Crystal Structure of Hybrid Dipeptide, (2-Carboxyethyl)-cytosine-1-yl-L-threonine Monohydrate

Mitsunobu DOI, Akiko ASANO and Toshimasa ISHIDA

Osaka University of Pharmaceutical Sciences, 4-20-1 Nasahara, Takatsuki, Osaka 569-1094, Japan

A hybrid molecule, (2-carboxyethyl)-cytosine-1-yl-L-threonine (C-Thr), was designed in a series of studies to investigate interactions between nucleic acids and peptides.¹⁻³ In the molecular recognition of protein for DNA or RNA, cooperation between the amino acids and nucleic bases is important in the local contact region. This dipeptide is a useful probe to investigate the function of the hydroxyl group of the amino acid; a structural analysis supports our interpretation of the functional group of the amino acids compared with the related peptides, *e.g.* (2-carboxyethyl)-cytosine-1-yl-L-threonine-L-tryptophan (C-Trp)² and L-tyrosine (C-Tyr).³ C-Thr (Fig. 1) was synthesized by coupling 1-(2-carboxyethyl)-*N*-(4,4-dimethoxytriphenyl)-cytosine and L-threonine benzyl ester using the conventional liquid-phase method. The benzyl group was hydrolyzed by 1 M NaOH, and the remaining protect group was treated with acetic acid. The residual product was extracted

using diluted acetic acid and diethylether. The monohydrate crystals were grown from an aqueous methanol solution. The crystal mounted on a glass fiber was stable during data collection, and the decay value, calculated from three standard reflections, was 5.3%. The crystal and experimental data are listed in Table 1. The structure was solved by a direct method using SHELXS-97 and was refined by a full-matrix least-squares method using SHELXL-97. Non-hydrogen atoms were refined anisotropically and the hydrogen atoms, calculated at idealized positions, were included in the calculation of the structure factors. The structure of C-Thr is shown in Fig. 2, and the final coordinates are listed in Table 2. The possible hydrogen bonds are listed in Table 3.

There are two independent C-Thr and two water molecules in the asymmetric unit (the respective molecules are represented as Molecules 1 and 2, or as a suffix of

Table 1 Crystal and experimental data

Formula: C ₁₁ H ₁₆ N ₄ O ₅ ·H ₂ O
Formula weight=302.29
Crystal system: triclinic
Space group: <i>P</i> ₁ <i>Z</i> =2
<i>a</i> =9.165(5) Å
<i>b</i> =9.942(4) Å
<i>c</i> =7.993(6) Å
α =110.73(4)°
β =102.44(5)°
γ =88.96(4)°
<i>V</i> =663.8(7) Å ³
<i>D</i> _x =1.512 g/cm ³
<i>F</i> (0 0 0)=320
μ (Cu K α)=1.060 mm ⁻¹
θ _{max} =65.34°
No. of reflections used=4490
No. of parameters=382
<i>R</i> =0.0670
<i>wR</i> =0.1845
(Δ / σ) _{max} =0.007
$\Delta\rho$ _{max} =0.436 e Å ⁻³
$\Delta\rho$ _{min} =-0.409 e Å ⁻³
Measurement: Rigaku AFC5R/RU200
Program system: ASC/AFC Diffractometer Control Software
Structure determination: direct method (SHELXS-97)
Refinement: full matrix least squares (SHELXL-97)

Fig. 1 Chemical structure of C-Thr.

Fig. 2 ORTEP drawing of C-Thr with atom labelings. The italic numbers represents independent Molecules 1 and 2, respectively. The thin line represents a hydrogen bond. The symmetry-translated cytosine base of Molecule 1 and O1W are illustrated by small balls and thin lines to show the interaction around the O12 atoms. The small italic labels represent symmetry-translated atoms.

Table 2 Final atomic coordinates and equivalent isotropic thermal parameters for non-hydrogen atoms

Molecule 1					Molecule 2			
Atom	x	y	z	$U_{eq}/\text{\AA}^2$	x	y	z	$U_{eq}/\text{\AA}^2$
N1	0.7010(3)	0.5162(3)	0.2579(4)	0.0292(6)	-0.0041(3)	0.3701(3)	-0.0223(4)	0.0295(6)
C2	0.6231(4)	0.3935(4)	0.1244(4)	0.0287(7)	0.0774(4)	0.4932(4)	0.1040(5)	0.0297(7)
O2	0.5166(3)	0.3981(3)	0.0058(3)	0.0356(6)	0.1851(3)	0.4898(3)	0.2206(4)	0.0424(7)
N3	0.6724(3)	0.2646(3)	0.1330(4)	0.0300(6)	0.0289(3)	0.6211(3)	0.0862(4)	0.0329(7)
C4	0.7878(4)	0.2513(4)	0.2632(5)	0.0304(7)	-0.0915(4)	0.6321(4)	-0.0392(5)	0.0313(8)
N4	0.8181(4)	0.1226(3)	0.2671(5)	0.0377(7)	-0.1181(4)	0.7575(4)	-0.0522(5)	0.0419(8)
C5	0.8707(4)	0.3790(4)	0.3897(5)	0.0328(8)	-0.1806(4)	0.5034(4)	-0.1517(5)	0.0330(8)
C6	0.8230(4)	0.5068(4)	0.3846(5)	0.0324(8)	-0.1323(4)	0.3786(4)	-0.1417(5)	0.0321(8)
C7	0.6460(4)	0.6582(4)	0.2633(5)	0.0315(8)	0.0503(4)	0.2292(4)	-0.0244(5)	0.0322(8)
C8	0.5021(4)	0.6870(4)	0.3297(5)	0.0323(7)	0.1872(4)	0.1907(4)	-0.1063(5)	0.0314(7)
C9	0.5213(4)	0.7299(3)	0.5356(4)	0.0281(7)	0.1574(4)	0.1680(3)	-0.3087(4)	0.0272(7)
O9	0.6428(3)	0.7310(3)	0.6371(3)	0.0376(6)	0.0330(3)	0.1867(3)	-0.3918(3)	0.0380(6)
N10	0.3945(3)	0.7670(3)	0.5924(4)	0.0300(6)	0.2709(3)	0.1295(3)	-0.3915(4)	0.0296(6)
C10	0.3793(4)	0.8108(4)	0.7820(4)	0.0295(7)	0.2516(4)	0.1134(4)	-0.5829(5)	0.0310(7)
C11	0.5038(4)	0.9243(4)	0.9137(5)	0.0289(7)	0.1771(4)	-0.0333(4)	-0.7186(5)	0.0322(8)
O10	0.5248(3)	1.0300(3)	0.8665(4)	0.0390(6)	0.1612(3)	-0.1319(3)	-0.6580(4)	0.0398(6)
O11	0.5701(3)	0.9129(3)	1.0607(4)	0.0375(6)	0.1357(4)	-0.0435(3)	-0.8806(4)	0.0508(8)
C12	0.3550(4)	0.6836(4)	0.8404(5)	0.0332(8)	0.3987(4)	0.1442(4)	-0.6282(5)	0.0361(8)
O12	0.4801(3)	0.5982(3)	0.8237(4)	0.0396(6)	0.4981(3)	0.0359(3)	-0.6125(4)	0.0461(7)
C13	0.2110(5)	0.5950(5)	0.7257(6)	0.046(1)	0.4700(7)	0.2916(5)	-0.5070(7)	0.060(1)
O1W	0.7427(4)	1.0379(4)	0.6594(5)	0.0560(8)				
O2W	0.9808(5)	0.8447(4)	0.5855(6)	0.067(1)				

$$U_{eq} = (1/3) \sum_i \sum_j U_{ij} (a_i^* a_j^*) (a_i a_j).$$

Table 3 Possible hydrogen bonds

D	H	A	D...A(Å)	H...A(Å)	$\angle D-H...A(^{\circ})$
N3_1	H3_1	O10_1 ⁱ	2.676(4)	1.818	176.2
N4_1	H4A_1	O11_1 ⁱ	2.890(4)	2.087	155.1
N4_1	H4B_1	O9_2 ⁱⁱ	2.868(5)	2.078	152.4
N10_1	H10_1	O10_2 ⁱⁱⁱ	3.016(5)	2.231	151.7
O12_1	H12_1	O2_1 ^{iv}	2.829(5)	1.859	178.7
O1W	H11W	O10_1	2.873(6)	1.937	178.1
O1W	H12W	O12_2 ⁱⁱⁱ	2.760(5)	1.801	178.0
N3_2	H3_2	O10_2 ⁱⁱⁱ	2.678(4)	1.827	170.1
N4_2	H4A_2	O11_2 ⁱⁱⁱ	2.842(5)	2.035	155.9
N4_2	H4B_2	O9_1 ^v	2.872(5)	2.037	163.6
N10_2	H10_2	O10_1 ⁱⁱ	3.153(5)	2.336	158.8
O12_2	H12_2	O11_1 ^{vi}	2.684(4)	1.730	179.6
O2W	H21W	O1W	2.898(6)	1.939	178.9
O2W	H22W	O10_2 ^{vii}	2.880(7)	1.940	178.1

i) x,y-1,z-1, ii) x+1,y,z+1, iii) x,y+1,z+1, iv) x,y,z+1, v) x-1,y,z-1, vi) x,y-1,z-2, vii) x+1,y+1,z+1

the molecule number in this paper). The conformations of C-Thr molecules seem to be similar to each other from the figure, and the molecules are approximately related by the pseudo centro-symmetry. However, the N1-C7 rotation of the cytosine bases are opposite in Molecules **1** and **2**: C2-N1-C7-C8 (Molecule **1/2**)=71.2(4)/-73.7(4)°. In the amino acid moiety, the bonds of N10-C10 and C10-C11 are also rotated in a different way: C9-N10-C10-C11=47.5(5)/-82.6(4)° and N10-C10-C11-O10=49.0(4)/-10.7(5)°.

The distances of N3...O10 and N4...O11 suggest an interaction between the imino/amino and the carboxyl groups (Table 3). In consideration of the donor-acceptor combination, the bond lengths of C11-O10 and C11-O11 suggest ionization of the carboxyl group: C11-O10=1.266(6)/1.260(6) Å and C11-O11=1.244(4)/1.236(5) Å. The bond angles around N3, which are larger than

the ideal angle, also support protonation at the N3 position; $\angle C2-N3-C4=124.5(3)/125.0(3)^{\circ}$. Therefore, the molecules of C-Thr are zwitterionic and the hydrogen bonds of N3-H...O10 and N4-H...O11 are formed between the C-Thr molecules (Table 3). This interaction mode is commonly observed in hybrid peptides.^{2,3} The threonine side chain of Molecule **1**, a secondary hydroxyl group (O12_1), is hydrogen-bonded to the O2 atom of the neighboring cytosine base (Fig. 2). Contrary to this, the hydroxyl group of Molecule **2** interacts with O11 of the carboxyl group of Molecule **1** and O1W (Table 3). In the crystal of C-Tyr, which has a phenolic hydroxyl group, an interaction between the cytosine base and the side chain was observed.³ Structural studies of related compounds suggest that only the hydroxyl group of the amino acid side chain directly interacts with the cytosine base.

This study was partially supported by the Grant-in-Aid for Scientific Research 07672427 from the Ministry of Education, Science, Sports and Culture.

References

1. T. Ishida, H. Iyo, H. Ueda, M. Doi and M. Inoue, *J. Chem. Soc., Chem. Commun.*, 297 (1990).
2. M. Doi, M. Tarui, M. Ogata, A. Asano and T. Ishida, *Acta Crystallogr.*, **C54**, 1941 (1998).
3. M. Doi, H. Miyako, A. Asano and T. Ishida, *Anal. Sci.*, **15**, 109 (1999).

(Received July 19, 1999)

(Accepted November 4, 1999)