Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

A crystallographic titration of the dipeptide L-isoleucyl-L-isoleucine

Carl Henrik Görbitz

Department of Chemistry, University of Oslo, PO Box 1033 Blindern, N-0315 Oslo, Norway

Correspondence e-mail: c.h.gorbitz@kjemi.uio.no

Received 24 February 2004
Accepted 14 June 2004

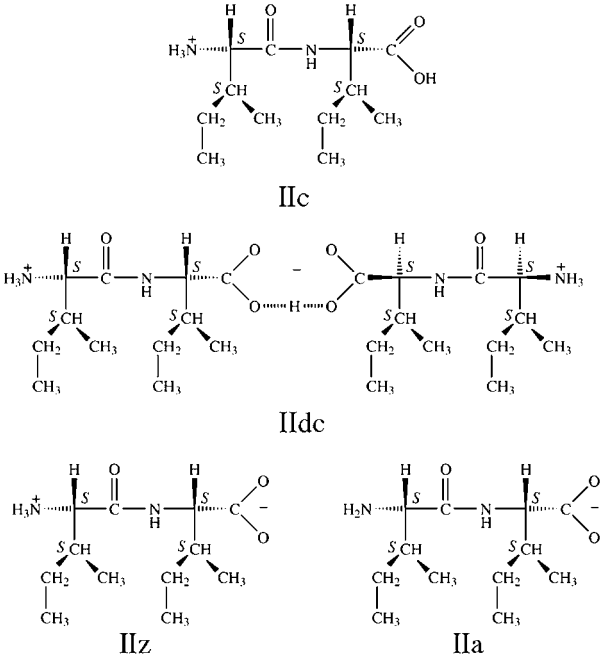
The dipeptide L-Ile-L-Ile has been crystallized in four different protonation states, including as a cation, a zwitterion and an anion, as well as a dimeric cation in which two peptide molecules, linked by a strong –COO···H···OOC– hydrogen bond, share an overall positive charge of +1. This unprecedented series of crystal structures exhibits differences in hydrogen-bonding capabilities, conformational properties and molecular packing arrangements. The crystallization of L-Ile-L-Ile as an anion was achieved by slow evaporation from an NH_3(l) solution to give the first ever ammonium salt of a peptide (or amino acid).

1. Introduction

A systematic investigation of the structures of hydrophobic dipeptides has led to the discovery of seven isostructural compounds with hydrophobic pores of variable diameters (Görbitz & Gundersen, 1996; Görbitz, 2002*b*, 2003). This structural family has been referred to as the VA class after its very first member L-Val-L-Ala (Görbitz & Gundersen, 1996) and includes dipeptides with L-Ala, L-Val and L-Ile residues. In a hypothetical model of the equivalent L-Ile-L-Ile (II) structure (Fig. 1), only modest steric conflict is present (H···H contacts > 1.4 Å); this conflict could potentially be alleviated by small modifications of the molecular geometry. Crystallization, however, resulted in thick rods and not the thin needles typical for VA-class crystals. Subsequent crystal structure determination revealed that zwitterionic II (IIz) had been obtained as a dihydrate in the orthorhombic space group $P2_12_12_1$. The cocrystallized water molecules play an important role in the hydrogen-bond network of the IIz structure, and the idea was born to keep water away during crystallization in a second attempt to grow II in a hexagonal modification. The solubility is low in most organic solvents, and one of the first alternatives tested was thus NH_3(l). Liquid ammonia is an extremely good solvent for amino acids and peptides, but its properties as a base meant that there were two possible outcomes of a successful crystallization: The specimens could have been assembled not only from peptide molecules in the desired zwitterionic state, but also from peptide anions in a previously unobserved ammonium salt. Crystals were produced, and the subsequent investigation showed that the latter alternative (IIa) had occurred. This crystal structure was only the second with a peptide anion, the first being (*R*)-phenylglycyl-(*R*)-phenylglycine [in the (*S*)-1-phenylethylamine salt (Akazome *et al.*, 1997)]. Even for simple amino acids, structural information on anions is very limited indeed; Tl^+·DL-Cys^- (Freeman & Moore, 1977) and Na^+·3,5-

dinitro-L-Tyr$^{-}\cdot H_2O$ (Cody *et al.*, 1979) are the only monoanions, while dianions have been obtained for $Ca^{2+}\cdot$L-Glu$^{2-}\cdot 3H_2O$ (Einspahr & Bugg, 1974) and $2K^{+}\cdot$3,3,3′,3′-tetramethyl-D-cystine$^{2-}\cdot 3H_2O$ (Faggiani *et al.*, 1984).

With the zwitterions and the anion at hand, interest was triggered for investigating the cation of II, in order to obtain the first series of structures describing all three classical protonation states of a peptide. Crystal structures of two protonation states (as zwitterion and cation) were known for only nine peptides (all dipeptides), of which three were devoid of Gly residues [Ala-Ala (Fletterick *et al.*, 1971) – Ala-Ala·HCl (Tomuka *et al.*, 1969); Ala-Phe·2-propanol (Görbitz, 1999) – Ala-Phe·HCl·$2H_2O$ (Cotrait & Barrans, 1974); Leu-Ala (Görbitz, 1997) – Leu-Ala 1,5-naphthalenedisulfonate (Sudbeck *et al.*, 1994)]. Surprisingly, the following experiments produced not only II as a regular cation in the nitrate salt (IIc), but also a structure in which a pair of peptide molecules have jointly taken up an H atom that serves to connect them by a strong –COO···H···OOC– hydrogen bond, essentially a dimeric cation (IIdc). This paper thus presents four crystal structures of II, in which the peptide is present in four different protonation states (see scheme below). All the amino acids discussed have the L configuration.



2. Experimental

2.1. Crystal preparation

Crystals of IIz were grown by slow diffusion of 2-propanol into an aqueous solution of the peptide. Crystals of IIa were obtained by dissolving a few milligrams of the peptide in about 1 ml of NH_3(l). Dry ice was used as a cooling agent, and therefore the procedures had to be carried out rapidly in order to avoid the formation of NH_3(s) (freezing point 195.5 K; dry ice is 194.7 K). Care was taken to prevent water vapor from coming into contact with the sample. The solution was placed inside a specially constructed bomb container with an adjustable vent and left for slow evaporation in a cold room at 250 K. Evaporation to dryness was completed in less than 24 h. Crystals of IIdc were formed by slow evaporation of acetonitrile into an aqueous solution containing approximately equimolar amounts of II and HNO_3. Crystals not used for data collection were stored in a sealed Petri dish with some mother liquor in a freezer at 250 K, since slow decay was observed at room temperature. A few days later the plates had transformed into small rods. Data collection for one of these rods showed them to be IIc. Another few days later, the remaining rods, still in the freezer, had disappeared and been replaced by flat crystals (dimensions up to 1.0 × 1.0 × 0.2 mm) that were shown to be IIz, a transfer that had probably been initiated by water being condensed when the cold sample of IIc was inspected under a microscope.

2.2. Data collection

Crystals of IIa, IIc and IIdc are unstable at room temperature, but no special measures were needed during crystal mounting and data collection. More than a hemisphere of reciprocal space was collected by a combination of five sets of exposures for IIz and three sets of exposures for IIa, IIc and IIdc. Exposure times were 30–120 s. The crystal-to-detector distance was fixed at 5.0 cm.

2.3. Structure determination and refinement

Coordinates for H atoms were refined for the carboxylate groups of IIdc and IIc, water molecules in IIz, and NH_4^+, $-NH_3^+$ and >N—H groups in IIz and IIa. Other H atoms were positioned geometrically and refined with constraints. U_{iso} values for H atoms were fixed at $1.5U_{eq}$ (water molecules, ammonium

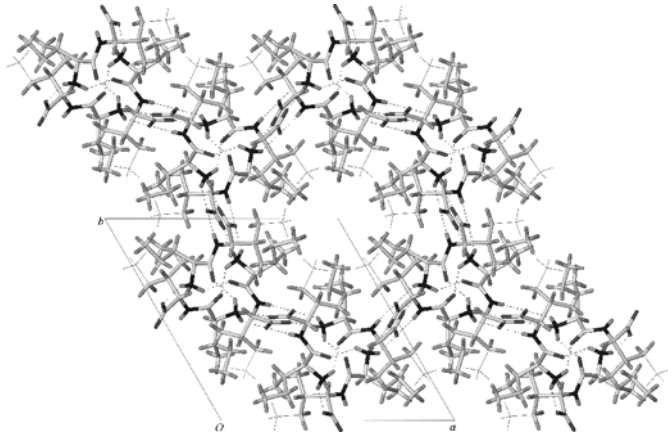


Figure 1
Hypothetical crystal structure of L-Ile-L-Ile in space group $P6_1$ derived from the structure of L-Ile-L-Val by adding a methyl group (shown in the line drawing) to the L-Val side chain. The structure could have been derived from L-Val-L-Ile in a similar manner.

Table 1
Experimental details.

	IIc	IIdc	IIz	IIa
Crystal data				
Chemical formula	$C_{13.08}H_{25}N_3O_6$	$C_{13.80}H_{27.20}N_{3.40}O_{4.70}$	$C_{12}H_{28}N_2O_5$	$C_{12}H_{27}N_3O_3$
M_r	320.36	315.99	280.36	261.37
Cell setting, space group	Trigonal, $P3_121$	Orthorhombic, $P2_12_12$	Orthorhombic, $P2_12_12_1$	Triclinic, $P1$
a, b, c (Å)	16.0682 (6), 16.0682 (6), 11.5317 (9)	16.2957 (19), 23.426 (2), 9.5842 (9)	7.6341 (3), 12.5335 (6), 16.3923 (7)	6.7765 (15), 9.718 (2), 11.882 (3)
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	81.580 (3), 80.484 (3), 88.350 (3)
V (Å^3)	2578.4 (2)	3658.7 (7)	1568.45 (12)	763.4 (3)
Z	6	8	4	2
D_x (Mg m^{-3})	1.238	1.147	1.187	1.137
Radiation type	Mo $K\alpha$	Mo $K\alpha$	Mo $K\alpha$	Mo $K\alpha$
No. of reflections for cell parameters	4175	3710	14 734	2001
θ range (°)	2.3–26.4	2.1–25.1	2.0–35.0	2.1–26.4
μ (mm^{-1})	0.10	0.09	0.09	0.08
Temperature (K)	105 (2)	105 (2)	105 (2)	105 (2)
Crystal form, color	Rod, colorless	Plate, colorless	Block, colorless	Block, colorless
Crystal size (mm)	0.25 × 0.06 × 0.05	0.55 × 0.42 × 0.05	1.10 × 0.45 × 0.25	0.55 × 0.24 × 0.22
Data collection				
Diffractometer	Siemens SMART CCD	Siemens SMART CCD	Siemens SMART CCD	Siemens SMART CCD
Data collection method	Sets of exposures each taken over 0.3° ω rotation	Sets of exposures each taken over 0.3° ω rotation	Sets of exposures each taken over 0.3° ω rotation	Sets of exposures each taken over 0.3° ω rotation
Absorption correction	Empirical (using intensity measurements)	Empirical (using intensity measurements)	Empirical (using intensity measurements)	Empirical (using intensity measurements)
T_{min}	0.968	0.804	0.848	0.779
T_{max}	0.995	0.996	0.978	0.982
No. of measured, independent and observed reflections	11 629, 3401, 2352	27 666, 3657, 3276	21 646, 6861, 6294	6449, 3063, 2165
Criterion for observed reflections	$I > 2\sigma(I)$	$I > 2\sigma(I)$	$I > 2\sigma(I)$	$I > 2\sigma(I)$
R_{int}	0.116	0.043	0.022	0.045
θ_{max} (°)	26.4	25.1	35.0	26.4
Range of h, k, l	$-20 \Rightarrow h \Rightarrow 19$ $-20 \Rightarrow k \Rightarrow 18$ $-11 \Rightarrow l \Rightarrow 14$	$-15 \Rightarrow h \Rightarrow 19$ $-27 \Rightarrow k \Rightarrow 27$ $-10 \Rightarrow l \Rightarrow 11$	$-12 \Rightarrow h \Rightarrow 12$ $-19 \Rightarrow k \Rightarrow 20$ $-26 \Rightarrow l \Rightarrow 26$	$-8 \Rightarrow h \Rightarrow 8$ $-11 \Rightarrow k \Rightarrow 12$ $-14 \Rightarrow l \Rightarrow 14$
Refinement				
Refinement on	F^2	F^2	F^2	F^2
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.068, 0.181, 1.06	0.047, 0.129, 1.14	0.029, 0.080, 1.05	0.044, 0.108, 0.98
No. of reflections	3401	3657	6861	3063
No. of parameters	245	449	210	395
H-atom treatment	Mixture of independent and constrained refinement	Mixture of independent and constrained refinement	Mixture of independent and constrained refinement	Mixture of independent and constrained refinement
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.0975P)^2]$, where $P = (F_o^2 + 2F_c^2)/3$	$w = 1/[\sigma^2(F_o^2) + (0.0659P)^2 + 1.4323P]$, where $P = (F_o^2 + 2F_c^2)/3$	$w = 1/[\sigma^2(F_o^2) + (0.0556P)^2]$, where $P = (F_o^2 + 2F_c^2)/3$	$w = 1/[\sigma^2(F_o^2) + (0.0564P)^2]$, where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$	<0.0001	0.007	0.001	0.005
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.35, −0.35	0.21, −0.31	0.32, −0.16	0.32, −0.23
Absolute structure	Flack (1983)	Flack (1983)	Flack (1983)	Flack (1983)
Flack parameter	−1.8 (19)	−4.9 (15)	−0.2 (4)	−0.7 (14)

Computer programs used: *SMART* (Bruker, 1998), *SAINT-Plus* (Bruker, 2001), *SHELXTL* (Bruker, 2000), *SADABS* (Sheldrick, 1996).

ions, amino groups, methyl groups) or $1.2U_{eq}$ (other) of the bonded atom. Free rotation was permitted for all amino and methyl groups.

For IIz and IIa, the refinement revealed neither conformational nor positional disorder. For IIc, diffuse electron density from cocrystallized solvent molecules inside channels was modeled by the introduction of three free C atoms. Attempts to use one or two acetonitrile molecules with fixed linear geometry resulted in a significantly higher R factor. The electron density may be due to water molecules, but the use of C atoms rather than O atoms was retained because of a significantly lower R factor. The peptides of IIdc are also well defined, but there is disorder for nitrate ions (four positions) and cocrystallized acetonitrile solvent molecules (three positions). The geometries of these ions and molecules were constrained by *SHELXTL* SAME 0.005 0.005 commands. Some diffuse electron density not accounted for was modeled as an O atom (of a water molecule) with an occupancy of

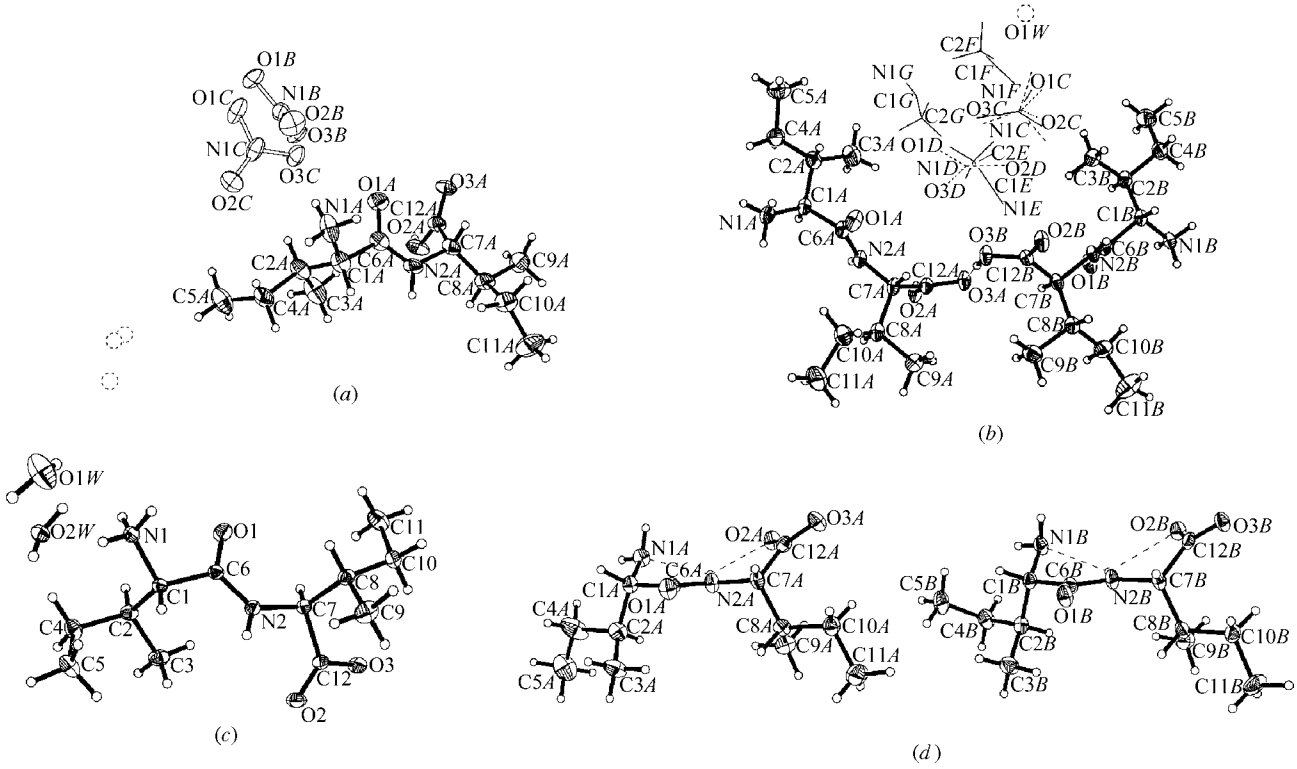


Figure 2
(*a*) The asymmetric unit of IIc with a fully ordered peptide cation and a nitrate anion that is distributed over two mutually exclusive positions *B* and *C*, each with occupancy 0.50. Disordered cocrystallized solvent has been modeled by the refinement of three isotropic C atoms, depicted here as dotted spheres of arbitrary size. (*b*) The asymmetric unit of IIdc. The two peptides are related by a pseudo-twofold rotation axis. There are two nitrate anion sites called *C* and *D*, with three and one different nitrate positions, respectively. The two least populated positions for site C are shown with dashed lines and no atom names. Line drawings are used for the three acetonitrile solvent molecules, while the position of an O atom used to model remaining diffuse electron density is shown as a dotted sphere of arbitrary size. (*c*) The molecular structure of IIz dihydrate. (*d*) The structures of the peptide anion *A* (left) and *B* (right) in IIa. Displacement ellipsoids are shown at the 50% probability level, with H atoms as spheres of arbitrary size

0.39 (3) and U_{iso} = 0.28 (2) Å^2, but the true source of this electron density is unclear and may be an additional organic solvent molecule.

Experimental data and refinement results are summarized in Table 1.[1]

2.4. Database searches

Structures were retrieved from the Cambridge Structural Database (CSD; Version 5.25 of November 2003; Allen, 2002) by means of the program *ConQuest*1.6.

3. Results and discussion

3.1. Molecular structure

The four peptide structures are depicted in Fig. 2. Form IIc, which has a charged amino group and a neutral C-terminal carboxyl group, constitutes the acidic starting point of the 'crystallographic titration' (Coleman & Sprang, 1998) of II. The crystallographic twofold symmetry of IIc (space group $P3_121$) is broken locally by the nitrate ion, since two symmetry-related positions cannot be occupied simultaneously because of steric conflict. Accordingly, there are two well defined nitrate positions, each with occupancy 0.500.

Fig. 2(*b*) shows the asymmetric unit of IIdc. The two peptides, both with positively charged *N*-terminal amino groups, are related by a pseudo-twofold rotation axis and are linked by a very short –COO···H···OOC– hydrogen bond. The observation of such a dimeric cation for a peptide is preceded only by the structure of bis[(*S*)-leucyl-(*S*)-(α-methyl-4-carboxyphenylglycine)]·HCl·H_2O (Coudert *et al.*, 1996). Among amino acids, dimeric cations occur in Gly complexes with at least seven different anions, known for their unique ferroelectric (or also pyroelectric) behavior (Němec *et al.*, 1996, and references therein). Other early structures include DL-homocystine·2COOH (Bigoli *et al.*, 1981), *S*,*S*′-methylene-bis(Cys)·HCl (Bigoli *et al.*, 1982) and bis(Phe)·HCOOH (Görbitz & Etter, 1992*a*). More recently, the structure of bis(Ala)·HNO_3 has been reported (Silva *et al.*, 2001), but the work leading to a firm establishment of the dimeric cation as an independent protonation state for amino acids, at least in the crystalline phase, has been provided by a series of 12 structures presented by Rajaram and co-workers (Ravikumar *et al.*, 2001, 2002; Pandiarajan *et al.*, 2001, 2002*a*,*b*; Sridhar *et al.*, 2001, 2002*a*,*b*; references therein). All but one have been crystallized as a nitrate (six) or a perchlorate (five), which suggests that these two anions are most efficient in promoting

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: DE5006). Services for accessing these data are described at the back of the journal.

Table 2
Torsion angles (°) for the structures of II and for two other, related peptides.

The label in parentheses (*A* or *B*) identifies the peptide molecule in the asymmetric unit. IV: L-Ile-L-Val (Görbitz, 2003); VI: L-Val-L-Ile (Görbitz, 2003).

Torsion	IIz	IIa (*A*)	IIa (*B*)	IIc	IIdc (*A*)	IIdc (*B*)	IV	VI
N1—C1—C6—N2 (ψ_1)	144.70 (5)	−32.0 (4)	−35.2 (4)	133.8 (3)	120.7 (3)	128.5 (2)	170.7 (2)	165.5 (3)
C1—C6—N2—C7 (ω_1)	168.96 (5)	175.3 (3)	177.8 (3)	−172.2 (3)	−173.7 (2)	178.6 (2)	−177.7 (2)	−177.6 (3)
C6—N2—C7—C12 (φ_2)	−142.51 (6)	−137.1 (4)	−138.9 (4)	−58.0 (4)	−133.1 (3)	−131.2 (3)	−139.2 (2)	−139.5 (3)
N2—C7—C12—O2 (ψ_T)	−12.99 (8)	−28.9 (5)	−29.2 (5)	−57.3 (4)	−51.8 (4)	−52.9 (3)	−25.4 (3)	−24.5 (3)
N1—C1—C2—C4 ($\chi_1^{1,1}$)	−62.94 (6)	−69.4 (4)	−55.8 (4)	−71.0 (4)	−60.9 (3)	−59.3 (3)	57.5 (3)	55.3 (3)
N1—C1—C2—C3 ($\chi_1^{1,2}$)	171.90 (5)	167.1 (4)	178.9 (3)	165.2 (3)	175.2 (3)	176.6 (2)	−69.2 (2)	−70.0 (4)
C1—C2—C4—C5 (χ_1^2)	−54.94 (8)	169.5 (4)	−56.3 (5)	172.3 (3)	164.7 (3)	169.8 (3)	157.3 (2)	–
N2—C7—C8—C10 ($\chi_2^{1,1}$)	−161.08 (5)	−174.3 (3)	−175.7 (3)	−65.7 (4)	−63.1 (3)	−58.2 (3)	60.2 (3)	63.0(4)
N2—C7—C8—C9 ($\chi_2^{1,2}$)	73.77 (6)	59.9 (4)	58.1 (4)	171.8 (3)	172.9 (2)	176.4 (2)	−64.3 (3)	−62.4(4)
C7—C8—C10—C11 (χ_2^2)	64.71 (7)	173.1 (3)	169.0 (3)	163.9 (3)	165.0 (3)	165.0 (3)	–	165.5 (4)

the formation of dimeric cations. Notably, dimeric cations are absent in a long series of amino acid sulfates and salts with various carboxylic acids studied by the same research group.

The nitrate ion of IIdc is distributed over four positions, namely three close to each other with occupancies of 0.402 (9), 0.323 (9) and 0.188 (9) (site *C*), and one at about 2 Å separation with an occupancy of 0.072 (3) (site *D*). There are three acetonitrile positions denoted *E*, *F* and *G*. Positions *F* and *G* are close to the twofold axis of the $P2_12_12$ space group, and as for the nitrate ion of IIc, the application of the twofold rotation leads to prohibitive steric conflict. The occupancies for positions *F* and *G* could therefore not be higher than 0.5 and refined to 0.438 (6) and 0.401 (9), respectively. The reduction from 0.5 probably reflects loss of solvent during crystal mounting. Site *E*, with an occupancy of 0.928 (3), overlaps with nitrate site *D* (constrained combined occupancy 1.000).

After full deprotonation of the carboxyl group to the carboxylate, the normal peptide zwitterion is reached, obtained in the crystallize phase as the dihydrate IIz (Fig. 2*c*). The peptide titration is completed by deprotonation of the charged amino group to neutral $-NH_2$. The structure of IIa has two negatively charged peptide anions and two ammonium ions in the asymmetric unit. The two peptides are shown in Fig. 2(*d*) (the full asymmetric unit, which also contains two ammonium cations, could not be shown in a single drawing without the overlap of atoms).

3.2. Molecular conformations

Torsion angle values for all molecules of II are listed in Table 2. It is obvious that II displays extended flexibility and explores a substantial part of the available conformational space in these four structures. We have previously used the torsion angle $\theta = C_1^\beta—C_1^\alpha\cdots C_2^\alpha—C_2^\beta$ to describe the relative orientation of the two side chains in dipeptides (Görbitz, 2001*b*). With very few exceptions, $|\theta| > 90°$, meaning that side chains usually point in opposite directions. The structure of IIz in Fig. 2(*c*) is a good example; the molecule is in a fairly extended conformation and the value for θ is 174.5°. Both peptides in IIdc have main-chain conformations similar to that of IIz, with θ = 174.1° for *A* and 179.1° for *B*, a feature that may not be easily perceived from Figs. 3 and 4 because of the different orientations of the molecules.

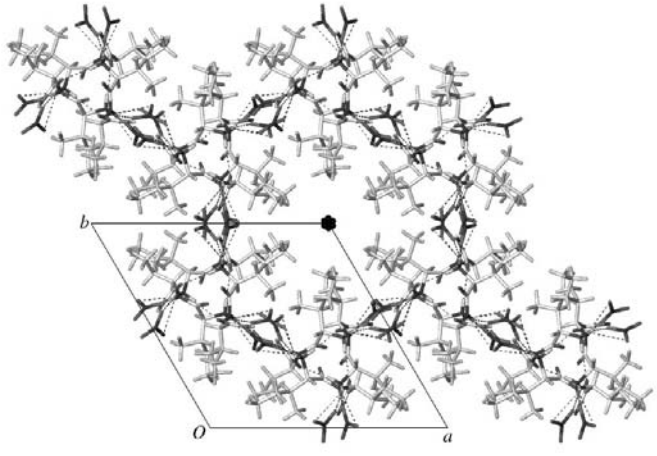


Figure 3
The unit cell and crystal packing of IIc, viewed along the *c* axis. Positions for disordered solvent molecules appear as small spheres near the trigonal axis. Hydrogen bonds are shown as dashed lines.

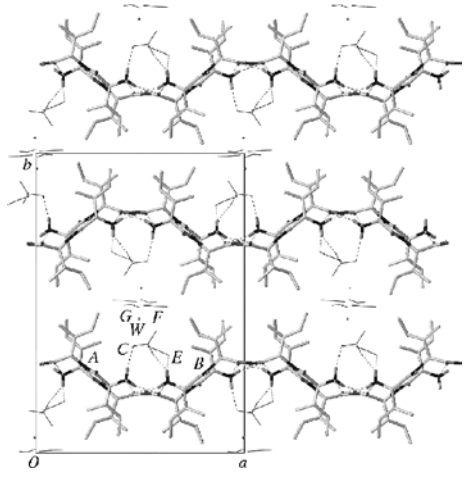


Figure 4
The unit cell and crystal packing of IIdc, viewed along the *c* axis. Peptide *B* is shown in a darker tone. Nitrate anions and solvent molecules, shown in the line drawings, have been labeled according to Fig. 5. Only one of the *C*-site nitrate positions is included. The low-occupancy nitrate position *D* and all H atoms not involved in hydrogen bonds have been omitted for clarity.

The conformation of (IIc) is slightly more folded than that for IIz (and IIdc); this difference is brought about by a rotation of φ_2 from −142.51 (6) to −58.0 (4)° (Table 2). The value for θ in this case is −95.2°.

As IIz is deprotonated to IIa, this rather homogenous picture changes dramatically. The side chains are flipped to the same side of the peptide plane for both molecules in the asymmetric unit, giving $\theta = -11.1°$ for *A* and −10.5° for *B*. The seven other known dipeptides with $|\theta| < 45°$ all have ψ_1 values in the range 111–163° and φ_2 values in the range 48–60°; the configurations of these dipeptides have the effect of pointing the peptide N—H bond in the opposite direction to the N-terminal amino group and the C-terminal carboxylate group, thus facilitating hydrogen bonding. A completely new peptide conformation, rendered possible by the uncharged N-terminal group, is observed for IIa, where ψ_1 is −32.0 (4) and −35.2 (4)° for *A* and *B*, respectively, and φ_2 is −137.1 (4) and −138.9 (4)°. The N—H peptide bond then points in the same direction as the two terminal groups and uniquely forms an intramolecular three-center interaction with both (Fig. 2*d*).

There are no unusual conformations for the 12 *sec*-butyl Ile side chains in the four structures shown in Figs. 2–5. Seven have the common *gauche*−/*trans*,*trans* combination for $\chi^{1,1}/\chi^{1,2},\chi^2$ (Ponder & Richards, 1987; Benedetti *et al.*, 1983); the remaining five are distributed between the three other low-energy conformations *gauche*−/*trans*,*gauche*− (two), *gauche*+/*trans*,*trans* (two) and *gauche*+/*trans*,*gauche*+ (one). The two molecules in the asymmetric unit of IIdc are very similar, while the two molecules of IIa differ at χ_1^2, which is *trans* for molecule *A* and *gauche*− for molecule *B* (Table 2 and Fig. 2*d*).

3.3. Crystal packing

The exceptional crystal packing arrangement of IIc is shown in Fig. 3. Elements from the VA class, as exemplified by the model shown in Fig. 1, are readily recognized, but the symmetry is trigonal rather than hexagonal. The most important modification, however, is the insertion of nitrate ions into the hydrogen-bonded ring surrounding a hydrophobic column. The hydrophobic volume thus undergoes the increase required to encompass twelve Ile side chains (not including translation), with room remaining for a central channel with van der Waals diameter 4.0 Å [Val-Val: 4.3 Å; Ile-Val and Val-Ile: 3.3 Å (Görbitz, 2003)]. The associated increases in cell dimensions are from $a = b = 14.8003$ (6) and $c = 10.2809$ (3) Å for Ile-Val, which has the largest unit cell of the VA class, to 16.0682 (6) and 11.5317 (9) Å for IIc. As for the VA class, solvent molecules are trapped within the channels in an unstoichiometric ratio, and deterioration of crystal quality with time can be attributed to solvent loss in the same manner as observed for Val-Ala (Görbitz & Gundersen, 1996).

The structure of IIdc (Fig. 4) is divided into hydrophilic layers and hydrophobic layers. The two side chains of a specific peptide are in alternate (but structurally identical) layers, a common-arrangement previously classified as an L1) type of structure (Görbitz & Etter, 1992*b*). Nitrate ions reside inside conspicuous channels parallel to the *c* axis, together with solvent molecules that may diffuse out of the crystal at room temperature, thus causing the observed slow decay of the crystals. The peptide main chains form a single hydrogen-bonded wave-shaped sheet. The pseudo-twofold rotation axis relating the two peptide molecules in the asymmetric unit runs parallel to the *b* axis in Fig. 4, at $x = 0.5$ and $z = 0.25$. A second element of pseudosymmetry is the rotation axis parallel to the *c* axis at $c = 0.25$ and $y = 0.25$.

The crystal packing arrangement of IIz dihydrate in Fig. 5 is characterized by the three-dimensional hydrogen-bond network that incorporates hydrophobic columns parallel to the 7.6341 (3) Å *a* axis. The central channels with hexagonal symmetry observed for the VA class are missing for IIz. The structure instead resembles the tetragonally symmetric structures of Ala-Ala (Fletterick *et al.*, 1971) and Abu-Ala (Görbitz, 2002*c*), with four independent side-chain contributions to each hydrophobic column, classified as a C3 structure (Görbitz & Etter, 1992*b*). Most of all, however, IIz is

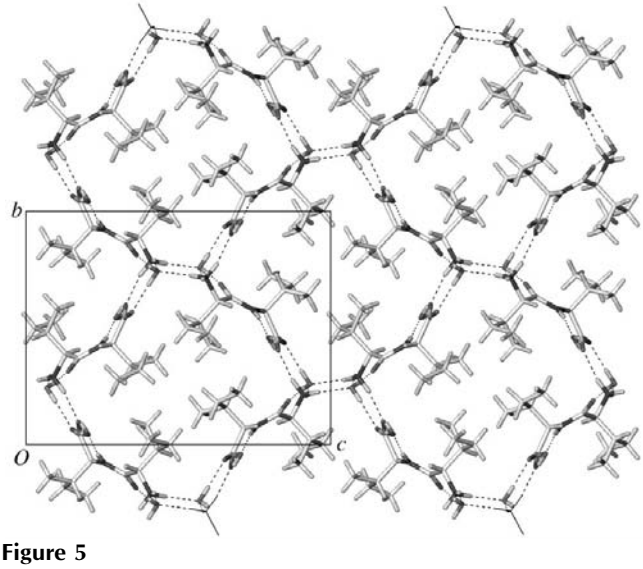


Figure 5
The unit cell and crystal packing of IIz, viewed along the *a* axis.

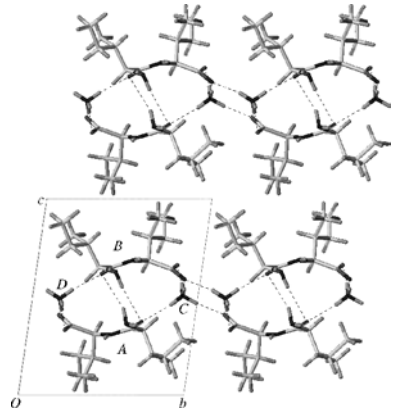


Figure 6
The unit cell and crystal packing of IIa, viewed along the *c* axis. Labels have been included for peptide anions *A* and *B*, as well as ammonium cations *C* and *D*.

remarkably similar to the structure of Phe-Ala dihydrate (Görbitz, 2001*a*). The combined bulk of the two side chains is about the same for Ile-Ile and Phe-Ala, with eight C atoms (4 + 4 and 7 + 1). The calculated density is of IIz (1.187 g cm^{-3}), which is comparable to the density of Ile-Val (1.210 and 1.182 Mg m^{-3} with and without including solvent in the channels; Görbitz, 2003) and slightly lower than that for Ala-Ala (1.276 Mg m^{-3}; Fletterick *et al.*, 1971).

The triclinic unit cell of IIa has an almost perfect twofold screw axis at the center of the unit cell when viewed along the *a* axis and closely mimics the $P2_1$ space group (Fig. 6). The monoclinic symmetry is broken only by the above-mentioned χ_1^2 difference between *A* and *B*, deviations from 90° for the β and γ cell angles, which would correspond to α and γ in a monoclinic system [β = 80.484 (3)° and γ = 88.350 (3)°; Table 1]. The side chains do not form hydrophobic columns, like those in the crystal structure of IIz, but rather form thick hydrophobic layers. The special aspect of the structure is that both side chains of a specific molecule are located in the same layer, an L2 type of structure (Görbitz & Etter, 1992*b*). This particularity is rendered possible by the low θ angles discussed previously and is usually observed for mixed LD peptides. Among regular LL peptides, only Val-Phe dihydrate (Görbitz, 2002*a*), Ala-Trp (Emge *et al.*, 2000) and Phe-Pro (Panneerselvam & Chacko, 1989) share this packing pattern. In all four structures, the hydrophilic region is constructed from two more or less well defined hydrogen-bonded sheets, as seen for IIa in Fig. 6. The instability of the IIa crystals is probably due to the absorption of water vapor, with the subsequent conversion to IIz (as a powder) and release of NH_3(g).

3.4. Intermolecular interactions

Individual strands in the left-handed double helices of the VA class (Görbitz, 2003), visible as triangles in Fig. 1, are held together by head-to-tail –$NH_3^+\cdots{}^-$OOC– chains (Suresh & Vijayan, 1985) and –$NH_3^+\cdots$O=C< hydrogen bonds that cross-link the two strands. A second –$NH_3^+\cdots{}^-$OOC– interaction, together with C^α—H$\cdots{}^-$OOC– and >N—H$\cdots{}^-$OOC hydrogen bonds, connect adjacent helices. In the double helices of IIc, shown in detail in Fig. 7, only the head-to-tail chain has been retained, with cross-linking being provided by a second $NH_3^+\cdots{}^-$OOC– interaction as well as a weak C^α—H$\cdots$O=C< hydrogen bond. The hydrogen-bond pattern is dominated by nitrate receptors, which interact with the whole palette of donors including –COOH, –NH_3^+, >N—H and C^α—H at the interface between double helices. Therefore, all three direct peptide–peptide contacts of the VA class have been lost, and nitrate ions instead serve the role of helix connectors. The new layout of the double helices has led to a *c*-axis length of 11.5317 (9) Å, which is significantly longer than that of Val-Ile [10.3140 (13) Å; Görbitz, 2003] and is thus the highest value found for the VA class.

Neighboring nitrate ions in Fig. 7 are almost coplanar, and the N$\cdots$N distance is just 3.369 Å, which would correspond to a normal van der Waals separation. In view of the electrostatic repulsion between negative charges this contact may appear to be somewhat unexpected, but a CSD search revealed a substantial number of nitrate ions with N$\cdots$N separations of about 3.15 Å and upwards.

One amino H atom of each peptide in IIdc is accepted by a nitrate ion or a cocrystallized acetonitrile molecules. Other hydrogen bonds between peptides are shown in Fig. 8. *A* and *B* peptides related by pseudo-twofold rotation generate hydrogen-bonded tapes in the *z* direction with peptide carbonyl acceptors and >N—H and C^α—H donors. Traditional head-to-tail chains involving just one type of species run parallel to the *a* axis. The remaining amino H atoms take part

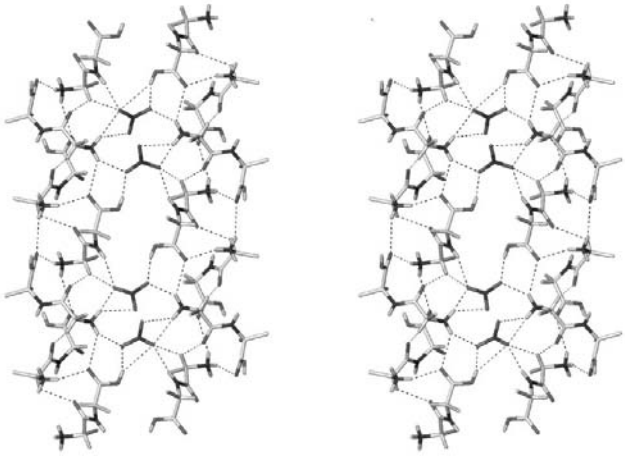


Figure 7
Stereoview of nitrate anions connecting two peptide columns in IIc. When one nitrate anion is in the *B* position, the adjacent nitrate must be in the *C* position to avoid steric conflict. For the nitrate pair this yields two alternative arrangements, both shown in this illustration. The view is along the twofold axis in space group $P3_121$ (the short unit-cell diagonal), while the threefold screw axis is vertical.

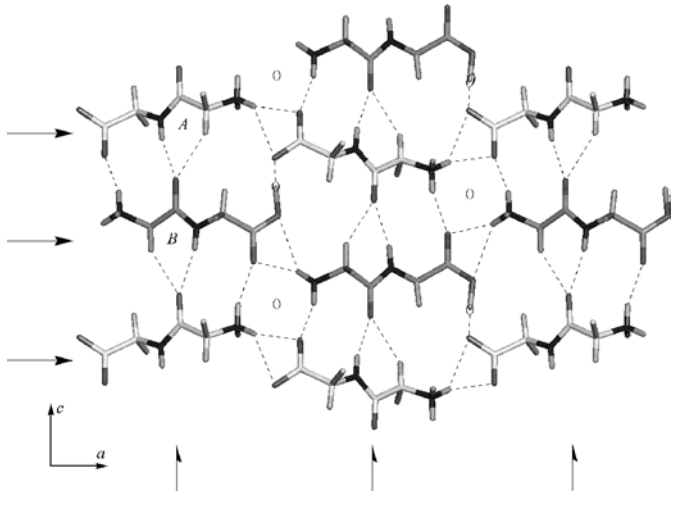


Figure 8
Hydrogen bonds between peptide molecules in the crystal structure of IIdc. Peptide *B* molecules are shown in a slightly darker tone; side chains have been terminated at C^β for clarity. The view is along the *b* axis, the same direction as the pseudo-twofold rotation axes indicated by (). Arrows identify true twofold rotation axes parallel to the *a* axis and pseudo-twofold rotation axes parallel to the *c* axis.

Table 3
Hydrogen-bond geometry (Å, °).

D—H···A	D—H	H···A	D···A	D—H···A
IIc				
O2A—H5A···O1B^{i}	0.79 (5)	2.01 (5)	2.745 (5)	153 (5)
O2A—H5A···O1C^{i}	0.79 (5)	1.80 (5)	2.543 (5)	156 (5)
N1A—H1A···O3A^{ii}	0.91	2.17	3.051 (4)	161
N1A—H2A···O1B^{iii}	0.91	2.29	2.914 (6)	126
N1A—H2A···O1C^{iii}	0.91	2.05	2.824 (6)	142
N1A—H2A···O3A^{iv}	0.91	2.53	3.258 (4)	137
N1A—H3A···O2C	0.91	2.52	3.187 (8)	130
N1A—H3A···O3B	0.91	2.26	3.146 (8)	165
N1A—H3A···O3C	0.91	1.82	2.726 (9)	172
N2A—H4A···O3B^{ii}	0.88	2.14	2.981 (8)	161
N2A—H4A···O3C^{ii}	0.88	2.18	3.028 (4)	161
C1A—H11A···O1A^{ii}	1.00	2.42	2.999 (4)	116
C1A—H11A···O3B^{ii}	1.00	2.45	3.262 (9)	138
IIdc				
O3B—H5B···O3A	1.12 (3)	1.31 (3)	2.429 (3)	173 (3)
N1A—H1A···N1E^{v}	0.89	2.23	2.952 (5)	138
N1A—H1A···O2C^{vi}	0.89	2.24	2.745 (7)	115
N1A—H2A···O2A^{v}	0.89	2.00	2.841 (3)	156
N1A—H2A···O3A^{v}	0.89	2.47	3.188 (3)	139
N1A—H3A···O2B^{vi}	0.89	2.02	2.843 (3)	154
N2A—H4A···O1B^{v}	0.86	2.21	3.033 (3)	161
C1A—H11A···O1B^{v}	0.98	2.36	3.244 (3)	150
N1B—H1B···O3C^{vii}	0.89	1.90	2.782(15)	173
N1B—H1B···O1D^{vii}	0.89	2.20	2.88 (3)	133
N1B—H2B···O2B^{vii}	0.89	2.02	2.865 (3)	158
N1B—H2B···O3B^{vii}	0.89	2.50	3.223 (3)	139
N1B—H3B···O2A^{vii}	0.89	2.00	2.828 (3)	155
N2B—H4B···O1A^{vii}	0.86	2.16	2.989 (3)	162
C1B—H11B···O1A^{vii}	0.98	2.33	3.215 (3)	150
C2F—H3F···O1C	0.96	1.91	2.827 (3)	158
IIz				
N1—H1···O2^{ix}	0.928 (11)	1.814 (11)	2.7293 (7)	168.7 (10)
N1—H2···O2W	0.835 (11)	2.001 (12)	2.7920 (7)	157.8 (11)
N1—H3···O2W^{x}	0.903 (10)	1.959 (10)	2.8502 (7)	169.2 (10)
N2—H4···O1W^{xi}	0.831 (11)	2.173 (11)	2.9905 (8)	167.9 (9)
O1W—H11W···O3^{ix}	0.811 (16)	1.997 (16)	2.7968 (7)	168.5 (15)
O1W—H12W···O2^{xii}	0.801 (17)	2.364 (15)	2.9487 (8)	130.6 (14)
O2W—H21W···O1^{xiii}	0.772 (12)	1.969 (13)	2.7271 (7)	167.5 (12)
O2W—H22W···O3^{ix}	0.870 (12)	1.792 (13)	2.6582 (7)	173.6 (11)
IIa				
N1A—H1A···O1B^{xiii}	0.93 (5)	2.40 (5)	3.323 (5)	174 (4)
N1A—H2A···O1A^{xiii}	0.89 (5)	2.56 (5)	3.410 (5)	162 (4)
N2A—H4A···O2A	0.85 (4)	2.26 (4)	2.666 (4)	109 (3)
N2A—H4A···N1A	0.85 (4)	2.27 (4)	2.698 (5)	112 (3)
N1B—H1B···O1A	0.93 (5)	2.43 (5)	3.318 (5)	158 (4)
N1B—H2B···O1B^{xiii}	0.85 (5)	2.68 (5)	3.451 (4)	153 (4)
N2B—H4B···O2B	0.73 (4)	2.44 (5)	2.681 (4)	102 (4)
N2B—H4B···N1B	0.73 (4)	2.28 (4)	2.700 (5)	118 (4)
N1C—H1C···O1A	0.86 (2)	2.01 (2)	2.862 (4)	173 (4)
N1C—H2C···O2B	0.88 (2)	1.89 (2)	2.768 (4)	174 (4)
N1C—H3C···O3A^{xiv}	0.87 (2)	1.93 (2)	2.777 (4)	166 (4)
N1C—H4C···O3B^{xiii}	0.85 (2)	2.03 (2)	2.863 (4)	165 (4)
N1D—H1D···O1B	0.86 (2)	2.02 (2)	2.882 (4)	172 (4)
N1D—H2D···O2A	0.88 (2)	1.93 (2)	2.779 (4)	164 (4)
N1D—H3D···O3B	0.89 (2)	1.90 (2)	2.764 (4)	164 (4)
N1D—H4D···O3A	0.86 (2)	1.98 (2)	2.829 (4)	170 (4)

Symmetry codes: (i) $1-x, -x+y, -z+\frac{1}{3}$; (ii) $1-y, x-y, z+\frac{1}{3}$; (iii) $y, x, 1-z$; (iv) $1-x+y, 1-x, z+\frac{2}{3}$; (v) $x+\frac{1}{2}, -y+\frac{3}{2}, -z$; (vi) $x+\frac{1}{2}, -y+\frac{3}{2}, 1-z$; (vii) $x-\frac{1}{2}, -y+\frac{3}{2}, 1-z$; (viii) $x-\frac{1}{2}, -y+\frac{3}{2}, -z$; (ix) $-x, y-\frac{1}{2}, -z+\frac{1}{2}$; (x) $x-\frac{1}{2}, -y+\frac{1}{2}, -z$; (xi) $1-x, y+\frac{1}{2}, -z+\frac{1}{2}$; (xii) $1-x, y-\frac{1}{2}, -z+\frac{1}{2}$; (xiii) $x+1, y, z$; (xiv) $x, y-1, z$; (xv) $x-1, y, z$; (xvi) $x, y+1, z$.

in ···H—N—H···O—C—OH···O—C—O··· chains (N—H···O—C—OH is three-centered). The carboxyl–carboxylate interaction is very short; the O···O distance is 2.429 (3) Å and the H atom is positioned almost symmetrically between the O atoms (Table 3). The *anti–anti* geometry has previously been observed for only about 35 structures in the CSD, with a minimum O···O distance of 2.437 (2) Å at 30 K for the much studied hydrogen bond in potassium hydrogen dichloromaleate (Olovsson *et al.*, 2001, and references therein). Among interactions with the much more plentiful *syn–syn* geometry only six are shorter, including, interestingly, three amino acid nitrates: bis(Lys)·2HCl·HNO_3 (Srinivasan *et al.*, 2001*a*), bis(Phe)·HNO_3 (Srinivasan *et al.*, 2001*b*) and bis(Pro)·HNO_3 (Pandiarajan *et al.*, 2002*b*). The latter has an O···O distance of 2.414 (3) Å, which is the shortest carboxyl–carboxylate interaction overall in the CSD.

The three-dimensional hydrogen-bond network of IIz includes only a single direct peptide–peptide interaction (N1—H1···O2), which gives rise to a zigzag head-to-tail hydrogen-bonded chain parallel to the *b* axis. There are no hydrogen bonds between water molecules, but a total of seven reasonably linear peptide–water interactions are present.

The intramolecular hydrogen bonds involving the peptide amide groups of IIa have been discussed above; geometry parameters are listed in Table 3. It is interesting to see how the rest of the hydrogen-bond pattern is completely dominated by strong interactions between the ammonium ions and carboxylate or carbonyl acceptors. The two *N*-terminal $-NH_2$ groups participate only in surprisingly weak interactions, with the peptide carbonyl groups as acceptors (Table 3). Since the carbonyl groups in IIz and IIa are involved in hydrogen bonds with N—H donors, there is no room for C^{α}—H···O=C interactions such as those found for (IIc) and (IIdc).

4. Summary

The crystal structures of the dipeptide Ile-Ile in four different protonation states illustrate the conformational flexibility of the peptide main chain during two titration steps, with an unprecedented formation of an intramolecular three-center hydrogen bond between the amide >N—H group and the terminal $-COO^{-}$ and $:NH_2-$ acceptors in the anionic state. The crystal-packing patterns vary tremendously, the hexagonal structure for the cation nitrate being the most curious solution to the quest for aggregation of hydrophobic groups into well defined regions with concomitant formation of an acceptable hydrogen-bond network. Hydrogen bonding is dominated by the charged groups, either as part of the peptide chain ($-NH_3^{+}$ and $-COO^{-}$) or as counter-ions for the peptide in the cationic and anionic state (NO_3^{-} and NH_4^{+}). The observation of a dimeric cation for Ile-Ile (average charge +0.5 for each molecule) is the first report of this protonation state for a regular peptide, having been known previously only for a series of amino acid complexes.

References

Akazome, M., Matsuno, H. & Ogura, K. (1997). *Tetrahedron Asym.* **8**, 2331–2336.
Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.

Benedetti, E., Morelli, G., Némethy, G. & Scheraga, H. A. (1983). *Int. J. Peptide Protein Res.* **22**, 1–15.
Bigoli, F., Lanfranchi, M., Leporati, E., Nardelli, M. & Pellinghelli, M. A. (1981). *Acta Cryst.* B**37**, 1258–1265.
Bigoli, F., Lanfranchi, M., Leporati, E., Nardelli, M. & Pellinghelli, M. A. (1982). *Acta Cryst.* B**38**, 498–502.
Bruker (1998). *SMART.* Version 5.054. Bruker AXS Inc., Madison, Wisconsin, USA.
Bruker (2000). *SHELXTL.* Version 6.10. Bruker AXS Inc., Madison, Wisconsin, USA.
Bruker (2001). *SAINT-Plus.* Version 6.22. Bruker AXS Inc., Madison, Wisconsin, USA.
Cody, V., Langs, D. A. & Hazel, J. P. (1979). *Acta Cryst.* B**35**, 1829–1835.
Coleman, D. E. & Sprang, S. R. (1998). *Biochemistry*, **37**, 14376–14385.
Cotrait, M. & Barrans, Y. (1974). *Acta Cryst.* B**30**, 1018–1023.
Coudert, E., Acher, F. & Azerad, R. (1996). *Tetrahedron Asym.* **7**, 2963–2970.
Einspahr, H. & Bugg, C. E. (1974). *Acta Cryst.* B**30**, 1037–1043.
Emge, T. J., Agrawal, A., Dalessio, J. P., Dukovic, G., Inghrim, J. A., Janjua, K., Macaluso, M., Robertson, R. R., Stiglic, T. J., Volovik, Y. & Georgiadis, M. M. (2000). *Acta Cryst.* C**56**, e469–e471.
Faggiani, R., Howard-Lock, H. E., Lock, C. J. L. & Martins, M. L. (1984). *Can. J. Chem.* **62**, 1127–1133.
Flack, H. D. (1983). *Acta Cryst.* A**39**, 876–881.
Fletterick, R. J., Tsai, C.-C. & Hughes, R. E. (1971). *J. Phys. Chem.* **75**, 918–922.
Freeman, H. C. & Moore, C. J. (1977). *Acta Cryst.* B**33**, 2690–2692.
Görbitz, C. H. (1997). *Acta Cryst.* C**53**, 736–739.
Görbitz, C. H. (1999). *Acta Cryst.* C**55**, 2171–2177.
Görbitz, C. H. (2001*a*). *Acta Cryst.* C**57**, 575–576.
Görbitz, C. H. (2001*b*). *Chem. Eur. J.* **7**, 5153–5159.
Görbitz, C. H. (2002*a*). *Acta Cryst.* B**58**, 512–518.
Görbitz, C. H. (2002*b*). *Acta Cryst.* B**58**, 849–854.
Görbitz, C. H. (2002*c*). *Acta Cryst.* C**58**, o533–o536.
Görbitz, C. H. (2003). *New J. Chem.* **27**, 1789–1793.
Görbitz, C. H. & Etter, M. C. (1992*a*). *Acta Cryst.* C**48**, 1317–1320.
Görbitz, C. H. & Etter, M. C. (1992*b*). *Int. J. Peptide Protein Res.* **39**, 93–110.
Görbitz, C. H. & Gundersen, E. (1996). *Acta Cryst.* C**52**, 1764–1767.
Němec, I., Císařová, I. & Mička, Z. (1996). *J. Solid State Chem.* **140**, 71–82.
Olovsson, I., Ptasiewicz-Bak, H., Gustafsson, T. & Majerz, I. (2001). *Acta Cryst.* B**57**, 311–316.
Pandiarajan, S., Sridhar, B. & Rajaram, R. K. (2001). *Acta Cryst.* E**57**, o466–o568.
Pandiarajan, S., Sridhar, B. & Rajaram, R. K. (2002*a*). *Acta Cryst.* E**58**, o74–o76.
Pandiarajan, S., Sridhar, B. & Rajaram, R. K. (2002*b*). *Acta Cryst.* E**58**, o862–o864.
Panneerselvam, K. & Chacko, K. K. (1989). *Acta Cryst.* C**45**, 106–109.
Ponder, J. W. & Richards, F. M. (1987). *J. Mol. Biol.* **193**, 775–791.
Ravikumar, B., Sridhar, B. & Rajaram, R. K. (2001). *Acta Cryst.* E**57**, o682–o684.
Ravikumar, B., Sridhar, B. & Rajaram, R. K. (2002). *Acta Cryst.* E**58**, o123–o125.
Sheldrick, G. M. (1996). *SADABS.* University of Göttingen, Germany.
Silva, M. R., Paixão, J. A., Beja, A. M. & da Veiga, L. A. (2001). *Acta Cryst.* C**57**, 838–840.
Sridhar, B., Srinivasan, N. & Rajaram, R. K. (2001). *Acta Cryst.* E**57**, o1004–o1106.
Sridhar, B., Srinivasan, N., Dalhus, B. & Rajaram, R. K. (2002*a*). *Acta Cryst.* E**58**, o779–o781.
Sridhar, B., Srinivasan, N. & Rajaram, R. K. (2002*b*). *Acta Cryst.* E**58**, o1372–o1374.
Srinivasan, N., Sridhar, B. & Rajaram, R. K. (2001*a*). *Acta Cryst.* E**57**, o888–o890.
Srinivasan, N., Sridhar, B. & Rajaram, R. K. (2001*b*). *Acta Cryst.* E**57**, o916–o918.
Sudbeck, E. A., Etter, M. C. & Gleason, W. B. (1994). *Chem. Mater.* **6**, 1192–1199.
Suresh, C. G. & Vijayan, M. (1985). *Int. J. Pept. Protein Res.* **26**, 311–328.
Tokuma, Y., Ashida, T. & Kakudo, M. (1969). *Acta Cryst.* B**25**, 1367–1373.