STRUCTURAL SCIENCE
CRYSTAL ENGINEERING
MATERIALS

Volume 72 (2016)

Supporting information for article:

Structure-forming units of amino acid maleates. Case study of l-valinium hydrogen maleate

Denis Rychkov, Sergey Arkhipov and Elena Boldyreva

S1. Experimental details

S1.1. Materials

L-Valine (Fluka, 99%) and maleic acid (Himreaktiv, 98%) were used without preliminary purification.

S1.2. Crystallization

To test preliminary if a new salt can be formed, procedure described in (Rychkov et al., 2014) was used. Formation of a new phase was detected using IR spectroscopy (Figure S2) and powder XRD analysis (Figure S1). Subsequent crystallisation using slow-evaporation technique was used to grow transparent well-shaped block type crystals. Crystals of L-valine and maleic acid were also grown using the same technique.

S1.3. X-Ray powder diffraction

All the powder samples were characterized by powder XRD analysis (STOE STADI-MP diffractometer, Cu Kα radiation (λ = 1.54060 Å), MYTHEN 1K detector, 2Θ ranging from 5° to 51.8°, step 1.05°, time on step 10 seconds with an operating potential of 40 kV and a current of 40 mA. Data were processed using Win XPOW program of STOE company [Stoe& Cie. WinXPOW. Stoe&Cie GmbH, Darmstadt, Germany, 2002].

S1.4. Crystal structure determination

The crystal structure of L-valinium hydrogen maleate has not been reported prior to this work. The crystal structures of L-valine and maleic acid were re-visited and refined again because either the quality of data deposited for them at CCDC database was not sufficient to be used for energy calculations, or data have been collected at different experimental temperatures what can also affect energy calculations‡. In the present work diffraction data were collected at room temperature (298 K) using a Stoe IPDS II diffractometer, Mo Kα radiation (λ=0.71073 Å). All structures were solved using Olex2 (Dolomanov et al., 2003, 2009) with the ShelXS (Sheldrick, 1997, 2008) and refined with ShelXL (Sheldrick, 2008) refinement package using Least Squares minimization. All H atoms were initially located in a difference Fourier map. The positions of all H atoms were subsequently geometrically optimized and refined using a riding model, with the following assumptions and restraints:

For **L-valine**: N—H=0.89 Å and Uiso(H)=1.2Ueq(N) for NH$_3$ group, C—H=0.98 Å Uiso(H)=1.2Ueq(C) for C-H group, C—H=0.96 Å and with Uiso(H)=1.5Ueq(C) for CH$_3$ group.

For **L-valinium hydrogen maleate**: N—H=0.89 Å and Uiso(H)=1.2Ueq(N) for NH$_3$ group, O—H=0.82 Å and Uiso(H)=1.5Ueq(O) for OH groups of maleate anion and L-valinium catione, C—H=0.98 Å Uiso(H)=1.2Ueq(C) for C-H group in L-valinium catione, C—H=0.96 Å with Uiso(H)=1.5Ueq(C) for CH$_3$ groups, C—H=0.93 Å Uiso(H)=1.2Ueq(C) for C-H groups in maleic acid.

For **maleic acid**: O—H=0.82 Å and Uiso(H)=1.5Ueq(O) for OH groups and C—H=0.93 Å Uiso(H)=1.2Ueq(C) for C-H groups.

‡ Maleic acid: "MALIAC", R-factor (%) 20, (Shahat, 1952); "MALIAC2", R-factor (%) 8.2, (Gupta & Mahata, 1975), "MALIAC11", R-factor (%) 5(James & Williams, 1974), "MALIAC12", R-factor (%) 3.87, T=180 K, (Day et al., 2006); "MALIAC13", R-factor (%) 2.61, T=180 K, Polymorph II, (Day et al., 2006);

‡ L-valine: "LVALIN", R-factor (%) 12.6, (Torii & Iitaka, 1970); "LVALIN01", R-factor (%) 3.4, T=120K, (Dalhus et al., 1996); "LVALIN02", R-factor (%) 3.72, T=270K (Wang et al., 2002) and "LVALIN03" (Gong et al., 2005) does not contain atom coordinates.

S1.5. IR spectroscopy

Samples were prepared by mixing equimolar amounts (about 1 micromole for each substance) and co-grinding in an agate mortar. Thereafter, about half of the mixture was removed from the mortar and this part was termed as "mixture". A small drop of water was added to the remaining half of the sample and co-grinding was carried out again. The resulting sample was allowed to dry under ambient conditions for 10-15 hours. This sample was termed "product". The FTIR ATR spectra from the "mixture" and the "product" were recorded using a DigiLab Excalibur 3100, Varian spectrometer equipped with a MIRacle ATR accessory in the range 600–4000 cm^{-1} with resolution of 2 cm^{-1} without any sample preparation.

S1.1. SHG measurements

The powder second harmonic generation (SHG) test was carried out for L-valinium hydrogen maleate samples (particle size 100–200 μm) using Kurtz and Perry technique. "Standa" STA-01-7 pulsed laser with a wavelength of 1062 nm was used as radiation source, with pulse duration of 0.6 ns, a pulse repetition rate of 1 kHz, and an average power of 100 mW. The laser beam was directed onto the sample placed in a thermostat. Quartz crystals were used for calibrating the SHG intensity. A backscattered second harmonic signal (531 nm) was directed through a collimator to a slit of a MDR-2 monochromator, detected by a FEU photomultiplier, and accumulated in a computer for the time necessary to achieve the required accuracy of measurements.

S1.2. Humidity measurements

Humidity was measured using an Ebro TFH 620 hygrometer (Germany), which allows one to measure relative humidity and temperature simultaneously. All measurements were done with an air probe.

S2. Computational details

S2.1. DFT calculations

All structures at ambient pressure were optimised using density functional theory with dispersion correction (DFT-D) coupled to the plane-wave pseudopotential (Kresse & Joubert, 1999) methods, as implemented in CASTEP (Clark et al., 2005) 5.5. version. The dispersion correction scheme of Tkachenko and Scheffler (Tkatchenko & Scheffler, 2009) was used in all calculations. Treatment of the electronic exchange and correlation was handled by the generalized gradient approximation (Grimme, 2006) (GGA) formalized by Perdew, Burke, and Ernzerhof (PBE) (Perdew et al., 1996; Grimme, 2006). Pseudopotentials that were generated using CASTEP were used; the plane-wave cutoff energy used throughout was 775 eV for L-valinium hydrogen maleate-, L-valine, and maleic acid, which ensured convergence of lattice parameters and total energies less than 2 meV per atom. Brillouin zone sampling was obtained using Monkhorst-Pack grids 2x2x1 for L-valinium hydrogen maleate-, 2x3x1 for L-Valine, and 2x2x2 for maleic acid, respectively. The structures were relaxed (using the Broyden, Fletcher, Goldfarb, and Shannon (BFGS) method (Fischer & Almlof, 1992), to allow both atomic coordinates and unit cell vector to be optimized simultaneously, while constraining space group symmetry. Convergence criteria were applied: maximum change in system energy = 2 x 10^{-5} eV, maximum root-mean-square (RMS) force = 0.0025 eV A, maximum RMS displacement = 0.002 A, maximum RMS stress = 0.01GPa. In order to determine lattice energy of experimental structures, single-point energy (SPE) calculations were performed on the preliminary optimized experimental structures of L-valine, maleic acid and L-valinium hydrogen maleate, and on isolated molecules taken from optimized experimental structures; to diminish intermolecular interaction between molecules in neighbour unit cells, the unit cell was enlarged twice. Initial atom coordinates were used from single crystal X-ray diffraction data obtained in this work.

S2.2. PIXEL calculations

Atom coordinates for calculations were taken from X-ray diffraction data obtained in this work, with normalization of H-bond positions as implemented in the CLP computer program package. Structure of L-valinium maleate was used to calculate the molecular electron density using GAUSSIAN09 (Frisch et al., 2009) with the MP2/6-31G(d,p) level of theory. The electron density was then used for lattice energy calculations using the PIXEL method (Gavezzotti, 2011). The condensation level was set to 4 and the cut-off at 30 Å, resulting in 531 molecules in the cluster. Calculations were also carried out for pairs of molecules, identifying significant values of pair-wise interactions (i.e. more than 2.5 kJ/mol). The output file from these calculations contains the total lattice energy and its

Coulombic, polarization, dispersion and repulsion components as well as most significant pair-wise interactions. Some underestimation of lattice energy for the new salt was observed, which is known for PIXEL (see PIXEL manual), resulting in lattice energies of -243.2 kJ/mol for L-valine, -108.0 kJ/mol for maleic acid, (-351.2 kJ/mol total) and -353.4kJ/mol for L-valinium hydrogen maleate. However, despite the differences in the absolute energy of the values calculated by different techniques, the salt was shown to be somewhat more stable than initial components also according to PIXEL calculations; the ratio of the relative contributions of different motifs (not the absolute values) calculated by different methods also agreed reasonably. To estimate the energies of the motifs, 4 molecules of L-valine and 4 molecules of maleic acid were used. All pair-wise interactions (3 "A-A" interactions, 3 "B-B" interactions and 4 "A-B" interactions, where A is L-valine and B is maleic acid) were summed in each motif. Interactions between molecules of different motifs were completely neglected in these calculations, what is obviously a large oversimplification of the model and can be improved in the extension of the present work.

Table S1 $C_2^2(12)$ motifs in different amino acid maleates according to the latest literature data adapted and updated from(Arkhipov et al., 2015).18 out of 25 known structures of amino acid maleates contain $C_2^2(12)$ motif or its variation.

| №	Compound name, Maleates (M)	CCDC Refcodes	Comments related to $C_2^2(12)$ motif	Solubilities in water:Amino acids/maleic acid, g/100ml
1	(GlyH$^+$)*M$^-$	RENBAN	All these structures have $C_2^2(12)$ chains	25/78.8
2	(L-AlaH$^+$)*M$^-$	BOQTEG		16.65/78.8
3	(L-PheH$^+$)*M$^-$	EDAXIQ		2.96/78.8
4	(DL-PheH$^+$)*M$^-$	VAGVIJ		1.42/78.8
5	(DL-ValH$^+$)*M$^-$	QURSUR		7.04/78.8
6	(L-SerH$^+$)*M$^-$	REZPET		25*/78.8
7	(DL-SerH$^+$)* M$^-$	REZPAP		5.0/78.8
8	(DL-MetH$^+$)*M$^-$	MOCXUX		3.38/78.8
9	(SarH$^+$)*M$^-$	MIYBAX01		8.9/78.8
10	(L-ValH$^+$)*M$^-$	--------		8.85/78.8
11	(L-HisH$_2$$^{2+}$)*(M$^-$)$_2$	TENVOZ	$C_2^2(12)'$	4.16/78.8
12	(DL-ThrH$^+$)*M$^-$	ETEYOR	$C_2^2(12)''$	20.0/78.8
13	(β-AlaH$^+$)*M$^-$	EDASUX	$C_2^2(13)$because the backbone of β-alanine is one atom longer then the backbones of common amino acids	54.5/78.8
14	(BacH$^+$)*M$^-$		$C_2^2(14)$- backbone of baclofen is two atoms longer then the backbones of	0.43/78.8

			common amino acids	
15	$(L\text{-}LysH^+)*M^{-\dagger}$	XADTOL	The formation of $C_2^2(12)$ chains is	100/78.8
16	$(L\text{-}ArgH^+)*M^-*2H_2O^\dagger$	GIHGEK	impossible because the carboxylic group	5.0/78.8
17	$(DL\text{-}ArgH^+)*M^{-\dagger}$	*	of this amino acid is deprotonated, while	18.2/78.8
18	$(L\text{-}HisH^+)_2*(M^-)_2^\dagger$	XADTIF	the group of the side chain is protonated.	4.16/78.8
19	$(L\text{-}HisH^+)*M^-*H_2O$	TENVUF		4.16/78.8
20	$(L\text{-}HisH^+)_2*(M^-)_2*3H_2O$	VAZJUD		4.16/78.8
21	$(BetH^+)*M^-$	NASQED01	$C_2^2(12)$ chains cannot be formed because amino group of betaine is completely methylated	5.0/78.8
22	$(L\text{-}IleH^+)_2*(M^-)_2*H_2O$		$C_{2n}^{2n}(12n)$, $n = 2$	4.12/78.8
23	$(L\text{-}LeuH^+)_3*(M^-)_3$		$C_{2n}^{2n}(12n)$, $n = 3$	2.43/78.8
24	$L\text{-}Met*L\text{-}MetH^+*M^-$	**	$C_3^3(17)$ is formed by L-Met•••L-MetH$^+$ (case 22) or L-Nva•••L-NvaH$^+$ (case 23) di-cations	5.0/78.8
25	$L\text{-}Nva*L\text{-}NvaH^+*M^-$			4.87/78.8

[†] Maleates that were crystallized from acetonitrile. All other maleates were crystallized from water solutions.

* For $(DL\text{-}ArgH^+)*M^-$ atom coordinates were taken from (Ravishankar et al., 1998); ** the atomic coordinates of L-Met L-MetH$^+$*M$^-$ were not found in the papers by (Natarajan et al., 2008, 2010) or in the Cambridge Structural Database (Version 5.35, updates to February 2014;(Groom & Allen, 2014)), but in Fig. 3 in (Natarajan et al., 2008) a $C_3^3(17)$ motif can be clearly seen.

For additional information about structural motives in maleates see (Arkhipov et al., 2015). Structures were first published in (1) - (Rajagopal et al., 2001a), (2) - (Alagar et al., 2001b), (3) - (Alagar et al., 2001c), (4) - (Alagar et al., 2003), (5) - (Alagar et al., 2001a) , (6) and (7) - (Arkhipov et al., 2013), (8) - (Alagar et al., 2002), (9) - (Ilczyszyn et al., 2003), (10) - this paper, (11) - (Fleck et al., 2013), (12) - (Rajagopal et al., 2004), (13) - (Rajagopal et al., 2001c), 14 - (Báthori & Kilinkissa, 2015), (15) and (18) - (Pratap et al., 2000), (16) - (Sun et al., 2007),(16) and (17) - (Ravishankar et al., 1998), (19) - (Fleck et al., 2013), (20) - (Gonsago et al., 2012), (21) - (Haussühl & Schreuer, 2001), (22), (23) and (25) (Arkhipov et al., 2015), (24) - (Natarajan et al., 2010).

a)
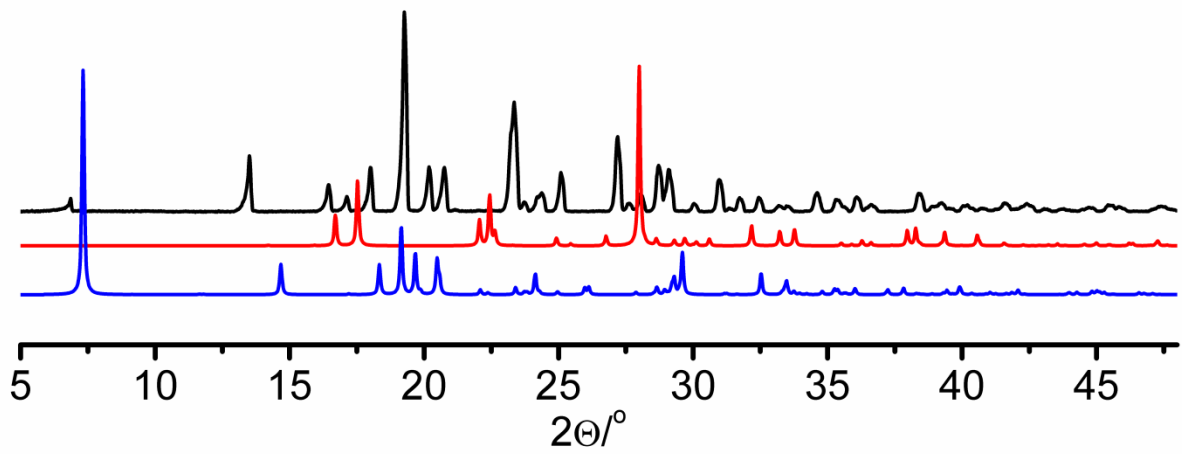
b)
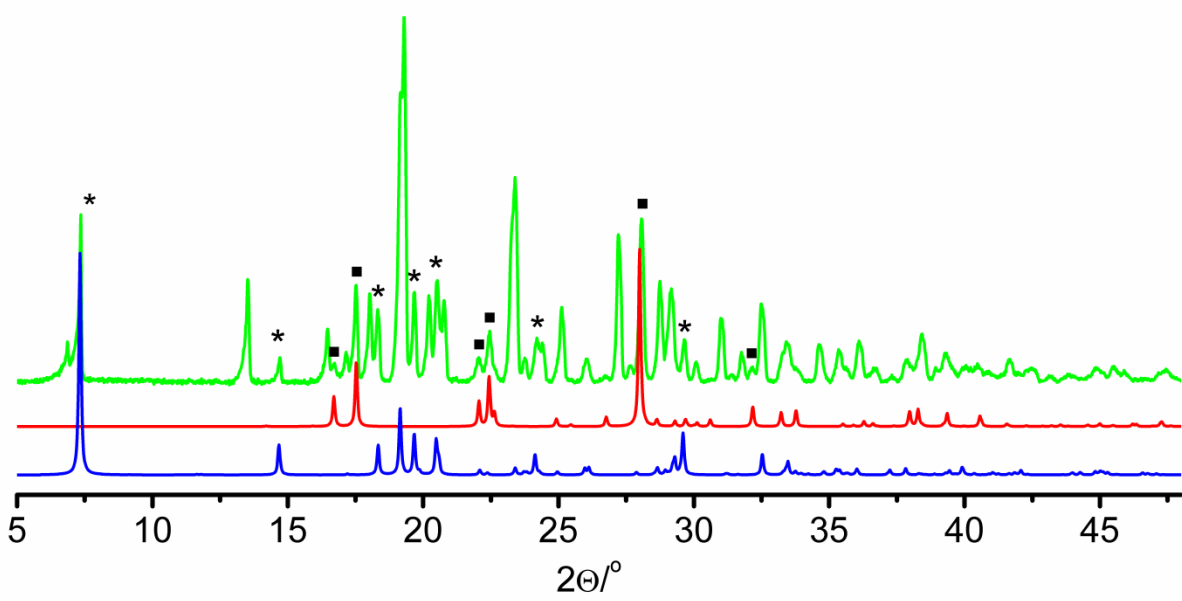
Figure S1   a) Powder diffraction pattern of the sample after liquid assisted grinding (LAG); a new phase has been formed without visible peaks of the starting substances as individual phases: L-valinium hydrogen maleate, experimental (Black), L-valine, calculated from single-crystal data obtained in this work (Blue) and maleic acid, calculated from single-crystal data obtained in this work (Red). For a comparison of the theoretical and experimental powder diffraction patterns of L-valinium hydrogen maleate see Fig.S3; b) Powder diffraction patterns of a sample ground without adding water on air with relative humidity of 50% (peaks from initial compounds are present): the sample ground without adding water (Green), L-Valine (Blue), maleic acid (Red). The peaks of initial substances in the XRPD pattern of the sample ground without water are marked with "star" symbol (*) for L-valine and "square" symbol (■) for maleic acid.
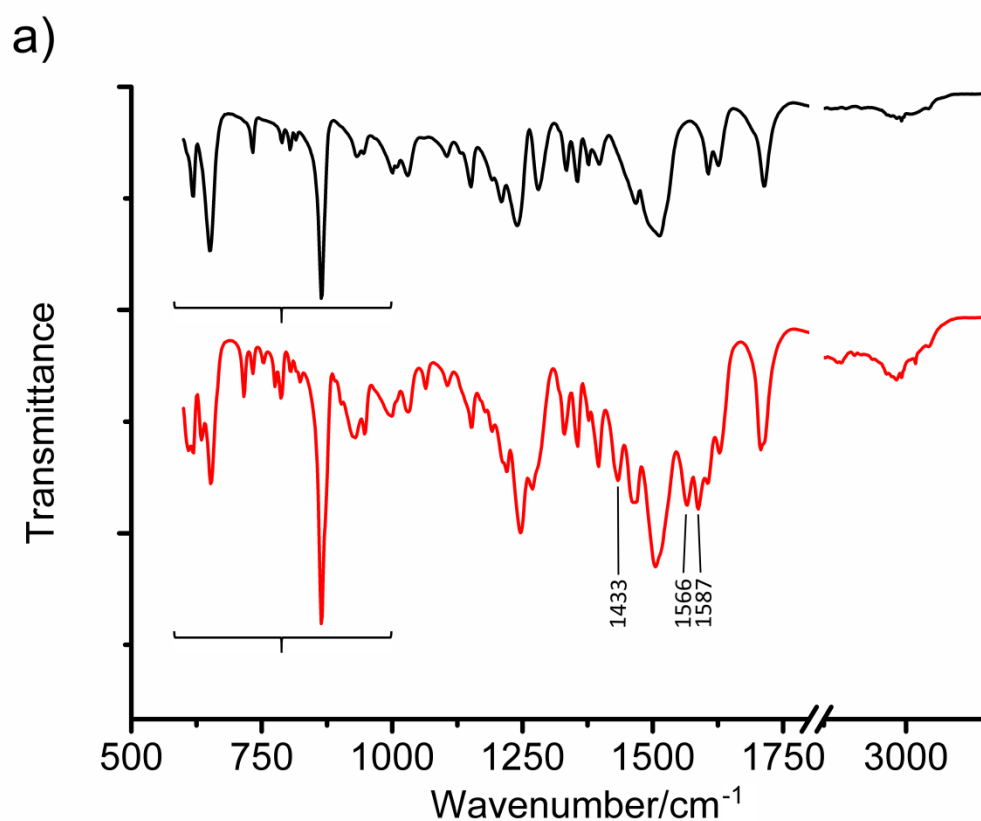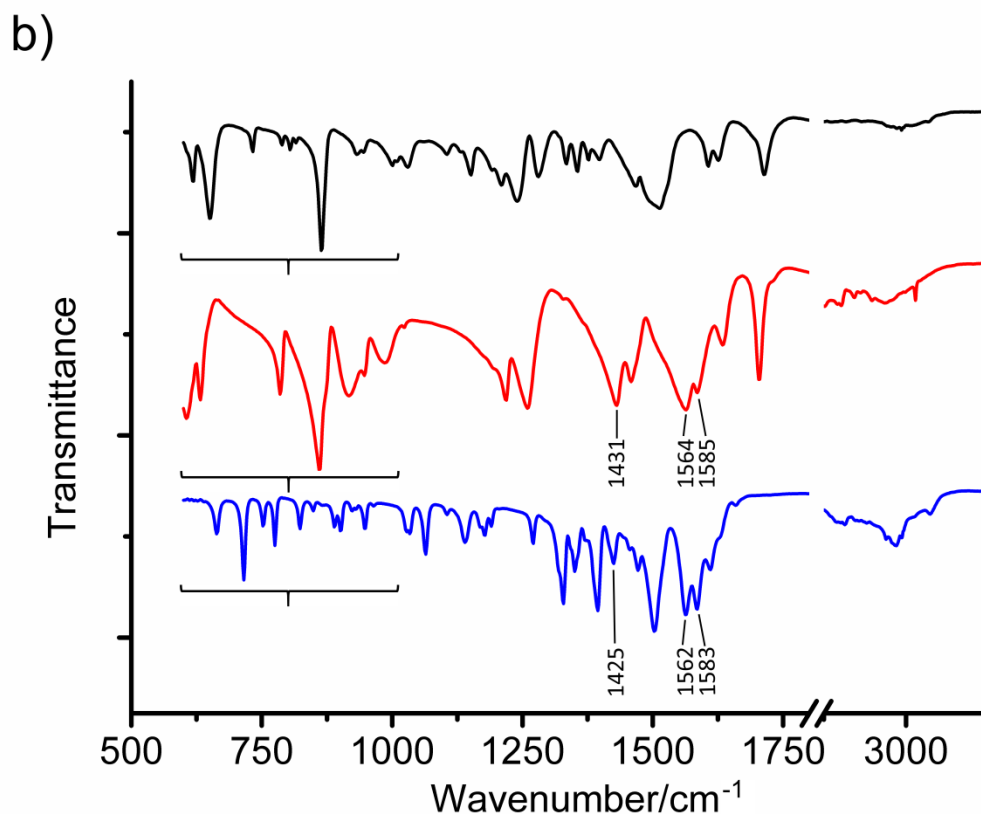
Figure S2 a) Infra-red spectra of the initial mixture of L-valine and maleic acid (Red), and of the new salt, L-valinium hydrogen maleate (Black); bands from initial compounds and from phase of L-valinium hydrogen maleate are present in physical mixture. b) Infra-red spectra of the individual components: L-valine (Blue), maleic acid (Red) and a new salt, L-valinium hydrogen maleate (Black). Plots are staggered along the y axis for clarity. Most significant differences in IR spectra are highlighted. For additional information and band assignment see Table S2

Table S2 A comparison of IR frequencies of L-Valine, Maleic acid, their physical mixture and L-Valinium hydrogen maleate.

L_Valine	Maleic Acid	Physical mixture of L_valine and Maleic acid -	L-valinium hydrogen maleate	Band assignment
-	605	609	619	COO bend + CCC bend
-	**632**	**634**	-	CO tor + CC tor
663	-	653	650	COO Bend
715	**-**	**715**	-	CC str + CO bend
-	-	732	732	CC str + COO bend
752	**-**	**754**	-	CC str + CO bend
775	**-**	**775**	-	COO bend
-	785	787	788	CO str
-	-	806	804	OH tor
823	**-**	**823**	-	OH tor
-	860	864	864	CH tor
903	**-**	**902**	-	CC str

-	916	929	931	OH bend
949	**947**	**947**	-	CN str
-	985	999	1001	CC str
1033	1024	1031	1030	CH$_3$ bend
1064	**-**	**1064**	-	CH$_3$ bend
1107	-	1107	1105	CH bend
1139	**-**	**1153**	-	CC str
1178	**-**	**1180**	-	CH bend
-	1219	1220	1209	CH bend
-	1259	1246	1240	CH bend
1271	-	1269	1280	CH bend
1328	-	1331	1334	CC str + OH bend
1350	-	1356	1356	CH$_3$ bend
1375	-	1377	1377	CH bend
1394	-	1396	1398	OH bend + CH bend
1425	**1431**	**1433**	-	COO str + OH bend + CH bend
1458	1458	1467	1467	CO bend + OH bend
1471	-	-	-	COO str + OH bend + CH bend
1502	-	1504	1512	COO str + OH bend + CH bend
1562	**1564**	**1566**	-	CC str + COO str
1583	**1585**	**1587**	-	CC str + COO str
1610	-	1604	1607	CC str + COO str +OH bend
-	1633	1628	1626	CC str + COO str +OH bend
-	1703	1707	1714	CO str

Due to the low intensity of characteristic bands after 2000 cm^{-1} IR band comparison and assignment was performed in the region 600cm^{-1} - 1700cm^{-1}. Most significant differences in IR spectra in the range 600cm^{-1} - 1700cm^{-1} are highlighted bold. Band assignment was performed using (Bellamy, 1980; Maçôas et al., 2001; Kumar, 2011).
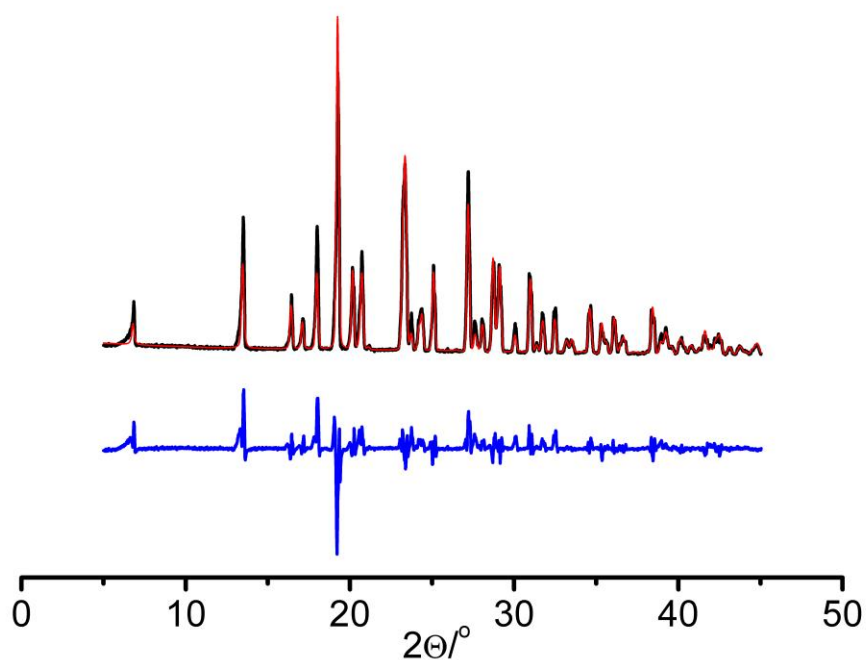
Figure S3 Powder diffraction patterns, calculated from single crystal diffraction data (Red) and measured experimentally (Black), and difference plot (Blue) proving that the powder sample is in fact pure L-valinium hydrogen maleate. Rietveld refinement was done using FullProf (Rodriguez-Carvajal, 1990)

Experimental details. Crystal structure determination.

Experiments were carried out at 298 K with Mo $K\alpha$ radiation. H-atom parameters were constrained.

	lval_all	lval_mal_all	Maleic acid
Crystal data			
Chemical formula	$C_5H_{11}NO_2$	$C_5H_{12}NO_2 \cdot C_4H_3O_4$	$C_4H_4O_4$
M_r	117.15	233.22	116.07
Crystal system, space group	Monoclinic, $P2_1$	Monoclinic, $P2_1$	Monoclinic, $P2_1/c$
a, b, c (Å)	9.6697 (16), 5.2749 (9), 12.063 (2)	5.7793 (8), 7.5974 (11), 12.9136 (17)	7.1511 (8), 10.1107 (11), 7.6405 (10)
β (°)	90.803 (14)	93.13 (1)	119.405 (8)
V (Å^3)	615.22 (18)	566.16 (14)	481.26 (10)
Z	4	2	4

μ (mm^{-1})	0.10	0.12	0.15
Crystal size (mm)	$0.35 \times 0.25 \times 0.1$	$0.3 \times 0.2 \times 0.1$	$0.25 \times 0.2 \times 0.15$
Data collection			
No. of measured, independent and observed [$I > 2\sigma(I)$] reflections	4649, 2516, 1333	3587, 2276, 1193	3499, 979, 541
R_{int}	0.042	0.045	0.071
$(\sin \theta/\lambda)_{max}$ (Å^{-1})	0.625	0.625	0.625
Refinement			
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.037, 0.071, 0.78	0.040, 0.065, 0.77	0.042, 0.107, 0.84
No. of reflections	2516	2276	979
No. of parameters	151	150	75
No. of restraints	1	1	0
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.13, −0.15	0.13, −0.16	0.16, −0.21

Computer programs: STOE *X-AREA*, STOE *X-RED*, *SHELXS97* (Sheldrick, 2008), *SHELXL2014* (Sheldrick, 2015), *Mercury* (Macrae *et al.*, 2006), *OLEX2* (Dolomanov *et al.*, 2009) and *publCIF* (Westrip, 2010).