, 1998; Dahaoui *et al.*, 1999).

2. Experimental

Crystals were grown from commercially available L-tryptophan by cooling a hot solution of the amino acid in a mixture of propan-2-ol and a small amount of formic acid. The X-ray data set was collected within 48 h at beamline F1 of HASYLAB/DESY, Hamburg, Germany, on a κ-axis diffractometer equipped with a Bruker SMART 1K CCD detector. The temperature was maintained at 100 (1) K using an Oxford Cryosystems nitrogen-gas-stream cooling device. The measurement strategy was planned with *ASTRO*, and the *SMART* and *SAINT* routines (Bruker, 1997–2001) were used for data collection and integration. 62 915 reflections were measured up to a resolution of $\sin\theta/\lambda = 1.38$ Å^{-1} (or $d = 0.36$ Å), thus giving 12 305 unique reflections. Further details of the crystal data and the experimental conditions are given in Table 1.[1]

Table 1
Experimental details.

Crystal data	
Chemical formula	$C_{11}H_{12}N_2O_2{\cdot}CH_2O_2$
M_r	250.25
Cell setting, space group	Orthorhombic, $P2_12_12_1$
a, b, c (Å)	5.2951 (3), 8.1213 (4), 27.256 (2)
V (Å^3)	1172.1 (1)
Z	4
D_x (Mg m^{-3})	1.418 (1)
Radiation type	Synchrotron
Wavelength (Å)	0.54
No. of reflections for cell parameters	20 000
θ range (°)	2.6–48.0
$F(000)$	528.0
μ (mm^{-1})	0.07
Temperature (K)	100 (1)
Crystal form, color	Block, colorless
Crystal size (mm)	0.40 × 0.35 × 0.30
Data collection	
Diffractometer	Bruker SMART 1000 CCD detector at DESY HASYLAB beamline F1 on κ-axis diffractometer
Data collection method	ω and φ scans
Absorption correction	None
No. of measured, independent and observed reflections	62 915, 12 305, 10 905
Criterion for observed reflections	$I > 2.5\,\sigma(I)$
Completeness overall, up to $\sin\theta/\lambda = 1.2$ Å (%)	87.1, 97.9
Redundancy	5.1
R_{int}	0.030
θ_{max} (°)	48.2
Range of h, k, l	$-14 \Rightarrow h \Rightarrow 7$ $-22 \Rightarrow k \Rightarrow 10$ $-65 \Rightarrow l \Rightarrow 64$
Refinement	
Refinement on	F
$R_1(F)$, $wR(F)$, $R_{all}(F)$, S†	0.022, 0.023, 0.026, 1.24
No. of reflections (N_{ref})	10 905
No. of parameters (N_{var})	541
N_{ref}/N_{var}	20.19
H-atom treatment	Mixture of independent and constrained refinement
Weighting scheme	Based on measured s.u.'s
$(\Delta/\sigma)_{max}$	<0.0001
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.21, −0.21

† $R_{int}(F^2) = \sum |F_o^2 - F_o^2(\text{mean})| / \sum F_o^2$, $wR(F) = [\sum w||F_o| - k|F_c||^2 / \sum w|F_o|^2]^{1/2}$, $w = 1/\sigma^2$, $R_1(F) = \sum ||F_o| - |F_c|| / \sum |F_o|$, $R_{all}(F)$ using all reflections, $S = [\sum ||F_o| - k|F_c||^2)/(n_o - m_{var})]^{1/2}$.

3. Density models and refinement strategy

For the application of the aspherical-atom multipole formalism according to the method of Hansen & Coppens (1978), the program package *XD* (Koritsánszky *et al.*, 1995) was used. The starting atomic parameters were taken from the spherical-atom refinement (*SHELXL*97; Sheldrick, 1997).

The multipole formalism allows modeling of the aspherical fraction of $\rho(\mathbf{r})$ using atom-centered multipole functions,

$$\rho_{atom}(\mathbf{r}) = \rho_{core}(r) + P_v\kappa'^3\rho_{valence}(\kappa' r) + \rho_{deformation}, \qquad (1)$$

with

$$\rho_{deformation} = \sum_{l=0}^{4} \kappa''^3 R_l(\kappa''\mathbf{r}) \sum_{m=0}^{l} P_{lm\pm} Y_{lm\pm}(\theta, \varphi).$$

The core and the spherical valence density of the heavy atoms are composed of Hartree–Fock wavefunctions expanded over Slater-type basis functions. For the deformation terms, single-zeta orbitals with energy-optimized Slater exponents were used (Clementi & Roetti, 1974). The quantity $\sum_{\mathbf{H}} w_{\mathbf{H}}[|F_{obs}(\mathbf{H})| - k|F_{calc}(\mathbf{H})|]^2$ was minimized using the statistical weight $w_{\mathbf{H}} = \sigma^{-2}[F_{obs}(\mathbf{H})]$. The populations of C, N and O atoms were refined up to hexadecapoles, and bond-directed dipoles were applied to H atoms. A quadrupolar term was included for the H atom that is involved in a strong hydrogen bond. In addition to the positional, displacement and multipole population parameters, 17 expansion/contraction parameters, κ', were refined in the least-squares procedure. *m* symmetry was assigned to the atoms in the five-membered ring of the indole group, to atoms C12 and C18 of the carboxy/carboxylate groups, and to atom C10. Finally, *mm*2 symmetry was assigned to the remaining C atoms of the six-membered ring of the indole group. (For the atomic numbering scheme see Fig. 1.) Bonds to H atoms were set to the standard neutron distances (*International Tables for Crystallography*, 1992, Vol. C). For the formic acid H17 atom,

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: LC5002). Services for accessing these data are described at the back of the journal.

which is involved in the strong hydrogen bond, the distance was set to 1.05 Å, after Steiner & Saenger (1994; these authors plotted O···O *versus* O—H distances based on neutron diffraction data). Charge transfer was allowed between the two molecules because of the strong hydrogen bond; however, the amount of transfered charge is negligible (see §5; for the agreement factors see Table 1). The Hirshfeld (1976) test showed that the highest difference of mean-squares displacement amplitudes was 0.0005 Å^2, indicating a proper deconvolution of thermal motion and electron density. No significant electron density was found in the residual density maps (Fig. 2) that include all reflections (no resolution cutoff).

4. Theoretical calculations

The program package *GAUSSIAN*98 (Frisch *et al.*, 1998) was used for a single-point density-functional calculation based on the experimental geometry with the basis set B3LYP/6-311++G(3*d*,3*p*). Full topological analysis was performed with the program *AIMPAC* (Cheeseman *et al.*, 1992).

5. Results and discussion

The crystal structure was recently determined and refined spherically (Hübschle *et al.*, 2002). The asymmetric unit consists of a zwitterionic tryptophan molecule and the neutral formic acid molecule. There is a strong O—H···O hydrogen bond from the formic acid group to one of the O atoms of the carboxylate group of the tryptophan molecule. Bond lengths and angles from the present multipole refinement do not differ significantly from those determined using the spherical-atom model and will not be discussed further.

The static deformation density distributions in the planes of the indole and carboxylate groups (in Fig. 3) describe qualitatively the expected covalent and lone-pair regions. The electrostatic potential is shown in Fig. 4 in a similar orientation as is shown in Fig. 1. There is a large negative segment around the carboxylate–formic acid interaction region, including the formal C═O group of the formic acid molecule. Negative potential is also seen above and below the five-membered ring of the indole group.

For a quantitative description of the electronic structure of (I), a full topological analysis was performed with the *XDPROP* program of the *XD* package (Koritsánszky *et al.*, 1995). (3,−1) bond critical points (BCPs) were found for all covalent bonds. In addition, two (3,+1) ring critical points

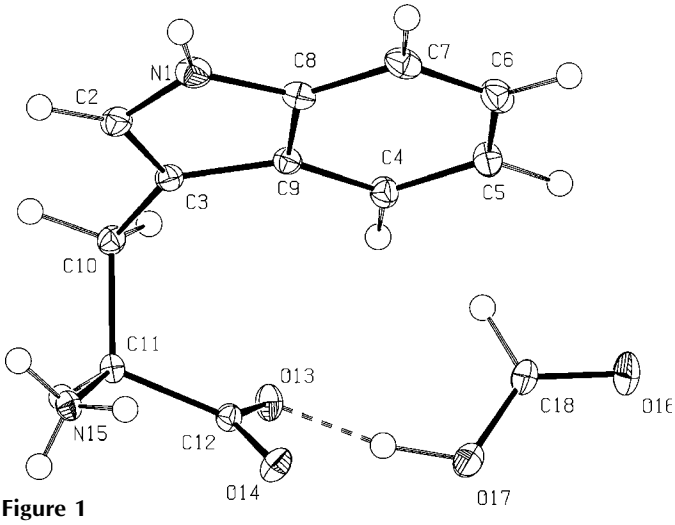

Figure 1
ORTEP (Burnett & Johnson, 1996) representation of the title compound (I) (50% probability displacement ellipsoids) with the atomic numbering scheme.

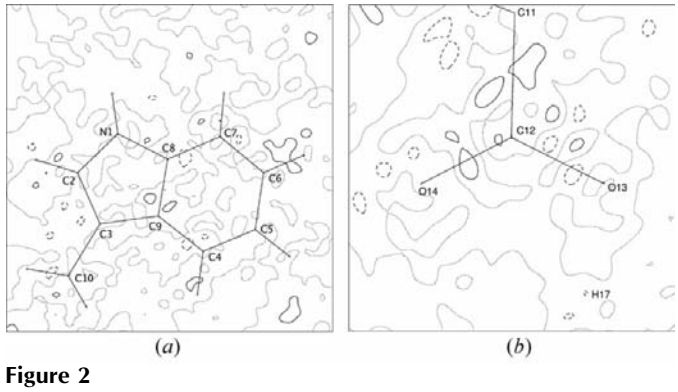

Figure 2
Residual maps in the (*a*) indole-ring and (*b*) carboxy-group planes. Positive, negative and zero contours are represented by solid, dotted and dashed lines, respectively. Contour intervals are drawn at 0.1 e Å^{-3}.

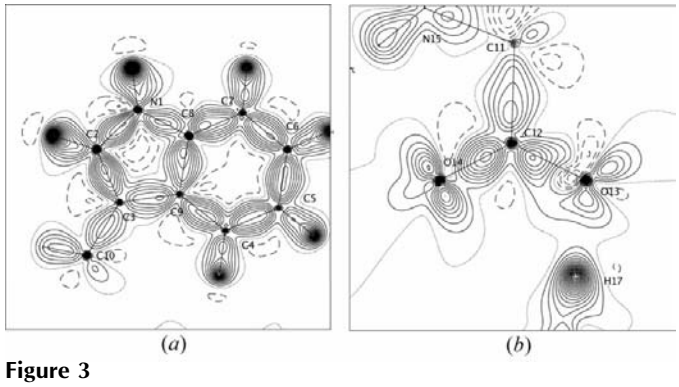

Figure 3
Static deformation density maps in the (*a*) indole-ring and (*b*) carboxy-group planes. Contour intervals are drawn at 0.1 e Å^{-3}; for further explanation see Fig. 2.

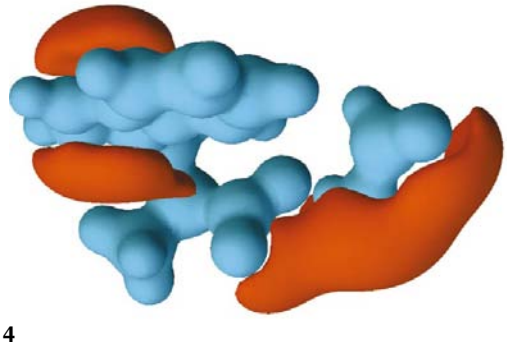

Figure 4
Three-dimensional representation of the electrostatic potential calculated from experimental charge density. Two isopotential surfaces are shown: blue: 0.5 e Å^{-1}; red: −0.15 e Å^{-1}.

Table 2
(*a*) Topological properties at critical points of ρ.

d is the distance from the first atom of the bond to the bond critical point. ε is the bond ellipticity, $\varepsilon = (\lambda_1/\lambda_2) - 1$, where λ_1 and λ_2 are the two negative curvatures. First line: values from a theoretical B3LYP/6-311++G(3*df*,3*pd*) calculation based on the experimental geometry. Second line: experimental results. Third line: results from an HF/6-311++G** calculation on the free (non-zwitterionic) tryptophan molecule (Matta & Bader, 2003).

Bond (Å)	$\rho(\mathbf{r})$ (e Å^{-3})	$\nabla^2\rho(\mathbf{r})$ (e Å^{-5})	d (Å)	ε	Bond (Å)	$\rho(\mathbf{r})$ (e Å^{-3})	$\nabla^2\rho(\mathbf{r})$ (e Å^{-5})	d (Å)	ε
O13—C12	2.55	−18.5	0.828	0.10	C7—C8	2.11	−22.1	0.667	0.21
1.2728 (3)	2.64 (4)	−26.9 (2)	0.740	0.01	1.4011 (4)	2.13 (3)	−17.2 (1)	0.686	0.21
	2.08	−1.54	0.901	0.16		2.14	−24.1	0.657	0.23
	[2.71 (9)	−33.6 (45)	0.791 (17)]†						
O14—C12	2.74	−16.1	0.810	0.11	C8—N1	2.09	−19.4	0.495	0.13
1.2404 (3)	2.94 (5)	−43.3 (3)	0.796	0.17	1.3730 (4)	2.22 (1)	−23.7 (2)	0.582	0.12
	2.95	2.46	0.793	0.06		2.07	−12.1	0.454	0.04
	[2.83 (11)	−35.6 (36)	0.772 (15)]						
N1—C2	2.11	−20.6	0.870	0.18	C3—C10	1.75	−15.7	0.750	0.05
1.3769 (5)	2.35 (4)	−28.9 (2)	0.805	0.30	1.4989 (4)	1.70 (3)	−12.9 (1)	0.749	0.07
	2.04	−11.0	0.922	0.07		1.76	−17.4	0.772	0.05
C2—C3	2.22	−23.3	0.715	0.33	C11—C10	1.64	−13.5	0.802	0.03
1.3749 (4)	2.29 (4)	−24.7 (1)	0.738	0.37	1.5397 (4)	1.61 (3)	−8.3 (1)	0.783	0.04
	2.29	−25.7	0.724	0.44		1.71	−16.5	0.792	0.04
						[1.68 (8)	−11.7 (21)	0.778 (16)]	
C3—C9	1.94	−18.3	0.725	0.16	N15—C11	1.60	−12.6	0.950	0.04
1.4426 (3)	1.95 (3)	−13.2 (1)	0.710	0.14	1.4891 (3)	1.57 (3)	−7.1 (2)	0.852	0.07
	1.94	−19.9	0.716	0.17		1.88	−21.2	0.913	0.02
						[1.68 (5)	−10.8 (21)	0.853 (17)]	
C8—C9	2.09	−21.3	0.730	0.20	C12—C11	1.74	−15.5	0.730	0.09
1.4183 (3)	2.06 (3)	−15.1 (1)	0.735	0.17	1.5281 (3)	1.64 (3)	−9.2 (1)	0.792	0.14
	2.17	−24.7	0.737	0.24		1.81	−19.2	0.838	0.05
						[1.75 (6)	−13.0 (22)	0.759 (19)]	
C4—C9	2.10	−21.9	0.682	0.17	O13···H17	0.57	2.7	1.039	0.02
1.4046 (3)	2.09 (3)	−20.3 (1)	0.719	0.29	1.4479 (3)	0.69 (10)	4.8 (3)	0.944	0.06
	2.12	−23.5	0.693	0.21	–	–	–	–	–
C4—C5	2.17	−22.9	0.699	0.22	O16—C18	2.87	−14.4	0.798	0.12
1.3872 (4)	2.21 (2)	−21.5 (1)	0.689	0.21	1.2202 (5)	2.93 (5)	−37.4 (3)	0.730	0.21
	2.21	−24.9	0.700	0.28	–	–	–	–	–
C5—C6	2.08	−21.4	0.696	0.18	O17—C18	2.39	−13.9	0.847	0.05
1.4106 (5)	2.16 (1)	−20.3 (1)	0.706	0.08	1.2908 (5)	2.69 (5)	−32.4 (3)	0.748	0.04
	2.11	−23.3	0.694	0.21	–	–	–	–	–
C6—C7	2.15	−22.5	0.684	0.23	O17—H17	2.11	−56.2	0.829	0.01
1.3899 (5)	2.19 (2)	−21.3 (1)	0.700	0.21	1.05	2.36 (2)	12.1 (9)	0.704	0.03
	2.21	−24.8	0.680	0.28	–	–	–	–	–
5 ring	0.35	8.2			6 ring	0.15	3.72		
	0.37 (1)	7.6 (1)				0.16 (1)	3.0 (1)		

(*b*) Mean differences $\delta(\rho)$ and $\delta(\nabla^2\rho)$ for the non-hydrogen bonds.

Quantity	$\delta(\rho)$ (e Å^{-3})	$\delta(\nabla^2\rho)$ (e Å^{-5})	n
$(1/n)\sum$ |[exp −*th*(DFT)]|	0.067	7.1	18
$(1/n)\sum$ |[exp −*th*(DFT)]|	0.065	3.6	10 (indole)
$(1/n)\sum$ |[exp −*th*(HF)]|	0.070	6.7	10 (indole)
$(1/n)\sum$ |[*th*(DFT) − *th*(HF)]|	0.042	3.7	10 (indole)

† Mean values of Flaig *et al.* (2002).

were found in the centers of the five- and six-membered rings of the indole group. The topological descriptors are summarized in Table 2(*a*), which also lists the corresponding values from the DFT calculation and further results from Matta & Bader's (2003) HF calculation on the free tryptophan amino acid. The $\rho(\mathbf{r}_{BCP})$ values on the bond critical points provide information on the covalent bond strength. For the C—C bonds, the aromatic $\rho(\mathbf{r}_{BCP})$ values in the indole ring are the strongest, being on average 2.14 (10) e Å^{-3}, which can be compared with an average of 1.65 (5) e Å^{-3} in the three weaker single bonds C3—C10, C10—C11 and C11—C12. Among these bonds, the C3—C10 bond next to the indole ring has the highest $\rho(\mathbf{r}_{BCP})$ value. A comparable trend is seen for the N—C bonds, which average $\rho(\mathbf{r}_{BCP})$ = 2.29 (9) e Å^{-3} in the indole ring; the weaker C—N bond to the protonated atom N15 has a much smaller value, $\rho(\mathbf{r}_{BCP})$ = 1.57 (3) e Å^{-3}. Among the four C—O bonds, the C12—O13 and C18—O17 bonds, which are involved in the strong hydrogen bond, have $\rho(\mathbf{r}_{BCP})$ values that are remarkably smaller than the other two C—O bonds.

The (3,−1) critical point of the strong hydrogen bond between atoms H17 and O13 has a $\rho(\mathbf{r}_{BCP})$ value of 0.69 (10) e Å^{-3} and a positive Laplacian of 4.8 (3) e Å^{-5} (see Table 3), indicating a closed-shell interaction. Espinosa and coworkers (Espinosa, Lecomte & Molins, 1999; Espinosa, Souhassou *et al.*, 1999) discussed an exponential relation between these topological descriptors and the H···O hydrogen distance. The magnitude of 0.69 (10) e Å^{-3} for $\rho(\mathbf{r}_{BCP})$ fits this relation well, while the Laplacian is lower than would be predicted by Espinosa and coworkers. The O···O distance [2.49 (1) Å] is similar to that [2.502 (4) Å] reported for a hydrogen bond in benzoylacetone (Madsen *et al.*, 1998), for which a negative Laplacian and hence a covalent contribution was derived.

The topological data for the weak interactions are summarized in Table 3. As already stated by Hübschle *et al.* (2002), there is a weak N1—H1···C interaction with a neighboring indole fragment. This bifurcated hydrogen bond has two BCPs, with H1···C7 = 2.33 and H1···C8 = 2.44 Å. From geometric criteria, the C2—H2···O16 contact could also be considered as a weak hydrogen bond; however, in contrast to the other entries in Table 3, no critical point could be located. This result is in line with the findings of

Table 3
Intra- and intermolecular contacts indicating possible hydrogen bonds, and their bond topological parameters.

$\rho(\mathbf{r}_{BCP})$ in e Å^{-3}, $\nabla^2\rho(\mathbf{r}_{BCP})$ in e Å^{-5}, distances in Å and angles in °.

D—H···A	Symmetry code	$\rho(\mathbf{r}_{BCP})$	$\nabla^2\rho(\mathbf{r}_{BCP})$	H···A	D···A	D—H	D—H···A
N15—H15A···O13	$1+x, y, z$	0.12 (2)	2.7 (1)	2.00	2.9273 (4)	1.00	157
N15—H15B···O16	$1+x, y-1, z$	0.17 (3)	4.0 (1)	1.81	2.7900 (6)	1.00	168
N15—H15C···O14	$1-x, y-\frac{1}{2}, \frac{3}{2}-z$	0.16 (3)	3.4 (1)	1.90	2.7469 (4)	1.00	145
O17—H17···O13	x, y, z	0.69 (1)	4.8 (3)	1.45	2.4870 (4)	1.05	169
N1—H1···C7	$\frac{1}{2}+x, \frac{1}{2}-y, 2-z$	0.05 (2)	1.1(1)	2.33	3.2910 (4)	1.00	161
N1—H1···C8	$\frac{1}{2}+x, \frac{1}{2}-y, 2-z$	0.06 (1)	1.0 (1)	2.45	3.3291 (4)	1.00	152
C2—H2···O16	$1+x, y-1, z$	–	–	2.41	3.1169 (6)	1.08	122

Table 4
Atomic properties (Å^3, e) derived from zero-flux surface partitioning for L-tryptophan formic acid.

Atom	V_{tot}	V_{001}	N	N_{001}	Q	Q_{001}
O13	15.84	14.76	8.83	8.83	−0.83	−0.83
O14	17.41	15.98	8.92	8.91	−0.92	−0.91
O16	19.14	16.72	8.78	8.77	−0.78	−0.77
O17	17.55	14.86	8.87	8.86	−0.87	−0.86
N1	12.90	12.83	7.98	7.98	−0.98	−0.98
N15	11.94	11.29	7.79	7.79	−0.79	−0.79
C2	10.95	10.81	5.81	5.81	0.19	0.19
C3	10.29	10.16	6.10	6.10	−0.10	−0.10
C4	11.67	11.29	5.98	5.98	0.02	0.02
C5	12.08	11.69	5.91	5.91	0.09	0.09
C6	13.00	11.98	5.96	5.95	0.04	0.05
C7	11.43	11.13	5.92	5.92	0.08	0.08
C8	8.80	8.77	5.90	5.90	0.10	0.10
C9	10.45	10.40	6.17	6.17	−0.17	−0.17
C10	9.12	8.61	5.93	5.93	0.07	0.07
C11	6.68	6.67	5.67	5.67	0.33	0.33
C12	6.17	6.17	4.87	4.87	1.13	1.13
C18	8.89	8.05	4.67	4.67	1.33	1.33
H1	3.99	3.97	0.85	0.85	0.15	0.15
H2	6.64	5.72	1.00	1.00	0.00	0.00
H4	5.35	5.29	0.86	0.86	0.14	0.14
H5	8.14	6.38	0.88	0.87	0.12	0.13
H6	7.27	6.67	0.88	0.88	0.12	0.12
H7	5.94	5.77	0.86	0.86	0.14	0.14
H11	7.22	5.48	0.88	0.87	0.12	0.13
H10A	6.93	5.76	0.96	0.95	0.04	0.05
H10B	5.40	5.40	0.96	0.96	0.04	0.04
H15A	3.27	3.20	0.72	0.72	0.28	0.28
H15B	3.51	3.38	0.71	0.71	0.29	0.29
H15C	4.04	3.56	0.71	0.71	0.29	0.29
H17	1.44	1.40	0.62	0.62	0.38	0.38
H18	6.18	6.15	1.04	1.04	−0.04	−0.04
Sum	289.63	270.28	131.99	131.91	0.01	0.09

Desiraju & Steiner (1999) that N—H···π hydrogen bonds are stronger than C—H···O interactions.

Concerning the question of how transferable atomic fragments are, the BCP data around the C′—C$_\alpha$—C$_\beta$ section, common to all amino acids, are of interest. These data can be compared with the experimental averages reported for 13 amino acids by Flaig *et al.* (2002), who showed that the statistical spreads of these averages are 0.05–0.11 e Å^{-3} for $\rho(\mathbf{r}_{BCP})$ and 2–5 e Å^{-5} for $\nabla^2\rho(\mathbf{r}_{BCP})$. A spread of the same magnitude was reported by Pichon-Pesme *et al.* (2000) for a number of experimental studies on pseudo-peptides, so this value is obviously the reproducibility range for these quantities. The corresponding data for the title compound agree with the 13 amino acid averages within 0.02–0.11 e Å^{-3} for $\rho(\mathbf{r}_{BCP})$ and 0.9–7.3 e Å^{-5} for $\nabla^2\rho(\mathbf{r}_{BCP})$, and hence almost within the above-mentioned statistical spreads.

The data summarized in Table 2(*a*) can be used further to compare experiment and theory. The experimental and theoretical data (based on the DFT calculation) for the title compound are compared in Table 2(*a*) (first and second lines). For the 18 non-hydogen bonds, the mean differences are 0.067 e Å^{-3} and 7.1 e Å^{-5} for $\rho(\mathbf{r}_{BCP})$ and $\nabla^2\rho(\mathbf{r}_{BCP})$. The large differences for $\nabla^2\rho(\mathbf{r}_{BCP})$ are mainly seen for the polar bonds, a well known finding that has been attributed to the limited flexibility of the radial functions in the multipole model (Volkov, Abramov *et al.*, 2000). Differences of almost the same magnitude are found if, on the other hand, the topological data of the title compound are compared with the results of the HF calculations of Bader & Matta (2003) (third line in Table 2*a*). Since their calculation was based on the free (and not zwitterionic) amino acid, the comparison should be restricted to the ten bonds of the indole fragment. The averages given in Table 2(*b*) indicate that the experimental *versus* theory difference, as well as the theory (DFT; this work) *versus* theory (HF; Matta & Bader, 2003) differences are practically in the same range.

5.1. Atomic volumes and charges

Following Bader's (1994) theory of atoms in molecules (AIM), a molecule can be partitioned into submolecular regions, fragments or single atoms, making use of the zero-flux surfaces of the electron gradient vector field $\nabla\rho(\mathbf{r})$. To evaluate the atomic volumes and charges, the program *TOPXD* (Volkov, Gatti *et al.*, 2000) was used. Fig. 5 shows the experimental gradient vector field, $\nabla\rho(\mathbf{r})$, in the plane of the formic acid group, while the numerical results are summarized in Table 4. The atomic volumes, V_{tot}, are defined by the interatomic boundaries in the crystal, while the volumes denoted V_{001} are based on a cutoff of ρ = 0.001 atomic units, which is commonly used in theory when isolated molecules are

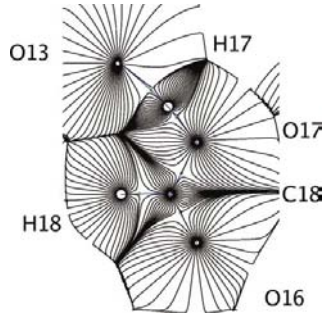

Figure 5
Gradient vector field $\nabla\rho(\mathbf{r})$ in the molecular plane of the formic acid solvent molecule containing the strong O—H···O hydrogen bond.

considered. This cutoff was also applied to the experimental densities (see V_{001} column), to allow a comparison between this experiment and the theoretical calculation of Matta & Bader (2003) on the free tryptophane molecule (see below). The sum of the V_{tot} volumes (multiplied by $Z = 4$) reproduces the unit-cell volume to within 1.2%. Since the sum of all atomic charges is zero, the integration procedure can be considered to have worked properly. In some cases, there are clear differences between V_{tot} and V_{001}, reaching more than 2 Å^3 for atoms O16 and O17; however, the charges are not affected, the Q and Q_{001} values being practically equal.

As expected, the strongest negative charges are found at the N and O atoms, and these charges are compensated mostly by the positive charges of carboxy atoms C12 and C18 and the H-atom charges.

As already mentioned, Bader & Matta (2003) have derived atomic properties from their HF calculation on the neutral tryptophan molecule, making a comparison within the indole fragment sensible. The experimental atomic volumes are systematically smaller than the theoretical volumes, especially for the indole N atom, which is about 18% smaller. Nevertheless, the comparably small volume of the C8 atom, neighbor to the indole N atom (experimental: 8.77 Å^3), is also the smallest C atom in theory (9.78 Å^3). In Fig. 6, the two types of volumes are plotted against each other and fitted linearly. The difference between this line and the bisector gives an impression of the difference between experiment and theory. Comparable differences between experiment and theory (from a DFT calculation) for C-atom volumes have already been reported (Lentz *et al.*, 2003). The charge of the indole ring is practically zero (experimental: −0.06*e*; theory: −0.08*e*).

For the fragment around $C_\alpha - C'$ (see Fig. 7), volumes and charges can be compared with the corresponding quantities derived for the amino and C-terminal ends of some dipeptides (Dittrich, 2002). The agreement between the title compound and the averages of the five dipeptides is within the statistical spread, which also indicates that charges and volumes in this fragment are reproduced within 0.1–0.2*e* and better than 1 Å^3. The overall charge of the fragment given in Fig. 7 is −1.07*e* for the five-dipeptide average and −1.08*e* for the title compound. This charge is almost completely compensated by the positive charges of the ammonium and C_α H atoms (sum +0.98*e*), and together with the $C_\beta H_2$ group (charge +0.15*e*), the title molecule can be subdivided into two neutral fragments, *viz.* the indole group (charge −0.06*e*) and the $C_\beta - C_\alpha - C'OO$ fragment (charge +0.04 *e*), so that the whole zwitterion is also neutral. The negligible charge of −0.02*e* is cancelled out by the overall charge of the formic acid molecule (0.02*e*), thus confirming that practically no charge transfer from the tryptophane zwitterion to the formic acid group has taken place.

The volume of the H17 atom, which is involved in the strong hydrogen bond, is remarkably small (V_{tot}/V_{001} = 1.44/1.40 Å^3). This fact is also evident in Fig. 5, which shows the gradient vector field in the solvent-molecule plane. In pentafluorobenzoic acid, where a strong hydrogen bond (H···O = 1.64 Å) exists between two carboxy-group dimers (Lentz *et al.*, 2003), the H-atom volume is 1.24 Å^3 and the charge (+0.68*e*) is even more positive than in the present case (+0.38*e*).

6. Conclusion

In this study, a full topological analysis of the experimental charge density, $\rho(\mathbf{r})$, was performed on the zwitterionic L-tryptophan linked to formic acid by a strong hydrogen bond. A detailed comparison of the quantitative topological results was performed with the various experimental and theoretical data derived in the field of previously investigated amino acids. For $\rho(\mathbf{r})$ and $\nabla^2\rho(\mathbf{r})$ values on the bond critical points, the experimental reliability, found to be 0.1 e Å^{-3} and 5 e Å^{-5}

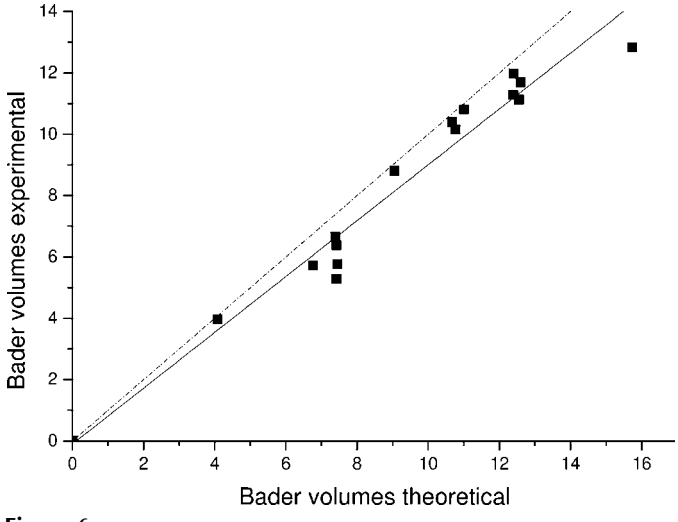

Figure 6
Theoretical *versus* experimental Bader volumes (Å^3), fitted linearly (solid line). The bisector (dashed line) illustrates the deviation from exact agreement.

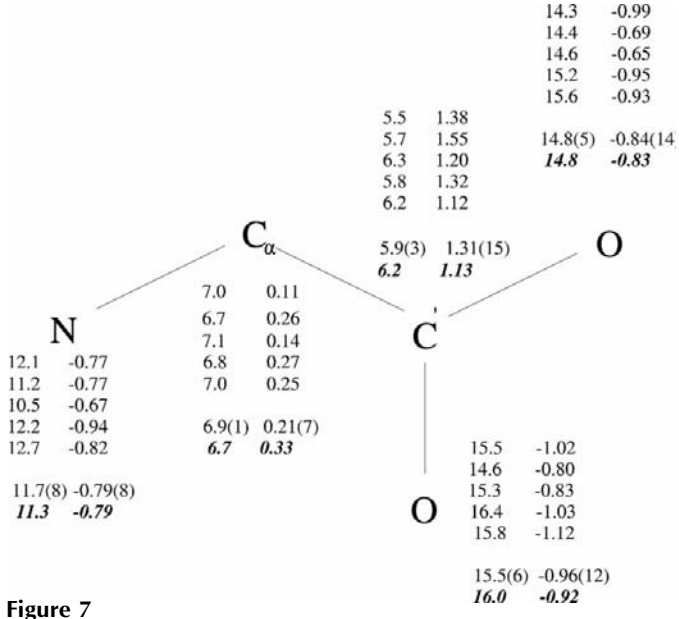

Figure 7
Atomic volumes (Å^3, left) and charges (*e*, right) around the $C_\alpha - C'$ bonds of five dipeptides (Gly L-Thr, Gly L-His, L-His-L-Ala, Gly DL-Phe and L-Phe-L-Pro; Dittrich, 2002), their averages and the corresponding quantities for the title compound (in bold italics).

for $\rho(\mathbf{r}_{BCP})$ and $\nabla^2\rho(\mathbf{r}_{BCP})$, respectively, is in line with previous findings of Flaig *et al.* (2002) and Pichon-Pesme *et al.* (2000). These values are of the same order as the differences between experiment and theory, and between different theoretical calculations (Table 2*b*). The larger disagreement for the Laplacian is mainly seen for the polar bonds.

The derived atomic volumes and charges (Fig. 7) agree well with corresponding values from several dipeptides, so it seems feasible to use atoms derived from the partitioning of single molecules for the construction of peptides.

The strong hydrogen bond is characterized by the following features: The charge density on its bond critical point is by far the highest compared with the other weak interactions in this structure. The value $\rho(\mathbf{r}_{BCP})$ = 0.69 (10) e Å^{-3}, which reaches almost 30% of the corresponding value for the covalent O17—H17 bond, fits the exponential relation derived by Espinosa and co-workers (Espinosa, Lecomte & Molins, 1999; Espinosa, Souhassou *et al.*, 1999). The Laplacian of this hydrogen-bond BCP is positive and hence indicates closed-shell interactions. Also, covalent contributions in hydrogen bonds with similar distances have been discussed by Madsen *et al.* (1998).

The strength of this hydrogen bond is also reflected in the small atomic volume of atom H17 and its high positive charge (0.38*e*). The volume (∼1.4 Å^3) is more than four times smaller than that of the H atoms not involved in short intermolecular contacts [mean volume 6.6 (9) Å^3]. The ammonium H atoms and atom H1, all of which are involved in weaker hydrogen bonds, have volumes between 3 and 4 Å^3, so these atomic properties seem sensitive to the strength of hydrogen bonding.

References

Bader, R. F. W. (1994). In *Atoms in Molecules: A Quantum Theory*, 2nd ed. Oxford: Clarendon Press.

Bakke, O. & Mostad, A. (1980). *Acta Chem. Scand. B*, **34**, 559–570.

Bruker (1997–2001). *ASTRO*, *SMART* and *SAINT*. Bruker AXS Inc., Madison, Wisconsin, USA.

Burnett, M. N. & Johnson, C. K. (1996). *ORTEP*III. Report ORNL-6895. Oak Ridge National Laboratory, Tennessee, USA.

Cheeseman, J., Keith, T. A. & Bader, R. F. W. (1992). *AIMPAC*. McMaster University, Hamilton, Ontario, Canada.

Clementi, E. & Roetti, C. (1974). *At. Nucl. Data Tables*, **14**, 177–478.

Dahaoui, S., Jelsch, C., Howard, J. A. K. & Lecomte, C. (1999). *Acta Cryst.* B**55**, 226–230.

Desiraju, G. R. & Steiner, T. (1999). In *The Weak Hydrogen Bond*. Oxford University Press.

Dittrich, B. (2002). PhD thesis, Freie Universität Berlin, Germany.

Espinosa, E., Lecomte, C. & Molins, E. (1999). *Chem. Phys. Lett.* **300**, 745–748.

Espinosa, E., Souhassou, M., Lachekar, H. & Lecomte, C. (1999). *Acta Cryst.* B**55**, 563–572.

Flaig, R., Koritsánszky, T., Dittrich, B., Wagner, A. & Luger, P. (2002). *J. Am. Chem. Soc.* **124**, 3407–3417.

Frisch *et al.* (1998). *GAUSSIAN*98. Revision A.7. Gaussian Inc., Pittsburgh, PA, USA.

Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.

Hirshfeld, F. L. (1976). *Acta Cryst.* A**32**, 239–244.

Hübschle, C. B., Dittrich, B. & Luger, P. (2002). *Acta Cryst.* C**58**, 540–542.

Jelsch, C., Pichon-Pesme, V., Lecomte, C. & Aubry, A. (1998). *Acta Cryst.* D**54**, 1306–1318.

Jelsch, C., Teeter, M. M., Lamzin, V., Pichon-Pesme, V., Blessing, R. H. & Lecomte, C. (2000). *Proc. Natl Acad. Sci.* **97**, 3171–3176.

Koritsánszky, T., Howard, S., Richter, T., Su, Z. W., Mallinson, P. R. & Hansen, N. K. (1995). *XD*. Freie Universität Berlin, Germany.

Lentz, D., Patzschke, M., Bach, A., Scheins, S. & Luger, P. (2003). *Org. Biomol. Chem.* **1**, 409–414.

Madsen, G. K. H., Iversen, B. B., Larsen, F. K., Kapon, M., Reisner, G. M. & Herbstein, H. (1998). *J. Am. Chem. Soc.* **120**, 10040–10045.

Matta, C. F. & Bader, R. F. W. (2000). *Proteins*, **40**, 310–329.

Matta, C. F. & Bader, R. F. W. (2002). *Proteins*, **48**, 519–538.

Matta, C. F. & Bader, R. F. W. (2003). *Proteins*, **52**, 360–399.

Pichon-Pesme, V., Lachekar, H., Souhassou, M. & Lecomte, C. (2000). *Acta Cryst.* B**56**, 728–737.

Pichon-Pesme, V. & Lecomte, C. (1998). *Acta Cryst.* B**54**, 485–493.

Sheldrick, G. M. (1997). *SHELXL*97. University of Göttingen, Germany.

Steiner, T. & Saenger, W. (1994). *Acta Cryst.* B**50**, 348–357.

Volkov, A., Abramov, Y., Coppens, P. & Gatti, C. (2000). *Acta Cryst.* A**56**, 332–339.

Volkov, A., Gatti, C., Abramov, Y. & Coppens, P. (2000). *Acta Cryst.* A**56**, 252–258.