From Rational Design of Drug Crystals to Understanding of Nucleic Acid Structures: Lamivudine Duplex

ARTICLE *in* CRYSTAL GROWTH & DESIGN · FEBRUARY 2010

Impact Factor: 4.89 · DOI: 10.1021/cg901103r

3 AUTHORS, INCLUDING:

Antonio Carlos Doriguetto
Universidade Federal de Alfenas
119 PUBLICATIONS **894** CITATIONS
SEE PROFILE

Javier Ellena
University of São Paulo
410 PUBLICATIONS **2,665** CITATIONS
SEE PROFILE

DOI: 10.1021/cg901103r

CRYSTAL GROWTH & DESIGN Article

2010, Vol. 10
676–684

From Rational Design of Drug Crystals to Understanding of Nucleic Acid Structures: Lamivudine Duplex

Felipe T. Martins,[†] Antônio C. Doriguetto,[§] and Javier Ellena*,[†]

[†]*Instituto de Física de São Carlos, Universidade de São Paulo, CP 369, 13560-970 São Carlos, São Paulo, Brazil, and* [§]*Departamento de Ciências Exatas, Universidade Federal de Alfenas, Rua Gabriel Monteiro da Silva 700, 37130-000 Alfenas, Minas Geraís, Brazil*

Received September 9, 2009

ABSTRACT: A DNA-like duplex of nucleosides is probable to exist even without the 5′-phosphate groups needed to assemble the chain backbone. However, double-stranded helical structures of nucleosides are unknown. Here, we report a duplex of nucleoside analogs that is spontaneously assembled due to stacking of the neutral and protonated molecules of lamivudine, a nucleoside reverse transcriptase inhibitor (NTRI) widely used in anti-HIV drug combinatory medication. The left-handed lamivudine duplex has features similar to those of *i*-motif DNA, as the face-to-face base stacking and the helix rise per base pair. Furthermore, the protonation pattern on alternate bases expected for a DNA-like duplex stabilized by pairing of neutral and protonated cytosine fragments was observed for the first time in the lamivudine double-stranded helix. This structure demonstrates that hydrogen bonds can substitute for covalent phosphodiester linkage in the stabilization of the duplex backbone. This interesting example of spontaneous molecular self-organization indicates that the 5′-phosphate group could not be a requirement for duplex assembly.

Introduction

Nucleoside reverse transcriptase inhibitors (NRTIs) are first-line drugs in the treatment of AIDS. Lamivudine, β-L-2′,3′-dideoxy-3′-thiacytidine (3TC), is one of the most commonly used NTRIs in the anti-HIV drug combinatory medication that is preferred over single-drug therapies.[1] It is also used in clinical treatment for hepatitis B virus (HBV) infection.[2] This drug is a 2′,3′-dideoxygenated cytidine analog with a sulfur atom in place of the methylene group at the 3′ position of the sugar ring. This alteration in the sugar ring creates an oxathiolane system, which forms an unusual β-L-stereochemistry in lamivudine.[3]

A recent study in the crystal engineering research field demonstrated that neutral and cationic lamivudine fragments can pair through their cytosine rings[4] in a hydrogen bonding pattern related to that of the C−C$^+$ base pair of *i*-motif DNA.[5−9] The *i*-motif of DNA comprises two intercalated parallel duplexes assembled by hemiprotonated C−C$^+$ base pairs, giving rise to a four-stranded structure in which each individual duplex exhibits an antiparallel orientation relative to the other double-stranded fiber.[5] Three hydrogen bonds between the neutral and protonated cytosine fragments are responsible for both the 3TC−3TC$^+$ pairing in the cocrystal of the drug with 3,5-dinitrosalicylate[4] and the C−C$^+$ of *i*-motif DNA (Figure 1a).[5] Given the structural similarities between the hemiprotonated C−C$^+$ base pair of *i*-motif DNA and the lamivudine dimer formed by positively charged and neutral molecules, we were interested to know whether this NRTI could adopt a DNA-like duplex structure by properly stacking the pairs. One might assume that formation of a double-stranded helix by common nucleosides or their analogs in isolation is probable even without the phosphate groups that make up the covalent backbone of alternating phosphate and pentose residues. A DNA-like lamivudine structure seems noteworthy because this drug is a 2′,3′-dideoxynucleoside derivative, missing the 3′ hydroxyl substituent that would covalently bond to the (lacking) phosphate moiety of the adjacent monomer in the strand. Furthermore, counterions could contribute to stabilize individual strands of lamivudine through noncovalent interactions. However, no DNA-like double helix structure in which each hydrogen-bonded pair of nucleosides is helically stacked on the top of one another has been reported, although they could provide insights into the physical basis of DNA structure assembly.

We have devoted a great deal of effort to obtain a structure in which the lamivudine could be self-assembled into a DNA-like duplex without phosphodiester linkages by stacking the 3TC−3TC$^+$ pairs in a helical fashion. Based on crystal engineering studies,[4,10] numerous solvent systems and various cocrystallizing agents were screened before the suitable conditions for crystallization of lamivudine duplex were found. Our first attempts to synthesize such a structure using only simple inorganic acids such as hydrochloric and phosphoric acid failed. Next, solvents and organic acids were employed in the synthesis strategy. We have successfully obtained a structure of lamivudine in which the 3TC−3TC$^+$ dimers are arranged into a double helix. Hydrogen-bonding contacts involving counterions and solvent molecules play the role of the missing 5′-phosphate groups.

Experimental Section

Synthesis of the Lamivudine−Lamivudine$^+$ Helix. Lamivudine was weighed (10 mg) and dissolved in isopropanol (5 mL) by heating the mixture at 308 K in a water bath under slow stirring for 5 min. The newly prepared solution was cooled to 298 K by incubation at room temperature, and a solution of maleic acid (1 mg in 250 μL of 280 mM HCl) was then added to it. The mixture was shaken at 298 K for 5 min and then kept standing for 7 days in the dark within a crystal growth apparatus at room temperature. The external symmetry of these crystals can be viewed in Figure 2, a scanning electron

*To whom correspondence should be addressed. E-mail: javiere@ifsc.usp.br. Phone: +55 16 3373 8096. Fax: +55 16 3373 9881.

 Published on Web 10/15/2009

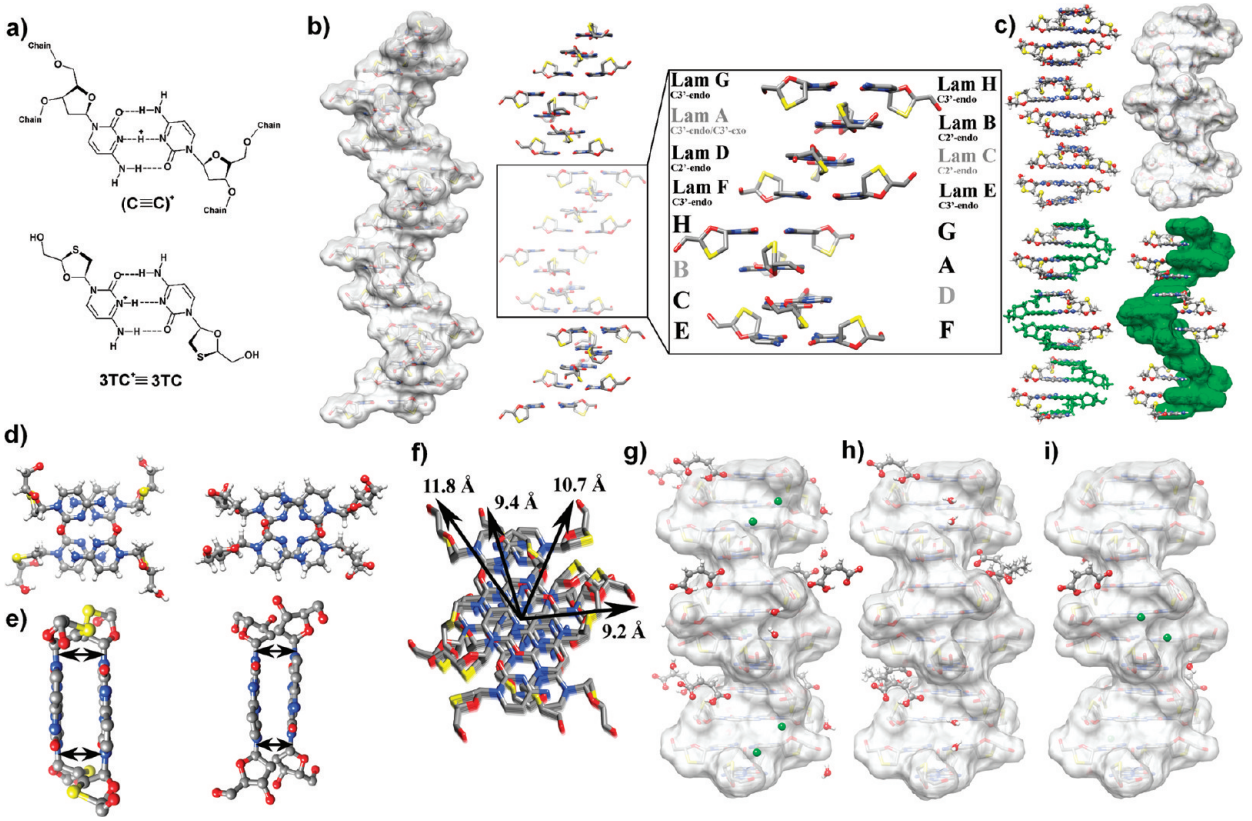

Figure 1. (a) Cytosine base pairing of *i*-motif DNA and lamivudine duplex and (b) transparent surface rendering and stick representation of lamivudine duplex. Lamivudine molecules of 2.5 unit cells are shown. The structure is self-assembled along the *b* axis. One full helical turn (lamivudine molecules of one unit cell) is highlighted. The hydrogen atoms were hidden for clarity. The puckering modes of the oxathiolane rings are relative to the fundamental nucleotides. Panel (c) shows a side view (ball-and-stick and surface drawings) of the structure in which major and minor grooves of similar depth are highlighted. On the bottom of the panel, one strand is green colored and represented as a surface rendering to illustrate the pairing responsible for the assembly of the lamivudine double-stranded helix. Lamivudine molecules of 1.5 unit cells are shown. Base stacking of the 3TC−3TC$^+$ (left) and C−C$^+$ (right) pairs viewed (d) along and (e) perpendicular to the helix axis of lamivudine duplex and d(C_4) *i*-motif DNA[6] (PDB code 190d), respectively. Hydrogens and phosphates of the d(C_4) fragment were hidden in the drawing (e), and the arrows indicate the face-to-face orientation of each base pair in relation to one another. Panel (f) shows a top view of lamivudine duplex along the *b* axis. The helical radii relative to the outermost hydroxyl groups are shown. Water and isopropanol molecules and chloride and maleate anions interact with (g) the hydroxymethylene branches of lamivudine, playing the role of the missing 5′-phosphate groups in stabilizing the duplex through the fiber periphery and the amino groups of the drug in (h) the major and (i) the minor grooves. Note the helical disposition of the solvents and anions in the grooves. Lamivudine molecules of 1.5 unit cells are shown.

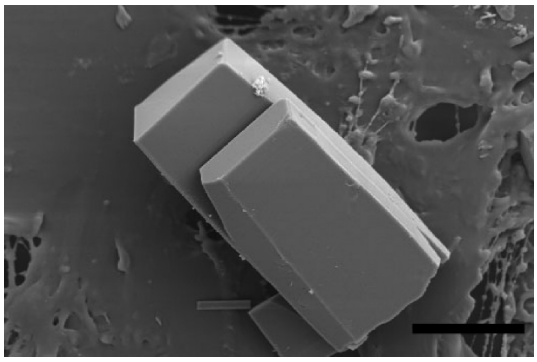

Figure 2. Crystals of the lamivudine double-stranded helix. Scale bar: 100 μm.

micrograph acquired using a LEO 435 VP microscope at 15 kV accelerating voltage, after the crystals were coated with colloidal gold in a SCD-040 Ion Sputter Balzer device.

Structure Determination of the Lamivudine−Lamivudine$^+$ Helix. A crystal measuring 0.48 × 0.21 × 0.18 mm^3 was selected for X-ray diffraction experiments at room (298 ± 1 K) and low (150.0 ± 0.2 K) temperatures using an Enraf-Nonius Kappa-CCD diffractometer. In the case of low-temperature data collection, a cold N_2 gas blower cryogenic device (Oxford Cryosystem, Oxford) was employed to freeze the crystal. The Mo Kα beam (λ = 0.71073 Å) from a sealed Mo tube was generated at 55 kV and 33 mA, and the resulting radiation was monochromatically filtered with graphite before it was passed through the crystal. The data collection geometry was strategically calculated by setting φ scans and ω scans with κ offsets, and a CCD camera of 95 mm was the detector on a κ-goniostat. The COLLECT program was used for monitoring the diffraction frame acquisition.[11] To deal with raw data, the HKL Denzo-Scalepack software package was used.[12]

Room-temperature X-ray diffraction data were used to solve a structure with some atoms distributed over at least two occupancy sites. Refinements adopting a split position atom model of this structure did not converge easily. Furthermore, they were very unstable and showed high correlation parameters. Fortunately, refinement of the structure on the basis of the low-temperature Bragg reflections converged more easily, giving satisfactory *R*-factors. Based on the low-temperature data, lower correlations between the refined atomic displacement parameters (ADPs) and site occupancy factors (SOFs) were achieved, and the split atom approach allowed the precise positioning of the atomic fractions. For this reason, only the structure determined using the low-temperature intensities is presented, although the room-temperature experiment conclusively demonstrates that no solid−solid phase transformation occurs between these temperatures because the cell parameters determined for lamivudine helix at low and room temperatures are similar.

The structure of the lamivudine helix was solved crystallographically using the direct methods of phase retrieval with SIR2004,[13] locating all of the C, N, O, S, and Cl atoms of the asymmetric unit from the electronic density map generated by Fourier synthesis. The starting model was refined by a full-matrix least-squares based on F^2 with SHELXL-97,[14] using anisotropic thermal displacement parameters for non-hydrogen atoms. Except for the hydrogen atoms of water molecules, those bonded to lamivudine, maleate, and isopropanol fragments were fixed in stereochemically predicted positions. However, both N−H protons of $3TC^+$ and O−H hydrogen atoms of 3TC, $3TC^+$, maleate, and isopropanol molecules were first localized from the difference Fourier map so that they were ideally positioned for covalent bonding to the corresponding nitrogen and oxygen atoms. The constrained interatomic distances were 0.93 Å (C_{sp^2}—H), 0.97 Å (C_{sp^3}—H in methylene groups without disordered occupancy sites of hydrogen atoms), 0.96 Å (C_{sp^3}—H in methylene and methyl groups with disordered occupancy sites of hydrogen atoms), 0.98 Å (C_{sp^3}—H in methine groups), 0.82 Å (O−H), and 0.86 Å (N−H). Similarly, hydrogen atoms of water molecules were fixed in positions located from the difference map, constraining their fractional coordinates during refinements. Concerning the hydrogen vibration model, the isotropic thermal displacement parameter of hydrogen atoms covalently bonded to nitrogen and carbon, except those of the methyl branches, was fixed to a value 20% greater than the equivalent isotropic parameter of the corresponding atom, whereas this percentage was set to 50% for hydrogen atoms bonded to oxygens and the isopropanol methyl groups.

Furthermore, five of the eight crystallographically independent lamivudine molecules (labeled as molecules A−H) in the asymmetric unit of the helix structure showed disordered sites for some of their atoms. For clarity, the atom labels ending in "a" correspond to lamivudine molecule A, those ending in "b" refer to molecule B, and so on. Four independent lamivudine molecules (B, D, E, F) were refined by splitting the oxygen and hydrogen atoms of their hydroxyl groups over two positions: the atomic fractions labeled as O5′b—H5′Ob, O5′d—H5′Od, O5′e—H5′Oe, and O5′f—H5′Of had a constrained 60% major site occupancy, whereas the atomic fractions labeled as O5′b′—H5′Ob′, O5′d′—H5′Od′, O5′e′—H5′Oe′ and O5′f′—H5′Of′ filled extra sites with a fixed 40% occupancy value. The methylene carbon atoms directly bonded to these hydroxyl groups were also expected to be disordered over two positions because the different conformational modes would affect the whole hydroxymethylene branches and not just the hydroxyl groups. Despite trial refinements constraining the primary site occupancy of these carbon atoms to 50%, extra occupancy sites were not found for the carbons C5′b, C5′d, C5′e, and C5′f . These methylene carbon atoms were not disordered due to the two occupancy sites of the hydroxyl oxygen atom oriented toward two of three surrounding bonding positions of the methylene carbon, excluding that of the C−C bond. This means the spatial arrangement of ligands on each methylene carbon only assumes a staggered conformation with three positions where two methylene hydrogen atoms and a disordered hydroxyl oxygen are attached. Therefore, neither rotation on the C−C bond axis nor slight displacements over two positions were observed for such methylene groups. To agree with this conformational model, we proposed that a methylene hydrogen atom filled two occupancy sites positioned toward those of the hydroxyl oxygen. Each disordered site of a split methylene hydrogen atom had an occupancy value equal to that of the opposite site filled with a hydroxyl group. The hydrogen atom fractions labeled as H5′xb, H5′xd, H5′xe, and H5′xf are in the constrained 60% occupancy major sites, whereas those labeled as H5′zb, H5′zd, H5′ze, and H5′zf filled the fixed 40% occupancy extra sites. The methylene hydrogen atoms labeled H5′yb, H5′yd, H5′ye, and H5′yf were not disordered over two occupancy positions in the lamivudine molecules B, D, E, and F, respectively. Finally, the lamivudine molecule A showed disordered positions for the methylene carbon and sulfur atoms of the five-membered oxathiolane ring. The sulfur atom and the methylene group were modeled over two positions constrained to 60% occupancy for the major sites (atomic fractions S3′a, C2′a, H2′xa, and H2′ya) and 40% occupancy for the extra positions (S3′a′, C2′a′, H2′xa′, and H2′ya′). The methyl branches of solvent isopropanol were refined as ideally disordered ones, with two 50% occupancy positions for each hydrogen atom. Crystal structure representations were prepared using MERCURY[15] and CHIMERA[16] softwares.

All data regarding the lamivudine double-stranded helix structure determination, except the structure factors, were deposited with the Cambridge Structural Data Base as a crystallographic information file (.cif file) under the code CCDC 734743. Copies of these files may be retrieved free of charge from The Director, CCDC, 12 Union Road, Cambridge, CB2 1EZ, UK, fax +44123−336−033, e-mail deposit@ccdc.cam.ac.uk or http:www.ccdc.cam.ac.uk.

Results and Discussion

Eight lamivudine conformers are found in the helix structure (Figure 1b). Four of them are protonated (labeled lamivudine A, D, E, and H) and the other four are neutral lamivudine fragments (labeled lamivudine B, C, F, and G). The pyrimidine-derived bases are present with an anti conformation relative to the oxathiolane system, similarly to cytidine residues of *i*-motif DNA.[5−8] A conformation resembling the C2′-endo of canonical nucleotides is observed for lamivudine molecules B, C, and D. The lamivudine units E, F, G, and H adopt an oxathiolane conformation similar to the C3′-endo puckering that occurs in the *β*-2′-deoxy-D-ribofuranose of cytidine in DNA. The lamivudine molecule A has two different modes of puckering in the oxathiolane ring, resembling the C3′-endo and C3′-exo puckered sugars. Disordered lamivudine molecules are also found in the crystal structure of lamivudine form I,[17] a 0.2-hydrate crystallizing in the $P2_12_12_1$ space group, because there are two modes of puckering in the five-membered ring similar to those of the conformer A present in lamivudine duplex. In the crystal structure of the oligonucleotide d(CCCC),[6] which is suitable for comparisons to lamivudine duplex because both structures are stabilized by hydrogen bonds between neutral and protonated cytosine fragments, the C4′-exo and C3′-endo puckers prevail in this *i*-motif DNA, although the C2′-endo puckering occurs in several sugar residues.

Similarly to lamivudine duplex, most cytidine sugars are present with the C3′-endo puckering mode in d(ACCCT), which is another *i*-motif DNA.[8] The conformation of the hydroxymethylene branch is also different for lamivudine molecules assembled into the double-stranded helical structure. The arrangement of the hydroxyl oxygen and methylene hydrogen atoms around the methylene sp^3-hybridized carbon helps to describe the hydroxymethylene conformation. On the positioning of these substituents attached to the methylene carbon, some questions make the description of hydroxymethylene conformation straightforward. These questions include the following: (1) With the oxathiolane plane taken as a reference, is a substituent in the axial position opposite the cytosine fragment? (2) Is the hydroxyl oxygen in the axial position? (3) If this oxygen atom is in an equatorial position, is it opposite the sulfur atom? Concerning the lamivudine A, the hydroxyl oxygen is in the equatorial position opposite the sulfur atom. At the same time, one of the two hydrogen atoms of the methylene group is in the axial position opposite the amino-pyrimidinone base. In lamivudine B, the hydroxyl group is disordered over two occupancy sites. The hydroxyl oxygens filling the major occupancy site (60%) are in the axial position on the same side of cytosine ring. The other hydroxyl oxygens (occupancy value of 40%) are in the equatorial position that is *cis* oriented relative to the sulfur atom of lamivudine unit B. In contrast, the hydroxyl oxygen of molecule C is in the equatorial position opposite the sulfur atom.

Table 1. Relevant Torsion Angles (deg) and Related Conformational Parameters of the Lamivudine Conformers Present in the Lamivudine Duplex

	torsion					
conformer	O1′−C1′−N1−C2	conformation[a]	O1′−C1′−C2′−S3′	conformation[b]	O1′−C4′−C5′−O5′	conformation[c]
A	156.9(4)	anti	26(1)[d]	C3′-endo	−6(1)	equatorial trans
			−23(1)[e]	C3′-exo		
B	143.6(4)	anti	−43.6(5)	C2′-endo	65.1(7)[f]	axial cis
					−166.4(8)[g]	equatorial cis
C	154.9(4)	anti	−48.0(4)	C2′-endo	−59.0(7)	equatorial trans
D	147.9(4)	anti	−42.6(5)	C2′-endo	−171.4(5)[f]	equatorial cis
					−61(1)[g]	equatorial trans
E	163.1(4)	anti	36.0(4)	C3′-endo	−172.5(5)[f]	equatorial cis
					64.2(7)[g]	axial cis
F	154.5(4)	anti	33.3(4)	C3′-endo	65.8(5)[f]	axial cis
					−179.9(5)[g]	equatorial cis
G	156.6(4)	anti	33.4(4)	C3′-endo	63.3(6)	axial cis
H	160.5(4)	anti	32.5(4)	C3′-endo	176.3(3)	equatorial cis

[a] Conformation of the cytosine base relative to the oxathiolane ring. [b] Puckered conformation of a oxathiolane ring compared with that in the fundamental nucleotides. [c] Conformation of the hydroxymethylene branch where the hydroxyl oxygen is either cis or trans oriented relative to the sulfur atom (when in an equatorial position) and the amino-pyrimidinone group (when in the axial position). [d] The fractions of the C2′a and S3′a atoms are in the major 60% occupancy sites. [e] The fractions of the C2′a′ and S3′a′ atoms are in the minor 40% occupancy sites. [f] The fraction of the O5′ atom is in the major 60% occupancy site. [g] The fraction of the O5′ atom is in the minor 40% occupancy site.

The axial position, in which there is a methylene hydrogen, is on the same side of the cytosine fragment of molecules C and D. In the last lamivudine conformer, the atomic fractions of the disordered hydroxyl oxygen are in the two equatorial positions. These two oxygen fractions are on the same (60% major occupancy) and opposite (40% minor occupancy) sides of the sulfur atom. Regarding molecules E and F, their hydroxyl groups also occupy two positions. In these conformers, one of the two hydroxyl oxygen fractions is in the equatorial position, on the same side of the sulfur atom. The other oxygen fraction is oriented axially relative to the cytosine ring. However, the major occupancy site (60%) for the oxygen fraction of molecule E is in the above-mentioned equatorial position, whereas the corresponding hydroxyl oxygen fraction of molecule F is in the axial position. In molecules G and H, the substituent in the axial position is *cis* located relative to the cytosine fragment. This substituent is a hydroxyl oxygen in molecule G and one of the two methylene hydrogen atoms in conformer H. The hydroxyl oxygen of the lamivudine H is equatorially placed on the same side of the sulfur atom. The most relevant torsion angles describing the conformation of lamivudine conformers and the crystal data are given in Tables 1 and 2, respectively. A scanning electronic micrograph of the lamivudine duplex crystals is displayed in Figure 2.

Four 3TC−3TC$^+$ base pairs give rise to one-half of a helical turn. The occurrence of eight lamivudine pairs spaced 3.2 Å apart in each entire turn of the lamivudine duplex measuring about 25.6 Å (Figure 1b,c) revealed that the helix rise per base pair is similar to the stacking distance in *i*-motif DNA (3.0−3.2 Å).[6−8] The least-squares plane of the lamivudine pairs is slightly bent by about 3° relative to the helical axis. Constraint of base pair tilts is also reported for *i*-motif DNA structures.[5,6] According to the labeling of lamivudine conformers, the base pairing occurs between the units A and B (Figure 3), C and D (Figure 4), E and F (Figure 5), and G and H (Figure 6). If we take the average of these four lamivudine pairs, the distances between the hydrogen bonding donors and acceptors of the two peripheral N−H$_{(3TC^+)}$···O$_{(3TC)}$ and N−H$_{(3TC)}$···O$_{(3TC^+)}$ and the central N$^+$−H···N hydrogen bonds measure 2.77(4), 2.89(5), and 2.83(1) Å, respectively. These measurements are slightly different than those of *i*-motif DNA. In the crystal structures of the d(CCCC)[6] and d(CCCT)[7] C-rich strands of *i*-motif DNA, which form intercalated parallel duplexes held together by hemiprotonated C−C$^+$ base pairs, the two hydrogen-bonding distances between the exocylic amino and carbonyl groups are similar. On average, the N···O distances measure 2.75(10) and 2.77(8) Å in d(CCCC)[6] and d(CCCT),[7] respectively, while the mean N$^+$···N distance of the central hydrogen bond is 2.76(8) Å in d(CCCC)[6] and 2.74(8) Å in d(CCCT).[7] The

Table 2. Crystal Data of the Lamivudine Helix at Room and Low Temperatures and Statistics for Structure Determination and Refinement

		lamivudine double helix	
temp (K)		298 ± 1	150.0 ± 0.2
unit cell dimensions	*a* (Å)	14.2817(5)	14.2219(2)
	b (Å)	26.0361(6)	25.6079(3)
	c (Å)	15.1994(5)	15.1754(2)
	β (deg)	117.3968(12)	117.4652(5)
volume (Å^3)		5017.9(3)	4903.9(1)
calcd density (g cm^{-3})		1.501	1.538
data collected		38 282	64 288
θ range for data collection (deg)		2.96−25.35	2.97−25.42
index ranges	*h*	−16 to 16	−17 to 17
	k	−30 to 31	−30 to 30
	l	−18 to 15	−18 to 18
unique reflns		17 481	17 829
symmetry factor (R_{int})		0.1041	0.0952
completeness to θ_{max}		96.6%	98.6%
unit cell content		8($C_8H_{11}N_3O_3S$)-8($C_8H_{12}N_3O_3S$)-4($C_4H_3O_4$)2(C_3H_8O)8(H_2O)4Cl	
3TC−3TC$^+$ pairs per helix turn		8	
crystal system		monoclinic	
space group		$P2_1$	
Refinement (low temp)			
abs. coeff. (mm^{-1})		0.335	
abs. correction method		Gaussian	
	T_{min}	0.884	
	T_{max}	0.964	
F (000)		2380	
params refined		1367	
goodness-of-fit on F^2		1.103	
final R factors for $I > 2\sigma(I)$		R1 = 0.0587, wR2 = 0.1518	
R factors for all data		R1 = 0.0789, wR2 = 0.1719	
largest diff. peak/hole (*e* Å^{-3})		0.778/−0.670	
absolute structure	factor	Flack parameter	
	refined	0.05(5)	
	data	8686 Friedel pairs	

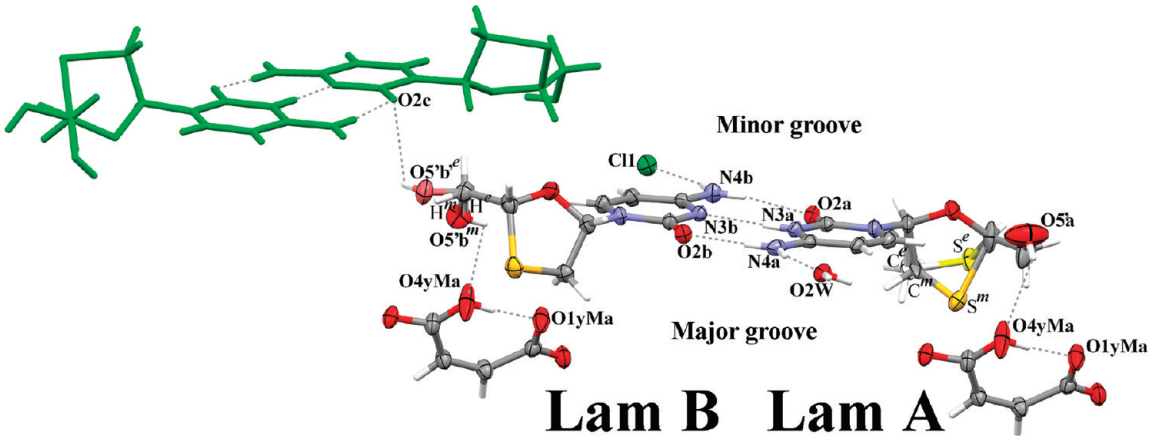

Figure 3. Hydrogen-bonding pattern of the lamivudine conformers A and B found in the double helix structure. The hydrogen atoms were drawn in a capped stick fashion, and thermal ellipsoids at the 30% probability level represent the other atoms. Green sticks are used to depict the lamivudine molecules on a neighboring duplex, which interact with the hydroxymethylene branch of a conformer in the framed pair whose atoms are represented as 30% probability ellipsoids. Only the labels of the atoms involved in hydrogen bonds are shown. Italic superscript lower-case letters *m* and *e* refer to the fractions of atoms disordered over two positions in the main (60% S.O.F.) and extra (40% S.O.F.) occupancy sites, respectively. Slightly highlighted ellipsoids represent carbon, oxygen, and sulfur atom fractions in the minor 40% occupancy sites.

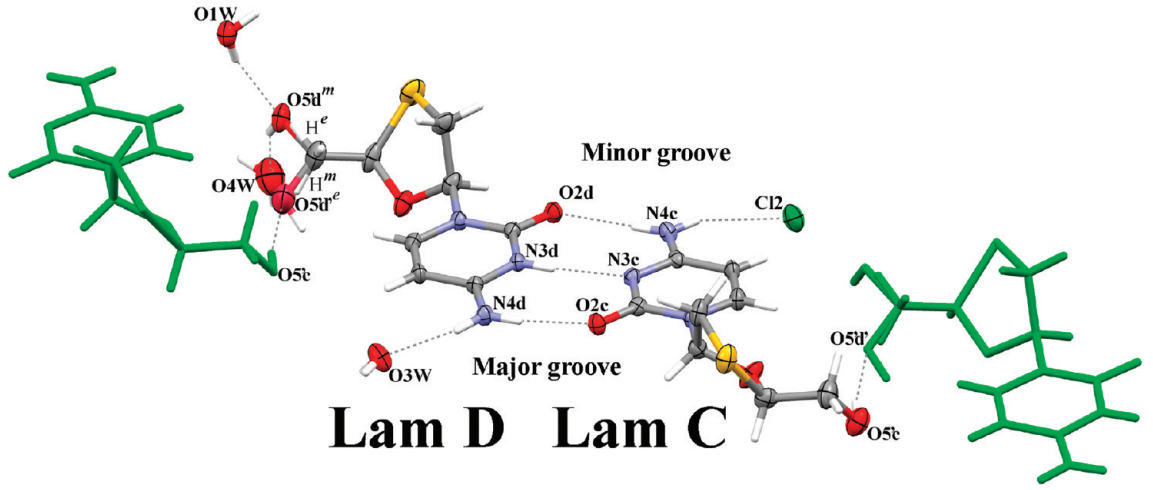

Figure 4. Hydrogen bonds of the lamivudine conformers C and D assembling the double helix structure. For drawing details, see the caption of Figure 3.

dimers A≡B, C≡D, E≡F, and G≡H are stacked face-to-face on top of each other in this order. Here, two important parallels are established between lamivudine duplex and *i*-motif DNA quadruplexes. First, the positive charge is disordered within the C−C$^+$ pairs of *i*-motif DNA due to transfer of the imino proton from one strand to the other (Figure 1a).[5,6] Contrarily, protonated and neutral cytosine fragments are alternated into each strand of lamivudine duplex. When describing the tetrameric structure of *i*-motif DNA, Gehring and co-workers[5] anticipated that the protonation on alternate cytosine bases would be preferred in a duplex structure over the proton movement occurring in a quadruplex because the distance between charged cytosines should be maximized. In agreement with this hypothesis, lamivudine duplex is present with protonated cytosine fragments alternately placed into each strand. This confirms that a duplex held together by pairing of protonated and neutral cytosine rings has their imino protons fixed on alternate bases. Second, lamivudine duplex and *i*-motif DNA are highly related concerning the stacking of their building blocks. Each cytosine base pair is stacked face-to-face with its neighbor in both the lamivudine duplex and the *i*-motif DNA quadruplex structures (Figure 1d,e).[5] In *i*-motif DNA, this is a consequence of intercalating two parallel duplexes in which each double-stranded structure is antiparallel relative to one another. Lamivudine duplex demonstrates that face-to-face stacking can also occur in a double-stranded helix held together by hemiprotonated cytosine−cytosine base pairing. Similar to Z-DNA, the lamivudine double helix structure is infinitely left-handed, and there is not another helical orientation except that. On the other hand, C-rich four-stranded structures of *i*-motif DNA are slightly right-handed.[6−8] Major and minor grooves resembling those of B-DNA were apparent on the surface of the lamivudine double helix (Figure 1c). Shapes of the lamivudine duplex grooves differ from those of *i*-motif DNA. In the C-rich quadruplex stabilized by hemiprotonated C−C$^+$ pairs, very wide grooves are created on two sides and very narrow grooves are at its ends,[6,8] while major and minor grooves of similar depth occur in lamivudine duplex. Furthermore, different helical radii are measured for the outermost hydroxyl groups of lamivudine

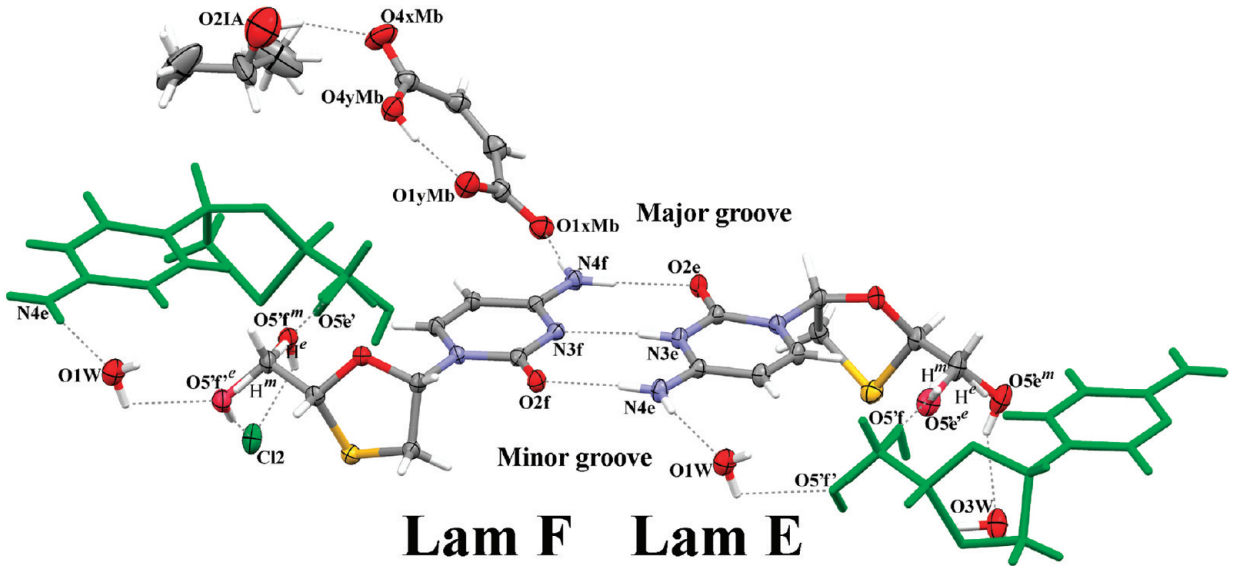

Figure 5. The conformers E and F present in the lamivudine duplex are hydrogen-bonded in the periphery of the fiber through the hydroxyl groups and in the major and minor grooves through the amine moieties. For drawing details, see the caption of Figure 3.

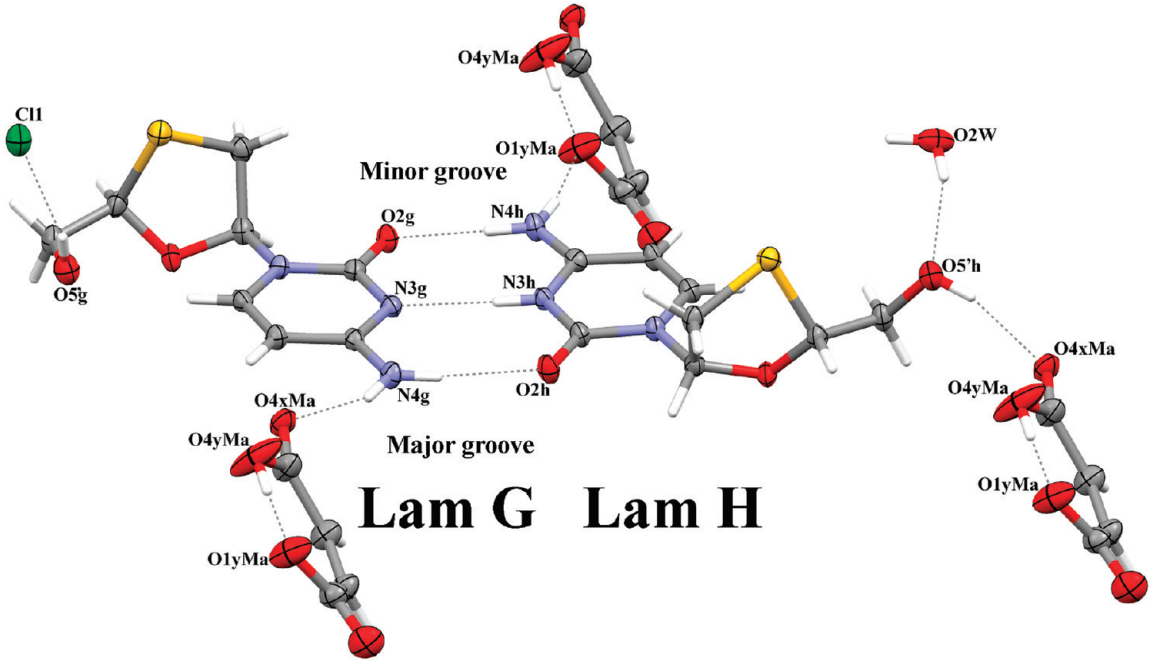

Figure 6. Scheme of hydrogen bonds of the lamivudine conformers G and H that compose the double-stranded helical structure. The hydrogen atoms were depicted as white sticks. The thermal ellipsoids at the 30% probability level represent the other atoms. Only the labels of the atoms engaged in hydrogen bonds were displayed.

units (Figure 1f), with distances of 11.8(2), 10.7(1), 9.4(1), and 9.2(1) Å from the helix center up to the hydroxyl oxygen atoms of lamivudine molecules H, B, E, and C, respectively. Chloride and partially ionized maleate anions are present in the duplex with the lamivudine pairs. Isopropanol and water molecules also crystallized in the structure. Noncovalent interactions play a crucial role in the stabilization of the lamivudine double helix (see in sequence). Other intermolecular contacts than base−base interactions contribute to the stabilization of the lamivudine duplex structure in the fiber periphery (Figure 1g). In the grooves of the helical structure, there are hydrogen bonds between the amino groups of all lamivudine subunits with hydrogen bond acceptors outside of the fiber backbone. Among the anions and solvents surrounding the grooves of the double helix structure, maleate and water molecules lie in the major groove (Figure 1h). Chloride anions, water, and maleate molecules lie in the minor groove of the helix surface (Figure 1i). Crystal structure analysis reveals that water molecules are hydrogen bonded to the amino groups of cytidine through their hydrogens, which are not involved in the $C-C^+$ pairing of *i*-motif DNA.[6] Resembling the lamivudine duplex structure, water molecules also surround the wide grooves of *i*-motif DNA. However, the minor grooves of lamivudine duplex are more hydrated than the narrow grooves of *i*-motif DNA, in which water molecules stay near phosphate groups.[6] Regarding the hydration pattern of lamivudine in their reported crystalline modifications, form I[17] (a 0.2-hydrate, see above), form III[18] (a 0.5-hydrate),

and the cocrystals with zidovudine[4] and 3,5-dinitrosalicylic acid[4] are stabilized through hydrogen bonds between water molecules and the cytosine rings of the drug. Likewise, this was observed in lamivudine duplex. On the other hand, form II,[17] the saccharinate salt[10] and the cocrystal with 4-quinolinone[4] are anhydrous phases of lamivudine that form layered structures. Base stacking gives rise to layers in lamivudine duplex. However, neither determined crystalline phase of lamivudine has a structure similar to that of lamivudine duplex in which the hydrogen-bonded 3TC−3TC$^+$ pairs are helically stacked on top of each other.

The hydroxyl groups of all lamivudine conformers are involved in hydrogen bonds. One of the two crystallographically independent molecules of maleate (labeled maleate A) is the hydrogen acceptor in three hydrogen bonds in which the hydroxyl groups of the conformers A, B, and H are hydrogen donors. In the case of molecule B, the hydroxyl group in the major 60% occupancy site is the hydrogen donor to maleate A, whereas the remaining fraction in the 40% occupancy site is involved in a hydrogen bond with a 3TC−3TC$^+$ pair of a neighboring lamivudine double helix. In the crystal, the hydroxymethylene branch of lamivudine unit B lies in the major groove of an adjacent duplex structure (Figure 7a). Thus, the positional disorder of the hydroxyl group of this conformer seems to be a result of the close packing of the two double-stranded helices. The hydrogen donation occurs from the extra part of the hydroxyl group to the carbonyl oxygen of a molecule C cytosine fragment. In this way, the hydrogen bond acceptor group of one of the three hydrogen bonds responsible for the C≡D pairing also interacts with a lamivudine molecule in an adjacent layer of another double helix structure. This observation allows us to state that the center groups pairing the cytosine fragments also can weakly bond to molecules accommodated in the major groove. The lamivudine H is also the hydrogen acceptor of a water molecule. Two water molecules interact with the main 60% occupancy hydroxyl group of the lamivudine molecule D through hydrogen bonds. One of these water molecules is the hydrogen acceptor from the disordered hydroxyl portion. The other is the hydrogen donor to the hydroxyl oxygen in the major occupancy position. The molecule D hydroxyl oxygens in the position with 40% occupancy are hydrogen-bonded to the lamivudine C on another double-stranded helix in the crystal. The hydroxyl group of molecule C acts as a hydrogen donor to the disordered hydroxyl oxygen of molecule D (Figure 7b). Another water molecule is the hydrogen acceptor for the high occupancy (60%) hydroxyl group of lamivudine E. The lower occupancy (40%) hydroxyl of lamivudine E interacts with the main hydroxyl group fraction of the lamivudine conformer F (Figure 7c). The high occupancy (60%) hydroxyl group of molecule F contributes to structural stabilization through two hydrogen bonds. In one of them, it is a hydrogen acceptor for the disordered hydroxyl group of molecule E. In the other hydrogen bond, it is a hydrogen donor to one chloride anion (Figure 5). Another chloride anion is hydrogen-bonded to the hydroxyl group of lamivudine G. In the case of molecule F, the fraction of the hydroxyl moiety in the minor occupancy site (40%) is also a hydrogen bond donor to the same chloride anion that interacts with the corresponding fraction in the 60% occupancy position. Therefore, the disorder in the hydroxyl group can be explained by the fact that both hydroxyl group conformations are equivalent in terms of orientation of the hydrogen bonds with one chloride anion (see also Table 3). However, the hydroxyl oxygens in the minor

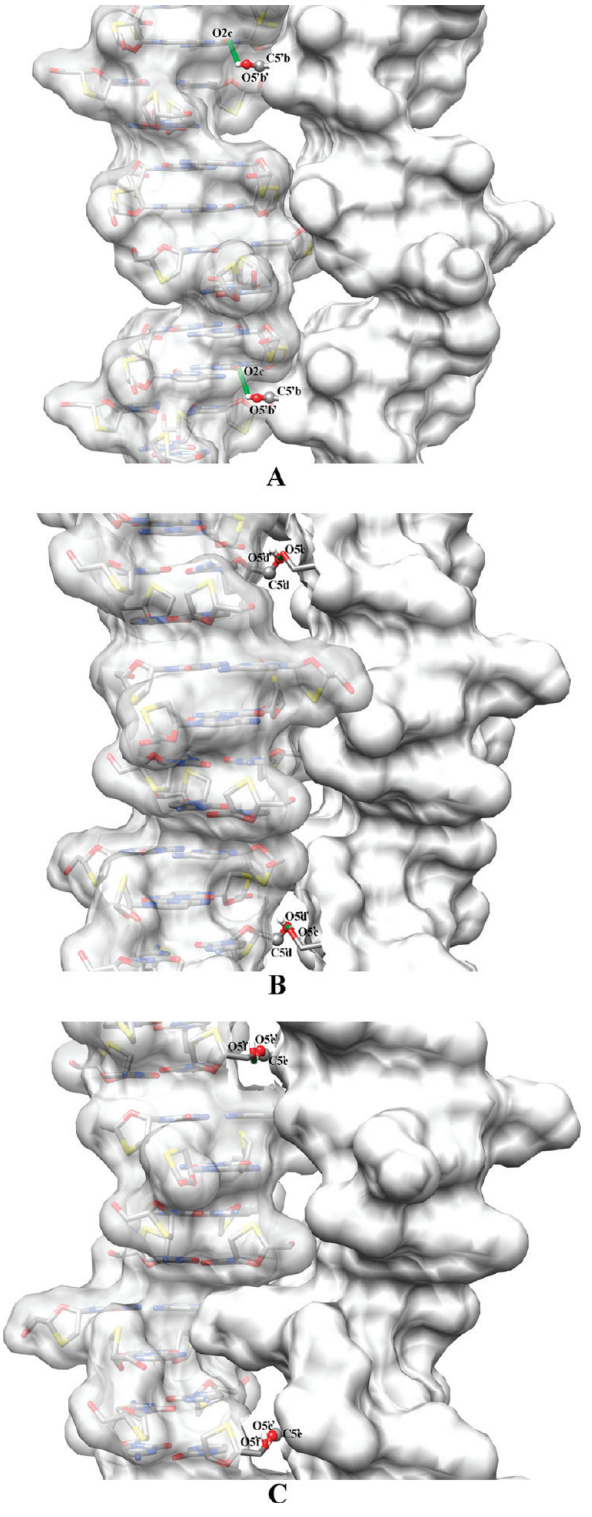

Figure 7. The hydroxyl oxygen and hydrogen atoms of some lamivudine conformers were disordered over two positions because the neighboring double-stranded helices were very close together in the crystal. (a) The hydroxyl groups in the 40% occupancy site of the lamivudine conformer B lie in the major groove, where they interact with a C≡D pair through one hydrogen bond. (b) The hydroxyl groups in the 40% occupancy site of the lamivudine molecule D are hydrogen-bonded to the corresponding groups of conformer C on a neighboring double-stranded helix. (c) The hydroxyl groups in the 40% and 60% occupancy sites of the lamivudine conformers E and F, respectively, connect two adjacent helices through a hydrogen bond.

Table 3. Hydrogen Bond Geometry in the Lamivudine Double-Stranded Helix

interaction	D−H···A[a]	D−H (Å)	H···A (Å)	D···A (Å)	D−H···A (deg)
A≡B pairing	$N_{4a}-H_{4Nxa}\cdots O_{2b}$	0.86	1.88	2.738(5)	176
	$N^{+}_{3a}-H_{3a}\cdots N_{3b}$	0.86	1.95	2.810(6)	173
	$N_{4b}-H_{4Nxb}\cdots O_{2a}$	0.86	2.08	2.928(5)	169
C≡D pairing	$N_{4c}-H_{4Nxc}\cdots O_{2d}$	0.86	1.98	2.834(5)	170
	$N^{+}_{3d}-H_{3d}\cdots N_{3c}$	0.86	1.98	2.832(6)	170
	$N_{4d}-H_{4Nxd}\cdots O_{2c}$	0.86	1.93	2.787(5)	173
E≡F pairing	$N_{4e}-H_{4Nxe}\cdots O2_{f}$	0.86	1.96	2.814(6)	175
	$N^{+}_{3e}-H_{3e}\cdots N_{3f}$	0.86	1.98	2.829(7)	172
	$N_{4f}-H_{4Nxf}\cdots O_{2e}$	0.86	2.00	2.858(6)	176
G≡H pairing	$N_{4g}-H_{4Nxg}\cdots O_{2h}$	0.86	2.06	2.920(6)	174
	$N^{+}_{3h}-H_{3h}\cdots N_{3g}$	0.86	1.98	2.836(7)	173
	$N_{4h}-H_{4Nxh}\cdots O_{2g}$	0.86	1.88	2.737(6)	178
		duplex−duplex			
$OH_{(extra)}$(B)−C≡D	$O_{5'b'}-H_{5'Ob'}\cdots O_{2c}$	0.82	2.72	3.299(7)	129
OH(C)−$OH_{(extra)}$(D)	$O_{5'c}-H_{5'Oc}\cdots O_{5'd'}$	0.82	2.15	2.957(8)	168
$OH_{(extra)}$(E)−OH(F)	$O_{5'e'}-H_{5'Oe'}\cdots O_{5'f}$	0.82	1.53	1.90(1)	104
		duplex−maleate			
OH(A)−maleate(A)	$O_{5'a}-H_{5'Oa}\cdots O_{4yMa}$	0.82	2.61	3.30(1)	143
OH(B)−maleate(A)	$O_{5'b}-H_{5'Ob}\cdots O_{4yMa}$	0.82	2.36	2.61(1)	98
NH_2(F)−maleate(B)	$N_{4f}-H_{4Nyf}\cdots O_{1xMb}$	0.86	1.99	2.842(6)	172
NH_2(G)−maleate(A)	$N_{4g}-H_{4Nyg}\cdots O_{4yMa}$	0.86	2.23	3.053(6)	160
OH(H)−maleate(A)	$O_{5'h}-H_{5'Oh}\cdots O_{4xMa}$	0.82	1.93	2.725(6)	164
NH_2(H)−maleate(A)	$N_{4h}-H_{4Nyh}\cdots O_{1yMa}$	0.86	2.18	3.014(6)	165
		duplex−water			
NH_2(A)−H_2O	$N_{4a}-H_{4Nya}\cdots O_{2w}$	0.86	1.98	2.817(7)	163
NH_2(D)−H_2O	$N_{4d}-H_{4Nyd}\cdots O_{3w}$	0.86	2.01	2.834(7)	161
OH(D)−H_2O	$O_{5'd}-H_{5'Od}\cdots O_{4w}$	0.82	2.06	2.846(10)	160
H_2O−OH(D)	$O_{1w}-H_{2w}\cdots O_{5'd}$	0.84	2.15	2.972(9)	163
NH_2(E)−H_2O	$N_{4e}-H_{4Nye}\cdots O_{1w}$	0.86	2.11	2.942(6)	162
OH(E)−H_2O	$O_{5'e}-H_{5'Oe}\cdots O_{3w}$	0.82	2.24	2.871(8)	134
H_2O−$OH_{(extra)}$(F)	$O_{1w}-H_{2w}\cdots O_{5'f}$	0.84	2.33	2.80(1)	115
H_2O−OH(H)	$O_{2w}-H_{4w}\cdots O_{5'h}$	0.85	1.98	2.786(6)	157
		duplex−chloride			
NH_2(B)−Cl^-	$N_{4b}-H_{4Nyb}\cdots Cl_1$	0.86	2.55	3.374(5)	161
NH_2(C)−Cl^-	$N_{4c}-H_{4Nyc}\cdots Cl_2$	0.86	2.49	3.304(5)	158
OH(F)−Cl^-	$O_{5'f}-H_{5'Of}\cdots Cl_2$	0.82	2.40	3.111(5)	146
$OH_{(extra)}$(F)−Cl^-	$O_{5'f'}-H_{5'Of'}\cdots Cl_2$	0.82	2.27	3.061(7)	163
OH(G)−Cl^-	$O_{5'g}-H_{5'Og}\cdots Cl_1$	0.82	2.35	3.144(4)	163
water−maleate	$O_{1w}-H_{1w}\cdots O_{1xMa}$	0.84	1.99	2.796(6)	161
water−chloride	$O_{2w}-H_{3w}\cdots Cl_1$	0.84	2.39	3.123(5)	146
	$O_{3w}-H_{5w}\cdots Cl_1$	0.84	2.37	3.196(4)	168
	$O_{3w}-H_{6w}\cdots Cl_2$	0.84	2.23	3.074(5)	177
	$O_{4w}-H_{7w}\cdots Cl_2$	0.88	2.38	3.130(6)	144
isopropanol−maleate	$O_{2IA}-H_{2OIA}\cdots O_{4xMb}$	0.82	1.88	2.668(11)	162
$maleate_{(intramolecular)}$	$O_{4yMa}-H_{4OMa}\cdots O_{1yMa}$	0.82	1.61	2.426(6)	177
	$O_{4yMb}-H_{4OMb}\cdots O_{1yMb}$	0.82	1.62	2.429(6)	169

[a] D indicates hydrogen donor; A indicates hydrogen acceptor.

occupancy site act as hydrogen bond acceptors for a water molecule in a hydrogen bond that is less geometrically oriented than that of the hydroxyl oxygens in the main occupancy site that form a hydrogen bond with lamivudine molecule E. At last, the amine groups of the lamivudine subunits are hydrogen-bonded to two crystallographically independent maleate anions labeled A (lamivudine G and H) and B (lamivudine F), chloride anions (lamivudine B and C), and water molecules (lamivudine A, D and E). The amine hydrogen atoms of cytosine fragments that are not involved in the hydrogen bonds responsible for the 3TC−$3TC^+$ pairing were involved in these hydrogen bonds.

Conclusions

The assembly of the lamivudine double helix structure is an interesting example of spontaneous molecular self-organization. Similarly, the self-association of guanosine 5′-monophosphate (5′-GMP) occurs in a neutral solution, giving rise to a right-handed quadruple helix of just one canonical nucleotide.[19] However, here we report a double-stranded helical structure of a nucleoside analog in the solid state for the first time. Stacking and hydrogen-bonding interactions between the cytosine bases of lamivudine were the decisive forces responsible for this interesting structural organization in which alternating neutral and protonated lamivudine conformers are placed into each helical strand. Other contacts such as hydrogen bonds involving duplex−duplex, duplex−anion (chloride and maleate), and duplex−water interactions were also important for the molecular assembly. The occurrence of the lamivudine double helix structure indicates that the 5′-phosphate group could not be a requirement for duplex assembly. Hydrogen bonds can play the role of covalent phosphodiester linkages in stabilizing the duplex. In summary, the expected protonation pattern on alternate bases was demonstrated for a DNA-like duplex stabilized by pairing of neutral and protonated cytosine fragments. Indeed, lamivudine duplex exhibits other similarities to *i*-motif DNA than the

self-assembly of the hemiprotonated cytosine−cytosine base pairs into multimeric helical structures, such as the face-to-face stacking of the cytosine rings and the helix rise per base pair. Last, we believe that this study provides insights into the understanding of DNA structure assembly.

Acknowledgment. We thank Altivo Pitaluga, Jr. (Fundação Oswaldo Cruz - FIOCRUZ, Manguinhos, Rio de Janeiro, Brazil), for the gift of lamivudine samples. We thank the Brazilian Research Council CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant 2007/07185-5) for research fellowships (A.C.D., J.E., and F.T.M.). We also thank FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais, Grants APQ-2011-5.02/07, APQ-6010-5.02/07, APQ-02685-09) and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Grant PNPD2008) for financial support.

Supporting Information Available: Crystallographic information in CIF format. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Coates, J. A.; Cammack, N.; Jenkinson, H. J.; Jowett, A. J.; Jowett, M. I.; Pearson, B. A.; Penn, C. R.; Rouse, P. L.; Viner, K. C.; Cameron, J. M. *Antimicrob. Agents Chemother.* **1992**, *36*, 733–739. (b) Kukhanova, M.; Liu, S. H.; Mozzherin, D.; Lin, T. S.; Chu, C. K.; Cheng, Y. C. *J. Biol. Chem.* **1995**, *270*, 23055–23059.

(2) Chang, C. N.; Doong, S. L.; Zhou, J. H.; Beach, J. W.; Jeong, L. S.; Chu, C. K.; Tsai, C. H.; Cheng, Y. C.; Liotta, D.; Schinazi, R. *J. Biol. Chem.* **1992**, *267*, 13938–13942.

(3) (a) Feng, J. Y.; Shi, J.; Schinazi, R. F.; Anderson, K. S. *FASEB J.* **1999**, *13*, 1511–1517. (b) Rezende, L. F.; Drosopoulos, W. C.; Prasad, V. R. *Nucleic Acids Res.* **1998**, *26*, 3066–3072. (c) Sarafianos, S. G.; Das, K.; Clark, A. D., Jr.; Ding, J.; Boyer, P. L.; Hughes, S. H.; Arnold, E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10027–10032.

(4) Bhatt, P. M.; Azim, Y.; Thakur, T. S.; Desiraju, G. R. *Cryst. Growth Des.* **2009**, *9*, 951–957.

(5) Gehring, K.; Leroy, J. L.; Gueron, M. *Nature* **1993**, *363*, 561–565.

(6) Chen, L.; Cai, L.; Zhang, X.; Rich, A. *Biochemistry* **1994**, *33*, 13540–13546.

(7) Kang, C. H.; Berger, I.; Lockshin, C.; Ratliff, R.; Moyzis, R.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11636–11640.

(8) Weil, J.; Min, T.; Yang, C.; Wang, S.; Sutherland, C.; Sinha, N.; Kang, C. *Acta Crystallogr., Sect. D* **1999**, *55*, 422–429.

(9) (a) Mergny, J. L.; Lacroix, L.; Han, X.; Leroy, J. L.; Helene, C. *J. Am. Chem. Soc.* **1995**, *117*, 8887–8898. (b) Singh, S.; Hosur, R. V. *J. Am. Chem. Soc.* **1995**, *117*, 12637–12638. (c) Leroy, J. L. *Nucleic Acids Res.* **2009**, *37*, 4127–4134.

(10) Banerjee, R.; Bhatt, P. M.; Ravindra, N. V; Desiraju, G. R. *Cryst. Growth Des.* **2005**, *5*, 2299–2309.

(11) *COLLECT Data Collection Software*; Nonius: Delft, 1998.

(12) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307–326.

(13) Burla, M. C.; Caliandro, R.; Camalli, M.; Carrozzini, B.; Cascarano, G. L.; De Caro, L.; Giacovazzo, C.; Polidori, G.; Spagna, R. *J. Appl. Crystallogr.* **2005**, *38*, 381–388.

(14) Sheldrick, G. M. *SHELXL-97: Program for Crystal Structure Analysis*; University of Göttingen: Göttingen, 1997.

(15) Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M. K.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. *Acta Crystallogr., Sect. B* **2002**, *58*, 389–397.

(16) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605–1612.

(17) Harris, R. K.; Yeung, R. R.; Lamont, R. B.; Lancaster, R. W.; Lynn, S. M.; Staniforth, S. E. *J. Chem. Soc., Perkin Trans. 2* **1997**, 2653–2659.

(18) Singh, G. P.; Srivastava, D.; Saini, M. B.; Upadhyay, P. R. International Publication No. WO2007119248, October 25 **2007**.

(19) Wu, G.; Kwan, I. C. M. *J. Am. Chem. Soc.* **2009**, *131*, 3180–3182.