Acta Crystallographica Section C
Crystal Structure Communications
ISSN 0108-2701

Cytosine-containing hybrid dipeptides: *N*-[2-(4-amino-2-oxo-1,2-dihydropyrimidin-1-yl)propionyl]-L-phenylalanine *N*-[2-(4-amino-2-oxo-1,2-dihydropyrimidin-1-yl)propionyl]-L-serine monohydrate and *N*-[2-(4-amino-2-oxo-1,2-dihydropyrimidin-1-yl)propionyl]-L-lysine

Mitsunobu Doi,* Yasuhiko Nakamoto and Akiko Asano

Osaka University of Pharmaceutical Sciences, 4-20-1 Nasahara, Takatsuki, Osaka 569-1094, Japan
Correspondence e-mail: doit@gly.oups.ac.jp

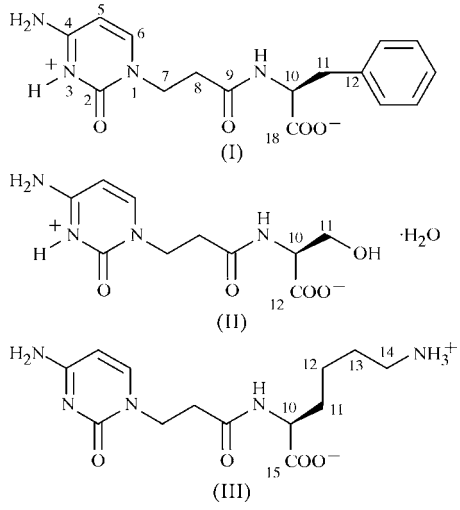
Received 18 April 2005
Accepted 17 August 2005
Online 10 September 2005

The title compounds are cytosine-incorporating hybrid dipeptides showing affinities for 9-ethyl-7-methylguanine (7mG). Four molecules of the L-phenylalanine (L-Phe) derivative, $C_{16}H_{14}N_4O_4$, are present in the asymmetric unit, with similar folded conformations but with slightly different torsion angles involving the L-Phe group. The L-serine (L-Ser) derivative crystallizes as a monohydrate, $C_{10}H_{14}N_4O_5\cdot H_2O$, the two independent molecules having extended conformations, whereas the two independent molecules of the L-lysine (L-Lys) system in the final compound, $C_{13}H_{21}N_5O_4$, are folded. The cytosine–cytosine base pair (pyrimidine N$\cdots$N interactions) was observed only for the L-lysine derivative. Conformational comparisons with previous structures of cytosine hybrid dipeptides may show the relationships between side-chain position and binding phases for 7mG.

Comment

Nucleic acid-incorporated peptides have been designed with the expectation that the incorporated amino acids could potentially strengthen interactions with complementary bases (Williams *et al.*, 1977; Voet, 1980). We have focused on the cytosine base and designed hybrid peptides (Tarui *et al.*, 1996). Cytosine-incorporated dipeptides have shown affinities for 9-ethyl-7-methylguanine (7mG) but not for 9-ethylguanine (Asano *et al.*, 2002). The incorporated amino acids affect the association constants between the dipeptides and 7mG (0.2–6.6 $\times$ 10^{-6} M^{-1}) and induce different binding phases at low and high peptide concentrations. To analyze the fundamental characteristics of cytosine hybrid dipeptides, the structures of the cytosine-containing dipeptides with L-phenylalanine (L-Phe), (I), L-serine (L-Ser), (II), and L-lysine (L-Lys), (III), have been determined.

Four molecules of (I) are present in the asymmetric unit (molecules *A*–*D*), with similar folded conformations (Fig. 1). Slight disorder is observed for the phenyl ring of *C*, with approximately 4% occupancy for the minor site. No significant difference is observed for the carboxyl C18*n*—O18*n* and C18*n*—O19*n* bond lengths (*n* = *A*–*D*), and a zwitterionic form with N3-quarternization is established for molecules *A*–*D*. Rotations about the C10*n*–C11*n* and C11*n*–C12*n* bonds mainly contribute to the different orientations of the pyrimidine and phenyl rings in the independent molecules (Table 1). The angles between the least-squares planes of the rings are 38.8 (2), 57.8 (2), 31.2 (3) and 27.8 (2)° for *A*, *B*, *C* (major part) and *D*, respectively, with intramolecular distances of, respectively, 4.401 (4), 4.683 (3), 4.392 (4) and 4.380 (3) Å between the centroids of the pyrimidine and phenyl rings, indicative of no intramolecular interactions between the aromatic rings. A similar folding was found in the structure of (2-carboxyethyl)cytosin-1-yl-L-tyrosine, (IV) (Doi, Miyako *et al.*, 1999). A complicated hydrogen-bonding scheme is



observed for (I) (Table 2). Atoms N3*n* and N4*n* of the cytosine base interact with carboxyl atoms O18*n* and O19*n* (Fig. 2). These pairings form eight-membered rings between molecules *A* and *D*(x, $y+1$, z), *B* and *C*($x+1$, y, $z-1$), *C* and *B*, and *D* and *A*($x-1$, $y-1$, $z+1$), with r.m.s. deviations of 0.25, 0.23, 0.17 and 0.15 Å, respectively. Atoms N4*n* of molecules *B*, *C* and *D* are hydrogen bonded to two acceptor atoms, but atom N4*A* (molecule *A*) has one acceptor atom (O19*D*). Consequently, molecule *A* has three hydrogen-bond donors, and molecules *B*, *C* and *D* have four (Table 2). These differences are related to acceptor atoms, because the numbers of hydrogen-bond acceptor atoms are 5, 4, 3 and 3 for molecules *A*, *B*, *C* and *D*, respectively. The independent molecules having slightly different conformations form this unique hydrogen-bonding network with no solvent molecules.

The two independent molecules of (II) (*A* and *B*) crystallize in a monohydrate form with an extended conformation (Fig. 3). In molecules *A* and *B*, the rotations of the N1*m*—C7*m* bonds are opposite to one another [74.2 (4) and −80.8 (4)° for

 DOI: 10.1107/S0108270105026259

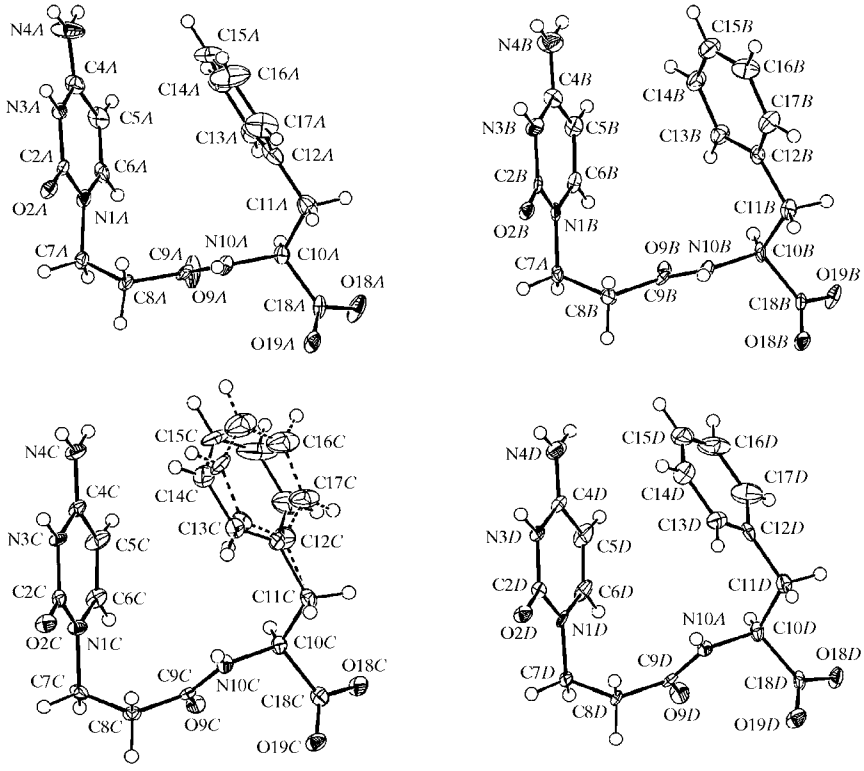


Figure 1
The structures of the molecules of (I), with displacement ellipsoids at the 50% probability level. The four independent molecules in the asymmetric unit are depicted and projected from a similar axis. The phenyl ring of molecule *C* is disordered over two sites, and the minor part has been drawn using dashed lines.

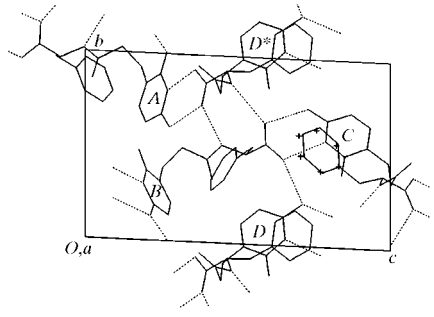


Figure 2
A packing diagram for (I) (*MERCURY*; Bruno *et al.*, 2002), viewed along (100). Molecules *A*–*D* are labeled at the cytosine base. Dotted lines represent hydrogen bonds. The minor part of the disordered phenyl ring is indicated by crosses (+). The asterisk (*) represents the symmetry operation (x, $1 + y$, z).

m = *A* and *B*, respectively; Table 3], and the dispositions of the L-Ser moieties are different for the base planes. The extended structures are induced by the *trans* positions of the C7m—C8m bond, *viz.* −178.1 (3) and −179.9 (3)° for *A* and *B*, respectively. The (2-carboxyethyl)cytosin-1-yl-L-threonine, (V), also crystallizes as a monohydrate, but the *syn* position of the C7—C8 bond results in an L conformation [N1—C7—C8—C9 = 112.5 (3) and 113.3 (3)°; Doi, Asano & Ishida, 1999]. The structures of (II) are similar to the extended structures of (2-carboxyethyl)cytosin-1-yl-L-tryptophan, (VI) [N1—C7—C8—C9 = −175.4 (4) and −179.4 (4)°; Doi *et al.*, 1998]. Hydrogen bonds between the cytosine base and the carboxyl group are formed between symmetry-related molecules *A* (N3*A*···O13*A*^{vii} and N4*A*···O12*A*^{vii}), indicating the zwitterionic form (Fig. 4 and Table 4); in addition, the C12—O12 and C12—O13 bond lengths are similar. Atom N4*A* also interacts with atom O13*B* of the adjacent molecule *B*. The hydroxy group of molecule *A* (O11*A*) interacts with the carboxyl group of molecule *B* (O12*B*) and a water molecule (O1*C*). Similar hydrogen bonds are observed for symmetry-related molecules *B*. The water molecule O1*C* bridges two molecules *A*, which translate along (100) (as O9*A*···O1*C*···O11*A*^{xi}), and the water molecule O1*D* similarly bridges molecules *B* (O9*B*^{xi}···O1*D*···O11*B*; symmetry code as in Table 4).

Two independent molecules, with a folded form (Fig. 5), are also present in the asymmetric unit of (III) (*A* and *B*). Molecules *A* and *B* are related by pseudosymmetry that fits 95% of the atoms of both molecules, but they are distinguished by some bond rotations (Table 5); the rotations of the N1o–C7o bonds induce different foldings for the base plane, *viz.* −77.2 (3) and 75.1 (3)° for o = *A* and *B*, respectively. Moreover, the Lys side chains are expanded over the base planes in a similar manner in *A* and *B*. These conformations are also different at the terminal ε-amino groups; the C12o—C13o—C14o—N14o torsion angles are −62.3 (3) and 73.6 (3)° for *A* and *B*, respectively. Cytosine–cytosine base pairings (N4*A*···N3*B*^{x} and N4*B*···N3*A*^{xiv}) are formed between molecules *A* and *B* (Table 6 and Fig. 6), a phenomenon that has not been observed in the structures of cytosine-hybrid dipeptides. This base pair indicates an un-ionized state of atoms N3o. The ε-amino groups (N14*A* and N14*B*) interact with carboxyl groups (Table 6), and ionized states are established for the carboxyl and ε-amino groups.

The independent molecules of each dipeptide have been fitted to the cytosine base (Fig. 7). The phenyl rings of (I) are located at one side of the cytosine base plane, but the side chains of (II) are separated from the base. The long side chains of (III) are located over the base, positioned on both sides of the base plane. When the positions of the amino acid side chain are defined as sites *O*, *P* and *S*, as shown in Fig. 7, the present and previously reported structures of cytosine-hybrid peptides are classified as listed in Table 7.

The side chains of (I), (III) and (IV) are located on site *O*, and the single folding to site *O*+ is observed for the aromatic analogs (I) and (IV). Compound (VI) also has an aromatic side chain, but the side-chain position is *S*±. In (VI), intermolecular π–π interactions are observed between cytosine bases and indole rings (Doi *et al.*, 1998). This interaction seems to affect the side-chain position of (IV). The side chains of (II) and (2-carboxyethyl)cytosin-1-yl-L-isoleucine, (VII) (Doi, Tsunemichi *et al.*, 1999), are folded to site *S*, but single folding to site *S*− is only observed for (VII). The side chains of (V) and (2-carboxyethyl)cytosin-1-yl-L-alanine, (VIII) (Doi, Tarui *et al.*, 1999), are expanded approximately perpendicular to the base plane (site *P*±). We can postulate relations between these folding positions and the associated constants with 7mG, but unfortunately find no clear pattern. However, the side-chain position may be related to binding phases, because two binding phases are observed for (I), (II), (IV) and (V). The hybrid dipeptides showing two binding phases interact with 7mG at an extra place(s) in addition to the cytosine base. The hydroxy groups of (II) and (V) separated from the cytosine base (sites *S* or *P*) are suitable for interaction with 7mG molecules. The folding forms of (I) and (IV) have space

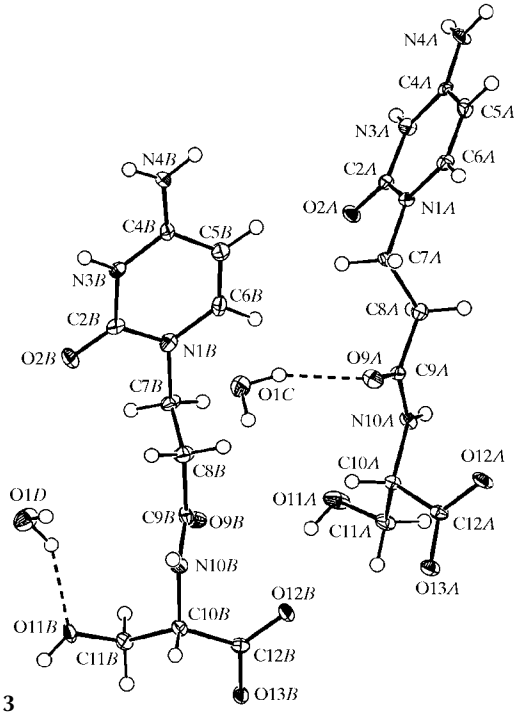


Figure 3
The structures of the molecules of (II), with displacement ellipsoids at the 40% probability level. Two independent molecules are present in the asymmetric unit, together with two water molecules (O1*C* and O1*D*); dashed lines represent hydrogen bonds.

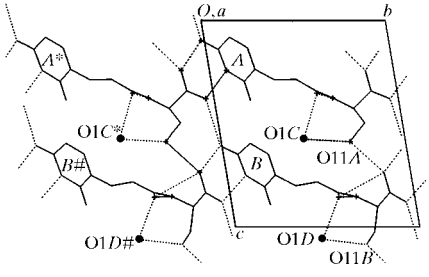


Figure 4
A packing diagram for (II) (*MERCURY*; Bruno *et al.*, 2002), viewed along (100). Molecules *A* and *B* are labeled at the cytosine base; dotted lines represent hydrogen bonds. The asterisks (*) and hashes (#) represent the symmetry operations ($x+1, y-1, z$) and ($x-1, y-1, z$), respectively. Crosses (+) represent overlapped atoms along the projection axis.

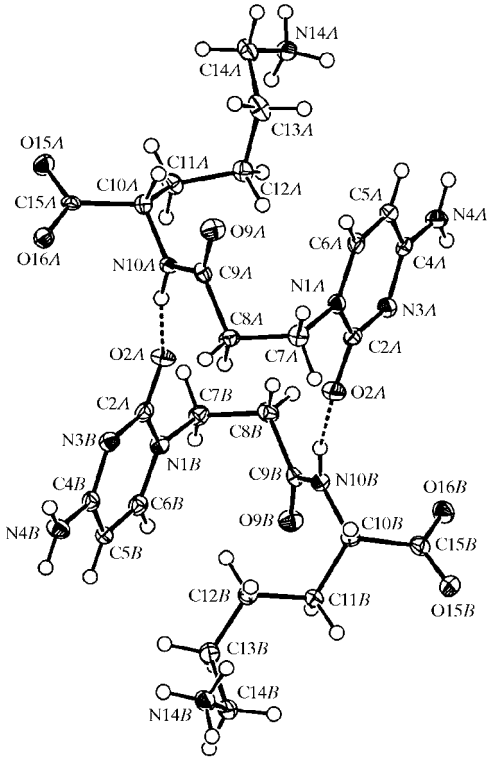


Figure 5
The structures of the two independent molecules of (III), with displacement ellipsoids at the 50% probability level; dashed lines represent hydrogen bonds.

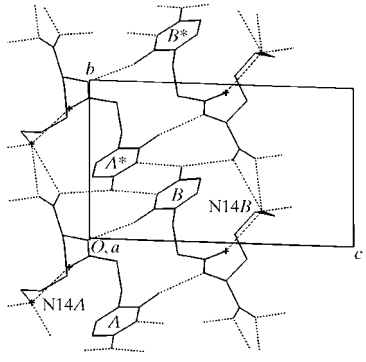


Figure 6
A packing diagram for (III) (*MERCURY*; Bruno *et al.*, 2002), viewed along (100). Molecules *A* and *B* are labeled at the cytosine base; dotted lines represent hydrogen bonds. Asterisks (*) represent the symmetry operation ($x+1, y+1, z$) and crosses (+) represent overlapped atoms along the projection axis.

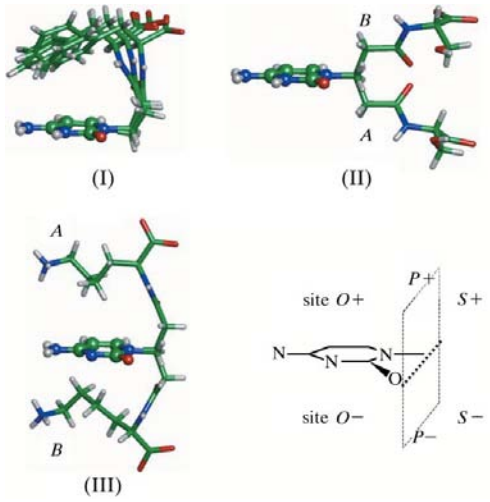


Figure 7
Fits (*iMol*; Rotkiewicz, 2004) to the cytosine base and the side-chain folding for the base; the minor part of disordered phenyl ring of (I) is not shown.

between the aromatic ring and the base, able to accept a 7mG molecule.

Experimental

The syntheses were carried out as described previously (Tarui *et al.*, 1996). L-Amino acids were used for the syntheses of (I), (II) and (III). Crystals of (I) and (II) were grown from aqueous hexyleneglycol solutions. Crystals of (III) were grown from an aqueous dimethylformamide solution. Each peptide (8–10 mg) was dissolved in the organic solvent (0.3–0.4 ml) with heating and 1–2 drops of water (50–100 µl) were added to the solution. Crystals were grown over periods of between two weeks and a month.

Compound (I)

Crystal data

$C_{16}H_{18}N_4O_4$
M_r = 330.34
Triclinic, $P\overline{1}$
a = 9.3587 (12) Å
b = 10.3445 (13) Å
c = 16.925 (2) Å
α = 92.079 (2)°
β = 95.775 (2)°
γ = 95.245 (2)°
V = 1621.7 (4) Å^3
Z = 4
D_x = 1.353 Mg m^{-3}
Mo $K\alpha$ radiation
Cell parameters from 2719 reflections
θ = 2.2–27.9°
μ = 0.10 mm^{-1}
T = 100 (2) K
Plate, colorless
0.25 × 0.20 × 0.04 mm

Data collection

Bruker SMART APEX diffractometer
ω scans
Absorption correction: empirical (using intensity measurements; *SADABS*; Sheldrick, 1996)
T_{min} = 0.832, T_{max} = 0.998
8913 measured reflections
6059 independent reflections
5788 reflections with $I > 2\sigma(I)$
R_{int} = 0.025
θ_{max} = 25.7°
$h = -9 \rightarrow 11$
$k = -12 \rightarrow 12$
$l = -20 \rightarrow 16$

Refinement

Refinement on F^2
$R[F^2 > 2\sigma(F^2)]$ = 0.064
$wR(F^2)$ = 0.150
S = 1.13
6059 reflections
896 parameters
H-atom parameters constrained
$w = 1/[\sigma^2(F_o^2) + (0.0753P)^2 + 0.9676P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$ = 0.011
$\Delta\rho_{max}$ = 0.33 e Å^{-3}
$\Delta\rho_{min}$ = −0.31 e Å^{-3}

Compound (II)

Crystal data

$C_{10}H_{14}N_4O_5\cdot H_2O$
M_r = 288.27
Triclinic, $P\overline{1}$
a = 4.863 (1) Å
b = 10.771 (2) Å
c = 12.228 (2) Å
α = 80.45 (3)°
β = 89.23 (3)°
γ = 88.61 (3)°
V = 631.4 (2) Å^3
Z = 2
D_x = 1.516 Mg m^{-3}
Mo $K\alpha$ radiation
Cell parameters from 1717 reflections
θ = 3.4–28.7°
μ = 0.13 mm^{-1}
T = 100 (2) K
Plate, colorless
0.25 × 0.25 × 0.06 mm

Table 2
Hydrogen-bond geometry (Å, °) for (I).

D—H···A	D—H	H···A	D···A	D—H···A
N3*A*—H3*A*···O18*D*^{i}	0.88	1.75	2.623 (6)	171
N4*A*—H4*AA*···O19*D*^{i}	0.88	1.94	2.811 (7)	173
N10*A*—H10*A*···O19*C*^{ii}	0.88	2.19	3.036 (6)	161
N3*B*—H3*B*···O18*C*^{iii}	0.88	1.73	2.565 (6)	159
N4*B*—H4*AB*···O19*C*^{iii}	0.88	2.14	2.965 (7)	156
N4*B*—H4*BB*···O2*A*^{iv}	0.88	2.42	2.934 (7)	118
N10*B*—H10*B*···O19*D*^{i}	0.88	2.05	2.885 (6)	159
N3*C*—H3*C*···O19*B*	0.88	1.74	2.601 (5)	167
N4*C*—H4*AC*···O18*B*	0.88	2.05	2.903 (6)	164
N4*C*—H4*BC*···O19*A*^{v}	0.88	2.28	3.090 (7)	154
N10*C*—H10*C*···O18*A*^{vi}	0.88	2.03	2.890 (5)	167
N3*D*—H3*D*···O19*A*^{vi}	0.88	1.75	2.615 (5)	168
N4*D*—H4*AD*···O18*A*^{vi}	0.88	2.06	2.920 (6)	165
N4*D*—H4*BD*···O19*B*	0.88	2.23	3.047 (6)	155
N10*D*—H10*D*···O18*B*^{iv}	0.88	2.10	2.945 (5)	162

Symmetry codes: (i) $x, y+1, z$; (ii) $x+1, y+1, z-1$; (iii) $x+1, y, z-1$; (iv) $x, y-1, z$; (v) $x-1, y, z+1$; (vi) $x-1, y-1, z+1$.

Table 1
Selected torsion angles (°) for (I).

C2*A*—N1*A*—C7*A*—C8*A*	−74.4 (5)	C2*C*—N1*C*—C7*C*—C8*C*	−78.4 (5)
N1*A*—C7*A*—C8*A*—C9*A*	−66.3 (5)	N1*C*—C7*C*—C8*C*—C9*C*	−61.1 (5)
C7*A*—C8*A*—C9*A*—N10*A*	150.0 (4)	C7*C*—C8*C*—C9*C*—N10*C*	124.4 (5)
C8*A*—C9*A*—N10*A*—C10*A*	−178.8 (4)	C8*C*—C9*C*—N10*C*—C10*C*	179.2 (4)
C9*A*—N10*A*—C10*A*—C11*A*	135.4 (5)	C9*C*—N10*C*—C10*C*—C11*C*	164.5 (4)
N10*A*—C10*A*—C11*A*—C12*A*	−53.9 (6)	N10*C*—C10*C*—C11*C*—C12*C*	−71.6 (5)
C10*A*—C11*A*—C12*A*—C13*A*	−47.3 (8)	C10*C*—C11*C*—C12*C*—C13*C*	92.9 (4)
C2*B*—N1*B*—C7*B*—C8*B*	−78.4 (5)	C2*D*—N1*D*—C7*D*—C8*D*	−80.6 (5)
N1*B*—C7*B*—C8*B*—C9*B*	−61.7 (5)	N1*D*—C7*D*—C8*D*—C9*D*	−64.2 (5)
C7*B*—C8*B*—C9*B*—N10*B*	148.0 (4)	C7*D*—C8*D*—C9*D*—N10*D*	124.2 (4)
C8*B*—C9*B*—N10*B*—C10*B*	179.9 (4)	C8*D*—C9*D*—N10*D*—C10*D*	178.4 (4)
C9*B*—N10*B*—C10*B*—C11*B*	141.4 (5)	C9*D*—N10*D*—C10*D*—C11*D*	163.8 (4)
N10*B*—C10*B*—C11*B*—C12*B*	−63.7 (6)	N10*D*—C10*D*—C11*D*—C12*D*	−59.8 (6)
C10*B*—C11*B*—C12*B*—C13*B*	−24.7 (7)	C10*D*—C11*D*—C12*D*—C13*D*	99.0 (6)

Table 3
Selected torsion angles (°) for (II).

C2*A*—N1*A*—C7*A*—C8*A*	74.2 (4)	C2*B*—N1*B*—C7*B*—C8*B*	−80.8 (4)
N1*A*—C7*A*—C8*A*—C9*A*	−178.1 (3)	N1*B*—C7*B*—C8*B*—C9*B*	−179.9 (3)
C7*A*—C8*A*—C9*A*—N10*A*	159.2 (3)	C7*B*—C8*B*—C9*B*—N10*B*	−160.2 (4)
C8*A*—C9*A*—N10*A*—C10*A*	177.3 (3)	C8*B*—C9*B*—N10*B*—C10*B*	−179.8 (3)
C9*A*—N10*A*—C10*A*—C11*A*	157.7 (4)	C9*B*—N10*B*—C10*B*—C11*B*	−67.9 (5)
N10*A*—C10*A*—C11*A*—O11*A*	−62.5 (5)	N10*B*—C10*B*—C11*B*—O11*B*	−59.8 (5)

Data collection

Bruker SMART APEX diffractometer
ω scans
Absorption correction: empirical (using intensity measurements; *SADABS*; Sheldrick, 1996)
T_{min} = 0.827, T_{max} = 0.993
3919 measured reflections
2336 independent reflections

2284 reflections with $I > 2\sigma(I)$
R_{int} = 0.022
θ_{max} = 25.7°
$h = -5 \rightarrow 4$
$k = -10 \rightarrow 13$
$l = -14 \rightarrow 14$

Refinement

Refinement on F^2
$R[F^2 > 2\sigma(F^2)]$ = 0.048
$wR(F^2)$ = 0.114
S = 1.17
2336 reflections
377 parameters
H atoms treated by a mixture of independent and constrained refinement

$w = 1/[\sigma^2(F_o^2) + (0.0615P)^2 + 0.1838P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$ = 0.007
$\Delta\rho_{max}$ = 0.32 e Å^{-3}
$\Delta\rho_{min}$ = −0.25 e Å^{-3}

Compound (III)

Crystal data

$C_{13}H_{21}N_5O_4$
M_r = 311.35
Triclinic, $P\overline{1}$
a = 6.9334 (9) Å
b = 8.2977 (10) Å
c = 13.2491 (16) Å
α = 90.780 (2)°
β = 93.498 (2)°
γ = 107.293 (2)°
V = 726.02 (16) Å^3

Z = 2
D_x = 1.424 Mg m^{-3}
Mo $K\alpha$ radiation
Cell parameters from 2571 reflections
θ = 2.5–26.8°
μ = 0.11 mm^{-1}
T = 100 (2) K
Plate, colorless
0.40 × 0.18 × 0.04 mm

Data collection

Bruker SMART APEX diffractometer
ω scans
Absorption correction: empirical (using intensity measurements; *SADABS*; Sheldrick, 1996)
T_{min} = 0.878, T_{max} = 0.996
7954 measured reflections

2970 independent reflections
2714 reflections with $I > 2\sigma(I)$
R_{int} = 0.025
θ_{max} = 26.4°
$h = -8 \rightarrow 8$
$k = -10 \rightarrow 10$
$l = -16 \rightarrow 16$

Table 4
Hydrogen-bond geometry (Å, °) for (II).

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
N3*A*—H3*A*···O13*A*^{vii}	0.88	1.88	2.747 (4)	167
N4*A*—H4*AA*···O12*A*^{vii}	0.88	1.93	2.805 (5)	170
N4*A*—H4*BA*···O13*B*^{viii}	0.88	2.06	2.928 (5)	168
N10*A*—H10*A*···O9*A*^{ix}	0.88	2.12	2.975 (5)	163
O11*A*—H11*A*···O12*B*	0.84	1.93	2.704 (5)	153
N3*B*—H3*B*···O13*B*^{x}	0.88	1.89	2.759 (4)	168
N4*B*—H4*AB*···O12*B*^{x}	0.88	1.91	2.789 (5)	174
N4*B*—H4*BB*···O13*A*^{iv}	0.88	2.08	2.960 (5)	175
N10*B*—H10*B*···O9*B*^{xi}	0.88	2.20	3.000 (5)	150
O11*B*—H11*B*···O12*A*^{xii}	0.84	2.01	2.759 (4)	148
O1*C*—H1*C*···O11*A*^{xi}	0.93 (5)	1.96 (6)	2.841 (5)	159 (5)
O1*C*—H2*C*···O9*A*	0.88 (6)	2.07 (6)	2.909 (4)	158 (6)
O1*D*—H1*D*···O11*B*	0.91 (8)	2.05 (8)	2.851 (5)	146 (7)
O1*D*—H2*D*···O9*B*^{xi}	0.80 (7)	2.17 (7)	2.951 (4)	164 (7)

Symmetry codes: (iv) $x, y-1, z$; (vii) $x+1, y-1, z$; (viii) $x, y-1, z-1$; (ix) $x+1, y, z$; (x) $x-1, y-1, z$; (xi) $x-1, y, z$; (xii) $x, y, z+1$.

Table 6
Hydrogen-bond geometry (Å, °) for (III).

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
N4*A*—H4*A*1···N3*B*^{x}	0.88	2.08	2.956 (4)	173
N4*A*—H4*A*2···O16*A*^{x}	0.88	2.14	2.857 (3)	138
N10*A*—H10*A*···O2*B*	0.88	2.06	2.907 (3)	161
N14*A*—H14*D*···O15*B*^{xiii}	0.91	1.86	2.769 (4)	177
N14*A*—H14*E*···O9*A*^{xi}	0.91	1.88	2.787 (4)	176
N14*A*—H14*C*···O16*A*^{x}	0.91	1.97	2.870 (4)	168
N14*A*—H14*C*···O15*A*^{x}	0.91	2.56	3.095 (3)	118
N4*B*—H4*B*1···N3*A*^{xiv}	0.88	2.07	2.939 (4)	171
N4*B*—H4*B*2···O16*B*^{xiv}	0.88	2.15	2.876 (3)	140
N10*B*—H10*C*···O2*A*	0.88	1.95	2.796 (3)	161
N14*B*—H14*I*···O15*A*^{xv}	0.91	1.80	2.694 (3)	165
N14*B*—H14*H*···O9*B*^{ix}	0.91	1.95	2.842 (4)	166
N14*B*—H14*J*···O15*B*^{xiv}	0.91	2.29	2.988 (3)	133
N14*B*—H14*J*···O16*B*^{xiv}	0.91	2.10	2.987 (4)	165

Symmetry codes: (ix) $x+1, y, z$; (x) $x-1, y-1, z$; (xi) $x-1, y, z$; (xiii) $x-1, y, z-1$; (xiv) $x+1, y+1, z$; (xv) $x+1, y, z+1$.

Table 5
Selected torsion angles (°) for (III).

C2*A*—N1*A*—C7*A*—C8*A*	−77.2 (3)	C2*B*—N1*B*—C7*B*—C8*B*	75.1 (3)
N1*A*—C7*A*—C8*A*—C9*A*	−57.6 (3)	N1*B*—C7*B*—C8*B*—C9*B*	56.5 (3)
C7*A*—C8*A*—C9*A*—N10*A*	133.1 (3)	C7*B*—C8*B*—C9*B*—N10*B*	−124.6 (3)
C8*A*—C9*A*—N10*A*—C10*A*	−176.5 (2)	C8*B*—C9*B*—N10*B*—C10*B*	−179.7 (2)
C9*A*—N10*A*—C10*A*—C11*A*	117.0 (3)	C9*B*—N10*B*—C10*B*—C11*B*	−44.8 (4)
N10*A*—C10*A*—C11*A*—C12*A*	−67.0 (3)	N10*B*—C10*B*—C11*B*—C12*B*	−58.1 (3)
C10*A*—C11*A*—C12*A*—C13*A*	−77.3 (3)	C10*B*—C11*B*—C12*B*—C13*B*	176.0 (2)
C11*A*—C12*A*—C13*A*—C14*A*	−72.9 (3)	C11*B*—C12*B*—C13*B*—C14*B*	75.1 (3)
C12*A*—C13*A*—C14*A*—N14*A*	−62.3 (3)	C12*B*—C13*B*—C14*B*—N14*B*	73.6 (3)

Table 7
The geometrical classification for the side-chain position of cytosine hybrid dipeptides.

Compound	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Amino acid	L-Phe	L-Ser	L-Lys	L-Tyr	L-Thr	L-Trp	L-Ile	L-Ala
Position	*O*+	*S*±	*O*±	*O*+	*P*±	*S*±	*S*−	*P*±

Refinement

Refinement on F^2	$w = 1/[\sigma^2(F_o^2) + (0.0707P)^2 + 0.0541P]$ where $P = (F_o^2 + 2F_c^2)/3$
$R[F^2 > 2\sigma(F^2)] = 0.038$	$(\Delta/\sigma)_{max} < 0.001$
$wR(F^2) = 0.100$	$\Delta\rho_{max} = 0.34$ e Å^{-3}
$S = 1.00$	$\Delta\rho_{min} = -0.19$ e Å^{-3}
2970 reflections	
399 parameters	
H-atom parameters constrained	

H atoms of hydrated water molecules of (II) were found from a difference Fourier map by considering hydrogen bonds. These H atoms were not restrained during refinement. All other H atoms were treated as riding atoms, with C—H distances of 0.95–1.00 Å, N—H distances of 0.88 (CONH) or 0.91 Å (NH_3), and O—H distances of 0.84 Å. In the absence of any significant anomalous scattering, the Flack (1983) parameters were meaningless (Flack & Bernardinelli, 2000). Hence, the Friedel pairs were merged prior to the final refinements, and the absolute structures were set by reference to the known chirality of the amino acid employed. A validation check suggested pseudosymmetry for (III), but the material chirality (L-Lys) was confirmed in the structure.

For all compounds, data collection: *SMART* (Bruker, 1998); cell refinement: *SMART*; data reduction: *SAINT-Plus* (Bruker, 1998); program(s) used to solve structure: *SHELXD* (Sheldrick, 1990*b*) for (I), and *SHELXS97* (Sheldrick, 1990*a*) for (II) and (III); program(s) used to refine structure: *SHELXL97* (Sheldrick, 1997); molecular graphics: *PLATON* (Spek, 2001) and *MERCURY* (Bruno *et al.*, 2002).

This study was partially supported by a Grand-in-Aid for High Technology Research from the Ministry of Education, Science and Culture, Japan.

Supplementary data for this paper are available from the IUCr electronic archives (Reference: GG1264). Services for accessing these data are described at the back of the journal.

References

Asano, A., Minoura, K., Ishida, T., Tarui, M. & Doi, M. (2002). *Chem. Lett.* **11**, 1136–1137.
Bruker (1998). *SAINT-Plus* (Version 5) and *SMART* (Version 5). Bruker AXS Inc., Madison, Wisconsin, USA.
Bruno, I. J., Cole, J. C., Edgington, P. R., Kessler, M. K., Macrae, C. F., McCabe, P., Pearson, J. & Taylor, R. (2002). *Acta Cryst.* B**58**, 389–397.
Doi, M., Asano, A. & Ishida, T. (1999). *Anal. Sci.* **15**, 1289.
Doi, M., Miyako, H., Asano, A. & Ishida, T. (1999). *Anal. Sci.* **15**, 109–110.
Doi, M., Tarui, M., Nomoto, N. & Ishida, T. (1999). *Anal. Sci.* **15**, 713–714.
Doi, M., Tarui, M., Ogata, M., Asano, A. & Ishida, T. (1998). *Acta Cryst.* C**54**, 1941–1943.
Doi, M., Tsunemichi, K., Asano, A., Tarui, M., Ishida, T. & Shiono, M. (1999). *Acta Cryst.* C**55**, 1865–1867.
Flack, H. D. (1983). *Acta Cryst.* A**39**, 876–881.
Flack, H. D. & Bernardinelli, G. (2000). *J. Appl. Cryst.* **33**, 1143–1148.
Rotkiewicz, P. (2004). *iMol*. URL: http://www.pirx.com.
Sheldrick, G. M. (1990*a*). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. M. (1990*b*). *SHELXD*. University of Göttingen, Germany.
Sheldrick, G. M. (1996). *SADABS*. University of Göttingen, Germany.
Sheldrick, G. M. (1997). *SHELXL97*. University of Göttingen, Germany.
Spek, A. L. (2001) *PLATON*. Utrecht University, The Netherlands.
Tarui, M., Nomoto, N., Hasegawa, Y., Minoura, K., Doi, M. & Ishida, T. (1996). *Chem. Pharm. Bull.* **44**, 1998–2002.
Voet, D. (1980). *J. Am. Chem. Soc.* **102**, 2071–2074.
Williams, G. J. B., Varghese, A. J. & Berman, H. M. (1977). *J. Am. Chem. Soc.* **99**, 3150–3154.