Acta Crystallographica Section C
Crystal Structure Communications
ISSN 0108-2701

A non-natural dinucleotide containing an isomeric L-related deoxynucleoside: dinucleotide inhibitors of anti-HIV integrase activity

M. Gary Newton,[a]* Charles F. Campana,[b] Guo-Chen Chi,[c] Doowon Lee,[a] Zhi-Jie Liu,[a] Vasu Nair,[c] James Phillips,[d] John P. Rose[a] and B.-C. Wang[a]

[a]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA, [b]Bruker AXS Inc., 5465 East Cheryl Parkway, Madison, Wisconsin 53711-5373, USA, [c]Department of Pharmaceutical and Biomedical Sciences, University of Georgia, Athens, GA 30602, USA, and [d]Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA
Correspondence e-mail: newton@secsg.uga.edu

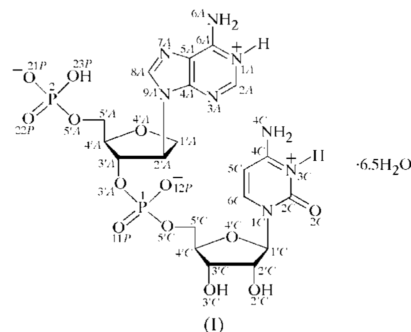
Received 4 March 2005
Accepted 15 June 2005
Online 23 July 2005

The first X-ray crystal structure of a non-natural dinucleotide, 5′-*O*-phosphoryl-1′-deoxy-2′-isoadenylyl-(3′ → 5′)-cytidine 6.5-hydrate (pIsodApC), $C_{19}H_{26}N_8O_{13}P_2 \cdot 6.5H_2O$, belonging to a family of dinucleotides that contain an isomeric nucleoside component, is described. A complex system of hydrogen bonds between water molecules and various sites on the dinucleotide was found. All H atoms were located from electron-density difference maps, which allowed identification of protonation sites. Compounds of this family have been found to bind at the active site of HIV integrase and to be inhibitors of this key viral enzyme. These dinucleotides are completely resistant to cleavage by exonucleases; an abnormal dihedral angle twist in an internucleotide phosphate bond revealed in the X-ray crystal structure may be contributing to this unusual stability towards nucleases.

Comment

The retroviral enzyme HIV-1 integrase incorporates HIV double-helical DNA into host chromosomal DNA (Dyda *et al.*, 1994; Mazumder *et al.*, 1996; Frankel & Young, 1998; Esposito & Craigie, 1999; Haren *et al.*, 1999; Nair, 2002, 2003). This viral enzyme initially catalyzes the excision of two terminal nucleotides at the 3′-end of each strand of viral DNA (3′-processing), leaving recessed ends that terminate with *xx*CA-OH. In the next steps (strand transfer and integration), attack of the terminal 3′-OH of the tailored HIV DNA on a specific internucleotide phosphodiester bond results in cleavage of host DNA, which is followed by integration of the HIV DNA into host DNA (Dyda *et al.*, 1994; Mazumder *et al.*, 1996; Frankel & Young, 1998; Esposito & Craigie, 1999; Haren *et al.*, 1999; Nair, 2002, 2003). The integration process is essential for the replication of HIV.

In designing inhibitors of this viral enzyme, it was suggested that residues immediately upstream of the dinucleotide cleavage site in the 3′-processing step may provide critical recognition/binding sites for HIV integrase. With this design



concept in mind, natural dinucleotides were investigated by Pommier and co-workers, and non-natural dinucleotides were synthesized and investigated by Nair and co-workers, as potential anti-HIV integrase inhibitors (Mazumder *et al.*, 1997; Taktakishvili, Neamati, Pommier & Nair, 2000; Taktakishvili, Neamati, Pommier, Pal & Nair, 2000; Taktakishvili *et al.*, 2001). The non-natural dinucleotides contain one natural D-nucleoside component and one isomeric L-related nucleoside moiety (Nair & Jahnke, 1995), which are joined together by an internucleotide phosphodiester bond and are 5′-phosphorylated. While both the natural and non-natural dinucleotides with the AC base components were found to be strong inhibitors of HIV-1 integrase (IC_{50} low μ*M* range for both the processing and strand-transfer steps), one remarkable aspect of the non-natural dinucleotides was that the internucleotide bond exhibited resistance to cleavage by mammalian 3′- and 5′-exonucleases (Taktakishvili, Neamati, Pommier & Nair, 2000; Taktakishvili, Neamati, Pommier, Pal & Nair, 2000; Taktakishvili *et al.*, 2001; Nair & Pal, 2004). This was not the case for the natural dinucleotides, which were substrates for 5′-exonucleases (Nair & Pal, 2004). The observed hypochromicity from the quantitative UV spectra of non-natural dinucleotides suggested the presence of base stacking in preferred conformations, which implied the presence of conformationally unusual internucleotide phosphate bonds,

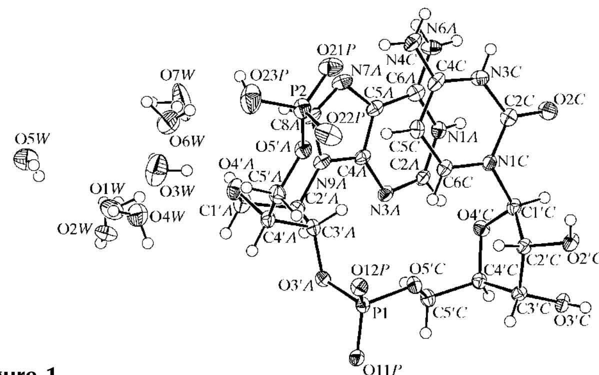


Figure 1
A plot of dinucleotide (I), showing the atom-numbering scheme; atom labels ending in *A* are associated with the adenosine portion and those ending in *C* are associated with the cytosine portion. Displacement ellipsoids are drawn at the 50% probability level.

 © 2005 International Union of Crystallography DOI: 10.1107/S0108270105019037

because of the spatial arrangement of the two sugar rings required to accommodate base stacking (Taktakishvili, Neamati, Pommier & Nair, 2000; Taktakishvili, Neamati, Pommier, Pal & Nair, 2000; Taktakishvili *et al.*, 2001). In order to obtain further structural information (conformation, internucleotide phosphate bond, base-stacking property) on this family of non-natural dinucleotides, we crystallized one of these compounds, the title compound, pIsodApC or (I), and examined its detailed X-ray crystallographic structure.

The numbering scheme for (I) is shown in the scheme. Atoms included in the adenosine portion have an *A* as the last letter of their label and those included in the cytosine portion have a *C* as the last letter. This coding allows easy recognition of the nucleotide components. All atom labels follow the usual chemical numbering for ribose, adenosine and cytosine. The O atoms of the seven (formally 6.5) water molecules are labeled O1*W*–O7*W*. Atom O7*W* occupies a diagonal twofold axis of the space group, is likely to be disordered and has an occupancy factor of 0.5. The H atoms attached to atom O7*W* are also assigned an occupancy factor of 0.5.

H atoms were located from electron-density difference maps, providing information about protonation sites: adenosine atom N1*A* and cytosine atom N3*C* are protonated. Both phosphate groups are monoanions, giving a neutral structure overall. All bond lengths and angles involving non-H atoms are within expected limits. Overall, the structure is formally dimeric, with two dinucleotide molecules being connected by hydrogen bonds from the water molecule at the O7*W* site to the O6*W* water molecule, which is also hydrogen bonded to the terminal phosphate group.

The seven water molecules form a very complex association with the dinucleotide structure. Table 1 lists all the water contacts of less than 3.25 Å between the dinucleotide and the water molecules and between water molecules. A summary of significant contacts of less than 3.0 Å (inter- and intramolecular) is as follows: for O1*W* contacts, N3*C*—O1*W* 2.636 (5) Å and O1*W*—O3*W* 2.686 (6) Å; for O2*W* contacts, O2*W*—O21*P* 2.776 (4) Å, O2*W*—O11*P* 2.854 (4) Å and N6*A*—O2*W* 2.769 (4) Å; for additional O3*W* contacts, O3*W*—O3′*C* 2.821 (5) Å and O3*W*—O21*P* 2.978 (5) Å. For O4*W* contacts, O4*W*—O22*P* 2.672 (5) Å and O4*W*—O12*P* 2.782 (5) Å; for O5*W* contacts, O5*W*—O12*P* 2.756 (4) Å, O5*W*—N4*C* 2.948 (5) Å and O6*W*—O5*W* 2.897 (5) Å; for additional O6*W* contacts, O6*W*—O23*P* 2.833 (5) Å and O7*W*—O6*W* 2.924 (6) Å. Finally, O7*W* (disordered) sits on a diagonal twofold axis and bridges between two symmetry-related molecules through an O6*W* water molecule.

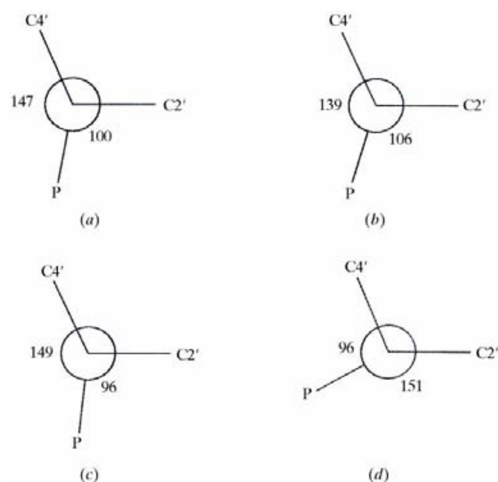


Figure 2
A comparison of the dihedral angles between (*a*) ApU1, (*b*) ApU2, (*c*) GpC and (*d*) (I), showing dihedral angle values in degrees.

Fig. 1 shows a plot of a single molecule of (I) with the seven associated water molecules. From this diagram, the overlap of the cytosine and adenosine rings can be readily observed. Best-plane calculations of the adenosine and cytosine rings were performed and the angle between these planes was found to be 12.6°. The mean deviation of atoms in the adenosine plane is 0.006 Å, and 0.016 Å for the cytosine ring. The atom contacts from the adenosine plane to the cytosine best plane range from 3.223 (atom C8*A*) to 4.122 Å (atom C2*A*), while the atom contacts from the cytosine ring to the adenosine best plane range from 3.096 (atom N4*C*) to 4.015 Å (atom O2*C*).

The dihedral angles along the connecting phosphate backbone were compared with equivalent dihedral angles in the structures of two previously reported normal dinucleotides, namely ApU (Seeman *et al.*, 1976) and GpC (Rosenberg *et al.*, 1976); both ApU and GpC are hydrated sodium salts. ApU had two independent molecules in the asymmetric unit and both units, labeled ApU1 and ApU2, were used in this comparison. The seven dihedral angles along the ribose–phosphate backbone around the C3′*A*—O3′*A*, O3′*A*—P1, P1—O5′*C*, O5′*C*—C5′*C* and O5′*C*—C4′*C* bonds were compared. All dihedral angles for these four structures, except for those around the C3′*A*—O3′*A* bond, were found to have approximately the same value [deviations range from 1–14° between (I) and the normal dinucleotides]. However, around the C3′—O3′ bond there is a 49° difference, on average, between the dihedral angles of (I) and ApU1, ApU2 and GpC. With the adenosine ring moved to C2′, some change in the twisting of the flexible phosphate backbone is clearly necessary to accommodate the adenosine–cytosine overlap. This abnormal twist may contribute to the unusual stability of (I) towards nucleases. Fig. 2 compares Newman projections of the C3′—O3′ bonds in ApU1, ApU2 and GpC (Figs. 2*a*–*c*) with (I) (Fig. 2 *d*).

Another notable difference between (I) and the ApU and GpC dinucleotides is the conformations of the ribose rings. On the pyrimidine base side (the 5′-side), ApU and GpC have envelope conformations, with atom C3′ in the flap position; in (I), the conformation is twisted around C2′—C3′. On the purine base side (the 3′-side), ApU and GpC have twisted conformations around C2′—C3′; in (I), the conformation is twisted around C1′—O4′.

Experimental

The title dinucleotide 5′-monophosphate, (I), was synthesized through the coupling of 4-*N*-2′,3′-*O*-tribenzoylcytidine and 6-*N*-benzoyl-5′-*O*-DMTr-isodeoxyadenosine (Wenzel & Nair, 1998), followed by 5′-phosphorylation (Chi *et al.*, 2004). Its structure was characterized by multinuclear NMR, high-resolution MS and quan-

titative UV spectroscopy (see below). Crystals were obtained by very slow evaporation of an aqueous solution of (I) in a refrigerator. ^{1}H NMR (D_2O): δ 8.28 (*s*, 1H), 8.23 (*s*, 1H), 7.50 (*d*, 1H, J = 8.0 Hz), 5.96 (*d*, 1H, J = 8.0 Hz), 5.49 (*d*, 1H, J = 4.5 Hz), 5.24 (*m*, 1H), 4.52 (*m*, 1H), 4.26 (*dd*, 1H, J = 11.0 and 6.0 Hz), 3.87–4.20 (*m*, 9H); ^{13}C NMR (D_2O): δ 159.3, 150.7, 148.6, 148.5, 146.3, 142.5, 117.7, 95.4, 89.1, 84.8, 82.8, 80.4, 74.1, 71.1, 69.0, 64.2, 63.7, 61.7, one C atom not observable; ^{31}P NMR (D_2O): δ 0.86, −0.44. FAB–HRMS: $[M + H]^+$ calculated for $C_{19}H_{27}N_8O_{13}P_2$: 637.1173, found: 637.1170; UV (H_2O): λ_{max} 264 nm (ε 19 800 dm^3 mol^{-1} cm^{-1}).

Crystal data

$C_{19}H_{26}N_8O_{13}P_2 \cdot 6.5H_2O$
M_r = 753.52
Tetragonal, $P4_12_12$
a = 12.566 (5) Å
c = 38.708 (5) Å
V = 6112 (4) Å^3
Z = 8
D_x = 1.638 Mg m^{-3}
Cu $K\alpha$ radiation
Cell parameters from 1009 reflections
θ = 2–50°
μ = 2.20 mm^{-1}
T = 100 (2) K
Needle, colorless
0.35 × 0.15 × 0.15 mm

Data collection

Bruker SMART 6000 CCD area-detector diffractometer
ω scans
Absorption correction: multi-scan (*SADABS*; Sheldrick, 1996)
T_{min} = 0.513, T_{max} = 0.734
23451 measured reflections
5272 independent reflections
5266 reflections with $I > 2\sigma(I)$
R_{int} = 0.025
θ_{max} = 67.0°
h = −11 → 14
k = −14 → 12
l = −42 → 42

Refinement

Refinement on F^2
$R[F^2 > 2\sigma(F^2)]$ = 0.047
$wR(F^2)$ = 0.124
S = 1.10
5272 reflections
474 parameters
H atoms treated by a mixture of independent and constrained refinement
$w = 1/[\sigma^2(F_o^2) + (0.0609P)^2 + 10.5537P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$ = 0.002
$\Delta\rho_{max}$ = 0.69 e Å^{-3}
$\Delta\rho_{min}$ = −0.49 e Å^{-3}
Absolute structure: Flack (1983), with 2147 Friedel pairs
Flack parameter: 0.04 (3)

Table 1
Hydrogen-bond geometry (Å, °).

D—H···A	D—H	H···A	D···A	D—H···A
O3W—H3WA···O3′C^{i}	0.85 (2)	1.99 (3)	2.821 (4)	167 (7)
O2′C—H2′O···O23P^{ii}	0.82	1.82	2.631 (4)	167
O3′C—H3′O···N7A^{iii}	0.82	2.04	2.767 (4)	147
O23P—H23P···O3W^{iv}	0.82	2.55	3.227 (6)	141
N1A—H1NA···O11P^{v}	0.74	1.97	2.692 (4)	166
N3C—H2NC···O1W^{vi}	0.86	1.79	2.636 (5)	169
N4C—H4AC···O5W^{vi}	0.86	2.21	2.948 (5)	144
N4C—H4BC···O21P	0.86	1.99	2.822 (5)	163
N6A—H6NA···O2W^{vi}	0.86	1.93	2.769 (4)	167
N6A—H6NB···O2′C^{vii}	0.86	2.13	2.967 (4)	163
O1W—H1WA···O3W	0.90 (2)	1.92 (5)	2.686 (6)	142 (6)
O2W—H2WA···O11P^{viii}	0.82 (2)	2.11 (4)	2.854 (4)	151 (7)
O2W—H2WB···O21P^{iv}	0.82 (2)	1.97 (2)	2.776 (4)	166 (6)
O3W—H3WB···O21P^{iv}	0.85 (2)	2.13 (2)	2.978 (5)	172 (6)
O4W—H4WA···O12P^{ix}	0.80 (2)	1.99 (2)	2.782 (4)	172 (7)
O4W—H4WB···O22P^{ix}	0.80 (2)	1.88 (2)	2.672 (5)	175 (7)
O5W—H5WA···N4C^{x}	0.80 (2)	2.45 (6)	2.948 (5)	122 (6)
O5W—H5WA···O1W	0.80 (2)	2.46 (4)	3.195 (5)	153 (7)
O5W—H5WB···O12P^{ix}	0.82 (2)	1.99 (3)	2.756 (4)	156 (6)
O6W—H6WA···O5W^{iv}	0.83 (2)	2.10 (3)	2.897 (4)	161 (7)
O6W—H6WB···O23P	0.84 (2)	2.00 (2)	2.833 (5)	172 (7)
O7W—H7WA···O6W	0.80	2.13	2.924 (6)	172
O7W—H7WB···O6W^{iv}	0.88	2.06	2.924 (6)	166

Symmetry codes: (i) $-x+\frac{1}{2}, y+\frac{1}{2}, -z+\frac{5}{4}$; (ii) $-x+\frac{3}{2}, y-\frac{1}{2}, -z+\frac{5}{4}$; (iii) $-y+\frac{1}{2}, x-\frac{1}{2}, z+\frac{1}{4}$; (iv) $y, x, -z+1$; (v) $-x+\frac{1}{2}, y-\frac{1}{2}, -z+\frac{5}{4}$; (vi) $x, y-1, z$; (vii) $y+\frac{1}{2}, -x+\frac{1}{2}, z-\frac{1}{4}$; (viii) $y-\frac{1}{2}, -x+\frac{3}{2}, z-\frac{1}{4}$; (ix) $-x+\frac{3}{2}, y+\frac{1}{2}, -z+\frac{5}{4}$; (x) $x, y+1, z$.

All H atoms were initially located in a difference Fourier synthesis. H atoms bonded to C, O and N atoms were constrained to ride on their parent atoms, with U_{iso}(H) = 1.2U_{eq}(C,O,N). Water molecules were restrained geometrically to an O—H distance of 0.82 Å and to a H···H distance of 1.3 Å. The H atoms on atom O7W were fixed with occupancy factors of 0.5. *PLATON* (Spek, 2003) indicates a number of short contacts between H atoms, namely H2′O···H23P = 1.70 Å, H4AC···H5WA = 1.63 Å and H23P···H6WB = 1.78 Å. The positions of the H atoms in question have been checked using omit maps and in each case the refined H-atom sites were consistent with the electron density seen in these maps. Although no alternative sites could be seen in the omit maps, it is possible that these H atoms were partially occupied, with additional alternative sites which are not visible by X-ray diffraction techniques.

Data collection: *PROTEUM* (Bruker, 2003); loop mounted and flash frozen (Hope, 1988); cell refinement: *PROTEUM*; data reduction: *PROTEUM* and *SADABS* (Sheldrick, 1996); program(s) used to solve structure: *SHELXTL* (Sheldrick, 1997); program(s) used to refine structure: *SHELXTL*; molecular graphics: *SHELXTL* and *PLATON* (Spek, 2003); software used to prepare material for publication: *SHELXTL*.

This project was supported by grant No. RO1 AI 43181 (VN) from the National Institutes of Health. The contents of this report are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Supplementary data for this paper are available from the IUCr electronic archives (Reference: HJ1051). Services for accessing these data are described at the back of the journal.

References

Bruker (2003). *PROTEUM*. Version 1.40. Bruker AXS Inc., Madison, Wisconsin, USA.
Chi, G. C., Neamati, N. & Nair, V. (2004). *Bioorg. Med. Chem. Lett.* **14**, 4815–4817.
Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R. & Davies, D. R. (1994). *Science*, **266**, 1981–1986.
Esposito, D. & Craigie, R. (1999). *Adv. Virus Res.* **52**, 319–333.
Flack, H. D. (1983). *Acta Cryst.* A**39**, 876–881.
Frankel, A. D. & Young, J. A. T. (1998). *Annu. Rev. Biochem*, **67**, 1–25.
Haren, L., Ton-Hoang, B. & Chandler, M. (1999). *Annu. Rev. Microbiol.* **53**, 245–281.
Hope, H. (1988). *Acta Cryst.* B**44**, 22–26.
Mazumder, A., Neamati, N., Ojwang, J. O., Sunder, S., Rando, R. F. & Pommier, Y. (1996). *Biochemistry*, **35**, 13762–13771.
Mazumder, A., Uchida, H., Neamati, N., Sunder, S., Jaworska Maslanka, M., Wickstrom, E., Zeng, F., Jones, R. A., Mandes, R. F., Chenault, H. K. & Pommier, Y. (1997). *Mol. Pharmacol.* **51**, 567–575.
Nair, V. (2002). *Rev. Med. Virol.* **12**, 179–193.
Nair, V. (2003). *Curr. Pharm. Des.* **9**, 2553–2565.
Nair, V. & Jahnke, T. S. (1995). *Antimicrob. Agents Chemother.* **39**, 1017–1029.
Nair, V. & Pal, S. (2004). *Bioorg. Med. Chem. Lett.* **14**, 289–291.
Rosenberg, J. M., Seeman, N. C., Day, R. O. & Rich, A. (1976). *J. Mol. Biol.* **104**, 145–167.
Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Parkkim, J. J. & Rich, A. (1976). *J. Mol. Biol.* **104**, 109–144.
Sheldrick, G. M. (1996). *SADABS*. University of Göttingen, Germany.
Sheldrick, G. M. (1997). *SHELXTL*. Version 5.10. Bruker AXS Inc., Madison, Wisconsin, USA.
Spek, A. L. (2003). *J. Appl. Cryst.* **36**, 7–13.
Taktakishvili, M., Neamati, N., Pommier, Y. & Nair, V. (2000). *Bioorg. Med. Chem. Lett.* **10**, 249–251.
Taktakishvili, M., Neamati, N., Pommier, Y. & Nair, V. (2001). *Bioorg. Med. Chem. Lett.* **11**, 1433–1435.
Taktakishvili, M., Neamati, N., Pommier, Y., Pal, S. & Nair, V. (2000). *J. Am. Chem. Soc.* **122**, 5671–5677.
Wenzel, T. & Nair, V. (1998). *Bioconjug. Chem.* **9**, 683–690.