Acta Cryst. (1996). B52, 999–1006

Structures of L-Valyl-L-Glutamine and L-Glutamyl-L-Valine

CARL HENRIK GÖRBITZ* AND PAUL HOFF BACKE

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

(Received 11 March 1996; accepted 22 May 1996)

Abstract

L-Val-L-Gln crystallizes in the orthorhombic space group $P2_12_12$ with $a = 16.419\,(3)$, $b = 15.309\,(3)$ and $c = 4.708\,(1)$ Å. The final $wR(F_o^2)$ is 0.100 for 2044 independent reflections, $R(F_o) = 0.050$ for 1475 reflections with $I > 2.0\sigma(I)$. L-Glu-L-Val crystallizes in the monoclinic space group $P2_1$ with $a = 6.487\,(2)$, $b = 5.505\,(2)$, $c = 16.741\,(4)$ Å and $\beta = 97.22\,(2)°$. The final $wR(F_o^2)$ is 0.111 for 1920 independent reflections, $R(F_o) = 0.047$ for 1576 reflections with $I > 2.0\sigma(I)$. Molecular geometries are normal, except for a unique eclipsed orientation of the charged amino group of L-Glu-L-Val. Dipeptides with a N-terminal hydrophobic residue and C-terminal hydrophilic residue are shown to have crystal packing patterns fundamentally different from those of dipeptides with the same types of residues in reversed order. Accordingly, the structure of L-Val-L-Glu [Eggleston (1984). *Acta Cryst.* C**40**, 1250–1252] is rather similar to L-Val-L-Gln, but different from its retroanalogue L-Glu-L-Val. Nevertheless, the pairing of hydrogen-bond donors and acceptors is the same for L-Val-L-Glu and L-Glu-L-Val, indicating very distinct hydrogen-bonding preferences. This is the first demonstration of such a coincidence among dipeptide structures. The differences between L-Val-L-Glu and L-Val-L-Gln structures stem from modifications of the molecular geometry and cell parameters due to the formation of an additional hydrogen bond from the extra donor in the L-Gln side chain.

1. Introduction

When molecules aggregate to form the systematic lattice of a crystal, it occurs in such a way so as to optimize hydrogen bonding, but under the condition that favourable hydrophobic interactions are also attained. For molecules with distinct hydrophobic as well as hydrophilic moieties, characteristic patterns are observed in which the hydrophobic groups segregate into two-dimensional layers or one-dimensional columns. In the course of our work on the crystal structures of dipeptides, we became interested in molecules with one hydrophobic and one hydrophilic residue. These compounds usually form reasonably complex two- or three-dimensional hydrogen-bond networks and utilize various mechanisms of aggregating the hydrophobic entities within the crystal. The present paper presents X-ray crystallographic studies of two mixed hydrophobic/hydrophilic dipeptides together with the results of database searches for related structures. A special emphasis is placed on the importance of the sequence of the two amino acid residues in dipeptides for crystal packing and hydrogen bonding.

1.1. *Nomenclature*

A hydrophilic amino acid residue has been denoted Hpi, a hydrophobic residue Hpo. The designation for common amino acids, L-Xaa, has been reduced to Xaa, implying the L-form.

2. Experimental

2.1. *Preparation*

Both compounds were obtained from Sigma and used as received. Crystals were grown by vapour diffusion of ethanol into 30 µl of an aqueous solution containing ~ 3 mg of the peptide.

2.2. *Data collection and refinement*

Experimental conditions and results from the refinements are given in Table 1. Both structures were solved by the direct methods program *MITHRIL* (Gilmore, 1984) and refined with *SHELXL* (Sheldrick, 1993). Positional parameters for H atoms bonded to N and O were refined, other H atoms were placed geometrically. Refinement then allowed the H atoms to move along the C—H bond direction with the C—H distance being kept constant for all H atoms on the same C atom. U_{iso} values were fixed at 1.2 × U_{eq} of the bonded atom, except that free variables for U_{iso} were refined for the amino group and two methyl groups in each molecule.

2.3. *Database studies*

Peptide structures were retrieved from the Cambridge Structural Database [(1995); Allen *et al.* (1991)] by means of the program *QUEST*. The main-chain and side-chain conformations were studied using the *GSTAT* program.

3. Results and discussion

*ORTEP*II (Johnson, 1976) drawings of the two molecules with the atomic numbering are shown in Fig. 1. Final atomic coordinates for heavy atoms are

Table 1. *Experimental details*

	L-Val-L-Gln	L-Glu-L-Val
Crystal data		
Chemical formula	$C_{10}H_{19}N_3O_4$	$C_{10}H_{18}N_2O_5$
Chemical formula weight	245.28	246.26
Cell setting	Orthorhombic	Monoclinic
Space group	$P2_12_12$	$P2_1$
a (Å)	16.419 (3)	6.487 (2)
b (Å)	15.309 (3)	5.505 (2)
c (Å)	4.708 (1)	16.741 (4)
β (°)		97.22 (2)
V (Å^3)	1183.4 (4)	593.1 (3)
Z	4	2
D_x (Mg m^{-3})	1.377	1.379
Radiation type	Mo $K\alpha$	Mo $K\alpha$
Wavelength (Å)	0.71069	0.71069
No. of reflections for cell parameters	25	25
θ range (°)	12.5–17.5	12.5–17.5
μ (mm^{-1})	0.107	0.111
Temperature (K)	120 (2)	120 (2)
Crystal form	Block	Plate
Crystal size (mm)	0.50 × 0.20 × 0.10	0.85 × 0.15 × 0.05
Crystal colour	Colourless	Colourless
Data collection		
Diffractometer	Nicolet P3	Nicolet P3
Data collection method	2θ scans	2θ scans
Absorption correction	None	None
No. of measured reflections	2044	1920
No. of independent reflections	2044	1920
No. of observed reflections	1475	1576
Criterion for observed reflections	$I > 2\sigma(I)$	$I > 2\sigma(I)$
θ_{max} (°)	30.0	30.0
Range of h, k, l	$0 \rightarrow h \rightarrow 23$	$-9 \rightarrow h \rightarrow 9$
	$0 \rightarrow k \rightarrow 21$	$-7 \rightarrow k \rightarrow 7$
	$0 \rightarrow l \rightarrow 6$	$-11 \rightarrow l \rightarrow 23$
No. of standard reflections	3	3
Frequency of standard reflections	Every 96 reflections	Every 96 reflections
Intensity decay (%)	<2.0	<2.0
Refinement		
Refinement on	F^2	F^2
$R(F)[F^2 > 2\sigma(F^2)]$	0.0502	0.0473
$wR(F^2)$	0.0999	0.1115
S	1.004	1.029
No. of reflections used in refinement	2044	1920
No. of parameters used	184	181
H-atom treatment	H atoms bonded to N and O refined	H atoms bonded to N and O refined
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.0432P)^2]$, where $P = (F_o^2 + 2F_c^2)/3$	$w = 1/[\sigma^2(F_o^2) + (0.0670P)^2]$, where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$	−0.016	0.006
$\Delta\rho_{max}$ (e Å^{-3})	0.314	0.336
$\Delta\rho_{min}$ (e Å^{-3})	−0.280	−0.281
Extinction method	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)
Computer programs		
Structure solution	*MITHRIL* (Gilmore, 1984)	*MITHRIL* (Gilmore, 1984)
Structure refinement	*SHELXL*93 (Sheldrick, 1993)	*SHELXL*93 (Sheldrick, 1993)

listed in Tables 2 and 3, molecular geometry is given in Tables 4 and 5.*

3.1. *Bond lengths and angles*

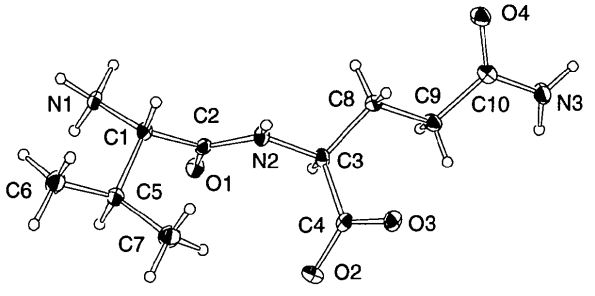
There are no remarkable values, but the >C10═O4 bond in the Gln side chain of Val-Gln is rather long, 1.250 (4) Å, since the O atom is unusual in participating in two hydrogen bonds (see below). Carbonyl double bonds are known to be sensitive to the number of H atoms accepted by the group (Taylor, Kennard & Versichel, 1984; Guo & Karplus, 1992). The carboxylate group is quite symmetric in both structures.

3.2. *Conformations*

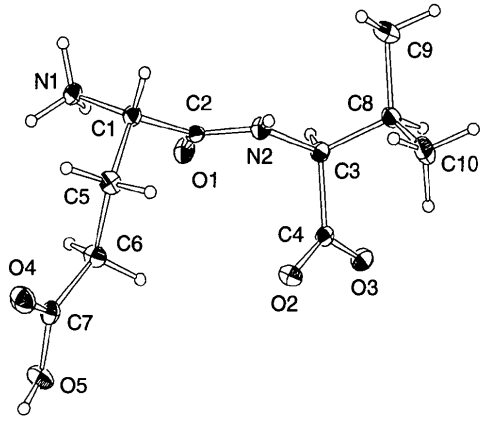
The dipeptide main chains occur in usual, slightly S-shaped conformations. The main difference lies in the orientation of the C-terminal carboxylate group, reflected

* Lists of atomic coordinates, anisotropic displacement parameters and structure factors have been deposited with the IUCr (Reference: HA0146). Copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

by the torsion angle N2—C3—C4—O2 (ψ_T), which is 75.0 (3)° for Val-Gln and −14.6 (3)° for Glu-Val. Both molecules have Val side chains, but with different conformations; in Val-Gln $\chi_1^{1,1}/\chi_1^{1,2}$ are *trans/gauche−*, while $\chi_2^{1,1}/\chi_2^{1,2}$ are *gauche+/trans* in Glu-Val. A survey of Val side chains in peptides (Görbitz & Gundersen, 1996) showed that these are the most common and least common rotamers, respectively, with relative frequencies 6:1. The orientation of the side chain was not found to depend significantly on the position of the residue within the peptide chain. From Fig. 1 it is clear that the Gln and Glu side chains occur in favourable *trans* conformations at the C^β—C^γ bond (χ_2), while orientations at the C^α—C^β bond are different with χ_1^2 *trans* in Val-Gln and χ_1^1 *gauche+* in Glu-Val. Data for other peptides with Gln and Glu residues are presented in Table 6. The number of observations is small, but the side-chain orientation appears to depend on the position of the residue within the peptide. Thus, the *gauche−* conformation, which is dominant for C-terminal residues and the only rotamer observed for central residues, is completely absent for N-terminal residues. It is not straightforward to explain these distributions, but consideration of molecular models suggests that a N-terminal *gauche−* conformation at χ^1 may cause the side chain



(*a*)



(*b*)

Fig. 1. Atomic numbering of (*a*) L-Val-L-Gln and (*b*) L-Glu-L-Val. Thermal ellipsoids for heavy atoms are shown at the 50% probability level. H atoms are shown as spheres of arbitrary size.

Table 2. *Fractional atomic coordinates and equivalent isotropic displacement parameters (Å²) for L-Val-L-Gln*

$U_{eq} = (1/3)\Sigma_i\Sigma_j U_{ij} a_i^* a_j^* \mathbf{a}_i.\mathbf{a}_j.$

	x	*y*	*z*	U_{eq}
O1	0.31741 (11)	0.40999 (11)	0.9818 (4)	0.0161 (4)
O2	0.26033 (10)	0.62396 (11)	1.2250 (4)	0.0157 (4)
O3	0.30745 (11)	0.73079 (11)	0.9517 (5)	0.0159 (4)
O4	0.57611 (11)	0.72086 (13)	0.5408 (5)	0.0180 (4)
N1	0.25476 (14)	0.31687 (14)	0.5356 (6)	0.0133 (5)
N2	0.31948 (13)	0.53211 (14)	0.7039 (5)	0.0125 (5)
N3	0.60105 (15)	0.78731 (16)	0.9591 (6)	0.0170 (5)
C1	0.23342 (14)	0.41140 (15)	0.5664 (6)	0.0113 (5)
C2	0.29506 (15)	0.45099 (15)	0.7707 (6)	0.0124 (5)
C3	0.36727 (15)	0.58679 (17)	0.8976 (6)	0.0124 (5)
C4	0.30746 (15)	0.65269 (16)	1.0352 (6)	0.0120 (5)
C5	0.14602 (14)	0.41997 (17)	0.6840 (6)	0.0143 (5)
C6	0.08314 (16)	0.38412 (18)	0.4741 (7)	0.0233 (7)
C7	0.12666 (16)	0.51472 (17)	0.7592 (7)	0.0207 (6)
C8	0.43809 (14)	0.62983 (16)	0.7409 (6)	0.0134 (5)
C9	0.49336 (15)	0.68044 (17)	0.9422 (6)	0.0153 (5)
C10	0.56044 (15)	0.73132 (18)	0.7984 (6)	0.0142 (6)

Table 3. *Fractional atomic coordinates and equivalent isotropic displacement parameters (Å²) for L-Glu-L-Va*

$U_{eq} = (1/3)\Sigma_i\Sigma_j U_{ij} a_i^* a_j^* \mathbf{a}_i.\mathbf{a}_j.$

	x	*y*	*z*	U_{eq}
O1	0.5331 (3)	0.3656 (4)	0.72762 (11)	0.0183 (4)
O2	0.0034 (3)	0.3366 (4)	0.69941 (11)	0.0175 (4)
O3	0.0320 (3)	0.0023 (4)	0.77703 (11)	0.0173 (4)
O4	0.2023 (3)	1.0091 (4)	0.47491 (12)	0.0216 (4)
O5	0.1749 (3)	0.6277 (4)	0.42939 (12)	0.0207 (4)
N1	0.7243 (3)	0.7552 (4)	0.68024 (14)	0.0154 (4)
N2	0.3058 (3)	0.5761 (4)	0.79311 (13)	0.0143 (4)
C1	0.5116 (4)	0.7950 (5)	0.70361 (15)	0.0141 (5)
C2	0.4498 (4)	0.5571 (5)	0.74246 (14)	0.0130 (5)
C3	0.2260 (4)	0.3515 (5)	0.82496 (14)	0.0135 (5)
C4	0.0741 (4)	0.2204 (5)	0.76151 (15)	0.0136 (5)
C5	0.3571 (4)	0.8725 (5)	0.63138 (15)	0.0163 (5)
C6	0.3150 (4)	0.6811 (5)	0.56535 (15)	0.0165 (5)
C7	0.2223 (4)	0.7934 (5)	0.48637 (15)	0.0151 (5)
C8	0.1283 (4)	0.3939 (5)	0.90308 (14)	0.0167 (5)
C9	0.2795 (5)	0.5258 (7)	0.96608 (16)	0.0280 (7)
C10	−0.0818 (4)	0.5214 (7)	0.89036 (16)	0.0234 (6)

to interfere with the need to appropriately position three acceptor groups around the main-chain $—NH_3^+$ amino group.

3.3. *Hydrogen bonds*

Hydrogen-bond parameters are given in Table 7. In both structures two of the three amino (N—)H's form head-to-tail hydrogen-bonded chains with the main-chain carboxylate group. The hydrogen-bonded sheets generated for each structure are shown in Fig. 2. In each case the third amino (N—)H is accepted by a group in the side chain of the hydrophilic residue. In Glu-Val this occurs by forming a three-centre hydrogen bond which is part of two eight-membered rings, Fig. 3 (see also Table 7). The unique feature of the pattern generated is the fully eclipsed orientation taken by the amino group. Only four out of 257 amino acid and peptide structures retrieved from the CSD (Allen *et al.*, 1991) had C′—C^α—N—H torsion angles with absolute values

Table 4. *Bond lengths (Å) and angles (°) for L-Val-L-Gln and L-Glu-L-Val*

	L-Val-L-Gln	L-Glu-L-Val
N1—C1	1.496 (3)	1.497 (3)
C1—C2	1.522 (4)	1.538 (4)
C1—C5	1.544 (3)	1.531 (4)
N2—C2	1.342 (3)	1.342 (3)
O1—C2	1.231 (3)	1.224 (3)
N2—C3	1.466 (3)	1.467 (3)
C3—C4	1.550 (4)	1.535 (3)
C3—C8	1.527 (3)	1.541 (3)
O2—C4	1.262 (3)	1.257 (3)
O3—C4	1.259 (3)	1.266 (3)
C5—C6	1.531 (4)	1.526 (4)
C5—C7	1.527 (3)	
C6—C7		1.514 (3)
O4—C7		1.208 (4)
O5—C7		1.328 (3)
C8—C9	1.524 (3)	1.530 (4)
C9—C10	1.509 (4)	
O4—C10	1.250 (4)	
N3—C10	1.323 (4)	
C8—C10		1.524 (4)
N1—C1—C2	106.9 (2)	106.5 (2)
N1—C1—C5	109.5 (2)	111.6 (2)
C2—C1—C5	110.9 (2)	113.1 (2)
O1—C2—N2	124.9 (2)	123.8 (2)
O1—C2—C1	120.4 (2)	120.2 (2)
N2—C2—C1	114.8 (2)	116.1 (2)
C2—N2—C3	122.8 (2)	118.0 (2)
N2—C3—C4	107.0 (2)	111.7 (2)
N2—C3—C8	110.7 (2)	112.5 (2)
C4—C3—C8	113.8 (2)	111.2 (2)
O2—C4—O3	123.5 (2)	125.9 (2)
O2—C4—C3	117.2 (2)	118.0 (2)
O3—C4—C3	119.2 (2)	116.1 (2)
C1—C5—C7	110.9 (2)	
C1—C5—C6	111.4 (2)	114.8 (2)
C6—C5—C7	110.5 (2)	
C5—C6—C7		111.5 (2)
O4—C7—O5		123.4 (3)
O4—C7—C6		124.3 (2)
O5—C7—C6		112.2 (2)
C3—C8—C9	111.9 (2)	111.1 (2)
C8—C9—C10	114.7 (2)	
O4—C10—N3	122.3 (3)	
O4—C10—C9	121.3 (3)	
N3—C10—C9	116.5 (3)	
C10—C8—C9		111.2 (2)
C10—C8—C3		114.0 (2)

< 15°. In the present structure any rotation of the amino group towards a more usual staggered conformation results in too short a contact with the side (O=C)—OH group of the Glu side chain and overall less favourable hydrogen-bond geometries. The peptide carbonyl groups participate only in weaker C^{α}—H···O=C< interactions.

3.4. *Crystal packing*

The crystal packing and unit cell for both title compounds are shown in Fig. 4. In addition to these two compounds there are ten other known structures of dipeptides with one hydrophobic and one hydrophilic residue. The 12 structures have been compared in Table 8. It can be seen that the Ala side chains form hydrophobic columns in three crystals. Our primary concern will be the nine remaining structures in which the side chains of other residues form hydrophobic layers.

Table 5. *Selected torsion angles (°) for L-Val-L-Gln, L-Glu-L-Val and L-Val-L-Glu**

	L-Val-L-Gln	L-Glu-L-Val	L-Val-L-Glu†
N1—C1—C2—N2 (ψ_1)	141.6 (2)	156.4 (2)	124.5
C1—C2—N2—C3 (ω_1)	169.2 (2)	173.7 (2)	175.2
C2—N2—C3—C4 (φ_2)	−101.2 (3)	−75.1 (2)	−81.7
N2—C3—C4—O2 (ψ_T)	75.0 (3)	−14.6 (3)	−30.7
N1—C1—C5—C6 ($\chi_1^1/\chi_1^{1.1}$)	−64.8 (3)	64.9 (3)	−59.6
N1—C1—C5—C7 ($\chi_1^{1.2}$)	171.7 (3)		175.6
C1—C5—C6—C7 (χ_1^2)		−161.0 (2)	
C5—C6—C7—O4 ($\chi_1^{3.1}$)		6.5 (4)	
N2—C3—C8—C9 ($\chi_2^1/\chi_2^{1.1}$)	−174.8 (2)	−53.5 (3)	−178.5
N2—C3—C8—C10 ($\chi_2^{1.2}$)		73.2 (3)	
C3—C8—C9—C10 (χ_2^2)	−174.8 (2)		−174.3
C8—C9—C10—O4 ($\chi_2^{3.1}$)	−10.7 (4)		0.3

* Eggleston (1984). † Atomic numbering as indicated in Scheme 1 (equivalent to L-Val-L-Gln).

Considering first the Hpo-Hpi group, one finds that all four structures share important structural features. Crystal packing diagrams for Leu-Glu (Eggleston &

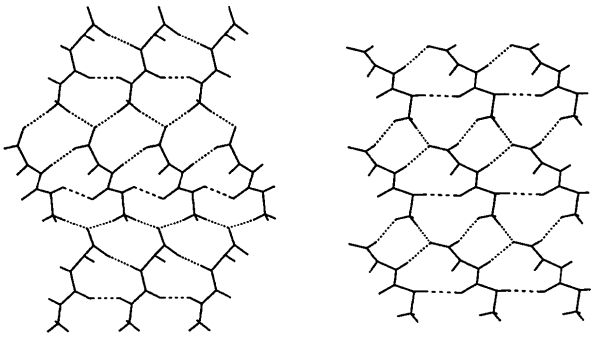


Fig. 2. Two-dimensional hydrogen-bonded sheets in the structures of (*a*) L-Val-L-Gln and (*b*) L-Glu-L-Val. Amino acid side chains have been omitted for clarity. Hydrogen bonds with N—H donors are shown as dotted lines and those with C^{α}—H donors are drawn with dashed lines.

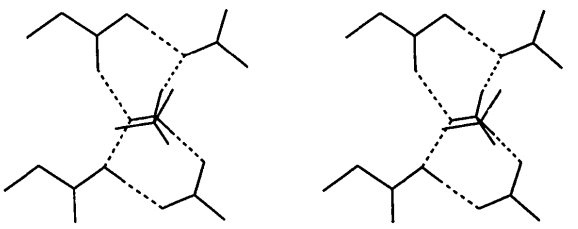


Fig. 3. Stereodiagram with details of the hydrogen bonding in L-Glu-L-Val showing the eclipsed N-terminal amino group and the three-centre hydrogen-bond interaction.

Table 6. *Orientations at the $C^{\alpha}-C^{\beta}$ bond (χ^1) of L-Glu and L-Gln side chains in peptides retrieved from the Cambridge Structural Database (Allen et al., 1991)*

Conformation	N-terminal	Central	C-terminal	Total
gauche+	2	0	1	3
trans	2	0	4	6
gauche−	0	7	8	15

Table 7. *Hydrogen-bond distances (Å) and angles (°) for L-Val-L-Gln, L-Glu-L-Val and L-Val-L-Glu**

D—H···A	*D*—H	H···*A*	D···*A*	*D*—H···*A*
L-Val-L-Gln				
N1—H1···O3^{i}	0.92 (4)	1.92 (4)	2.836 (3)	173 (3)
N1—H2···O3^{ii}	0.88 (4)	2.06 (4)	2.934 (3)	173 (3)
N1—H3···O4^{iii}	0.98 (3)	1.89 (3)	2.837 (3)	161 (3)
N2—H4···O2^{iv}	0.85 (3)	1.98 (3)	2.829 (3)	173 (3)
N3—H5···O2^{v}	0.87 (3)	2.21 (3)	3.072 (3)	175 (3)
N3—H6···O4^{vi}	0.89 (4)	2.10 (4)	2.950 (4)	159 (3)
C1—H11···O1^{iv}	0.90†	2.40	3.078 (3)	133
L-Glu-L-Val				
N1—H1···O1‡	0.93 (4)	2.10 (4)	2.649 (3)	117 (3)
N1—H1···O2^{vii}	0.93 (4)	2.13 (4)	2.923 (3)	143 (3)
N1—H2···O3^{viii}	0.94 (4)	1.83 (4)	2.765 (3)	179 (3)
N1—H3···O4^{ix}	0.95 (4)	2.43 (4)	3.020 (3)	121 (3)
N1—H3···O5^{x}	0.95 (4)	2.15 (4)	2.881 (3)	134 (3)
N2—H4···O3^{xi}	0.81 (4)	2.15 (4)	2.934 (3)	162 (3)
O5—H5···O2^{xii}	0.79 (4)	1.80 (4)	2.585 (3)	169 (4)
C1—H11···O1^{xi}	0.95†	2.47	3.168 (4)	130
L-Val-L-Glu§				
N1—H1···O1¶	0.85	2.66 (3)	3.171 (3)	120 (2)
N1—H1···O2	0.85	2.11	2.884	152
N1—H2···O3	1.00	1.86 (3)	2.806 (3)	157 (3)
N1—H3···O4	0.93	1.98 (2)	2.856 (3)	156 (2)
N1—H3···O5	0.93	2.73	3.205	113
N2—H4···O3	0.86	2.08 (3)	2.921 (3)	174 (2)
O5—H5···O2	0.83	1.83 (3)	2.601 (3)	154 (3)
C1—H11···O1	0.94	2.35	3.089	135

Symmetry codes: (i) $\frac{1}{2}-x, y-\frac{1}{2}, 1-z$; (ii) $\frac{1}{2}-x, y-\frac{1}{2}, 2-z$; (iii) $1-x, 1-y, z$; (iv) $x, y, z-1$; (v) $x+\frac{1}{2}, \frac{3}{2}-y, 2-z$; (vi) $x, y, z+1$; (vii) $x+1, y, z$; (viii) $x+1, y+1, z$; (ix) $1-x, y-\frac{1}{2}, 1-z$; (x) $1-x, y+\frac{1}{2}, 1-z$; (xi) $x, y+1, z$; (xii) $-x, y+\frac{1}{2}, 1-z$. *Eggleston (1984). †E.s.d.'s not meaningful due to the constrained refinement of the H atom. ‡Intramolecular hydrogen bond. §E.s.d.'s given when the parameter is listed in the original paper. ¶Eggleston (1984) also gives parameters for the long hydrogen bond N1—H1···O4 with d(H···O) = 2.83 (3) Å.

Hodgson, 1983*a*) and Leu-Tyr (Krause, Baures & Eggleston, 1993) have been depicted in Fig. 5. As for Val-Gln (Fig. 4), the hydrophobic layers are composed of side chains of the hydrophobic residues, as well as the hydrophobic ethylene or benzylene part of the hydrophilic side chain. The side-chain hydrogen-bond donor/acceptor interacts with molecules in the adjacent hydrophilic layer and generates a three-dimensional hydrogen-bond pattern. (For a more detailed discussion, see Görbitz & Etter, 1992.) The hydrogen-bonded sheet of Val-Gln is also found for Val-Glu (Eggleston, 1984).

Table 8. *Reported dipeptide structures with one hydrophobic and one hydrophilic residue*

Peptide*	Refcode	Space group	Packing†	Reference
Hpo-Hpi‡				
Ala-Asp	BURLIJ	$P2_12_12_1$	Columns	(*a*)
Ala-Ser	LALLSE	$P2_12_12_1$	Columns	(*b*)
Val-Gln	–	$P2_12_12$	Layers	Present work
Val-Glu	CIJGUX	$P2_12_12$	Layers	(*c*)
Leu-Glu	BOFZOL	$P2_12_12_1$	Layers	(*d*)
Leu-Tyr	JUKMEH	$P2_12_12_1$	Layers	(*e*)
Hpi-Hpo				
Asp-Ala	FUMTAI	$P2_1$	Columns	(*f*)
Glu-Val	–	$P2_1$	Layers	Present work
Tyr-Val.H_2O	CIHNUC	$P2_12_12_1$	Layers	(*g*)
His-Leu	JUKMOR	$P2_1$	Layers	(*h*)
Tyr-Leu.H_2O	VUZBIB	$P2_1$	Layers	(*i*)
Tyr-Phe.H_2O	CELTAO10	$P2_12_12_1$	Layers	(*j*)

(*a*) Eggleston & Hodgson (1983*b*); (*b*) Jones, Falvello & Kennard (1978); (*c*) Eggleston (1984); (*d*) Eggleston & Hodgson (1983*a*); (*e*) Krause, Baures & Eggleston (1993); (*f*) Görbitz (1987); (*g*) Ramakrishnan, Seshardi & Viswamitra (1984); (*h*) Krause, Baures & Eggleston (1993); (*i*) Ramakrishnan & Viswamitra (1988); (*j*) Murali & Subramanian (1987). *All amino acids are L isomers. †Aggregation pattern for hydrophobic groups. ‡Hpo is a hydrophobic residue, Hpi is a hydrophilic residue.

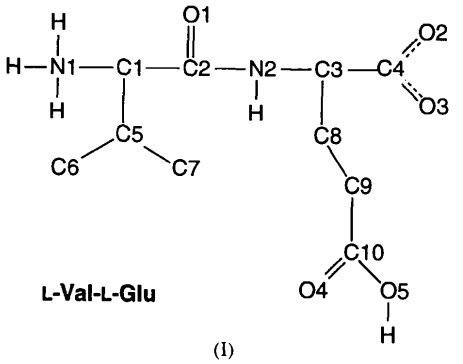
These two structures have been compared in a separate section below.

The layered structures in the Hpi-Hpo group are distinct from those of the Hpo-Hpi group in that the hydrogen-bond patterns are only two-dimensional, as illustrated for Glu-Val in Fig. 4 and Tyr-Phe.H_2O (Murali & Subramanian, 1987) and His-Leu (Krause, Baures & Eggleston, 1993) in Fig. 6. The main hydrophobic layer is entirely made up of hydrophobic residue side chains. Hydrophobic entities in the hydrophilic side chains form isolated hydrophobic columns, or, in the case of the large Tyr residues, an independent additional hydrophobic layer. The hydrogen-bonded sheet for Glu-Val is essentially identical to the sheets observed in the crystal structures of Gly-Leu (Patthabi, Venkatesan & Hall, 1974), Pro-Gly (Narasimhan & Chacko, 1982), Pro-Val (Narasimhan, Chacko & Swaminathan, 1982), Pro-Ile (Panneerselvam, Chacko & Veena, 1989) and even the rather hydrophilic structure of Glu-Glu (Eggleston & Hodgson, 1982*a*).

The different dimensionality of the hydrogen-bond networks in Val-Gln and Glu-Val is manifested by the typical differences in the macroscopic behaviour; while the crystals of Val-Gln are uniformly hard, the crystals of Glu-Val are easily cleaved along the hydrophobic layers.

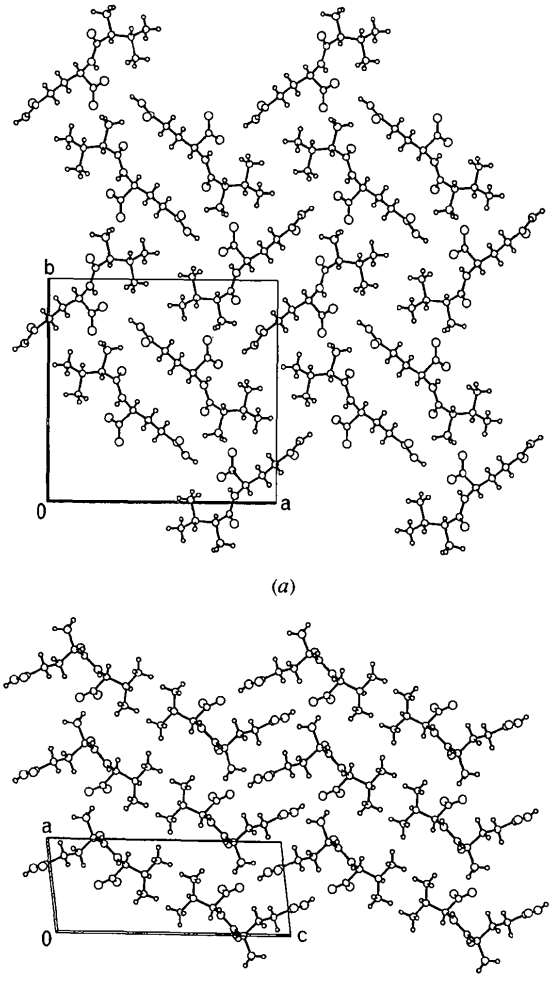
3.5. *Comparison between L-Glu-L-Val and L-Val-L-Glu*

It is of interest to compare the crystal structures of the two title compounds with the structure of the dipeptide Val-Glu (Eggleston, 1984), with the atomic numbering as indicated in Scheme I.



(I)

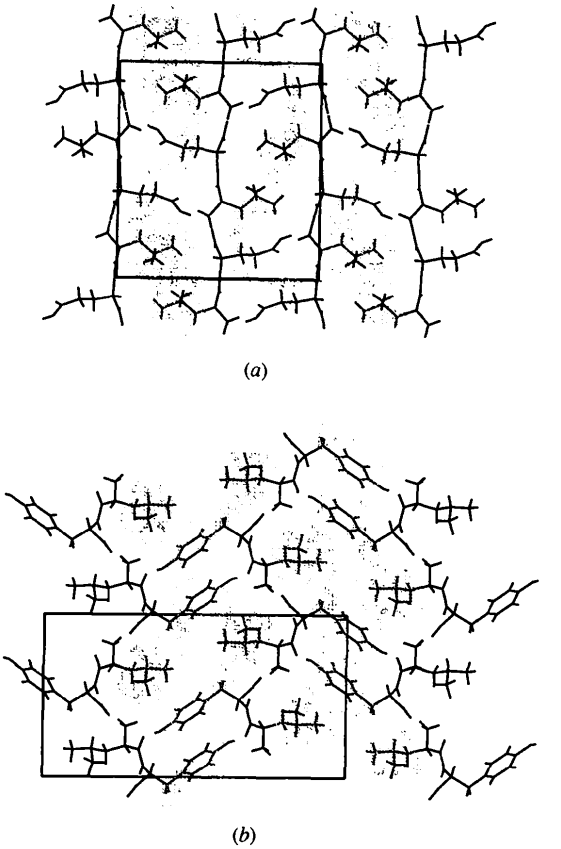
As for Glu-Val (and Val-Gln), O2 is the O atom which is *cis* to the N2 atom of the peptide bond. From the previous discussion on molecular packing, it is evident that the crystal structures of Glu-Val and its retroanalogue Val-Glu are rather different. It is then quite surprising to find from Table 7 that both crystals contain essentially the same hydrogen bonds, even down to the three-centre interaction for one of the amino H atoms. This result is in line with the work of Margaret C. Etter on the encoding and decoding of hydrogen-bond patterns of organic compounds, which led to the recognition of empirical 'Hydrogen Bond Rules' (Etter, 1990). The third rule states that: 'the best proton donor and acceptor remaining after intramolecular hydrogen-bond formation form intermolecular hydrogen bonds to one another'. This rule may tentatively be extended to say that the next best donor will form a hydrogen bond to the next best acceptor, and so on. However, since further interactions have to exist within progressively tighter steric constraints, it will certainly be difficult for all donors and acceptors to associate in strict rank-order (if indeed such an order can be unambiguously assigned). In the present structures the extended rule is obeyed for the three best donors, the side-chain carboxyl —OH and two amino H atoms, which all approach the best acceptor, the main-chain carboxylate group, but not for the last amino H atom, which is instead accepted by a side-chain carbonyl group. Still, the weaker peptide



(*a*)

(*b*)

Fig. 4. The packing arrangement viewed (*a*) along the *c* axis for L-Val-L-Gln and (*b*) the *b* axis for L-Glu-L-Val.



(*a*)

(*b*)

Fig. 5. Crystal packing of (*a*) L-Leu-L-Glu (Eggleston & Hodgson, 1983*a*) and (*b*) L-Leu-L-Tyr (Krause, Baures & Eggleston, 1993). To emphasize the layered nature of the structures in Figs. 5 and 6, a grey van der Waals surface is indicated for hydrophobic H atoms in peptide side chains.

bond >N—H donor manages to approach the carboxylate group. This pattern is used in both Glu-Val and Val-Glu, indicating distinct hydrogen-bonding preferences.

The database of reported dipeptide structures contains three other pairs with opposite sequences, but only one, Ala-Asp (Eggleston & Hodgson, 1983*b*) and Asp-Ala (Görbitz, 1987), with exactly the same number of hydrogen-bond donors and acceptors. In this case the hydrogen-bonding schemes are completely different, but the situation is not directly comparable to Val-Glu/Glu-Val, since the Asp-Ala dipeptide has been crystallized in an unusual protonation state with a main-chain —COOH group and side-chain —COO group. The two structures in the other two dipeptide pairs have a different number of hydrogen-bond donors and acceptors due to the variable number of cocrystallized solvent water molecules and a perfect match of hydrogen-bond types is not possible. Still, hydrogen bonds are largely identical in Gly-Asp.2H_2O (Eggleston & Hodgson, 1982*b*) and Asp-Gly.H_2O (Eggleston, Valente & Hodgson, 1981), except for a —$NH_3^+\cdots$OH— (side-chain) interaction in Asp-Gly, which is converted to a —$NH_3^+\cdots OH_2$ hydrogen bond in Gly-Asp. This similarity is not seriously disturbed by the extra water molecule in Gly-Asp, which donates both its H atoms to the C-terminal carboxylate group. Leu-Tyr (Krause, Baures & Eggleston, 1993) and Tyr-Leu.H_2O (Ramakrishnan & Viswamitra, 1988), on the other hand, have completely different hydrogen-bond types. It would obviously be of interest to obtain crystal structures of more retroanalogue dipeptide pairs to carry out further examinations of their hydrogen-bonding preferences.

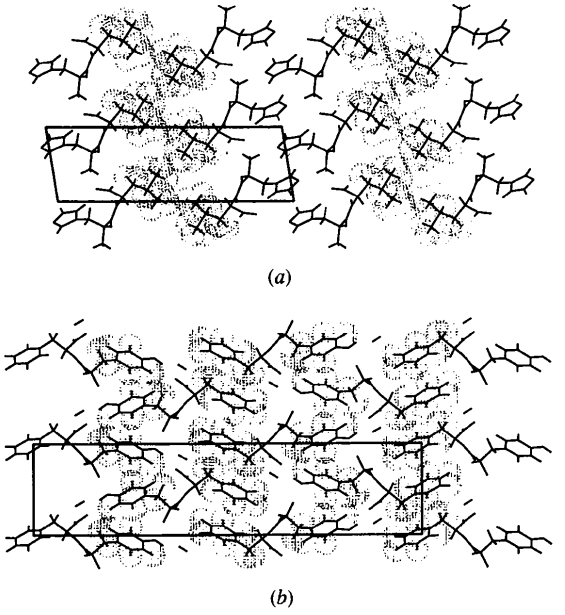
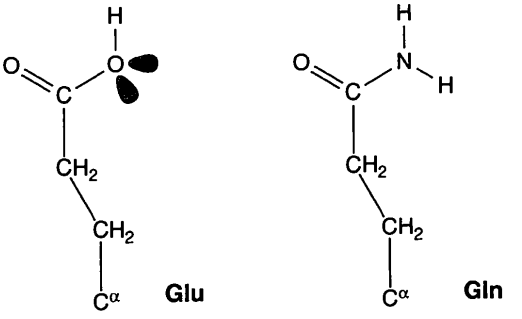


Fig. 6. Crystal packing of (*a*) L-His-L-Leu (Krause, Baures & Eggleston, 1993) and (*b*) L-Tyr-L-Phe.H_2O (Murali & Subramanian, 1987).

3.6. *Comparison between L-Val-L-Gln and L-Val-L-Glu*

Glu to Gln involves the exchange of —OH to —NH_2, *i.e.* a further acidic proton is added to the side chain, Scheme II.



(II)

Accordingly, structural modifications of the Val-Glu structure (Eggleston, 1984) are required in order to position a hydrogen-bond acceptor close to the amide group of the Gln side chain in the isomorphic Val-Gln structure, as is readily discerned from Fig. 7. The only acceptor directly available is the carbonyl group of the equivalent side chain translated along the *c* axis. This axis is shortened from 5.367 Å in Val-Glu to 4.708 Å in Val-Gln, among the smallest values observed for a crystallographic axis in dipeptide structures, to avoid too long a contact. At the same time there is a significant lengthening of the *b* axis, from 13.827 Å in Val-Glu to 15.309 Å in Val-Gln. The shift for the *a*

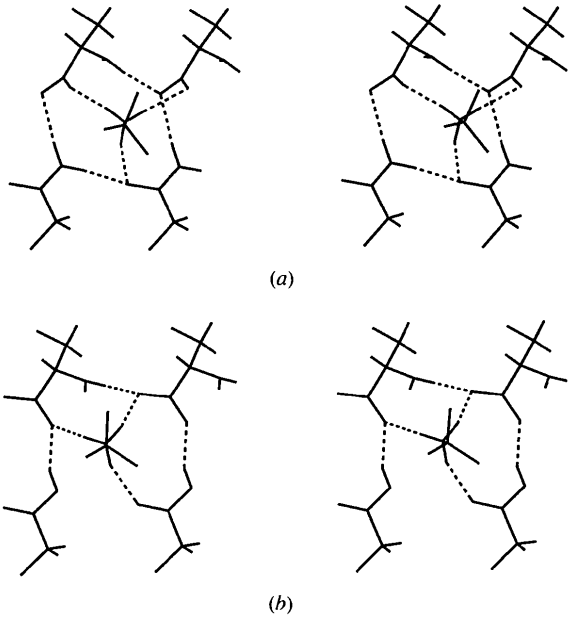


Fig. 7. Stereodiagram with details of the hydrogen-bonding patterns in (*a*) L-Val-L-Gln and (*b*) L-Val-L-Glu (Eggleston, 1984). For each figure the included molecular fragments (from the top) are: (1) the C-terminal carboxylate group and the peptide bond >N—H; (2) the N-terminal amino group; (3) the functional group of the side chain.

axis is more moderate (16.781/16.419 Å for Val-Glu/Val-Gln). Molecular geometry modifications include a more extended peptide chain conformation for Val-Gln than for Val-Glu, as reflected by the values for ψ_1 and φ_2 listed in Table 5, but the most salient reorientation takes place for the C-terminal carboxylate group. The N2—C3—C4—O2 (ψ_T) torsion angle is −30.7° for Val-Glu, but 74.9° for Val-Gln. As a result of this twisting the hydrogen-bonding scheme is changed; while both in-plane amino (N—)H's are accepted by the same carboxylate O atom in Val-Gln, the carboxylate O atoms in Val-Glu each accepts one amino (N—)H, Fig. 7.

4. Conclusions

Dipeptides with one hydrophilic and one hydrophobic residue tend to form crystals with distinct hydrophobic and hydrophilic layers. Two of the three N-terminal amino (N—)H's take part in the two-dimensional head-to-tail hydrogen-bond pattern in the hydrophilic layer, while the essential feature of these structures is the ability to use a functional group in the hydrophilic side chain as an acceptor for the third amino (N—)H. The packing of the dipeptides in the crystal is, however, radically dependent on the sequence of the two residues. Thus, if the hydrophilic residue is N-terminal, the hydrogen-bond pattern is three-dimensional, but two-dimensional if it is C-terminal.

The authors thank the referees for useful comments and suggestions.

References

Allen, F. H., Davies, J. E., Galloy, J. J., Johnson, O., Kennard, O., Macrae, C. F., Mitchell, E. M., Mitchell, G. F., Smith, J. M. & Watson, D. G. (1991). *J. Chem. Inf. Comput. Sci.* **31**, 187–204.

Cambridge Structural Database (1995). Cambridge Crystallographic Data Centre, 18 Union Road, Cambridge, England.

Eggleston, D. S. (1984). *Acta Cryst.* C**40**, 1250–1252.

Eggleston, D. S. & Hodgson, D. J. (1982*a*). *Acta Cryst.* B**38**, 1216–1220.

Eggleston, D. S. & Hodgson, D. J. (1982*b*). *Int. J. Pept. Protein Res.* **19**, 206–211.

Eggleston, D. S. & Hodgson, D. J. (1983*a*). *Acta Cryst.* C**39**, 75–78.

Eggleston, D. S. & Hodgson, D. J. (1983*b*). *Int. J. Pept. Protein Res.* **21**, 288–295.

Eggleston, D. S., Valente, E. J. & Hodgson, D. J. (1981). *Acta Cryst.* B**37**, 1428–1430.

Etter, M. C. (1990). *Acc. Chem. Res.* **23**, 120–126.

Gilmore, C. J. (1984). *J. Appl. Cryst.* **17**, 42–46.

Görbitz, C. H. (1987). *Acta Chem. Scand. B*, **41**, 679–685.

Görbitz, C. H. & Etter, M. C. (1992). *Int. J. Pept. Protein Res.* **39**, 93–110.

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

Guo, H. & Karplus, M. (1992). *J. Phys. Chem.* **96**, 7273–7287.

Johnson, C. K. (1976). Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA.

Jones, P. G., Falvello, L. & Kennard, O. (1978). *Acta Cryst.* B**34**, 1939–1942.

Krause, J. A., Baures, P. W. & Eggleston, D. S. (1993). *Acta Cryst.* B**49**, 123–130.

Murali, R. & Subramanian, E. (1987). *Int. J. Pept. Protein Res.* **29**, 187–192.

Narasimhan, P. & Chacko, K. K. (1982). *Cryst. Struct. Commun.* **11**, 2051–2056.

Narasimhan, P., Chacko, K. K. & Swaminathan, S. (1982). *Cryst. Struct. Commun.* **11**, 695–700.

Panneerselvam, K., Chacko, K. K. & Veena, R. (1989). *Int. J. Pept. Protein Res.* **33**, 191–194.

Patthabi, V., Venkatesan, K. & Hall, S. R. (1974). *J. Chem. Soc. Perkin Trans. 2*, pp. 1722–1727.

Ramakrishnan, B. & Viswamitra, M. A. (1988). *Acta Cryst.* C**44**, 1957–1959.

Ramakrishnan, B., Seshardi, T. P. & Viswamitra, M. A. (1984). *Acta Cryst.* C**40**, 1248–1250.

Sheldrick, G. M. (1993). *SHELXL93. Program for the Refinement of Crystal Structures*. University of Göttingen, Germany.

Taylor, R., Kennard, O. & Versichel, W. (1984). *Acta Cryst.* B**40**, 280–288.