Interaction of Indole Derivatives with Biologically Important Aromatic Compounds. Part 22.[1] Importance of Simultaneous Co-operation of Hydrogen-Bond Pairing and Stacking Interactions for Recognition of Guanine Base by a Peptide: X-Ray Crystal Analysis of 7-Methylguanosine-5'-phosphate–Tryptophanylglutamic Acid Complex

Toshimasa Ishida,*[,a] Hiromi Iyo,[a] Hitoshi Ueda,[a] Mitsunobu Doi,[a] Masatoshi Inoue,[a] Susumu Nishimura[b] and Kunihiro Kitamura[c]

[a] *Department of Physical Chemistry, Osaka University of Pharmaceutical Sciences, 2-10-65 Kawai, Matsubara Osaka 580, Japan*

[b] *Biology Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Tokyo 104, Japan*

[c] *Research Centre, Taisho Pharmaceutical Co. Ltd, 1-403 Yoshino-cho, Ohmiya, Saitama 330, Japan*

As a model to investigate the mode of recognition of the base guanine by peptides and proteins, the crystal structure of the 7-methylguanosine-5'-phosphate–tryptophanylglutamic acid complex was analysed by X-ray diffraction; this is the first crystal structure determination of a peptide–nucleotide complex. The complex crystals are stabilized by extensive hydrogen-bond formation in which three independent water molecules per complex pair participate. Both molecules are joined by the coupled contributions of the triple hydrogen bonds between the guanine base and the peptide backbone chain and of the prominent stacking interactions between the guanine base and the tryptophan indole side-chain, suggesting the importance of the coupling of hydrogen bonding and stacking interactions for recognition of the base to occur.

The ability of a protein selectively to recognize a specific nucleic acid base or base sequence of DNA or RNA is a fundamental biological process. Precise molecular recognition requires the strict structural complementarity of both the macromolecules, and formation of specific interactions between their constituent chemical groups.

As observed in the tertiary structure of double-stranded DNA or RNA, hydrogen bonding and aromatic stacking interactions could function as principal forces for such a selective binding, because they have the ability to select the most suitable partner among many possible candidates. This is, in fact, evidenced by the design of model receptors which selectively recognize a nucleic acid base by virtue of both the hydrogen bond and stacking interactions.[2]

In contrast with these synthetic compounds, there are very few reports concerning natural peptides which show high selectivity for binding with a nucleic acid base.[3] In particular, clear-cut presentations in the crystalline state are lacking, although such information is very important to our understanding of the essential forces and the mechanism for molecular recognition.

As a model for studying the selective recognition mechanism of the mRNA capped structure by cap-binding protein (termed initiation factor 4E),[4] we recently examined the interaction of a series of tryptophan-containing peptides with guanine nucleotides in the solution state by spectroscopic methods,[5] where coupling of hydrogen bonding and stacking interactions was suggested in order to account for the tight binding between both molecules. In efforts to dissect the characteristics of these interactions in the crystalline state, co-crystals of 7-methylguanosine-5'-phosphate (m7GMP) and tryptophanylglutamic acid (Trp-Glu) complex were obtained. This paper deals with the crystal and molecular structure as the first reported X-ray analysis of a peptide–nucleotide complex; preliminary results were reported previously.[6]

Experimental

Syntheses and Crystallization.—m7GMP formate was chemically synthesized from GMP according to the procedure described previously.[7] The Trp-Glu was synthesized from L-tryptophan and L-glutamic acid by the usual liquid-phase peptide condensation.

Equimolar quantities of m7GMP formate and Trp-Glu were dissolved in 30% aq. methanol. Transparent platelet crystals were obtained from the solution by slow evaporation at 20 °C. Thermal analysis of the crystals and the UV spectra of crystals dissolved in 50% aq. ethanol suggested that the crystals consist of an equimolar ratio of the respective components and contain three molecules of water of crystallization per complex pair.

Crystal Data and Intensity Collection.—Details of crystallographic data and the intensity-data-collection parameters are summarized in Table 1. The unit-cell dimensions were determined by a least-squares fit of 2θ angles for 25 reflections ($30° < 2\theta < 60°$) measured by graphite-monochromated Cu-Kα radiation (λ 1.5418 Å) on an automated Rigaku AFC-5 diffractometer. To prevent the release of water of crystallization during data collection, a single crystal with approximate dimensions 0.2 × 0.3 × 0.3 mm was sealed in a glass capillary containing some mother liquor. Intensity data were collected with the same diffractometer by employing an ω–2θ scanning technique, where the background was counted for 5 s at both extremes of the peak. Four standard reflections were monitored for every 100 reflection intervals and showed no significant time dependence. The observed intensities were corrected for Lorentz and polarization effects. No corrections for absorption or extinction effects were made.

Structure Solution and Refinement.—The structure was solved by direct methods by using the MULTAN87 program.[8] The positional parameters of the non-hydrogen atoms obtained were refined by full-matrix least-squares with isotropic thermal parameters and then by block-diagonal least-squares with anisotropic parameters. We calculated the positions of the geometrically reasonable hydrogen atoms, all of which, except for the ribose O(3'), phosphate O(1P), and the water O(3), were also verified on a difference Fourier map, and were included in

Table 1 Summary of crystal data and intensity collection details

Formula	$C_{11}H_{16}N_5O_8P \cdot C_{16}H_{19}N_3O_5 \cdot 3H_2O$
M_r	764.64
Space group	$P2_1$
a (Å)	6.558(1)
b (Å)	16.460(3)
c (Å)	15.395(4)
β (°)	96.70(2)
V (Å^3)	1650.5(5)
Z	2
D_m (g cm^{-3})	1.524(1)
D_c (g cm^{-3})	1.539
μ (cm^{-1})	14.89
F(000)	804
T of data collection (°C)	20
Scan speed in 2θ (° min^{-1})	3
Scan range in ω (°)	1.8 + 0.15tanθ
Data range measured (°)	$2 \leqslant \theta \leqslant 130°$
Data collected	$h, k, \pm l$
No. of unique data measured (M)	2924 [2882 for $F_o > \sigma(F_o)$, 2771 for $F_o > 3\sigma(F_o)$]
No. of variables (N)	634
R	0.048
R_w	0.064
S	0.996

subsequent refinements with isotropic thermal factors. The function minimized was $\Sigma w(|F_o| - |F_c|)^2$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factors, respectively. The weighting scheme used for the final refinement was w 0.264 43 $|F_o|$ 0.0 and $w = 1.0/[\sigma(F_o)^2 - 0.003\,99|F_o| + 0.001\,94|F_o|^2]$ for $|F_o| > 0.0$, where $\sigma(F_o)^2$ is the standard deviation of the reflection intensity on the basis of counting statistics. Final R $(= \Sigma(|F_o| - |F_c|)/\Sigma|F_o|)$, R_w $(= [\Sigma w(|F_o| - |F_c|)^2/w|F_o|^2]^{\frac{1}{2}})$, and S $(= [\Sigma w(|F_o| - |F_c|)^2/(M - N]^{\frac{1}{2}})$ are also given in Table 1. None of the positional parameters for non-hydrogen atoms shifted by more than one-fifth from their estimated standard deviations, and the residual electron density in the final difference Fourier map ranged from −0.20 to 0.25 e Å^{-3}. The final positional parameters for non-hydrogen atoms are listed in Table 2.* For all crystallographic computations, the UNICS system[9] was used, and the atomic scattering factors and terms of anomalous dispersion corrections were taken from ref. 10.

Molecular Orbital Calculations.—The quantum chemical parameters such as atomic charges and dipole moments were obtained by MNDO calculations.[11] The atomic co-ordinates used were derived from Table 2. The numerical calculations were performed on a MicroVAX II computer at the Computation Centre, Osaka University of Pharmaceutical Sciences.

Results and Discussion

Molecular Conformations and Dimensions.—Stereoscopic views of m7GMP and Trp-Glu molecules are present in Fig. 1, together with the atomic numbering used in this paper. The conformational torsion angles are given in Table 3. The atomic deviations (in angstroms) from their respective best planes, together with the dihedral angles between them, are deposited as supplementary material.

The Trp-Glu molecule exists as a zwitterion with the *N*-terminal end protonated and the *C*-terminal end deprotonated; the carboxy group of the Glu side-chain is in a neutral state. The (φ,ψ) backbone conformation is in the allowed torsion-angle range[12] and is similar to that in Tyr-Glu dipeptide,[13] where the ψ1, ω1 and φ2 torsion angles are 159.6°, −178.9°, and −74.1°, respectively. The conformation of the Trp side-chain is similar to those of L-tryptophan hydrochloride[14] and *N*-acetyl-L-tryptophan complexed with 1-methyl-3-carbamoylpyridinium.[15] The χ^1 and χ_1^2 torsion angles of 78.4° and 87.5° put this Trp side-

Table 2 Final atomic co-ordinates with e.s.d.s in parentheses[a]

Atom	x	y	z
m7GMP			
N(1)	0.903 2(5)	0.408 7(2)	0.814 2(2)
C(2)	0.896 3(6)	0.338 5(3)	0.766 6(2)
N(2)	0.891 3(6)	0.268 4(2)	0.808 0(2)
N(3)	0.897 3(5)	0.337 6(2)	0.679 6(2)
C(4)	0.893 4(5)	0.414 0(2)	0.645 5(3)
C(5)	0.895 6(6)	0.486 2(3)	0.689 1(3)
C(6)	0.901 8(5)	0.487 8(3)	0.781 8(3)
O(6)	0.904 8(5)	0.546 6(2)	0.830 3(2)
N(7)	0.884 8(5)	0.547 1(2)	0.626 7(2)
C(7m)	0.881 5(7)	0.635 1(3)	0.643 3(3)
C(8)	0.878 6(6)	0.513 6(3)	0.549 1(3)
N(9)	0.883 0(5)	0.432 2(2)	0.557 2(2)
C(1′)	0.884 4(6)	0.375 1(3)	0.481 7(2)
C(2′)	0.710 9(6)	0.312 9(2)	0.482 8(3)
O(2′)	0.756 6(5)	0.243 5(2)	0.434 1(2)
C(3′)	0.533 1(6)	0.357 6(3)	0.430 8(3)
O(3′)	0.379 5(6)	0.304 4(3)	0.392 6(3)
C(4′)	0.637 9(7)	0.405 1(3)	0.363 3(3)
O(4′)	0.844 4(5)	0.420 7(2)	0.405 0(2)
C(5′)	0.538 6(9)	0.482 6(3)	0.330 6(3)
O(5′)	0.483 3(6)	0.529 3(2)	0.403 7(2)
P(1)	0.366 8(2)	0.613 90(8)	0.398 84(7)
O(1p)	0.534 8(6)	0.676 8(3)	0.426 7(3)
O(2p)	0.206 2(5)	0.618 1(2)	0.451 5(2)
O(3p)	0.298 6(6)	0.625 5(2)	0.295 7(2)
Trp-Glu			
N(1)	0.735 6(6)	0.568 0(2)	0.994 7(2)
C(a)α	0.616 8(6)	0.500 9(3)	1.031 7(3)
C(1′)	0.752 3(6)	0.425 1(3)	1.035 0(2)
O(1)	0.901 6(4)	0.423 2(2)	0.992 9(2)
C(1)β	0.405 4(6)	0.486 5(3)	0.978 1(3)
N(1)ε^1	0.390 0(6)	0.333 1(2)	0.802 6(2)
C(1)δ^1	0.395 8(7)	0.356 7(3)	0.888 2(3)
C(1)γ	0.405 6(6)	0.439 3(3)	0.894 6(3)
C(1)ε^3	0.402 6(6)	0.547 1(3)	0.771 0(3)
C(1)ζ^3	0.391 3(6)	0.554 5(3)	0.680 1(3)
C(1)η^2	0.383 4(6)	0.485 7(4)	0.626 6(3)
C(1)ζ^2	0.386 7(6)	0.408 8(3)	0.659 7(3)
C(1)ε^2	0.395 8(6)	0.410 0(3)	0.751 0(3)
C(1)δ^2	0.404 4(5)	0.469 0(3)	0.806 9(3)
N(2)	0.696 4(5)	0.364 6(2)	1.081 1(2)
C(2)α	0.807 0(6)	0.287 7(2)	1.087 8(3)
C(2′)	0.678 7(6)	0.223 0(3)	1.034 1(3)
O(2)	0.742 3(6)	0.196 5(3)	0.968 3(2)
O(2′)	0.510 6(5)	0.202 1(2)	1.060 8(2)
C(2)β	0.865 0(6)	0.262 5(3)	1.183 2(3)
C(2)γ	1.028 3(6)	0.319 2(3)	1.228 7(3)
C(2)δ	1.226 4(6)	0.312 7(3)	1.188 2(3)
O(2)ε^1	1.286 7(5)	0.237 1(2)	1.180 8(3)
O(2′)ε^2	1.320 2(5)	0.370 3(2)	1.1647(2)
Water			
O(1)	0.551 4(6)	0.654 3(3)	0.170 0(3)
O(2)	0.873 9(6)	0.693 5(3)	0.361 4(4)
O(3)	0.947 1(9)	0.572 6(4)	0.169 9(4)

[a] The atomic numbering corresponds to that shown in Fig. 1.

* *Supplementary data* (see section 5.6.3 of Instructions for Authors, in the January issue). Anisotropic temperature factors of non-hydrogen atoms, full lists of atomic co-ordinates and thermal parameters for H- and non-H-atoms, bond lengths and angles and least-squares best planes have been deposited at the Cambridge Crystallographic Data Centre.

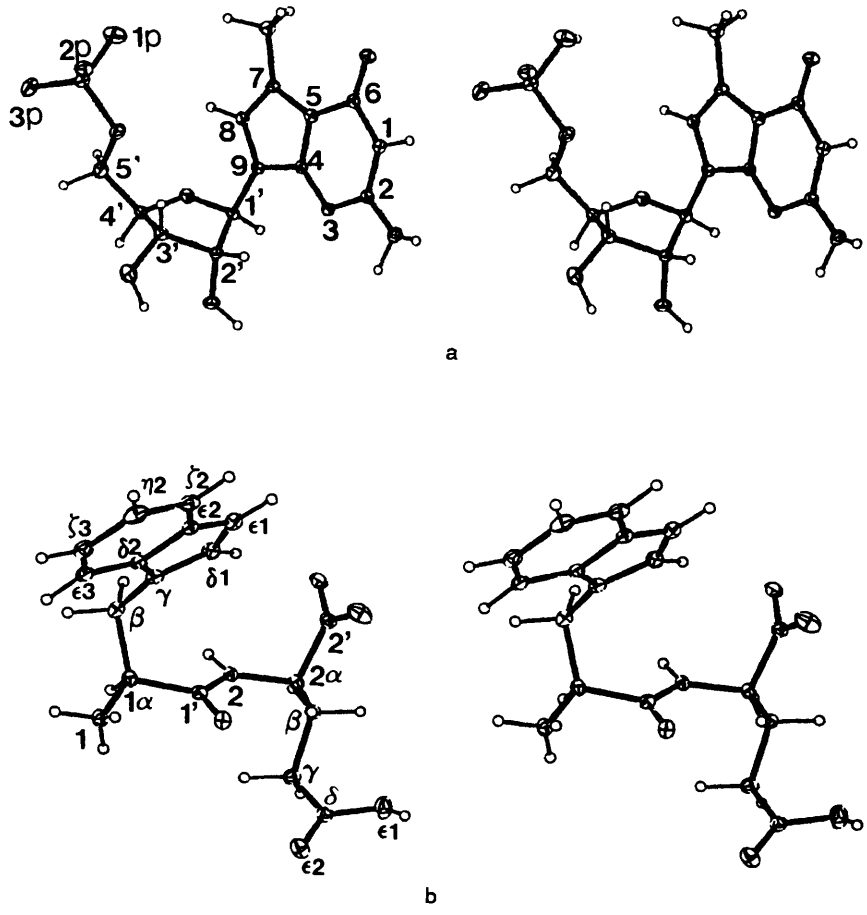


Fig. 1 Stereoscopic views of Trp-Glu (a) and m7GMP (b) molecule conformations, together with atomic numberings used in this paper

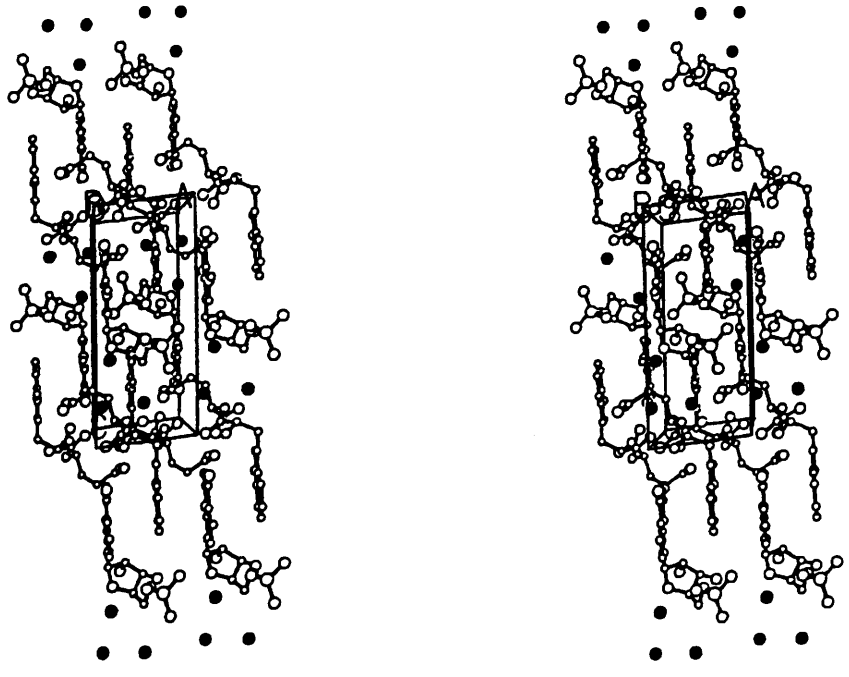


Fig. 2 A stereoscopic view of the crystal packing of the Tr-Glu–m7GMP complex. Filled circles represent water molecules of crystallization.

chain into one of the most frequently observed conformations in peptides and proteins.[16] The indole ring is planar within ±0.03 Å. The χ_2^1 and χ_2^2 torsion angles of the Glu side-chain fall in an energetically favourable range (−60°, −60°),[17] although the χ_2^3 angle of −50.0° deviates slightly from the usual range of ±15°.[18] The carboxy group of the Glu side-chain is in a neutral state and is almost planar.

The molecular conformation of m7GMP is *anti* about the glycosyl N(9)–C(1′) bond (χ_{CN} 172.5°), ribose puckering is C(3′)-*endo*-C(2′)-*exo* (ν_{max} 34.5°, P −0.1°), *gauche*,*gauche* about

Table 3 Conformational torsional angles of Trp-Glu and m7GMP molecules with e.s.d.s in parentheses

Bond sequence	Torsion angle (°)
Trp-Glu	
N(1)–C(1)α–C(1′)–N(2):ψ_1	165.3(4)
C(1)α–C(1′)–N(2)–C(2)α:ω_1	176.9(4)
N(1)–C(1)α–C(1)β–C(1)γ:χ_1^1	78.4(4)
C(1)α–C(1)β–C(1)γ–C(1)δ_1:χ_1^2	87.5(5)
C(1)α–C(1)β–C(1)γ–C(1)δ_2	−96.3(5)
C(1′)–N(2)–C(2)α–C(2′):φ_2	−105.7(4)
N(2)–C(2)α–C(2′)–O(2)	113.1(4)
N(2)–C(2)α–C(2′)–O(2′)	−65.5(4)
N(2)–C(2)α–C(2)β–C(2)γ:χ_2^1	−68.9(3)
C(2)α–C(2)β–C(2)γ–C(2)δ:χ_2^2	−64.4(4)
C(2)β–C(2)γ–C(2)δ–O(2)ε_1:χ_2^3	−50.0(4)
C(2)β–C(2)γ–C(2)δ–O(2)ε_2	129.8(5)
m7GMP	
C(8)–N(9)–C(1′)–C(2′)	−126.9(4)
C(4)–N(9)–C(1′)–C(2′)	55.7(3)
C(8)–N(9)–C(1′)–O(4′)	−10.2(3)
C(4)–N(9)–C(1′)–O(4′):χ_{CN}	172.5(4)
C(4′)–O(4′)–C(1′)–C(2′):ν_0	11.4(3)
O(4′)–C(1′)–C(2′)–C(3′):ν_1	−29.0(3)
C(1′)–C(2′)–C(3′)–C(4′):ν_2	34.5(3)
C(2′)–C(3′)–C(4′)–O(4′):ν_3	−29.0(3)
C(3′)–C(4′)–O(4′)–C(′):ν_4	112.2(3)
P (pseudorotation phase angle)	−0.11 34.5
ν_{max}	
O(3′)–C(3′)–C(4′)–C(5′):δ	86.6(4)
C(3′)–C(4′)–C(5′)–O(5′):γ	46.1(4)
O(4′)–C(4′)–C(5′)–O(5′):γ′	−73.8(4)
C(4′)–C(5′)–O(5′)–P(1):β	−177.7(3)
C(5′)–O(5′)–P(1)–O(1p):α	−110.2(3)

the exocyclic C(4′)–C(5′) bond (γ 46.1°, γ′ −73.8°), and *trans* about the C(5′)–O(5′) bond (β −177.7°). Although this conformation is one of the stable forms judging from the combination of (χ,*P*,γ) torsion angles in many nucleotides,[19] the C(3′)-*endo* sugar puckering found here for the m7GMP molecule is the first to be observed for this molecule; all of the m7GMP crystal structures hitherto reported show the C(2′)-*endo* puckering.[20] The phosphate group is in the monoanionic state, with one of three oxygen atoms protonated, and neutralizes the positive charge of the guanine base caused by the N(7)-methylation. The positive charge of the base makes the *gauche,gauche* conformation about the C(4′)–C(5′) bond more rigid: MNDO calculations showed that the protonation of the guanine N(7) atom increases the electropositivity of the 8-H atom [−0.014 e for guanine and 0.079 e for N(7)-protonated guanine], consequently strengthening the electrostatic interaction between the C(8) and O(5′) atoms [C(8)···O(5′) 3.232(5) Å].

The bond lengths and angles for the m7GMP and Trp-Glu molecules are deposited as supplementary material. The estimated standard deviations for the non-hydrogen atoms range from 0.004 to 0.007 Å for bond lengths and 0.2 to 0.3° for bond angles. Compared with related compounds,[13–15,20] these values are all in the usual range. The bonding parameters of the *N*-terminal C–N bond and *C*-terminal carboxy group of Trp-Glu correspond to their zwitterionic characters, while the carboxy group of the Glu side-chain has bond lengths and angles characteristic for protonation: N(1)–C(1)α 1.501(5) Å, C(2′)–O(2) 1.220(6) Å, C(2′)–O(2′) 1.269(5) Å, C(2)δ–O(2)ε^1 1.315 Å, and C(2)δ–O(2)ε^2 1.208(6) Å. The geometry of m7G is similar, as has been observed for protonated bases.[21] The geometrical data for the phosphate group show the monoanionic form (OPO_3H^-) where O(1) atom is protonated [P(1)–O(1p) 1.569(4) Å, P(1)–O(2p) 1.499(4) Å, and P(1)–O(3p) 1.478 Å];[19] the H-atom was not clearly identified on difference Fourier maps because of the relatively high thermal motion of the phosphate oxygen atoms.

Table 4 Hydrogen bonds and short contacts (<3.3 Å)

[1] Hydrogen bonds[a]

Donor (D)	Acceptor (A)	Symmetry operation[b]	D···A (Å)	H···A (Å)	D–H···A (°)
N(1)G	O(1)	*x*, *y*, *z*	2.762(4)	1.83(6)	164(6)
N(2)G	O(2)	*x*, *y*, *z*	3.001(5)	2.05(7)	160(3)
N(2)G	O(3p)G	1 − *x*, *y* − ½, 1 − *z*	3.028(5)	2.06(7)	179(6)
O(2′)G	O(2p)G	1 − *x*, *y* − ½, 1 − *z*	2.706(5)	1.74(7)	179(6)
O(3′)G	O(2′)G	*x*, *y*, *z*	2.675(5)		
O(1p)G	O(2)W	*x*, *y*, *z*	2.560(7)		
N(1)	O(6)G	*x*, *y*, *z*	2.901(5)	1.97(7)	179(6)
N(1)	O(2′)	1 − *x*, *y* + ½, 2 − *z*	2.809(5)	1.79(7)	156(6)
N(1)	O(3)W	*x*, *y*, *z* + 1	2.886(7)	1.97(7)	177(6)
N(1)ε_1	O(1)W	1 − *x*, *y* − ½, 1 − *z*	2.992(6)	2.20(7)	140(6)
N(2)	O(2)ε_2	*x* − 1, *y*, *z*	2.897(5)	1.92(7)	169(6)
O(2)ε_1	O(2′)	*x* + 1, *y*, *z*	2.555(5)	1.64(7)	176(6)
O(1)W	O(3p)G	*x*, *y*, *z*	2.732(5)	1.74(7)	178(6)
O(1)W	O(2)	1 − *x*, *y* + ½, 1 − *z*	2.787(6)	1.81(7)	179(6)
O(2)W	N(3)G	2 − *x*, *y* + ½, 1 − *z*	2.915(6)	1.92(7)	179(6)
O(2)W	O(2p)G	*x* + 1, *y*, *z*	2.740(6)	1.76(7)	180(6)
O(3)W	O(3p)G	*x* + 1, *y*, *z*	2.963(7)		
O(3)W	O(1)W	*x*, *y*, *z*	2.923(7)		

[2] Short contacts

Atom 1	Atom 2	Symmetry operation[c]	Distance (Å)
C(4)G	C(1)ζ_2	*x* + 1, *y*, *z*	3.218(6)
C(6)G	C(1)δ_2	*x* + 1, *y*, *z*	3.288(6)
O(6)G	O(1)	*x*, *y*, *z*	3.225(4)
C(8)G	O(5′)G	*x*, *y*, *z*	3.232(5)
O(2′)G	O(1p)G	1 − *x*, *y* − ½, 1 − *z*	3.225(5)
O(3′)G	C(7m)G	1 − *x*, *y* − ½, 1 − *z*	3.283(6)
O(3′)G	C(2)γ	*x* − 1, *y*, *z* − 1	3.220(6)
O(3′)G	C(2)δ	*x* − 1, *y*, *z* − 1	3.192(6)
O(1p)G	(C2′)G	1 − *x*, *y* + ½, 1 − *z*	3.174(6)
O(2p)G	C(8)G	*x* − 1, *y*, *z*	3.253(5)
O(3p)G	O(2)W	*x* − 1, *y*, *z*	3.270(6)
N(1)	O(1)	*x*, *y*, *z*	2.621(5)
N(2)	O(2′)	*x*, *y*, *z*	2.943(5)
N(1)ε_1	C(2)G	*x* − 1, *y*, *z*	3.222(5)
O(3)W	C(1)α	*x*, *y*, *z* − 1	3.088(7)

[a] Hydrogen atoms attached to O(3′)G, O(1p)G and O(3)W were not determined unequivocally because of their high thermal motions. The suffic letters G and W represent the m7GMP and water molecules, respectively. [b] This shows the symmetry operation of the acceptor atom with respect to the donor atom at *x*, *y*, *z*. [c] This shows the symmetry operation of atom 2 with respect to atom 1 at *x*, *y*, *z*.

Crystal Packing and Hydrogen Bonds.—A stereoscopic view of the crystal structure is shown in Fig. 2. It consists of alternate layers of m7GMP and Trp-Glu molecules stacked along the *a*-direction. The water molecules of crystallization are located among these layers and stabilize the molecule packing *via* hydrogen-bond formation.

The schematic hydrogen-bonding network formed in the crystal structure is shown in Fig. 3, the parameters for the hydrogen bonds and short contacts are given in Table 4. The short-contact pairs of N(1)···O(1) and N(2)···O(2′) could be characterized as ion pairs rather than as hydrogen bonds. Each of the three independent water molecules participates as

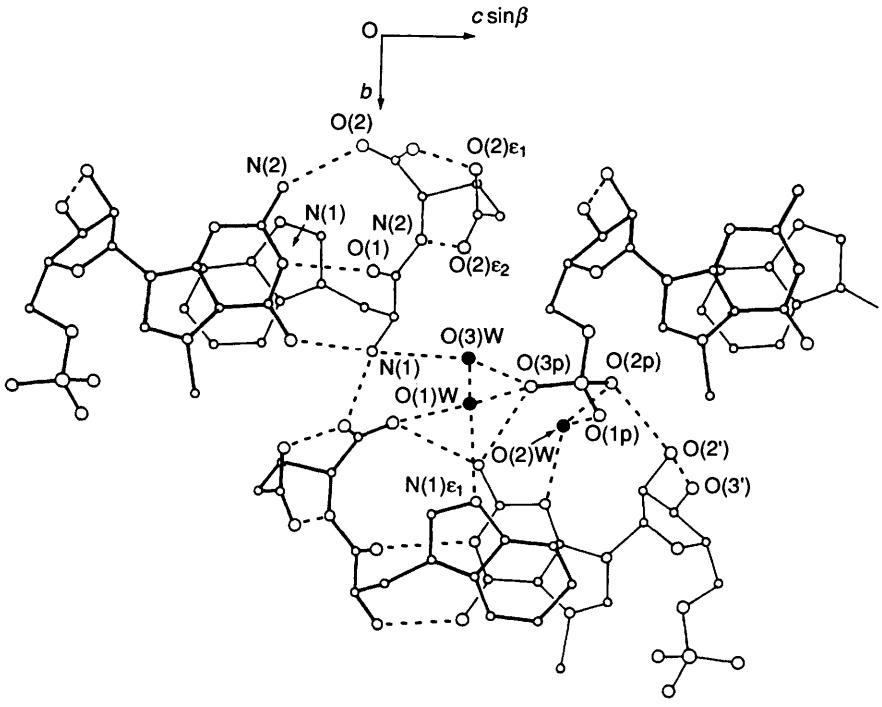


Fig. 3 Schematic hydrogen-bonding network formed in the complex crystal. The broken lines represent possible hydrogen bonds. The water molecules are indicated by the suffix letter W.

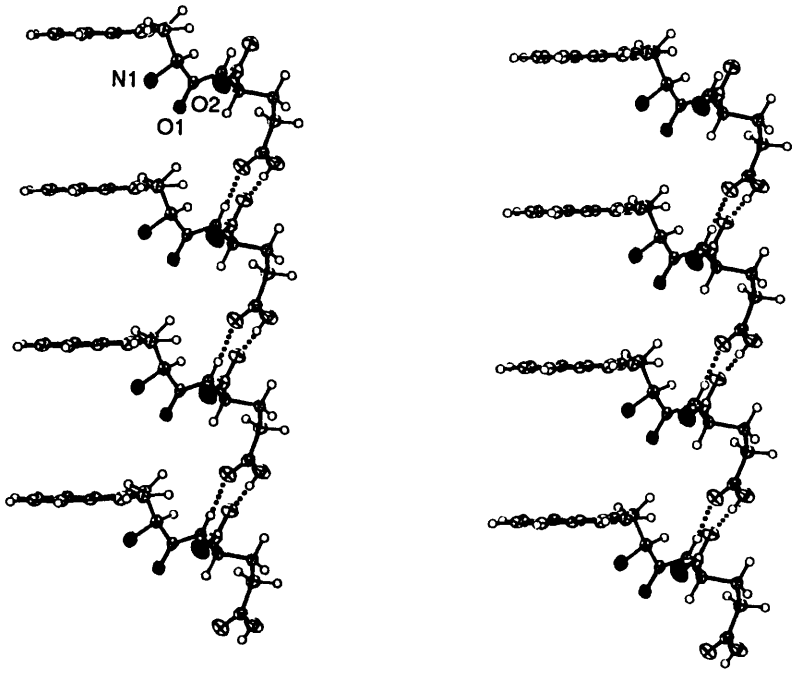


Fig. 4 A stereoscopic view of the Trp-Glu molecular arrangement which can bind the guanine base of m7GMP specifically. The dotted lines present hydrogen bonds.

an electron donor or electron acceptor in three to four hydrogen bonds with polar atoms of neighbouring m7GMP, Trp-Glu or water molecules. The O(3′)G and O(1p)G atoms participate in only one hydrogen bond, while the other atoms form two to three hydrogen bonds. This might be the reason why the former two atoms show relatively high thermal motions in the crystal. In addition to these hydrogen bonds *via* water molecules, the molecular packing of m7GMP and Trp-Glu is further stabilized by electrostatic, stacking and van der Waals interactions.

Interaction between m7GMP *and Trp-Gly Molecules.*—In the crystal structure, the Trp-Glu molecules form infinite columns along the *a*-direction, as shown in Fig. 4, where neighbouring Trp-Glu molecules are linked to each other by the hydrogen-bond pairs between the Glu carboxy side-chain and the main chain atoms: $O(2)\varepsilon^1 \cdots O(2')$ and $N(2) \cdots O(2)\varepsilon^2$ (Table 4). The Trp indole ring is almost at right angles to the direction of this column, and the N(1), O(1) and O(2) atoms of the Trp-Glu backbone chain are almost parallel to the indole ring. This spatial arrangement of aromatic ring and functional atoms provides a situation in which only the guanine and not the other three nucleic acid bases is able to interact specifically with the dipeptide. The interaction mode between the m7GMP and Trp-Glu molecules is stereoscopically shown in Fig. 5. The N(1),

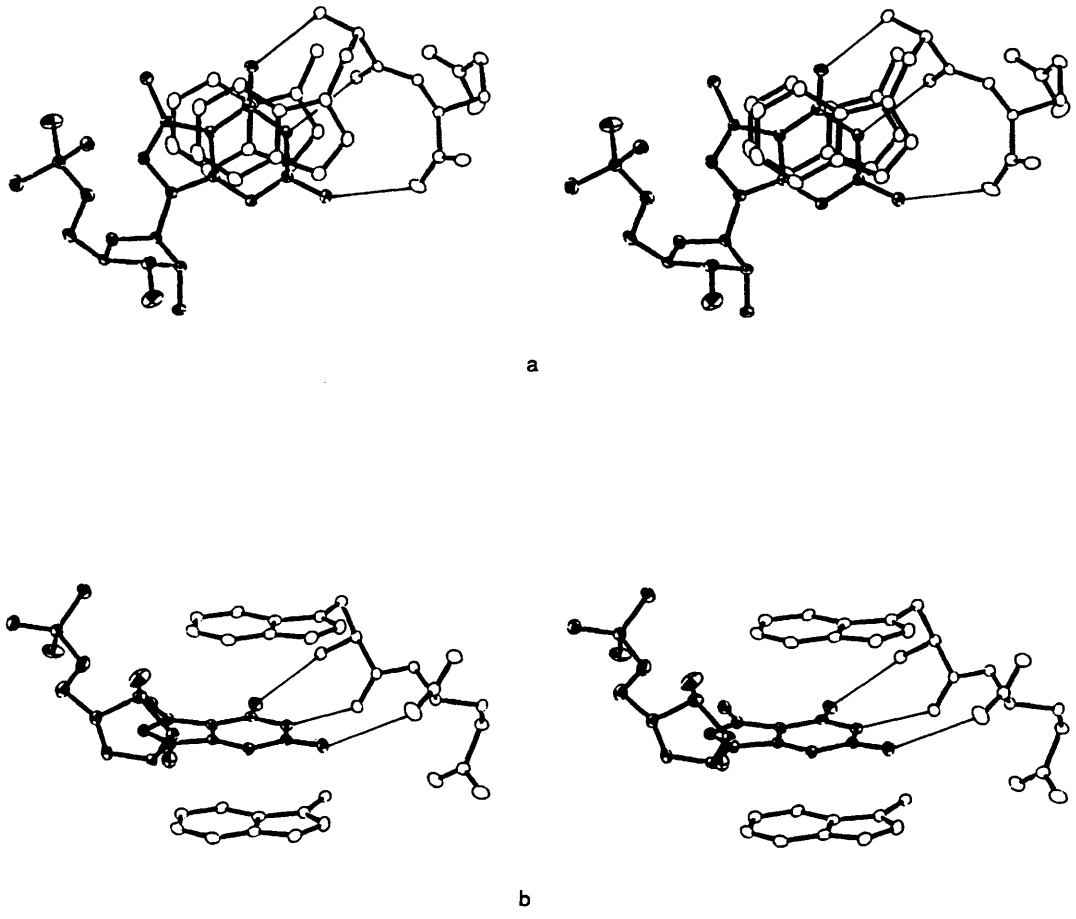
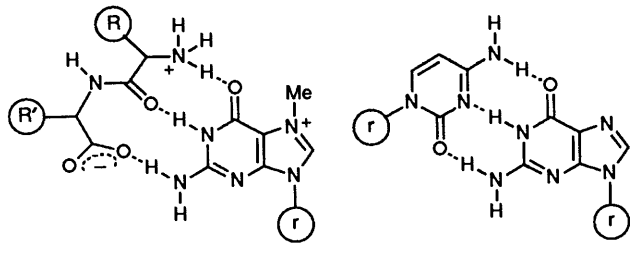


Fig. 5 Stereoscopic views of the simultaneous formation of stacking interactions and hydrogen bonds between the Trp-Glu and m7GMP molecules, viewed perpendicularly to the guanine base (a) and nearly parallel to the base (b)

O(1) and O(2) atoms participate in triple hydrogen-bond formation with the O(6), N(1) and N(2) atoms of m7GMP, respectively, just as in the usual Watson–Crick G:C base pairing (see Table 4).



In addition to this triple hydrogen-bond pairing, the guanine base is further fixed by a stacking interaction with the Trp indole ring, the guanine base being sandwiched between two neighbouring indole rings. The indole and guanine rings are arranged almost parallel, with a dihedral angle of 1.9(1)°, and their mean interplanar separation distances are 3.279 Å for the 'upper' stacking pair and 3.254 Å for the 'lower' one. An interplanar spacing less than the minimal van der Waals separation distance of 3.4 Å is typical for π–π charge-transfer interaction caused by the partial π-electron transfer from the HOMO (highest occupied molecular orbital) of the indole ring to the LUMO (lowest unoccupied molecular orbital) of the guanine base in the ground state. In addition, the stacking interaction is further stabilized by (a) the electrostatic interaction between the positively charged guanine base and the π-electron-rich indole ring, and (b) the dipole–dipole coupling interaction of both rings [the angle between their calculated dipole moments is 163.5°].

To the best of our knowledge, the present work provides the first X-ray study of a nucleotide–peptide complex, and the triple hydrogen-bond pairing of the guanine base with the peptide backbone chain is the first such interaction established by X-ray analysis. Since triple hydrogen bonding is usually observed for the Watson–Crick G:C base pair in DNA or RNA secondary structure, it could be regarded as a fundamental mode for the recognition of the guanine base by proteins, where the *N*(i) (donor), O(*i*) (acceptor) and O(*i* + 1) (acceptor) atoms of the peptide backbone chain are simultaneously hydrogen bonded to the guanine O(6), N(1) and N(2) atoms, respectively. In addition, the present complex structure also provides experimental evidence showing the importance of the stacking interaction with the aromatic amino acid side-chain in order that the guanine base be fixed tightly. Therefore it appears reasonable to assume that the co-operation between triple hydrogen bonding and stacking interactions is specific enough to select only the guanine base for bonding to the protein.

Acknowledgements

This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan (No. 61571037) and by the Encouragement Foundation of Pharmaceutical Research.

References

1 Part 21, T. Ishida, H. Ueda, K. Segawa, M. Doi and M. Inoue, *Arch. Biochem. Biophys.*, 1990, **278**, 217.

2 A. D. Hamilton and N. Pant, *J. Chem. Soc., Chem. Commun.*, 1988, 765; J. Rebek, Jr., B. Askew, P. Ballester, C. Buhr, S. Jones, D. Nemeth and K. Williams, *J. Am. Chem. Soc.*, 1987, **109**, 5033; J. Rebek, Jr., K. Williams, K. Parris, P. Ballester and K. S. Jeong, *Angew. Chem., Int. Ed. Engl.*, 1987, **26**, 1244; S. Goswami, A. D. Hamilton and D. V. Engen, *J. Am. Chem. Soc.*, 1989, **111**, 3425; B. Askew, P. Ballester, C. Buhr, K. S. Jeong, S. Jones, K. Parris, K. Williams and J. Rebek, Jr, *J. Am. Chem. Soc.*, 1989, **111**, 1082; K. Williams, B. Askew, P. Ballester, C. Buhr, K. S. Jeong, S. Jones and J. Rebek, Jr, *J. Am. Chem. Soc.*, 1989, **111**, 1090; A. D. Hamilton and D. Little, *J. Chem. Soc., Chem. Commun.*, 1990, 297.
3 C. Hélène and J. C. Maurizot, *CRC Crit. Rev. Biochem.*, 1981, **10**, 213; A. Laigle, L. Chinsky and P. Y. Turpin, *Nucleic Acids Res.*, 1982, **10**, 1707; C. H. Niu and S. Black, *J. Biol. Chem.*, 1979, **254**, 265.
4 S. M. Tahara, M. A. Morgan and A. J. Shatkin, *J. Biol. Chem.*, 1981, **256**, 7691; N. Sonenberg, M. A. Morgan, W. C. Merrick and A. J. Shatkin, *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 4843; G. M. Hellmann, L. Y. Chu and R. E. Rhoads, *J. Biol. Chem.*, 1982, **257**, 4056; N. R. Webb, R. V. J. Chari, G. DePillis, J. W. Kozarich and R. E. Rhoads, *Biochemistry*, 1984, **23**, 177.
5 H. Ueda, M. Doi, M. Inoue, T. Ishida, T. Tanaka and S. Uesugi, *Biochem. Biophys. Res. Commun.*, 1988, **154**, 199.
6 For preliminary results see T. Ishida, H. Iyo, H. Ueda, M. Doi and M. Inoue, *J. Chem. Soc., Chem. Commun.*, 1990, 217.
7 K. Kamiichi, M. Doi, M. Nabaei, T. Ishida and M. Inoue, *J. Chem. Soc., Perkin Trans. 2*, 1987, 1739.
8 T. Debaerdemaeker, G. Germain, P. Main, C. Tate and M. M. Woolfson, MULTAN87, a System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data; Universities of York, England, and Louvain, Belgium, 1987.
9 The Universal Crystallographic Computing, System-Osaka, The Computation Center, Osaka University, Osaka, Japan, 1979.
10 *International Tables for X-ray Crystallography*, Kynoch, Birmingham, England, 1974, vol. 4.
11 M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, 1977, **99**, 4899.
12 G. N. Ramachandran, C. Ramakrishnan and V. Sasisekharan, *J. Mol. Biol.*, 1963, **7**, 95.
13 J. Pandit, T. P. Seshadri and M. A. Viswamitra, *Acta Crystallogr., Sect C: Crystal Struct Commun.*, 1984, **40**, 169.
14 T. Takigawa, T. Ashida, Y. Sasada and M. Kakudo, *Bull. Chem. Soc. Jpn.*, 1966, **39**, 2369.
15 R. P. Ash, J. R. Herriott and D. A. Deranleau, *J. Am. Chem. Soc.*, 1977, **99**, 4471.
16 T. N. Bhat, V. Sasisekharan and M. Vijayan, *Int. J. Pept. Protein Res.*, 1979, **13**, 170; J. Janin, S. Wodak, M. Levitt and B. Maigret, *J. Mol. Biol.*, 1978, **125**, 357.
17 P. K. Ponnuswany and V. Sasisekharan, *Int. J. Pept. Protein Res.*, 1971, **3**, 9.
18 V. Cody, in *Chemistry and Biochemistry of the Amino Acids*, ed. G. C. Barrett, Chapman and Hall, London and New York, 1985, p. 625.
19 W. Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, 1984.
20 T. Ishida, M. Doi, H. Ueda, M. Inoue and G. M. Sheldrick, *J. Am. Chem. Soc.*, 1988, **110**, 2286; T. Ishida, M. Doi and M. Inoue, *Nucleic Acids Res.*, 1988, **16**, 6175.
21 R. Taylor and O. Kennard, *J. Mol. Struct.*, 1982, **78**, 1.

Paper 1/00264C
Received 18*th January* 1991
Accepted 27*th March* 1991