Isovaline monohydrate

Ray J. Butcher,[a]* Greg Brewer,[b] Aaron S. Burton[c] and Jason P. Dworkin[d]

[a]Department of Chemistry, Howard University, 525 College Street NW, Washington, DC 20059, USA, [b]Department of Chemistry, Catholic University of America, Washington, DC 20064, USA, [c]NASA Johnson Space Center, Astromaterial and Exploration Science Directorate, Houston, TX 77058, USA, and [d]Solar System Exploration Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

Correspondence e-mail: rbutcher99@yahoo.com

The title compound, $C_5H_{11}NO_2 \cdot H_2O$, is an isomer of the α-amino acid valine that crystallizes from water in its zwitterion form as a monohydrate. It is not one of the 20 proteinogenic amino acids that are used in living systems and differs from the natural amino acids in that it has no α-H atom. The compound exhibits hydrogen bonding between the water molecule and the carboxylate O atoms and an amine H atom. In addition, there are intermolecular hydrogen-bonding interactions between the carboxylate O atoms and amine H atoms. In the crystal, these extensive N—H···O and O—H···O hydrogen bonds lead to the formation of a three-dimensional network.

Related literature

The structure of the title compound or its salts have not been reported to the CCDC but there are reports of homoleptic coordination complexes of zinc(II) with isovaline, see: Strasdelt *et al.* (2001). For literature related to eighty amino acids that have been detected in meteorites or comets, see: Glavin & Dworkin (2009); Burton *et al.* (2012). For the role that crystallization plays in chiral separation, see: Blackmond & Klussmann (2007); Blackmond *et al.* (2008). For the role of the H atom on the α-C atom in enhancing the rate of racemization, see: Yamada *et al.* (1983). For the mechanism of racemization of amino acids lacking an α-H atom, see: Pizzarello & Groy (2011). For the role that crystallization can play in the enrichment of L-isovaline, see: Glavin & Dworkin (2009); Bada (2009); Bonner *et al.* (1979). For normal bond lengths and angles, see: Orpen (1993).
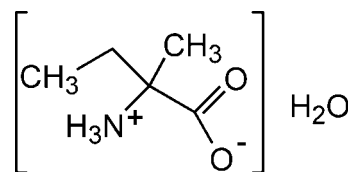
Experimental

Crystal data

$C_5H_{11}NO_2 \cdot H_2O$	$V = 736.10\ (12)$ Å^3
$M_r = 135.16$	$Z = 4$
Orthorhombic, $P2_12_12_1$	Cu $K\alpha$ radiation
$a = 5.9089\ (5)$ Å	$\mu = 0.84$ mm^{-1}
$b = 10.4444\ (10)$ Å	$T = 123$ K
$c = 11.9274\ (11)$ Å	$0.48 \times 0.08 \times 0.06$ mm

Data collection

Agilent Xcalibur (Ruby, Gemini) diffractometer	1662 measured reflections
Absorption correction: multi-scan (*CrysAlis PRO*; Agilent, 2012)	1204 independent reflections
$T_{min} = 0.383$, $T_{max} = 1.000$	1072 reflections with $I > 2\sigma(I)$
	$R_{int} = 0.072$

Refinement

$R[F^2 > 2\sigma(F^2)] = 0.056$	H atoms treated by a mixture of independent and constrained refinement
$wR(F^2) = 0.162$	
$S = 1.11$	
1204 reflections	$\Delta\rho_{max} = 0.37$ e Å^{-3}
91 parameters	$\Delta\rho_{min} = -0.27$ e Å^{-3}
3 restraints	

Table 1
Hydrogen-bond geometry (Å, °).

D—H···A	D—H	H···A	D···A	D—H···A
O1W—H1W1···O1^i	0.83 (2)	2.05 (2)	2.811 (3)	152 (4)
O1W—H1W2···O2^{ii}	0.83 (2)	1.96 (2)	2.787 (3)	171 (5)
N1—H1A···O2^i	0.91	1.84	2.745 (3)	177
N1—H1C···O1W	0.91	2.09	2.792 (4)	133
N1—H1B···O1^{iii}	0.91	1.98	2.832 (3)	156

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

Data collection: *CrysAlis PRO* (Agilent, 2012); cell refinement: *CrysAlis PRO*; data reduction: *CrysAlis PRO*; program(s) used to solve structure: *SHELXS97* (Sheldrick, 2008); program(s) used to refine structure: *SHELXL97* (Sheldrick, 2008); molecular graphics: *SHELXTL* (Sheldrick, 2008); software used to prepare material for publication: *SHELXTL*.

organic compounds

References

Agilent (2012). *CrysAlis PRO*. Agilent Technologies UK Ltd, Yarnton, England.

Bada, J. L. (2009). *Proc. Natl Acad. Sci.* **106**, E85.

Blackmond, D. G. & Klussmann, M. (2007). *Chem. Commun.* pp. 3990–3996.

Blackmond, D., Viedma, C., Ortiz, J., Torres, T. & Izuma, T. (2008). *J. Am. Chem. Soc.* **130**, 15274–15275.

Bonner, W. A., Blair, N. E., Lemmon, R. M., Flores, J. J. & Pollock, G. E. (1979). *Geochim. Cosmochim. Acta*, **43**, 1841–1846.

Burton, A. S., Stern, J. C., Elsila, J. E., Dworkin, J. P. & Glavin, D. P. (2012). *Chem. Soc. Rev.* **41**, 5459–5472.

Glavin, D. P. & Dworkin, J. P. (2009). *Proc. Natl Acad. Sci.* **106**, 5487–5492.

Orpen, G. A. (1993). *Chem. Soc. Rev.* **22**, 191–197.

Pizzarello, S. & Groy, T. L. (2011). *Geochim. Cosmochim. Acta*, **75**, 645–656.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Strasdelt, H., Busching, I., Behrends, S., Saak, W. & Barklage, W. (2001). *Chem. Eur. J.* **7**, 1133–1137.

Yamada, S., Hongo, C., Yoshioka, R. & Chibata, I. (1983). *J. Org. Chem.* **48**, 843–846.

supporting information

Isovaline monohydrate

Ray J. Butcher, Greg Brewer, Aaron S. Burton and Jason P. Dworkin

S1. Comment

The alpha amino acids are essential for life as they are the building blocks of all proteins and enzymes. Nature uses almost exclusively the *L* form of the nineteen naturally occurring chiral amino acids. Glycine is achiral. However, it is known that there are over eighty amino acids that have been detected in meteorites or comets (Glavin & Dworkin 2009; Burton *et al.*, 2012). One of these extraterrestrial non proteinogenic amino acids is isovaline. The majority of these amino acids show little or no enrichment of one enantiomer over the other. An intriguing question is the process that lead to the separation and enrichment of the *L* enantiomer over the *D*. There are several possible explanations for this including the role that crystallization plays (Blackmond & Klussmann, 2007). Only two of the twenty amino acids used biologically crystallize in a chiral space group, which allows for spontaneous separation of enantiomers, at the level of the crystal, from a racemic solution (Blackmond *et al.*, 2008). Isovaline, a non proteinogenic amino acid also allows for this separation of enantiomers at the level of the crystal as it crystallizes in the chiral space group, $P2_12_12_1$.

Another important aspect in the prebiotic chemistry of the amino acids is the role of racemization. All of the nineteen naturally occurring chiral amino acids have a hydrogen atom on the alpha carbon atom, which enhances the rate of racemization (Yamada *et al.*, 1983). However, little is known about the mechanism of racemization of amino acids lacking an alpha hydrogen atom (Pizzarello & Groy, 2011). The structure of isovaline given here can be used to examine the role that crystallization can play in the enrichment of *L* isovaline (Glavin & Dworkin 2009; Bada, 2009), and as a starting point in mechanistic studies of racemization mechanisms of isovaline (Bonner *et al.*, 1979).

In the structure of the title compound the amino acid is in the usual zwitterionic form. While the structure of the title compound or its salts have not been reported there are reports of homoleptic coordination complexes of isovaline with zinc(II) (Strasdelt *et al.* (2001). However, there are several important differences between these structures and the title compound. The metal complexes all crystallized in non-chiral space groups and, of course, when coordinated to a metal, the isovaline will not be in a zwitterionic form but, apart from the COO⁻ group, all the the bond lengths and angles are in the normal range for such compounds (Orpen, 1993). There is extensive N—H⋯O and O—H⋯O hydrogen bonding linking the zwitterions into a 3-D array.

S2. Experimental

A sample of the title compound was obtained from Acros Organics. Crystals of the title compound were grown from the slow evaporation of a *racemic* solution of the amino acid in water.

S3. Refinement

H atoms were placed in geometrically idealized positions and constrained to ride on their parent atoms with a C—H distances of 0.98 and 0.99 Å and an N—H distance of 0.91 \%A U_{iso}(H) = 1.2U_{eq}(C) and 0.96 Å for NH₃ [U_{iso}(H) = 1.5U_{eq}(C)]. The water H's were refined isotropically with O—H distances constrained to be 0.82 Å and the H—O—H
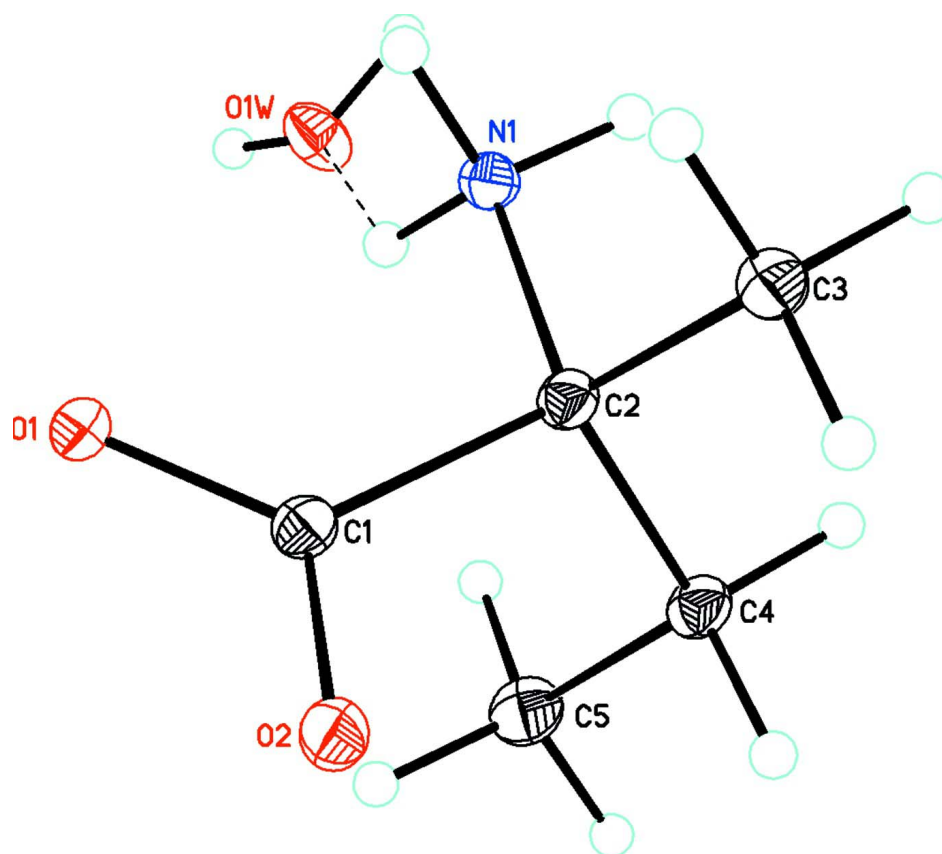
angle close to 104.5°.

Figure 1

Diagram of the title compound showing atom labeling. Atomic displacement parameters are at the 30% probablity level. Hydogen bonds are shown as dashed lines.
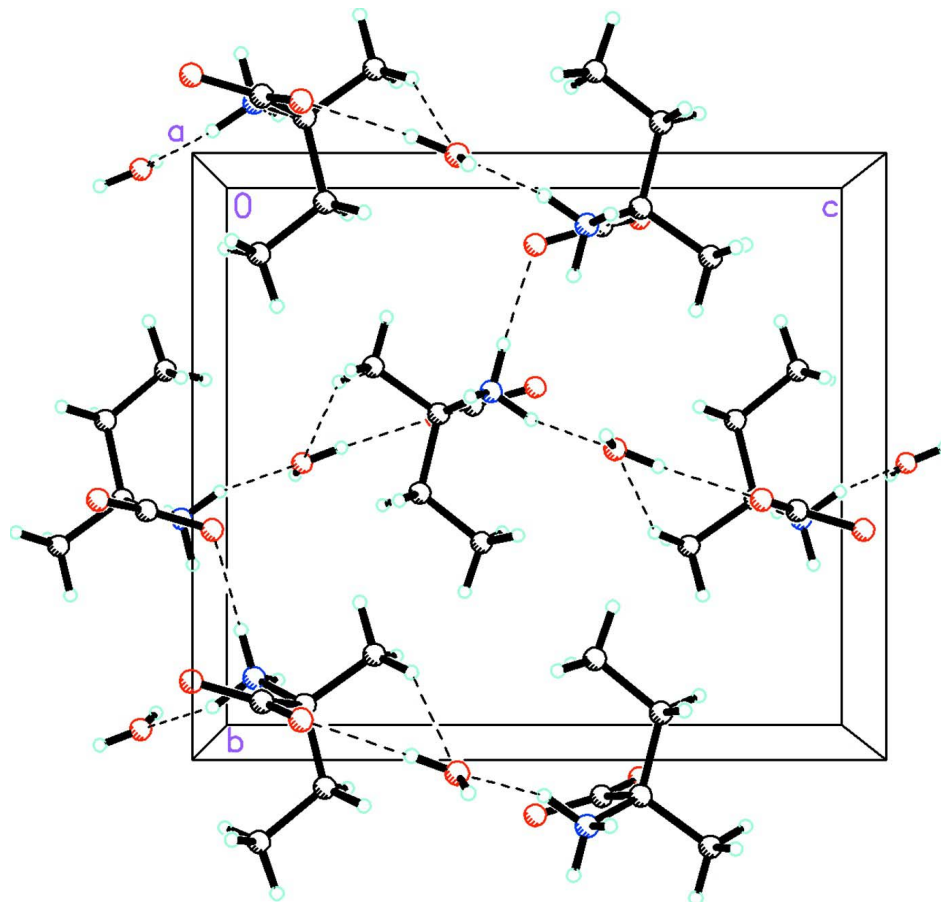
Figure 2

Packing diagram of the title compound viewed along the *a* axis showing tthe extensive N—H···O and O—H···O hydrogen bonds as dashed lines.

2-Azaniumyl-2-methylbutanoate monohydrate

Crystal data

$C_5H_{11}NO_2 \cdot H_2O$	$F(000) = 296$
$M_r = 135.16$	$D_x = 1.220$ Mg m^{-3}
Orthorhombic, $P2_12_12_1$	Cu $K\alpha$ radiation, $\lambda = 1.54184$ Å
Hall symbol: P 2ac 2ab	Cell parameters from 679 reflections
$a = 5.9089$ (5) Å	$\theta = 3.7$–$75.3°$
$b = 10.4444$ (10) Å	$\mu = 0.84$ mm^{-1}
$c = 11.9274$ (11) Å	$T = 123$ K
$V = 736.10$ (12) Å^3	Needle, colorless
$Z = 4$	$0.48 \times 0.08 \times 0.06$ mm

Data collection

Agilent Xcalibur (Ruby, Gemini) diffractometer	Absorption correction: multi-scan (*CrysAlis PRO*; Agilent, 2012)
Radiation source: Enhance (Cu) X-ray Source	$T_{min} = 0.383$, $T_{max} = 1.000$
Graphite monochromator	1662 measured reflections
Detector resolution: 10.5081 pixels mm^{-1}	1204 independent reflections
ω scans	1072 reflections with $I > 2\sigma(I)$

$R_{int} = 0.072$

$\theta_{max} = 75.4°, \theta_{min} = 5.6°$

$h = -7 \rightarrow 4$

$k = -12 \rightarrow 8$

$l = -14 \rightarrow 14$

Refinement

Refinement on F^2	Secondary atom site location: difference Fourier map
Least-squares matrix: full	
$R[F^2 > 2\sigma(F^2)] = 0.056$	Hydrogen site location: inferred from neighbouring sites
$wR(F^2) = 0.162$	
$S = 1.11$	H atoms treated by a mixture of independent and constrained refinement
1204 reflections	
91 parameters	$w = 1/[\sigma^2(F_o^2) + (0.093P)^2 + 0.2964P]$
3 restraints	where $P = (F_o^2 + 2F_c^2)/3$
Primary atom site location: structure-invariant direct methods	$(\Delta/\sigma)_{max} < 0.001$
	$\Delta\rho_{max} = 0.37$ e Å^{-3}
	$\Delta\rho_{min} = -0.27$ e Å^{-3}

Special details

Geometry. All e.s.d.'s (except the e.s.d. in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell e.s.d.'s are taken into account individually in the estimation of e.s.d.'s in distances, angles and torsion angles; correlations between e.s.d.'s in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell e.s.d.'s is used for estimating e.s.d.'s involving l.s. planes.

Refinement. Refinement of F^2 against ALL reflections. The weighted R-factor wR and goodness of fit S are based on F^2, conventional R-factors R are based on F, with F set to zero for negative F^2. The threshold expression of $F^2 > \sigma(F^2)$ is used only for calculating R-factors(gt) *etc*. and is not relevant to the choice of reflections for refinement. R-factors based on F^2 are statistically about twice as large as those based on F, and R- factors based on ALL data will be even larger.

Fractional atomic coordinates and isotropic or equivalent isotropic displacement parameters (Å^2)

	x	y	z	$U_{iso}*/U_{eq}$
O1	0.0058 (4)	0.37096 (19)	0.50201 (18)	0.0316 (5)
O2	−0.1536 (4)	0.4257 (3)	0.33811 (19)	0.0385 (6)
O1W	0.6455 (4)	0.4874 (3)	0.6172 (2)	0.0467 (7)
H1W1	0.776 (5)	0.465 (5)	0.601 (3)	0.070*
H1W2	0.661 (8)	0.517 (5)	0.682 (2)	0.070*
N1	0.4267 (4)	0.3850 (2)	0.4302 (2)	0.0297 (6)
H1A	0.5663	0.3957	0.3994	0.045*
H1B	0.4074	0.3014	0.4493	0.045*
H1C	0.4141	0.4348	0.4925	0.045*
C1	0.0133 (6)	0.4049 (2)	0.4006 (3)	0.0302 (6)
C2	0.2504 (5)	0.4229 (3)	0.3472 (2)	0.0294 (6)
C3	0.2766 (6)	0.3382 (3)	0.2440 (3)	0.0367 (7)
H3A	0.2503	0.2486	0.2648	0.055*
H3B	0.4301	0.3474	0.2139	0.055*
H3C	0.1664	0.3640	0.1869	0.055*
C4	0.2832 (6)	0.5657 (3)	0.3178 (3)	0.0344 (7)
H4A	0.4384	0.5779	0.2885	0.041*
H4B	0.1762	0.5888	0.2572	0.041*
C5	0.2477 (7)	0.6556 (3)	0.4151 (3)	0.0451 (9)
H5A	0.2565	0.7442	0.3886	0.068*
H5B	0.3650	0.6408	0.4716	0.068*

H5C	0.0984	0.6401	0.4482	0.068*

Atomic displacement parameters (Å^2)

	U^{11}	U^{22}	U^{33}	U^{12}	U^{13}	U^{23}
O1	0.0326 (11)	0.0293 (9)	0.0329 (10)	0.0009 (8)	0.0031 (9)	0.0061 (9)
O2	0.0268 (11)	0.0510 (13)	0.0376 (11)	0.0010 (11)	0.0007 (9)	0.0110 (11)
O1W	0.0379 (13)	0.0627 (17)	0.0396 (13)	0.0087 (13)	−0.0008 (11)	−0.0148 (12)
N1	0.0268 (12)	0.0270 (11)	0.0352 (14)	−0.0004 (9)	0.0016 (11)	0.0015 (11)
C1	0.0324 (15)	0.0214 (12)	0.0367 (15)	−0.0017 (11)	0.0027 (12)	0.0020 (11)
C2	0.0279 (15)	0.0275 (13)	0.0328 (14)	−0.0009 (12)	0.0009 (11)	0.0051 (13)
C3	0.0371 (18)	0.0392 (16)	0.0338 (16)	0.0040 (15)	−0.0005 (13)	0.0010 (14)
C4	0.0320 (15)	0.0317 (15)	0.0394 (15)	−0.0038 (13)	−0.0001 (13)	0.0085 (14)
C5	0.055 (2)	0.0260 (14)	0.054 (2)	−0.0030 (14)	0.0026 (18)	0.0020 (15)

Geometric parameters (Å, º)

O1—C1	1.261 (4)	C2—C4	1.545 (4)
O2—C1	1.255 (4)	C3—H3A	0.9800
O1W—H1W1	0.830 (19)	C3—H3B	0.9800
O1W—H1W2	0.832 (19)	C3—H3C	0.9800
N1—C2	1.490 (4)	C4—C5	1.507 (5)
N1—H1A	0.9100	C4—H4A	0.9900
N1—H1B	0.9100	C4—H4B	0.9900
N1—H1C	0.9100	C5—H5A	0.9800
C1—C2	1.550 (4)	C5—H5B	0.9800
C2—C3	1.524 (4)	C5—H5C	0.9800
H1W1—O1W—H1W2	102 (3)	C2—C3—H3B	109.5
C2—N1—H1A	109.5	H3A—C3—H3B	109.5
C2—N1—H1B	109.5	C2—C3—H3C	109.5
H1A—N1—H1B	109.5	H3A—C3—H3C	109.5
C2—N1—H1C	109.5	H3B—C3—H3C	109.5
H1A—N1—H1C	109.5	C5—C4—C2	114.2 (3)
H1B—N1—H1C	109.5	C5—C4—H4A	108.7
O2—C1—O1	126.2 (3)	C2—C4—H4A	108.7
O2—C1—C2	116.4 (3)	C5—C4—H4B	108.7
O1—C1—C2	117.4 (3)	C2—C4—H4B	108.7
N1—C2—C3	108.1 (3)	H4A—C4—H4B	107.6
N1—C2—C4	108.6 (3)	C4—C5—H5A	109.5
C3—C2—C4	111.3 (3)	C4—C5—H5B	109.5
N1—C2—C1	109.1 (2)	H5A—C5—H5B	109.5
C3—C2—C1	110.7 (3)	C4—C5—H5C	109.5
C4—C2—C1	108.9 (2)	H5A—C5—H5C	109.5
C2—C3—H3A	109.5	H5B—C5—H5C	109.5
O2—C1—C2—N1	−175.8 (3)	O1—C1—C2—C4	−114.2 (3)
O1—C1—C2—N1	4.2 (4)	N1—C2—C4—C5	−64.0 (3)

O2—C1—C2—C3	−57.0 (3)	C3—C2—C4—C5	177.0 (3)
O1—C1—C2—C3	123.0 (3)	C1—C2—C4—C5	54.7 (4)
O2—C1—C2—C4	65.8 (3)		

Hydrogen-bond geometry (Å, º)

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
O1*W*—H1*W*1···O1[i]	0.83 (2)	2.05 (2)	2.811 (3)	152 (4)
O1*W*—H1*W*2···O2[ii]	0.83 (2)	1.96 (2)	2.787 (3)	171 (5)
N1—H1*A*···O2[i]	0.91	1.84	2.745 (3)	177
N1—H1*C*···O1*W*	0.91	2.09	2.792 (4)	133
N1—H1*B*···O1[iii]	0.91	1.98	2.832 (3)	156

Symmetry codes: (i) *x*+1, *y*, *z*; (ii) −*x*+1/2, −*y*+1, *z*+1/2; (iii) *x*+1/2, −*y*+1/2, −*z*+1.