Symmetry, pseudosymmetry and packing disorder in the alcohol solvates of L-leucyl-L-valine

CARL HENRIK GÖRBITZ* AND EIRIN TORGERSEN

Department of Chemistry, University of Oslo, PO Box 1033, Blindern, N-0315 Oslo, Norway.
E-mail: c.h.gorbitz@kjemi.uio.no

Abstract

The dipeptide L-Leu-L-Val can crystallize as a hydrate in the hexagonal space group $P6_5$ [Görbitz & Gundersen (1996). *Acta Chem. Scand.* **50**, 537–543], but forms 1:1 solvates when methanol, ethanol and 2-propanol are used as precipitating agents. The structures of these complexes can be divided into hydrophobic and hydrophilic layers. The alcohol molecules are embedded in the hydrophobic layers, but with the hydroxyl groups as crucial parts of the hydrogen-bonding network. L-Leucyl-L-valine–methanol (1/1) crystallizes in the space group $P2_1$ with $Z = 2$. L-Leucyl-L-valine–ethanol (1/1) has very special systematic absences, and the structure could not be solved with direct methods. Unraveling the actual build-up of the crystal was an unusual process involving modeling with molecular graphics programs. The solution shows a structure which formally belongs to the space group $P2_1$, with four dipeptide molecules and four solvent molecules in the asymmetric unit ($Z = 8$). Additional non-crystallographic symmetry in three dimensions fixes all cell angles, including β, to 90°. L-Leucyl-L-valine–2-propanol (1/1) is structurally closely related to the ethanol solvate, but owing to a rare type of packing disorder the length of the *a* axis is halved ($Z = 4$, $P2_12_12_1$). The hydrogen-bonding pattern is still the same as in the ethanol solvate, which means that the hydrogen-bond periodicity along the *a* axis in the 2-propanol solvate is two unit-cell lengths.

1. Introduction

A good hydrogen-bond network is a prerequisite for the formation of high-quality peptide crystals. This can not be achieved if large hydrophobic side chains are evenly distributed throughout the structure. Instead, these moieties aggregate into rather well defined entities, as observed in protein structures. For dipeptides with two hydrophobic residues our previous (Görbitz, 1997, and references therein) and ongoing work have identified two different modes of packing and hydrogen bonding:

(i) Formation of hydrophobic columns in hexagonal space groups with a *c* axis of approximately 10 Å.

(ii) Formation of hydrophobic layers in monoclinic or orthorhombic space groups. Usually these crystals include co-crystallized organic solvent molecules.

L-Leu-L-Val forms hexagonal crystals when aceto-nitrile, acetone, 1-propanol and several other solvents are used as precipitating agents (Görbitz & Gundersen, 1996), but layered 1:1 solvates with methanol (LVM), ethanol (LVE) and 2-propanol (LV2P).

L-Leu-L-Val
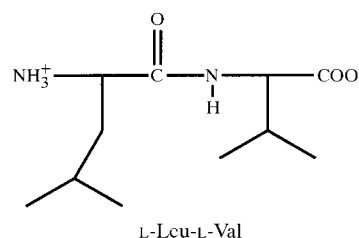
The resulting packing patterns are new for this class of compounds. Solving the structure of the LVE complex, with a rare type of pseudosymmetry giving four peptide molecules in the asymmetric unit, was a challenge in itself. A peculiarity of the LV2P structure is the apparent disorder of the amino and carboxylate groups. Disorder is rather common in peptide structures, but it is usually confined to side chains or solvent molecules. Very rarely is disorder observed for the charged N- and C-terminal groups, which are always fixed in place by an array of hydrogen bonds. The reason for this anomaly is discussed in detail.

2. Experimental

2.1. Preparation

L-Leu-L-Val was obtained from Sigma. The crystals were prepared by dissolving approximately 0.5 mg of the peptide in 30 µl of water, with subsequent diffusion of each of the alcohols into the aqueous solution. The crystals of LVM have limited stability in air, while crystals of the two other solvates are stable at room temperature.

2.2. Data collection

The data collections nominally covered over a hemisphere of reciprocal space by a combination of six sets of

exposures: two with the detector set at $2\theta = 29°$ and four with $2\theta = 55°$. Each set had a different φ angle for the crystal and each exposure covered $0.6°$ in ω. The crystal-to-detector distance was 5.0 cm. Coverage of the unique set is over 99% complete to at least $70°$ in 2θ.

All the LVE crystals tested were TLQS (twin-lattice quasi-symmetry) twins (Giacovazzo *et al.*, 1992). This was observed as a very small splitting of reflections at high 2θ angles. In the data reduction care was taken to make sure that both peaks were inside the integration box, which was slightly larger than usual (*x,y* size 2.0°, *z* size 0.80°). The twinning is then effectively treated as TLS (twin-lattice symmetry) since all reflections contain intensity from both twin components.

Experimental conditions are summarized in Table 1.

2.3. Structure determination

The structures of LVM and LV2P were solved routinely in space groups $P2_1$ and $P2_12_12_1$, respectively.

The unit-cell dimensions for LVE ($a = 11.0$, $b = 23.5$, $c = 12.3$ Å) are roughly similar to those of the orthorhombic cell of LV2P ($a = 5.2$, $b = 14.6$, $c = 22.4$ Å), except for a doubling of the a axis and the permutation of b and c. Reflections along the three axes in reciprocal space show the following systematic absences: $h00\ h = 4n + 2$ absent; $0k0\ k = 2n + 1$ absent; $00l\ l = 2n + 1$ absent. As this does not correspond to any of the 230 regular space groups, we suspected that some absences were due to pseudosymmetry rather than real symmetry. If reflections with h odd (which are systematically weak, indicating pseudotranslational symmetry) are removed from the data set [a-axis length halved, index transformation $h(\text{new}) = h(\text{old})/2$] we have $h00\ h = 2n + 1$ absent. As for LV2P the absences now correspond to $P2_12_12_1$ and an apparent solution was readily obtained with direct methods in this space group. An anisotropic
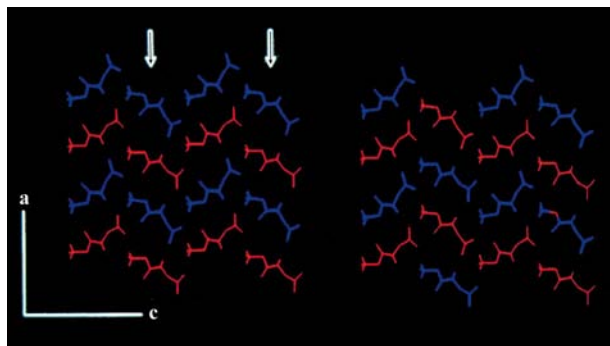

Fig. 1. The two alternative ways of stacking two dipeptide molecules with slightly different geometry in a two-dimensional hydrophilic sheet. The pattern to the left eventually proved to be correct (see text for the significance of the two arrows). The indication of the axes directions (view along the b axis) pertains to the final, correctly assigned $P2_1$ space-group setting of LVE. In Figs. 1 and 4, peptide side chains, C_α-bonded H atoms and the alcohol molecules have been removed for clarity.

refinement converged at an R factor around 0.13 with very elongated displacement ellipsoids for all atoms. Replacement of each ellipsoid by two independent atoms did not lower the R factor, but the new set of atoms clearly described two independent molecules with slightly different geometries.

Obtaining a structure from the complete data set was not trivial. We finally obtained a sensible structure suggestion from *SHELXS* (Sheldrick, 1990) in $P2_1$ with unique axis c (12.26 Å). Apparent solutions to the phase problem were not obtained in other space groups. However, upon anisotropic refinement two of the four peptide molecules in the asymmetric unit and two of the four ethanol molecules ended up with the same elongated ellipsoids as seen previously. The space group assignment was evidently not correct.

Further elaboration of the structure was then carried out using the molecular graphics program *SYBYL* (Tripos, 1996). The crystal structure of LV2P, with distinct hydrophilic and hydrophobic layers (see below), was used as a model and a search was started for alternative ways of packing two different peptide molecules (the solvent was disregarded at this stage) in a systematic three-dimensional pattern. The only constraint in this process was that there had to be an alternating sequence of unequal molecules along the a axis, since the pseudotranslational symmetry would otherwise not be observed. This manual procedure identified two alternative arrangements of a hydrophilic layer (Fig. 1). It may appear that there are several ways of stacking such a layer along the c axis by variable translations in the a and b directions, but all are identical after the proper unit-cell translations. Thus, there are only two possible structures, which were established to belong to the monoclinic space group $P2_1$ (as before), but with the 23.52 Å axis as unique. Both were generated from the refined $P2_1$ (unique axis 12.26 Å) structure by deleting all the disordered molecules and creating from the ordered peptide molecules **A** and **B** and ethanol molecules **E** and **F** new peptide molecules **C** and **D** and solvent molecules **G** and **H** by the proper symmetry transformations. Finally, the b and c axes were switched, with reversal of the c-axis direction to keep a right-handed system. One structure gave a significantly lower R factor when refined with *SHELXTL* (Sheldrick, 1994), and was on this basis assumed to be correct.

2.4. Structure refinement: LVM

The structure was refined with *SHELXTL* (Sheldrick, 1994). H atoms bonded to O or N were refined isotropically, other H atoms were placed geometrically and refined with constraints to keep all C—H distances and all C—C—H angles on one C atom the same, and to have U_{iso} values 1.2 times U_{eq} of the carrier atom, except that a free variable for U_{iso} was refined for each amino

Table 1. *Experimental details*

	LVM	LVE	LV2P
Crystal data			
Chemical formula	$C_{11}H_{22}N_2O_3.CH_4O$	$C_{11}H_{22}N_2O_3.C_2H_6O$	$C_{11}H_{22}N_2O_3.C_3H_8O$
Chemical formula weight	262.35	276.37	290.4
Cell setting	Monoclinic	Monoclinic	Orthorhombic
Space group	$P2_1$	$P2_1$	$P2_12_12_1$
a (Å)	5.2890 (1)	11.0112 (1)	5.1709 (1)
b (Å)	12.6877 (1)	23.5186 (1)	14.6136 (2)
c (Å)	11.4872 (1)	12.2636 (1)	22.4230 (4)
β (°)	95.8136 (5)	90.0121 (8)	—
V (Å^3)	766.871 (14)	3175.88 (4)	1694.40 (5)
Z	2	8	4
D_x (Mg m^{-3})	1.136	1.156	1.138
Radiation type	Mo $K\alpha$	Mo $K\alpha$	Mo $K\alpha$
Wavelength (Å)	0.71073	0.71073	0.71073
No. of reflections for cell parameters	8192	8192	7127
μ (mm^{-1})	0.084	0.085	0.082
Temperature (K)	150 (2)	150 (2)	150 (2)
Crystal form	Block	Block	Block
Crystal size (mm)	$0.50 \times 0.40 \times 0.35$	$0.60 \times 0.50 \times 0.25$	$0.55 \times 0.25 \times 0.10$
Crystal color	Colorless	Colorless	Colorless
Data collection			
Diffractometer	Siemens SMART CCD	Siemens SMART CCD	Siemens SMART CCD
Data collection method	Sets of exposures each taken over 0.6° in ω	Sets of exposures each taken over 0.6° in ω	Sets of exposures each taken over 0.6° in ω
Absorption correction	Multiscan (*SADABS*; Sheldrick, 1996)	Multiscan (*SADABS*; Sheldrick, 1996)	Multiscan (*SADABS*; Sheldrick, 1996)
T_{min}	0.959	0.950	0.952
T_{max}	0.971	0.979	0.980
No. of measured reflections	17 325	50 343	32 534
No. of independent reflections	12 154	38 561	12 521
No. of observed reflections	8119	28 178	9601
Criterion for observed reflections	$I > 2\sigma(I)$	$I > 2\sigma(I)$	$I > 2\sigma(I)$
R_{int}	0.0489	0.0388	0.0341
θ_{max} (°)	49.76	43.12	43.15
Range of h, k, l	$-11 \rightarrow h \rightarrow 8$	$-21 \rightarrow h \rightarrow 20$	$-9 \rightarrow h \rightarrow 7$
	$-27 \rightarrow k \rightarrow 24$	$-44 \rightarrow k \rightarrow 45$	$-28 \rightarrow k \rightarrow 28$
	$-23 \rightarrow l \rightarrow 23$	$-23 \rightarrow l \rightarrow 15$	$-43 \rightarrow l \rightarrow 36$
Refinement			
Refinement on	F^2	F^2	F^2
$R[F^2>2\sigma(F^2)]$	0.0656	0.0645	0.0559
$wR(F^2)$	0.1790	0.1441	0.1358
S	0.989	1.072	1.106
No. of reflections used in refinement	12 153	38 561	12 521
No. of parameters used	203	734	230
H-atom treatment	Mixture of independent and contrained refinement	Mixture of independent and contrained refinement	Mixture of independent and contrained refinement
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.0864P)^2]$ where $P = (F_o^2 + 2F_c^2)/3$	$w = 1/[\sigma^2(F_o^2) + (0.0434P)^2 + 0.3289P]$ where $P = (F_o^2 + 2F_c^2)/3$	$w = 1/[\sigma^2(F_o^2) + (0.0524P)^2 + 0.1584P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$	0.083	0.065	0.007
$\Delta\rho_{max}$ (e Å^{-3})	0.635	0.541	0.337
$\Delta\rho_{min}$ (e Å^{-3})	−0.282	−0.321	−0.225
Extinction method	None	None	None
Source of atomic scattering factors	*International Tables for Crystallography* (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4)	*International Tables for Crystallography* (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4)	*International Tables for Crystallography* (1992, Vol. C, Tables 4.2.6.8 and 6.1.1.4)
Computer programs			
Data collection	*SMART* (Siemens, 1995)	*SMART* (Siemens, 1995)	*SMART* (Siemens, 1995)
Cell refinement	*SAINT* (Siemens, 1995)	*SAINT* (Siemens, 1995)	*SAINT* (Siemens, 1995)
Data reduction	*SAINT* (Siemens, 1995)	*SAINT* (Siemens, 1995)	*SAINT* (Siemens, 1995)

Table 1 (*cont.*)

	LVM	LVE	LV2P
Structure solution	*SHELXS*86 (Sheldrick, 1990)	*SHELXS*86 (Sheldrick, 1990)	*SHELXS*86 (Sheldrick, 1990)
Structure refinement	*SHELXTL* (Sheldrick, 1994)	*SHELXTL* (Sheldrick, 1994)	*SHELXTL* (Sheldrick, 1994)
Preparation of material for publication	*SHELXTL* (Sheldrick, 1994)	*SHELXTL* (Sheldrick, 1994)	*SHELXTL* (Sheldrick, 1994)

and methyl group. Free rotation about C—C bonds was also permitted for methyl groups.

The hydroxyl group of the methanol molecule in LVM is disordered over two positions with occupancies 0.814 (6) and 0.186 (6). The O atom with the lower occupancy was refined isotropically. A riding refinement with tetrahedral covalent geometry was used for the associated H atom, which was oriented so as to point in the direction of the closest hydrogen-bond acceptor (O2A). The O-atom position with occupancy 0.186 (6) seems to be of little significance, and will not be discussed further.

2.5. *Structure refinement: LVE*

The structure was first refined in a traditional way as described for LVM, except that both C- and N-bonded H atoms were subject to geometry constraints. The crystal twinning was described by the twin-law matrix (1 0 0, 0 −1 0, 0 0 −1). The fractions of the two components refined to 0.561 (1) and 0.439 (1).

The systematically weak or unobserved $h00$ reflections result in a lower ratio of observed reflections to refined parameters and higher standard uncertainties in geometric parameters. The geometry of peptide molecule **A** is almost identical to that of peptide molecule **C** (see below), and the same is true for the pairs (**B**, **D**), (**E**, **G**) and (**F**, **H**). It was of interest to see if this circumstance could be used to reduce the number of refinement parameters without leading to higher values of $R(F)$ and $wR(F^2)$. Accordingly, refinement models incorporating various constraints were tested. *SHELXTL* (Sheldrick, 1994) has no special tools for handling non-crystallographic symmetry, but the sum of two parameters (usually occupancy factors) can be constrained to be 1, which was useful for our purpose. The constraints applied for equivalent atoms within each pair of molecules include:

(i) U^{ii} *values*: the atoms share the same set of U^{11}, U^{22} and U^{33} values.

(ii) y *coordinates*: for each atom pair $y_{\mathbf{X}} + y_{\mathbf{X}'} = 1$ (**X** = **A**, **B**, **E**, **F**; **X**' = **C**, **D**, **G**, **H**)

(iii) x *coordinates*: for the unconstrained refinement the sum $x_{\mathbf{X}} + x_{\mathbf{X}'}$ has an average value of 0.7485 and was constrained to be exactly 0.7485. (This was achieved in *SHELXTL* by translating the unit-cell origin and modifying the symmetry operator so that we could use the constraint $x_{\mathbf{X}} + x_{\mathbf{X}'} = 1$.)

We also tested a geometrical restraint (*SAME* in *SHELXTL*) requiring all bond lengths and angles involving C, N and O to be the same ($\sigma = 0.0001$) in the pairs (**A**, **C**), (**B**, **D**), (**E**, **G**) and (**F**, **H**).

2.6. *Structure refinement: LV2P*

The refinement strategy was generally similar to that used for LVM. There are two alternative positions for the carboxylate group, which after studies of the hydrogen-bond network were constrained to have equal occupancies [*versus* a free refinement which gives occupancies of 0.531 (11) and 0.469 (11) with no change for $R(F)$ or $wR(F^2)$]. Part A (Fig. 2) is geometrically related to the carboxylate groups in LVE **A** and **C**, while part B is related to LVE **B** and **D**. All four O atoms were refined anisotropically, while isotropic refinement was used for the two C atoms to keep standard uncertainties in geometric parameters as low as possible. With two carboxylate hydrogen-bond acceptor positions, there must also be two amino-group hydrogen-bond donor positions. Accordingly, the amino-N atom was split into N1A and N1B, both with occupancy 0.5.† In this case the labels 'A' and 'B' were chosen so as to emphasize the similarity with the hydrogen-bonding pattern in the LVE structure (see below). Additional splitting of C1A was tested in refinements, but not implemented in the final model since the refinement parameters did not improve.

3. Results and discussion

Refinement results are given in Table 1.‡ Additional results for several LVE refinement models are given in Table 2. The molecular structures are depicted in Fig. 2. The unit cells and crystal-packing patterns, with wave-like hydrophobic and hydrophilic layers, are shown in Fig. 3. Molecular geometry parameters are listed in Table 3.

Due to the LVE pseudotranslational symmetry (see above) and the high $2\theta_{max} = 86°$, many reflections are rather weak. Consequently, the resulting conventional R factor is not very low, around 0.065 for reflections with $I > 2\sigma(I)$, Table 1. Refinement with 24 758 independent reflections with d spacing greater than 0.62 Å ($2\theta_{max} \simeq$

---

† Failure to do so results in an unreasonably short N—H⋯O hydrogen bond.

‡ Supplementary data for this paper are available from the IUCr electronic archives (Reference: HA0173). Services for accessing these data are described at the back of the journal.

70°) gives $R(F) = 0.049$; 17 060 reflections with d spacing greater than 0.71 Å ($2\theta_{max} \simeq 60°$) give $R(F) = 0.039$. $wR(F^2)$ does not change, while the standard uncertainties in geometric parameters increase slightly (as they should). Without introduction of the *TWIN* command the LVE refinement converges at $R(F) = 0.174$ and $wR(F^2) = 0.461$.

3.1. *Symmetry and pseudosymmetry in the LVE structure*

The pseudotranslational symmetry along the a axis is observed directly from Fig. 3, relating **A** to **B** with a heavy-atom r.m.s. deviation of 0.130 Å, **C** to **D** with an r.m.s. deviation of 0.157 Å, and solvents **E** to **F** and **G** to **H**. All four peptide molecules in the asymmetric unit are present in each hydrophilic layer, the build-up of which is obtained from the left-hand pattern in Fig. 1 if red molecules are changed to white and blue molecules to yellow in stacks indicated by the arrows. The color coding is then **A** white, **B** yellow, **C** red and **D** blue. Neighboring hydrophilic layers are related by the crystallographic screw axis (b). Additionally, there is also a non-crystallographic screw axis parallel to the c axis at $x = \frac{3}{8}$ and $y = 0$, giving rise to $00l$ systematic absences, Fig. 4. It very accurately relates **A** to **C** and **B** to **D** (r.m.s. deviations 0.023 and 0.022 Å, respectively) and solvent molecules **E** to **G** and **F** to **H**. Finally, what may be referred to as a pseudo $2_{0.5}$ axis parallel to the a axis (translation $0.5/2 = 0.25$ of the unit-cell length between successive points) leads to the curious systematic absences observed for the $h00$ reflections. Fig. 5 shows a schematic view along the unique b axis with an indication of the symmetry and pseudosymmetry elements in the structure.

The deviation from real twofold symmetry along the c axis within a molecular layer was examined further by applying a regular twofold screw-axis operation to the coordinates of the **C**, **D**, **G** and **H** molecules, thus putting them approximately on top of the corresponding **A**, **B**, **E** and **F** molecules. This procedure shows conspicuous shifts for atoms in the solvent molecules. The maximum separation is 0.21 Å between C2E and C2H. Solvent molecules **G** and **H** move in opposite directions relative to **E** and **F**, and primarily along the a axis, but also along the c axis.

For the peptide molecules the non-crystallographic symmetry is very accurate, with 0.01 Å sliding motions parallel to the c axis of **A** relative to **C** and in the opposite direction for **B** relative to **D**. Additional small deviations occur only for C5 in the L-Leu side chain of **A** relative to **C** and for the peptide carbonyl O1 and L-Val side chain C9 of **B** relative to **D**. The marginally different intermolecular contacts for **A** and **C**, and for **B** and **D**, (Fig. 5) result in some very modest differences between equivalent torsion angles. The largest deviation is observed for C1—C2—C3—C5 ($\chi_1^{2,2}$ in the L-Leu side chain), which is 164.5 (2)° in **A** and 166.3 (2)° in **C**.
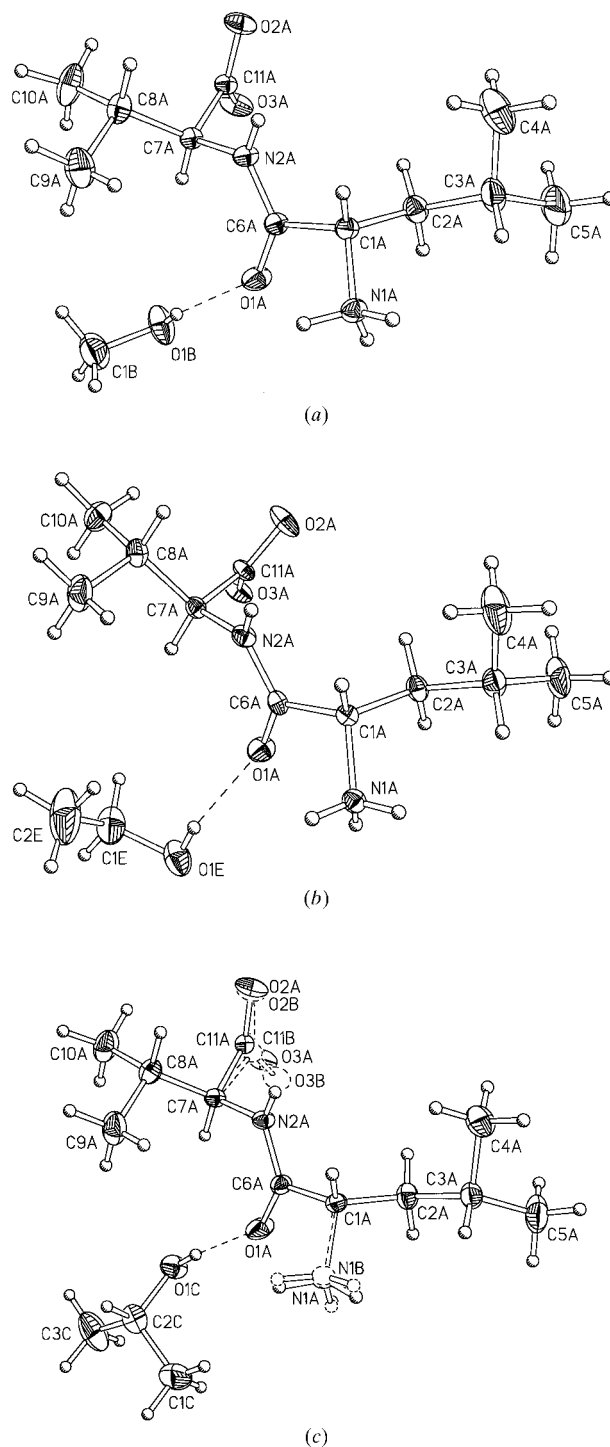

Fig. 2. Atomic numbering of (*a*) the L-Leu-L-Val methanol solvate (LVM), (*b*) the ethanol solvate (LVE, molecules **A** and **E**) and (*c*) the 2-propanol solvate (LV2P). Displacement ellipsoids for heavy atoms are shown at the 50% probability level. H atoms are shown as spheres of arbitrary radii. The least-populated position of the methanol O1B atom in LVM has not been shown. For LV2P the amino and carboxylate groups are both disordered over two positions with 50% occupancy.

3.2. *Molecular geometry*

L-Leu-L-Val has normal bond lengths and angles in the LVM and LV2P complexes. The exception is the LV2P N1−C1−C6 angle, Table 3, for which the difference between part A and part B of the disordered amino group is overestimated due to the refinement strategy, with two alternative positions for N1, but only one for C1 and C6. The peptide main chains are in a common extended dipeptide conformation. Unlike any of the four independent molecules in the L-Leu-L-Val·$\frac{3}{4}$H$_2$O structure (Görbitz & Gundersen, 1996), the most common rotamers are observed for the L-Leu side chains ($\chi_1^1/\chi_1^{2,1},\chi_1^{2,2}$ = *gauche$^-$/trans,gauche$^-$*) as well as for the L-Val side chains ($\chi_2^{1,1}/\chi_2^{1,2}$ = *trans/gauche$^-$*).

Bond lengths for the four independent dipeptide molecules in LVE are very similar, with the exception of the carboxylate group, Table 3. Bond angles, on the other hand, vary more between **A** and **B** or **D** than between **A** and **C**. All four peptide molecules have conformations rather similar to those observed for L-Leu-L-Val in the LVM and LV2P structures, but **B** and **D** have unique peptide bonds which are among the most twisted observed in linear peptide systems with $\omega_{\mathbf{B}}$ = 162.9 (2) and $\omega_{\mathbf{D}}$ = 162.3 (2)$^\circ$.

3.3. *The refinement models for LVE*

From Table 2 it can be seen that the introduction of U^{ii} constraints gives a substantial reduction in the number of refinement parameters from 734 to 618. The standard uncertainties in geometric parameters become smaller at the expense of very modest increases for $R(F)$ and $wR(F^2)$. The effect of the *SAME* command, which forces almost identical bond lengths and bond angles

(but not torsion angles) within the molecular pairs (**A**, **C**), (**B**, **D**), (**E**, **G**) and (**F**, **H**) (essentially the mean values for the two molecules after normal refinement), is even larger, reducing average bond-length s.u.'s to 0.015 Å. The introduction of this restraint could be justified by the fact that geometrical parameters are affected by the molecular environment in the crystal ('crystal forces') and when the environments are (more or less) the same, as in the LVE crystal, bond lengths and bond angles should also be similar. The *SAME* command thus appears to be an attractive option for this
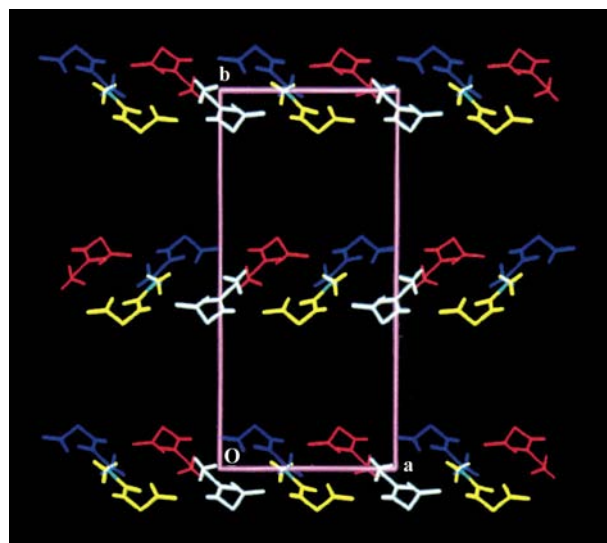

Fig. 4. View along the c axis for LVE. Note the non-crystallographic screw axis at $x = \frac{3}{8}$, $y = 0$.
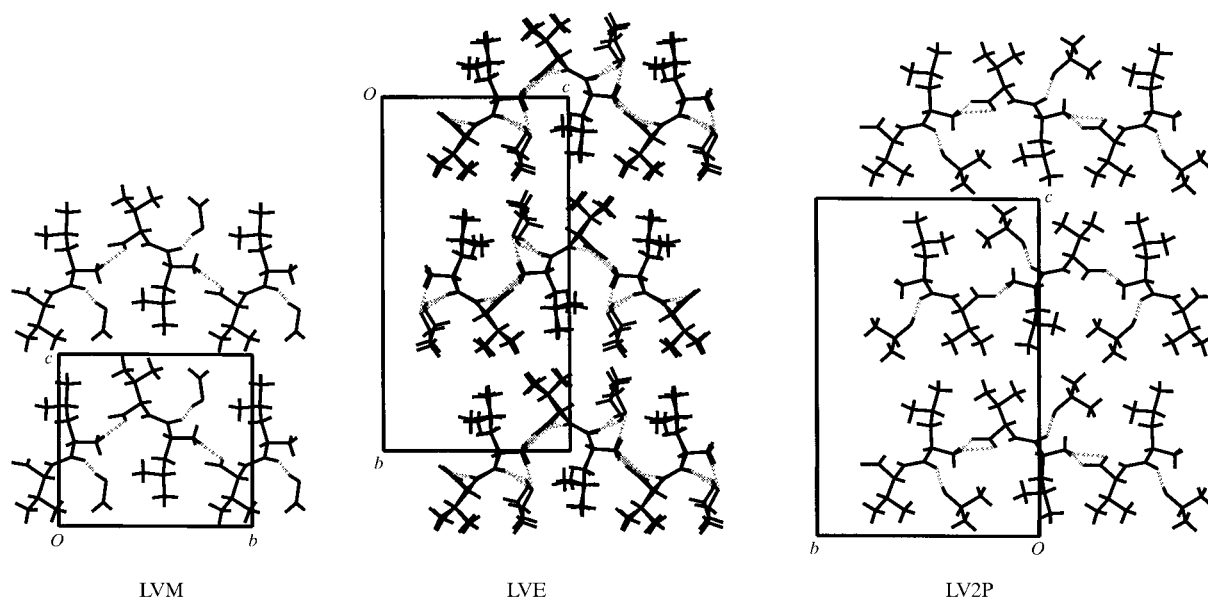

LVM

LVE

LV2P

Fig. 3. The packing arrangements viewed along the a axes with hydrogen bonds shown as dotted lines.

Table 2. *LVE refinement results with various constraints and geometry restraints*

Constraints and/or restraints	No. of parameters	$R(F)$	$wR(F^2)$	S.u.†
None	734	0.0645	0.1441	0.0033
U^{ii} values	618	0.0649	0.1449	0.0028
SAME restraint	734	0.0646	0.1445	0.0015
U^{ii} values + *SAME* restraint	618	0.0649	0.1451	0.0015
U^{ii} values + x coordinates	578	0.0682	0.1534	0.0028
U^{ii} values + y coordinates	578	0.0654	0.1463	0.0025
U^{ii} values + x and y coordinates	538	0.0687	0.1547	0.0022

† Average value for the s.u.'s of the heavy-atom bond lengths.

Table 3. *Selected geometric parameters (Å, °)*

	LVM **A**	LVE **A**	LVE **B**	LVE **C**	LVE **D**	LV2P **A**	LV2P **B**
N1—C1	1.4863 (11)	1.488 (2)	1.492 (2)	1.493 (2)	1.499 (2)	1.4984 (14)	1.4999 (14)
N2—C6	1.3316 (11)	1.331 (3)	1.340 (3)	1.332 (3)	1.342 (3)	1.3348 (9)	—
O1—C6	1.2394 (11)	1.240 (2)	1.233 (2)	1.233 (2)	1.228 (2)	1.2414 (9)	—
O2—C11	1.2495 (11)	1.248 (2)	1.261 (2)	1.241 (3)	1.257 (3)	1.256 (4)	1.260 (3)
O3—C11	1.2599 (11)	1.263 (2)	1.247 (2)	1.269 (2)	1.268 (3)	1.268 (3)	1.245 (3)
C7—C11	1.5454 (12)	1.547 (3)	1.539 (3)	1.544 (3)	1.537 (3)	1.5457 (13)	1.5563 (13)
C3—C4	1.532 (2)	1.511 (5)	1.523 (4)	1.521 (4)	1.534 (4)	1.5279 (14)	—
O1—C†	1.384 (2)	1.424 (4)	1.446 (4)	1.389 (4)	1.409 (3)	1.4354 (12)	—
N1—C1—C6	108.01 (7)	108.42 (15)	108.08 (15)	107.63 (15)	108.0 (2)	103.8 (2)	110.1 (2)
C1—C6—N2	114.36 (7)	115.3 (2)	114.7 (2)	114.5 (2)	114.6 (2)	115.98 (6)	—
C6—N2—C7	124.17 (7)	124.3 (2)	122.1 (2)	123.6 (2)	121.3 (2)	121.98 (6)	—
N2—C7—C11	107.16 (7)	107.7 (2)	110.5 (2)	107.5 (2)	110.1 (2)	110.14 (8)	106.01 (8)
O2—C11—O3	125.37 (9)	125.2 (2)	124.0 (2)	125.3 (2)	124.5 (2)	123.8 (2)	125.8 (2)
C2—C3—C5	108.59 (15)	109.6 (3)	109.1 (2)	109.8 (2)	110.4 (2)	110.69 (9)	—
C9—C8—C10	110.99 (12)	111.4 (3)	110.9 (3)	110.4 (3)	110.3 (3)	110.11 (8)	—
O1—C—C†	—	112.2 (3)	112.5 (3)	112.9 (4)	112.3 (2)	109.91 (9)	—
N1—C1—C6—N2 (ψ_1)	153.02 (8)	151.0 (2)	151.1 (2)	151.5 (2)	150.2 (2)	156.5 (2)	145.9 (2)
C1—C6—N2—C7 (ω_1)	169.87 (8)	162.9 (2)	168.1 (2)	162.3 (2)	168.4 (2)	179.37 (6)	—
C6—N2—C7—C11 (φ_2)	−116.54 (9)	−87.6 (3)	−85.2 (2)	−88.4 (3)	−84.8 (3)	−126.74 (9)	−117.65 (9)
N2—C7—C11—O2 (ψ_T)	−64.79 (10)	−54.8 (3)	−56.1 (2)	−54.2 (3)	−55.2 (3)	−71.5 (3)	−91.6 (3)
N1—C1—C2—C3 (χ_1^1)	−73.34 (11)	−73.5 (3)	−76.9 (2)	−72.5 (2)	−75.3 (2)	−82.5 (2)	−71.2 (2)
C1—C2—C3—C4 ($\chi_1^{2,1}$)	−78.59 (13)	−72.6 (3)	−77.4 (3)	−71.6 (3)	−78.7 (3)	−77.88 (9)	—
C1—C2—C3—C5 ($\chi_1^{2,2}$)	160.12 (13)	164.2 (3)	160.0 (2)	166.5 (2)	159.3 (2)	160.31 (9)	—
N2—C7—C8—C9 ($\chi_2^{1,1}$)	−62.30 (11)	−64.1 (3)	−66.8 (3)	−65.8 (3)	−67.3 (3)	−61.79 (9)	—
N2—C7—C8—C10 ($\chi_2^{2,2}$)	174.33 (11)	170.7 (2)	169.4 (2)	171.5 (2)	170.8 (2)	176.30 (7)	—

† In solvent molecule. For LVE these are **E**, **F**, **G** and **H**. For LV2P the O1B—C2B—C1B angle is given.

type of structure, either alone or combined with constraints for the anisotropic displacement parameters.

The introduction of *x*-coordinate constraints is incompatible with the observed relative shifts in solvent-molecule positions along the *a* axis (see above), causing a significant increase in $wR(F^2)$ (Table 2). Comparatively modest changes are observed with the introduction of a corresponding *y*-coordinate constraint, but some differences between bond lengths in equivalent molecules become unreasonably large. In summary, coordinate constraints were deemed unsuitable for the LVE structure.

3.4. *Hydrogen bonds*

The solvent molecules in the three structures are completely embedded in the hydrophobic layers, but with the hydroxyl groups constituting crucial parts of the hydrogen-bond networks. Hydrogen-bond parameters are given in Table 4 and the hydrogen-bond patterns in the hydrophilic layers of each structure are shown in Fig. 6.

In the LVM structure two of the amino-N H atoms are donated to carboxylate groups, one to an O2A atom and one to an O3A atom. The third amino-N H atom is pointing straight into the hydrophobic layer and can not find a peptide acceptor. Instead, this H atom is accepted by the methanol hydroxyl group, illustrating why the layered hydrophobic dipeptide structures include co-crystallized polar organic solvent molecules. The hydroxyl group is also a hydrogen-bond donor in O—H···O=C interactions, but this function is less important than its accepting ability since the peptide carbonyl group is frequently involved in C_α—H···O=C interactions when no alternative stronger donor is available. Hydrophobic dipeptides have thus also been crystallized as dimethyl sulfoxide solvates (Mitra & Subramanian,

1994; Mitra *et al.*, 1996). Special structures include L-Met-L-Met (Stenkamp & Jensen, 1975), where the third amino-N H atom is not used in hydrogen bonding at all, and L-Leu-L-Ala tetrahydrate (Görbitz, 1997) where the organic solvent molecule is replaced by a group of four water molecules.

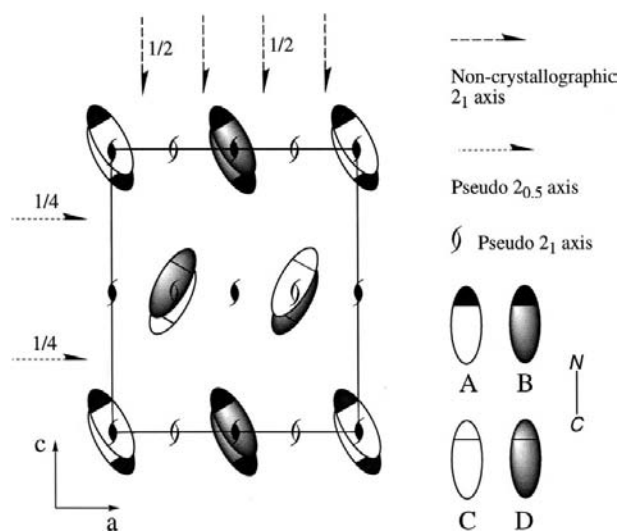Hydrogen bonds in LVE are essentially the same as in LVM, but with one important modification in that both $-NH_3^+\cdots{}^-OOC-$ interactions have O2 acceptor atoms for the **B** and **D** carboxylate groups. This very modest change in hydrogen-bonding pattern is responsible for the doubling of the *a* axis in the LVE structure compared to LVM.

The carboxylate group of LV2P has two alternative orientations, C_A and C_B, and there are also two alternatives, N_A and N_B, for the amino group. This leaves two possible hydrogen-bond patterns (see below) which happen to be equivalent to those illustrated in Fig. 1. By inspecting the hydrogen-bond geometries for each pattern is was clear that one, corresponding to that already observed for LVE, was more favorable. It follows that any deviation from a systematic, alternating sequence of carboxylate and amino orientations along the *a* axis leaves unreasonable hydrogen-bond geometries. The striking similarity between the hydrogen-bond patterns in LVE and LV2P is evident from Table 4 and Fig. 6. The carboxylate groups of LVE ($\mathbf{A} \simeq \mathbf{C}$, $\mathbf{B} \simeq \mathbf{D}$) are asymmetric with respect to the covalent C—O bond lengths, and even this feature is repeated in LV2P (Table 3).

Fig. 5. Schematic view along the *b* axis for LVE (solvent molecules omitted). There are non-crystallographic 2_1 screw axes parallel to the *c* axis relating molecules with almost identical geometries [(**A**, **C**) and (**B**, **D**)]. Pseudo 2_1 screw axes parallel to the *b* axis relate molecules with slightly different geometries, as do the pseudo $2_{0.5}$ screw axes parallel to the *a* axis (see text for details).
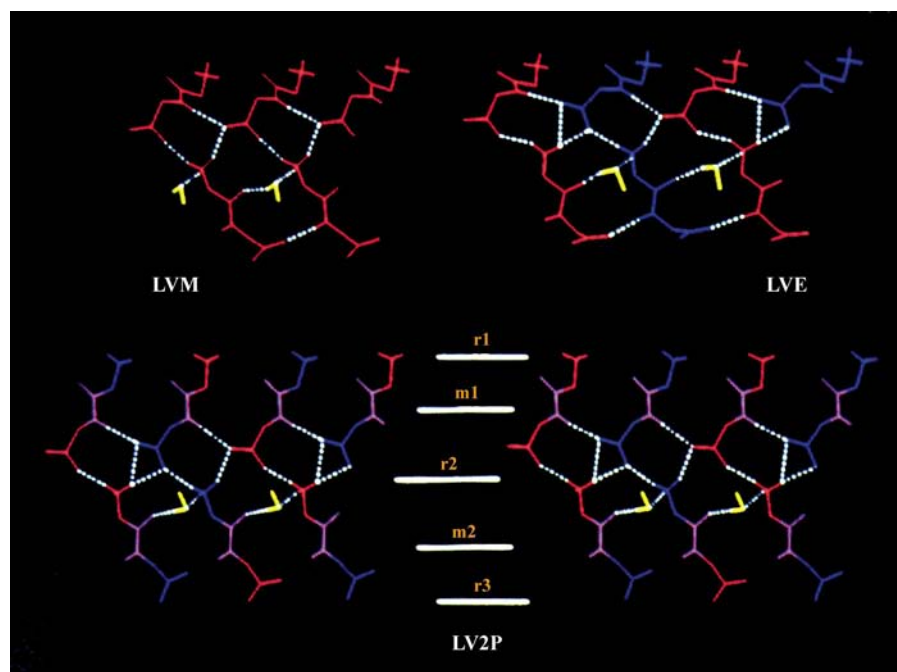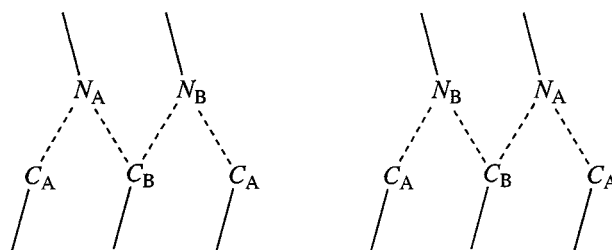

Fig. 6. Hydrogen-bonding patterns in the three crystal structures. As before, peptide side chains have been left out, but the alcoholic hydroxyl groups and the C atoms connected to them are now included (yellow). For LVM (top left) all molecules have A geometry (red), while for LVE (top right) the geometry and hydrogen-bonding interactions are either of type A (red) or type B (blue). For LV2P (bottom) N- and C-terminal ends have either A or B geometry, but the central part (magenta) is common to all molecules. Two molecular layers, *m1* and *m2*, have been identified, as well as the hydrogen-bonding regions *r1*, *r2* and *r3*.

Table 4. *Hydrogen-bond distances (Å) and angles (°)*

D—H$\cdots$O	D—H	H$\cdots$O	$D\cdots$O	D—H$\cdots$O
LVM				
N1A—H1A$\cdots$O2A[i]	0.87 (2)	1.87 (2)	2.731 (1)	173 (2)
N1A—H2A$\cdots$O3A[ii]	0.93 (2)	1.91 (2)†	2.759 (1)	150 (2)
N1A—H3A$\cdots$O1B[iii]	0.92 (2)	1.99 (2)	2.822 (1)	150 (2)
N2A—H4A$\cdots$O3A[iii]	0.81 (2)	2.05 (2)	2.848 (1)	168 (2)
O1B—H1B$\cdots$O1A	0.72 (3)	2.01 (3)	2.676 (1)	156 (3)
LVE‡				
N1A—H1A$\cdots$O2C[iv]	0.884 (15)§	1.957 (16)	2.721 (2)	144¶
N1A—H2A$\cdots$O2D[v]	0.884 (15)	2.019 (16)	2.892 (2)	169
N1A—H2A$\cdots$O3D[v]	0.884 (15)	2.570 (12)	3.067 (2)	116
N1A—H3A$\cdots$O1F[vi]	0.884 (15)	2.035 (14)	2.807 (2)	145
N1B—H1B$\cdots$O2D[v]	0.863 (16)§	1.934 (16)	2.770 (2)	163
N1B—H2B$\cdots$O3C[v]	0.863 (16)	1.927 (16)††	2.741 (2)	157
N1B—H3B$\cdots$O1E	0.863 (16)	2.031 (16)	2.831 (3)	154
N2A—H4A$\cdots$O3B[vi]	0.68 (3)§	2.03 (3)	2.712 (2)	177
N2B—H4B$\cdots$O3A	0.75 (3)§	2.11 (3)	2.853 (2)	174
O1E—H1E$\cdots$O1A	0.83 (4)	2.00 (4)	2.744 (2)	150 (3)
O1F—H1F$\cdots$O1B	0.79 (4)	2.13 (3)	2.848 (3)	151 (3)
LV2P				
N1A—H1A$\cdots$O2A[vii]	0.915 (15)§	1.813 (14)	2.709 (5)	166¶
N1A—H2A$\cdots$O2B[viii]	0.915 (15)	2.092 (14)	2.999 (5)	171
N1A—H2A$\cdots$O3B[viii]	0.915 (15)	2.362 (9)	2.989 (3)	126
N1A—H3A$\cdots$O1C[ix]	0.915 (15)	2.118 (14)	2.942 (5)	149
N1B—H1B$\cdots$O2B[vii]	0.903 (13)§	1.928 (13)	2.801 (5)	162
N1B—H2B$\cdots$O3A[viii]	0.903 (13)	2.016 (13)‡‡	2.846 (3)	152
N1B—H3B$\cdots$O1C[ix]	0.903 (13)	1.912 (14)	2.774 (5)	159
N2A—H4A$\cdots$O3B[ix]	0.827 (14)	1.973 (14)	2.796 (2)	173.4 (13)
N2A—H4A$\cdots$O3A[ix]	0.827 (14)	2.113 (14)	2.932 (2)	170.5 (12)
O1C—H1C$\cdots$O1A	0.87 (2)	1.83 (2)	2.682 (1)	167 (2)

Symmetry codes: (i) $2 - x$, $\frac{1}{2} + y$, $1 - z$; (ii) $1 - x$, $\frac{1}{2} + y$, $1 - z$; (iii) $x + 1$, y, z; (iv) $x - 1$, $y + 1$, z; (v) $x, y + 1, z$; (vi) $x - 1, y, z$; (vii) $-x, y - \frac{1}{2}, \frac{1}{2} - z$; (viii) $1 - x, y - \frac{1}{2}, \frac{1}{2} - z$; (ix) $x - 1, y, z$. † Distance to O2A[ii] = 2.90 (3) Å. ‡ Hydrogen bonds with **C** and **D** as donor molecules are equivalent to those listed for molecules **A** and **B**. § S.u.'s may be underestimated owing to the refinement procedure. ¶ S.u.'s not given when unreasonable owing to the refinement procedure. †† Distance to O2C[v] = 3.091 (15) Å. ‡‡ Distance to O2A[viii] = 2.775 (12) Å.

3.5. *LV2P packing disorder and the short a axis*

The hydrogen-bond pattern in LV2P is almost identical to the pattern in LVE. Why then, do the *a* axes in these two structures have different lengths? The reason lies in the fact that while the peptide molecules in LVE have either an A conformation at the N- and C-terminal ends, N_A–C_A, or a B conformation, N_B–C_B, the peptide molecules in LV2P can adopt four different conformations by combinations of A and B geometry: N_A–C_A, N_A–C_B, N_B–C_A, and N_B–C_B. The effect of this difference is profound: for LVE the hydrogen-bond pattern in one amino–carboxylate interaction region determines directly the arrangements in the adjacent regions, and consequently the complete structure, while in the LV2P structure the adjacent layer can always take two different positions. Thus, in Fig. 6 the different molecular conformations in layer *m1* give two possible arrangements in the hydrogen-bonding region *r1*, even if the arrangement in region *r2* remains the same. Similarly, there are two alternatives for hydrogen bonding in region *r3*, of which only one is shown. This disorder effectively cuts the length of the LV2P *a* axis in two, to only one-half of the hydrogen-bond periodicity in the *x* direction.†

4. Conclusions

Direct methods did not provide the right solution for the L-Leu-L-Val ethanol solvate (LVE), but the results could be combined with molecular graphics modeling in order to obtain the correct structure. The four peptide molecules in the asymmetric unit can be divided into two pairs related by pseudotranslational symmetry. The two molecules within each pair are related by a non-crystallographic but almost perfect twofold screw axis. Alternative structure refinements were explored since it appears reasonable to restrain the values of equivalent bond lengths and bond angles for each pair to be almost the same, thereby reducing the standard uncertainties

† To be certain than a doubling of the *a* axis had not been overlooked in our data collection for LV2P, the crystal was put back on the diffractometer and small set of new frames was collected. Careful scrutiny of reciprocal space revealed no trace of any additional, previously undetected Bragg peaks.

for the calculated geometric parameters. The molecules can also share the same set of U^{ii} values ($i = 1, 2, 3$), giving a significant reduction in the number of parameters required to refine the structure at the expense of an insignificant increase for $wR(F^2)$. The results of all these refinements were very similar.

Hydrogen bonding in the structure of L-Leu-L-Val 2-propanol solvate is identical to that in LVE, but an unusual type of disorder affects the arrangement of the molecules in the crystal. As a result of this, the hydrogen-bond periodicity is twice as long as the unit-cell length.

References

Giacovazzo, C., Monaco, H. L., Viterbo, D., Scordari, F., Gilli, G., Zanotti, G. & Catti, M. (1992). *Fundamentals of Crystallography*. Oxford University Press.

Görbitz, C. H. (1997). *Acta Cryst.* C**53**, 736–739.

Görbitz, C. H. & Gundersen, E. (1996). *Acta Chem. Scand.* **50**, 537–543.

Mitra, S. N., Govindasamy, L. & Subramanian, E. (1996). *Acta Cryst.* C**52**, 2871–2873.

Mitra, S. N. & Subramanian, E. (1994). *Biopolymers*, **34**, 1139–1143.

Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.

Sheldrick, G. M. (1994). *SHELXTL*. Version 5.03. Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.

Sheldrick, G. M. (1996). *SADABS. Program for Absorption Correction*. University of Göttingen, Germany.

Siemens (1995). *SMART and SAINT Area-Detector Control and Integration Software*. Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.

Stenkamp, R. E. & Jensen, L. H. (1975). *Acta Cryst.* B**31**, 857–861.

Tripos (1996). *SYBYL*. Version 6.3. Tripos Inc., 1699 s. Henley Road, St. Louis, MO 63144, USA.