2003

Effect of Sequence on the Conformation of DNA Holliday Junctions

Franklin A. Hays

Jeff Vargason
George Fox University, jvargason@georgefox.edu

P. Shing Ho
hops@onid.orst.edu

Recommended Citation
Previously published in Biochemistry, 2003, volume 42, pp. 9586-9597 http://pubs.acs.org/doi/abs/10.1021/bi0346603

Effect of Sequence on the Conformation of DNA Holliday Junctions

Franklin A. Hays, Jeffrey M. Vargason,‡ and P. Shing Ho*

Department of Biochemistry and Biophysics, ALS 2011, Oregon State University, Corvallis, Oregon 97331

ABSTRACT: Structures of the DNA sequences d(CCGGCGCCGG) and d(CCAGTACbr^{5}UGG) are presented here as four-way Holliday junctions in their compact stacked-X forms, with antiparallel alignment of the DNA strands. Thus, the *ACC*-trinucleotide motif, previously identified as important for stabilizing the junction, is now extended to PuCPy, where Pu is either an adenine or guanine, and Py is either a cytosine, 5-methylcytosine, or 5-bromouracil but not thymine nucleotide. We see that both sequence and base substituents affect the geometry of the junction in terms of the interduplex angle as well as the previously defined conformational variables, J_{roll} (the rotation of the stacked duplexes about their respective helical axis) and J_{slide} (the translational displacement of the stacked duplexes along their respective helical axis). The structures of the GCC and parent ACC containing junctions fall into a distinct conformational class that is relatively undistorted in terms of J_{slide} and J_{roll}, with interduplex angles of 40−43°. The substituted ACbr^{5}U structure, however, is more akin to that of the distorted methylated ACm^{5}C containing junction, with J_{slide} ($\geq$2.3 Å) and a similar J_{roll} (164°) opening the major groove-side of the junction, but shows a reduced interduplex angle. In contrast, the analogous d(CCAGTACTGG) sequence has to date been crystallized only as resolved B-DNA duplexes. This suggests that there is an electronic effect of substituents at the pyrimidine Py position on the stability of four-stranded junctions. The single-crystal structures presented here, therefore, show how sequence affects the detailed geometry, and subsequently, the associated stability and conformational dynamics of the Holliday junction.

The exchange of genetic information across double-helical DNA through recombination is an important process in DNA biochemistry and has been implicated in the mechanisms of DNA repair, replication restart, and viral integration (*1*). The critical role of a four-stranded junction as a DNA intermediate in recombination was proposed nearly 40 years ago by Holliday (*2*). The structure of this four-stranded complex has been elucidated in detail through a series of single-crystal structures as complexes with proteins (*1*), and more recently, as isolated DNA constructs (*3*). All of the DNA junctions that have been crystallized to date are decanucleotides with a common ACC trinucleotide motif (or its methylated variant ACm^{5}C)[1] at the $N_6N_7N_8$ positions (represented by the sequence d(CCGGT*ACC*GG), where the underlined nucleotides are the ACC junction core). Not surprisingly, these structures have many common structural features. All adopt the compact antiparallel stacked-X form in which pairs of the duplex arms stack collinearly into nearly continuous double-helices (broken only at the crossover point of the junction on the inside strand of each pair, Figure 1). The structure with a methylated analogue of the ACC core shows that the substituent group perturbs the stacked-X geometry. Here, we present the single-crystal structures as Holliday junctions of the two sequences d(CCGGC*GCC*GG) and d(CCAGT*AC*br^{5}UGG) that do not contain the ACC junction motif at the core trinucleotide (underlined) positions.

The compact stacked-X form of the Holliday junction was first proposed from gel mobility assays (*4*, *5*) and fluorescence resonance energy transfer studies in solution (*6*−*8*) to form under high salt conditions in nonhomologous sequence constructs that immobilize the four-way junction at specific sites. The general model from these studies is that high concentrations of cations help to screen the negatively charged phosphates of the backbone allowing the four arms, which would be splayed out from the junction in the open-X form (Figure 1a), to pair and coaxially stack to form two sets of semicontinuous double-helices (Figure 1b). The solution studies also showed that these pairs of stacked duplexes are related across the junction crossover by a right-handed interduplex angle of ~60°. This general geometric relationship was confirmed by atomic force microscopy (*9*), but the most detailed molecular details of this conformation have come only from single-crystal X-ray diffraction studies.

The structure of the Holliday junction has now been determined in a number of nucleic acid constructs, including RNA-DNA complexes (*10*, *11*) and DNA constructs with true inverted repeat patterns (*12*), mismatched base pairs (*13*), the cross-linking drug psoralen (*14*), and methylated cy-

† This work was funded by grants from the National Institutes of Health (R1GM62957A) and the National Science Foundation (MCB0090615). The X-ray diffraction facilities were supported by the Proteins and Nucleic Acids Facility Core of the Environmental Health Sciences Center at OSU (NIEHS ES00210) and a grant from the Murdock Charitable Trust. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-Eng-38. Use of BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under Grant RR07707.

* To whom correspondences should be addressed. Phone: (541) 737-2769. Fax: (541) 737-0481. E-mail: hops@onid.orst.edu.

‡ Current address: Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709.

[1] Abbreviations: br^{5}U, 5-bromouracil; m^{5}C, 5-methylcytosine; FRET, fluorescence resonance energy transfer; Pu, purine; Py, pyrimidine.

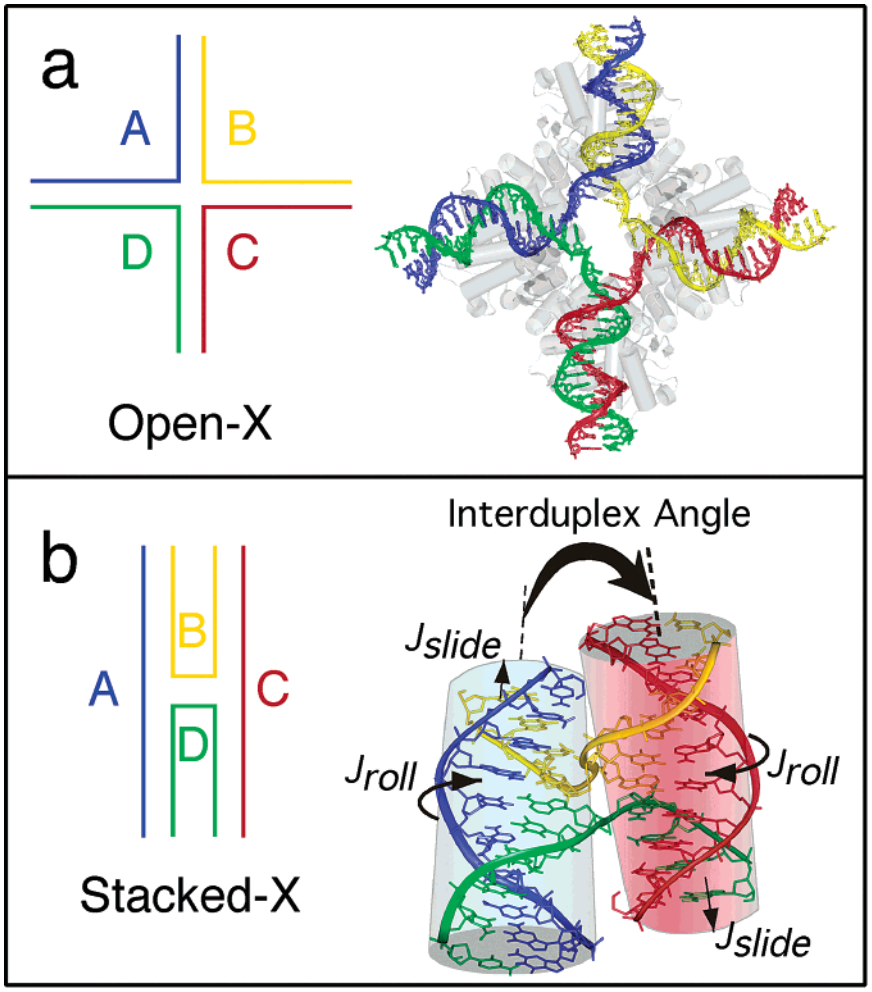

FIGURE 1: Open-X (a) and stacked-X (b) forms of the Holliday junction. The phosphodeoxyribose backbones of the DNAs are traced with ribbons and color-coded to distinguish between strands of the junction. (a) The open-X form modeled from the DNA in the single-crystal structure of the Cre-loxP complex (*35*). The protein is shown in gray, with the α-helices rendered as cylinders. (b) The compact stacked-X structure of the reference d(CCGGTACCGG) single-crystal structure (*12*). The coaxially stacked arms in each column of pseudo-continuous duplexes are shown as cylinders. The conformational parameters of the junction are shown as the twist of the stacked duplexes across the junction (the interduplex angle), the translation of each set of stacked duplexes along their respective helix axes (J_{slide}), and the rotation of the duplexes about the axes (J_{roll}).

tosines (*15*). The DNA structures show that the geometry of the compact stacked-X junction (Figure 1b) was basically that determined in solution, with the coaxially stacked duplexes related by a right-hand rotation across the junction crossover. In all but one case so far, the DNA junction structures were determined from decanucleotide sequences with a core ACC trinucleotide (or its methylated analogue, ACm^{5}C) at the $N_6N_7N_8$ positions, as in the sequences d(CCGGG*ACC*GG) (*13*), d(CCGGT*ACC*GG) (*12*), and d(TCGGT*ACC*GA) (*16*). The one exception to date is the structure of the psoralen cross-linked sequence d(CCGCT*-AGCGG), which was proposed to form a junction because of the destabilization of the duplex by the drug cross-linked thymine base (T*) (*14*). Otherwise, the crystal structures demonstrate the significance of this core ACC trinucleotide, particularly through the interactions from the cytosine base at the C_8 cytosine to the phosphate backbone at the junction crossovers.

The methylated junction seen in d(CCGGTACm^{5}CGG) first showed that, in addition to the interduplex angle, the geometry of the junction can also be affected along and around the helix axes of the collinear duplex arms (*15*). The methyl substituent was shown to disrupt the direct hydrogen bonding interactions between the cytosine C_8 base and the phosphate backbone of the crossover. Furthermore, the accessibility of the major and minor grooves of the stacked duplex arms were shown to depend on the rotation of these duplexes along their helix axes (J_{roll}), and the arms were seen to slide along their axes (J_{slide}) to affect the overall symmetry of the complex. Similar distortions to J_{roll} were seen to be affected by the types of cation interactions localized at the junction (*16*).

The interduplex angle relating the orientation of the stacked duplex arms across all of these ACC-type junctions are relatively constant, at ∼40°, which appears to be more shallow than the 60° seen in solution (*17*). However, a study of immobilized junctions that incorporated the ACC trinucleotide motif showed that this angle, as well as many details concerning the position of the crossover, did not result from crystal lattice distortions but can directly be attributed to the sequence at the junction (*18*). This raises the question of how sequence variations would affect the structure and geometry of the junction. From solution studies, it has been shown that the sequence around an immobilized junction

Table 1: Data Collection and Refinement Statistics

	d(CCGGCGCCGG)	d(CCAGTACBr^{5}UGG)	d(CCAGTACTGG)
Data			
unit cell parameters	$a = 66.5$ Å, $b = 24.2$ Å, $c = 37.0$ Å, $\beta = 110.0°$	$a = 64.9$ Å, $b = 22.6$ Å, $c = 38.1$ Å, $\beta = 106.8°$	$a = b = 33.4$ Å, $c = 87.7$ Å
space group	*C*2	*C*2	$P6_122$
total reflections	50581	13013	32690
unique reflections	5137	4011	2143
resolution (Å)	50−1.6	50−1.8	50−2.0
completeness[a]	78.7% (48.0%)	82.0% (34.9%)	94.6% (53.2%)
$<I/\sigma>$[a]	22.1 (2.7)	23.8 (4.1)	25.3 (3.3)
R_{merge}[a,b]	4.9% (27.8%)	4.1% (18.9%)	4.8% (12.4%)
Refinement			
resolution (Å)	20−1.7	20−1.9	20−2.0
R_{cryst} (R_{free})[c]	22.9% (26.0%)	21.5% (24.3%)	23.6% (26.8%)
DNA atoms	404	404	202
solvent atoms	116	89	67
RMSD bond lengths[d] (Å)	0.007	0.005	0.015
RMSD bond angles[d] (deg)	1.24	0.88	1.75
conformation	junction	junction	B-DNA

[a] Values in parentheses refer to the highest resolution shell. [b] $R_{merge} = \sum_{hkl}\sum_i|I_{hkl,i} - <I>_{hkl}|/\sum_{hkl}\sum_i|I_{hkl,i}|$ where I_{hkl} is the intensity of a reflection and $<I>_{hkl}$ is the average of all observations of this reflection and its symmetry equivalents. [c] $R_{cryst} = \sum_{hkl}|F_{obs} - kF_{calc}|/\sum_{hkl}|F_{obs}|$. $R_{free} = R_{cryst}$ for 10% of reflections that were not used in refinement (*34*). All refinements were performed targeting maximum likelihood. [d] Root-mean-square deviation of bond lengths and angles from ideal values.

defines which arms pair in the coaxially stacked duplexes (*19*−*21*). In addition, single-molecule FRET studies show that the exchange between coaxially stacked arms is dynamic, with the junction undergoing rapid conformational rearrangements in a sequence-dependent manner (*22*). However, the strong apparent requirement of the ACC trinucleotide for their crystallization has not allowed us to study the effect of sequence at the critical positions on the conformational features of the Holliday junction.

The question is thus raised as to whether the ACC core is an absolute requirement for the formation of DNA Holliday junctions, particularly in crystals. The sequence d(CCGGCGCCGG) had earlier been solved as standard B-DNA duplexes but recently had been reported to crystallize in a lattice similar to that of the mismatched d(CCGGGACCGG) junction (*23*). We confirm in this study that indeed this DNA crystal form is that of a four-stranded junction, indicating therefore the ACC trinucleotide sequence motif is not an absolute requirement for crystallization of the junction. We can thus compare the effects of N_6 being either an adenine or guanine on the four-stranded conformation. In addition, we show that the sequence d(CCAGTACBr^{5}UGG), but not the analogous d(CCAGTACTGG) sequence, crystallizes as a junction, showing that the substitiuent at this Py_8 position strongly influences the formation and geometry of the Holliday junction.

MATERIALS AND METHODS

All deoxyoligonucleotides were synthesized on an Applied Biosystems DNA synthesizer in the Center for Gene Research and Biotechnology at Oregon State University using phosphoramidite chemistry, with the trityl-protecting group left intact at the 5′-terminal nucleotide for subsequent purification by reverse-phase HPLC. The purified DNA was deprotected by treatment with 3% acetic acid for 15 min, neutralized with ammonium hydroxide, and desalted on a Sigma G-25 sephadex column. Samples were lyopholized and stored at −80 °C, then resuspended in deionized double-distilled water prior to crystallization. All crystals were grown using the sitting drop vapor diffusion method.

Crystallization and Structure of d(CCGGCGCCGG). Crystals were grown at 25 °C from solutions containing 0.5 mM DNA, 25 mM sodium cacodylate buffer (pH 7.0), 7.5 mM $CaCl_2$, and 0.1 mM spermine tetrahydrochloride in the crystallization drop and equilibrated against a reservoir solution of 28% (v/v) 2-methyl-2,4-pentanediol (MPD). These are very different conditions as compared to those that yielded crystals of the B-DNA duplex form of the sequence (*31*) (B-DNA duplex crystals were grown at 4 °C by microdialysis, with solutions containing 2 mM DNA, 10 mM Tris-HCl, pH 7.0, 150 mM MgCl and equilibrated against the same buffer with 24% MPD).

A thin diamond plate crystal measuring 0.25 × 0.25 × 0.1 mm was used for data collection and found to be in the monoclinic *C*2 space group, with unit cell dimensions $a = 66.51$ Å, $b = 24.18$ Å, $c = 37.00$ Å, and $\beta = 110.01°$. X-ray diffraction data were collected to 1.7 Å resolution at liquid nitrogen temperatures using CuKα radiation from a RUH3R generator with an RAXIS-IV image plate detector. The volume of the unit cell in this space group indicated that there are two strands of DNA in the asymmetric unit of the crystal lattice. Thus, a search model was constructed using two strands of the d(CCGGTACCGG) Holliday junction (*12*), one crossover and one noncrossover, for molecular replacement using EPMR (*24*) against 2.5 Å data. A distinct structural solution with two unique DNA strands adjacent to the crystallographic 2-fold axis yielded a correlation coefficient of 59% and $R_{cryst} = 52\%$. Subsequent refinement in CNS (*25*) using rigid body refinement, followed by simulated annealing, several rounds of positional and individual *B* factor refinement, and addition of solvent produced final values of $R_{cryst} = 22.9\%$ and $R_{free} = 26.0\%$ (Table 1). The coordinates and structure factors have been deposited in the Protein Data Bank (*26*) with accession number 1P4Y.

Crystallization and Structure Determination of d(CCAGTACTGG) and d(CCAGTACBr^{5}UGG). Crystals of d(CCAGTACTGG) were grown at 20 °C from solutions containing 0.6 mM DNA in 5 mM Tris-HCl buffer (pH 7.5), 25 mM calcium acetate, and 7% MPD and equilibrated against 30% MPD. A single crystal measuring 0.2 × 0.2 × 0.3 mm was

used for data collection under liquid nitrogen temperatures on BIO-CARS beamline 14-BMC at the Advanced Photon Source, Argonne National Labs, with 1.0 Å radiation. Crystals of d(CCAGTACbr^{5}UGG) were grown at room temperature from solutions containing 0.6 mM DNA in 5 mM Tris-HCl buffer (pH 7.5), 120 mM calcium acetate, and 16% MPD, and equilibrated against 20% MPD. A 1.9 Å data set was collected on this crystal at liquid nitrogen temperatures using CuKα radiation from the in-house RUH3R generator with an RAXIS-IV image plate detector. Crystals of the d(CCAGTACTGG) are in the hexagonal space group $P6_122$ with unit cell dimensions $a = 33.38$ Å, $b = 33.38$ Å, and $c = 87.67$ Å, while those of d(CCAGTACbr^{5}UGG) are in the monoclinic space group $C2$ with unit cell dimensions $a = 64.84$ Å, $b = 22.58$ Å, $c = 38.07$ Å, and $\beta = 106.78°$.

The structure of d(CCAGTACTGG) was solved by molecular replacement against a 2.0 Å data set using a previously solved B-DNA duplex model in a similar hexagonal lattice. Resulting solutions from an EPMR search yielded a correlation coefficient of 64.2% and $R_{cryst} = 47.4\%$. Subsequent refinements in CNS using rigid body refinement, simulated annealing followed by standard positional and individual B factor refinement, and addition of solvent resulted in final values of $R_{cryst} = 23.6\%$ and $R_{free} = 26.8\%$.

The lattice of the d(CCAGTACbr^{5}UGG) crystal was similar to previously crystallized DNA junctions, suggesting that this br^{5}U analogue of the d(CCAGTACTGG) sequence could in fact be a four-stranded complex. Thus, the structure of this sequence was solved by molecular replacement using the two unique strands of the d(CCGGTACm^{5}CGG) junction structure (*15*) as the initial search model in an EPMR search. The search resulted in a position and orientation of the model having a correlation coefficient of 74% and R factor of 44.5% for the best solution. An initial round of rigid body refinement and simulated annealing in CNS yielded values of $R_{cryst} = 42.5\%$ and $R_{free} = 43.5\%$. Omit maps calculated with the crossover phosphates removed and after simulated annealing of the edited model showed distinct $F_o - F_c$ density consistent with the crossovers of the junction and discontinuities in phosphodiester backbones along the stacked DNA duplexes. Thus, the structure of d(CCAGTACbr^{5}UGG) was refined as a DNA Holliday junction with two strands, one crossover and one noncrossover, in the asymmetric unit, with the two additional related strands generated by a crystallographic 2-fold axis to form a complete four-stranded complex. Refinement was carried out in CNS using rigid body and simulated annealing routines followed by standard positional and individual B factor refinement producing a final $R_{cryst} = 21.5\%$ and $R_{free} = 24.3\%$ after addition of solvent. The coordinates and structure factors have been deposited in the Protein Data Bank (*26*) with accession number 1P4Z for d(CCAGTACTGG) and 1P54 for d(CCAGTACbr^{5}UGG).

All data was reduced using the HKL suite of programs (*27*). Root-mean-square-deviation (RMSD) values were calculated using the algorithm (*28*) implemented in the program ProFit v2.2 (http://www.bioinf.org.uk/software/profit/). Structural analysis was performed with CURVES 5.2 (*29*) and X3DNA (*30*).

RESULTS

The crystal structures presented here show that d(CCGGCGCCGG) and d(CCAGTACbr^{5}UGG) both crystallize as four-stranded Holliday junctions in the compact antiparallel stacked-X form, with the former nearly identical to that of the parent d(CCGGTACCGG) junction (*12*) and the latter most closely related to that of the methylated sequence d(CCGGTACm^{5}CGG) (*15*). The sequence d(CCAGTACTGG), however, was determined to be resolved B-DNA duplexes, now in two unrelated crystal forms. Thus, this study shows that the $A_6C_7C_8$ trinucleotide motif identified as important for stabilizing the junction can accommodate other nucleotide bases at the 6th and 8th positions. In addition, the two bases have their own distinct effect on the conformation of the detailed intramolecular interactions, and consequently, the overall geometries relative to previously determined structures. Thus, the resulting junction structures will be interpreted through comparisons between the current and the previously reported structures in their respective structural classes, with the d(CCGGTACCGG) structure serving as the reference for the unsubstituted and d(CCGGTACm^{5}CGG) for the substituted junctions. These structures therefore allow us to define each component of the junction geometry, including the interduplex angle, rolling of the stacked duplex columns relative to each other (J_{roll}) and sliding of one junction relative to the other along their helical axes (J_{slide}), and how they can be affected by the sequences at the core of the junction.

Structure of d(CCGGCGCCGG) as a Holliday Junction. The single-crystal structure of the sequence d(CCGGC<u>GC-C</u>GG) (which we will call the GCC junction), originally solved as a B-DNA duplex (*31*) under different crystallization conditions, was recently reported to crystallize in a crystal lattice that is isomorphous with that of a DNA four-way junction (*23*). We show here that, under our crystallization conditions, the sequence indeed is a Holliday junction. The complete junction is generated by 2-fold symmetry applied to the two unique strands (one outside noncrossover and one inside crossover strand) of the asymmetric unit. This four-stranded complex (Figure 2a) has four arms—two longer six base pair arms (from the C_1 to G_6 of one outside strand paired with G_{10} to C_5 of one crossover strand) and two short four base pair arms (pairing the C_7 to G_{10} of the same outside strand to G_4 to C_1 of an alternative crossover strand). Each six base pair duplex domain is stacked over a four base pair duplex to form the near continuous 10 base pair helices of the stacked-X junction (Figure 2b), with the crossover occurring between nucleotides G_6 and C_7 of the inside crossover strands (Figure 2c).

In nearly all respects, this GCC junction is identical to that of the reference d(CCGGTACCGG) structure (ACC junction, Figure 3, RMSD = 0.426 Å for all common DNA atoms). The stacked double-helical arms are related by an interduplex angle of 40.0° in the GCC junction as compared to the 41.4° angle in ACC (Table 2). Furthermore, there is no sliding of the duplexes across the junction crossover in GCC ($J_{slide} \approx 0$ Å) relative to the ACC junction. The exposure of the major and minor grooves are also nearly identical between the sequences, with J_{roll} values = 159° in both junction structures.

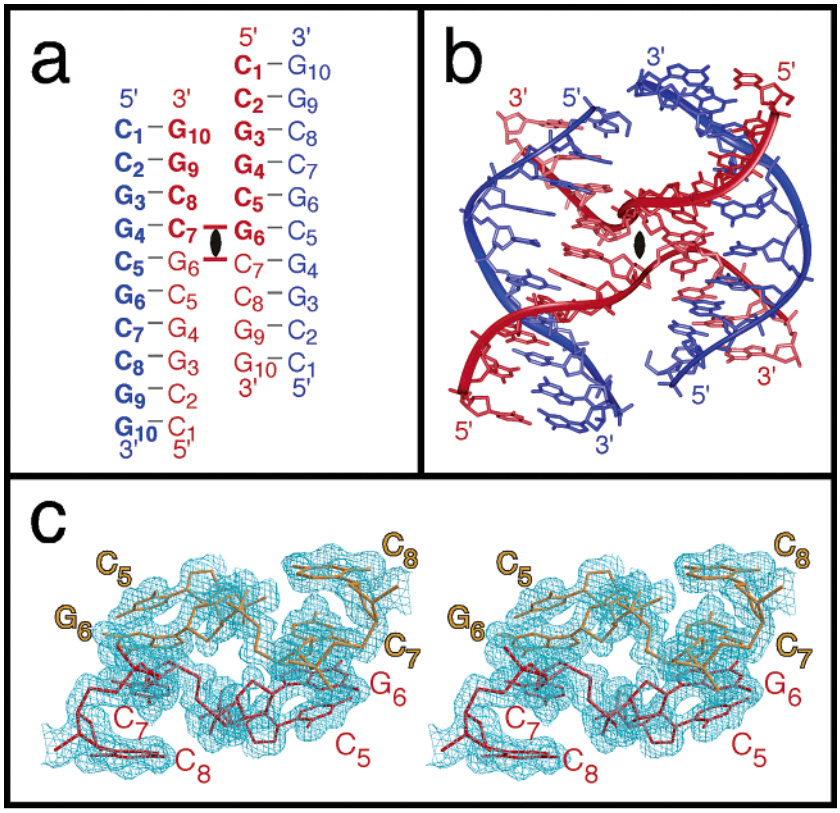

FIGURE 2: Crystal structure of d(CCGGCGCCGG) as a DNA Holliday junction. (a) Sequence topology of the d(CCGGCGCCGG) junction. The four-stranded antiparallel stacked-X Holliday junction is generated by applying the crystallographic 2-fold symmetry to the two unique strands (bold). Strands are numbered from 1 to 10 in the 5′ to 3′ direction, with the inside crossing strands colored red and the outside noncrossing strands colored blue. (b) The atomic structure of d(CCGGCGCCGG). Chemical bonds in the structure are rendered as sticks, and the phosphodeoxyribose backbones are rendered as a solid ribbon (colors and strand designations are as in panel a; figure rendered with InsightII from MSI/Biosym, Inc.). (c) Electron density map. The $2F_o - F_c$ map (contoured at 1σ) shows the discontinuity in electron density between nucleotides G_6 and C_7 in the stacked DNA duplexes, but bridging between adjacent stacked duplexes forms the junction crossovers (panel created with Bobscript (*36*)).

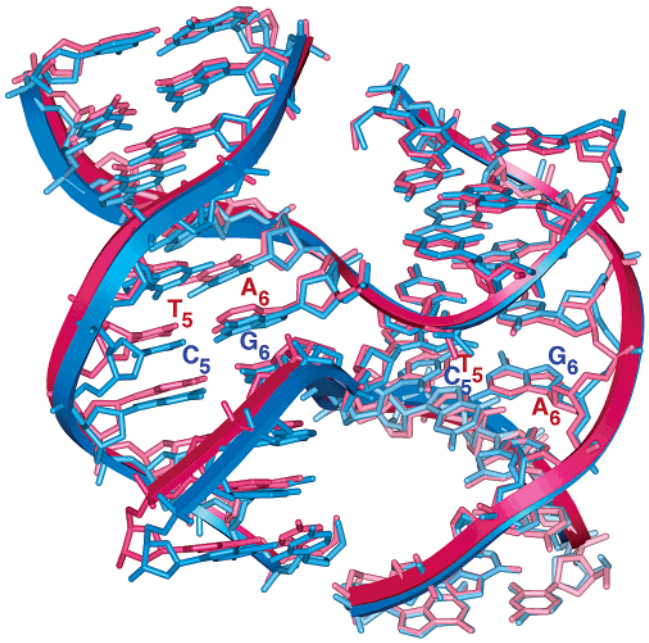

FIGURE 3: Superposition of d(CCGGCGCCGG) and d(CCGGTACCGG) junctions. The structures of the two junctions are superimposed using the common atoms of the stacked arms on the right of the junction in this figure. In superposition of the overall structure, RMSD = 0.426 Å for all common DNA atoms.

Table 2: Comparison of Global Conformational Geometry Values for d(CCGGCGCCGG) (GCC) and d(CCAGTACBr^{5}UGG) (ACBr^{5}U) with Previous Junction Structures of d(CCGGTACCGG) (ACC (*12*)) and d(CCGGTACm^{5}CGG) (ACm^{5}C (*15*))

junction[a]	interduplex angle (deg)	J_{slide}[b]	J_{roll}[c] (deg)
ACC	41.4		159.6
ACm^{5}C	41.3	~1.7 Å per duplex	170.4
GCC	40.0	0 Å	159.8
ACBr^{5}U	38.8	~1.2 Å per duplex	164.0

[a] Trinucleotide core region corresponding to N_6-N_7-N_8 positions of the general sequence d(CCN$_3$N$_4$N$_5$*N*$_6$*N*$_7$*N*$_8$GG). [b] Relative to d(CCGGTACCGG) junction, as estimated by the slide seen for one set of duplexes when the opposing stacked duplexes are superimposed with those of the ACC junction. Values are for the duplexes on both sides of the junction, sliding symmetrically. If one set of stacked duplexes are superimposed, then the opposing stacked double-helices would show a slide of twice the value reported in this table (i.e., Figure 7). [c] Angle between vectors extended from the center of the junction to points that bisect the phosphates on the two outside strands that complement the nucleotides at the gap of the junctions (see, i.e., Figure 4).

One interesting difference between the GCC junction and the other DNA junctions to date is that the width of the minor groove (as measured by the phosphorus to phosphorus distance, Table 3) is wider at the C_5•G_6/G_6•C_5 and G_6•C_5/C_7•G_4 dinucleotide steps (18.5 and 19.5 Å, respectively, as compared to the averages of 17.5 Å (0.26 Å SD) and 18.8 Å (0.35 Å SD) for ACN-type junctions). In addition, we see that the G_6•C_5/C_7•G_4 dinucleotide steps are underwound by 3° as compared to the analogous A_6•T_5/C_7•G_4 dinucleotide step in the junctions (Table 4) and significantly underwound

Table 3: Minor Groove Widths (Å) Measured as the Phosphorus P to P Distance for the Phosphodiester Linkage between Dinucleotides[a]

dinucleotide	ACC-J	GCC-J	ACm^{5}C-J	ACbr^{5}U-J	ACT-B *C*2	ACT-B *P*6$_1$22	GCC-B
ref	*12*	this paper	*15*	this paper	*32*	this paper	*31*
C_1·G_{10}/C_2·G_9	16.9	16.9	16.3	17.2	18.1	18.3	18.1
C_2·G_9/Pu_3·Py_8	17.4	17.2	17.1	17.0	16.7	17.0	17.4
Pu_3·Py_8/G_4·C_7	17.2	17.5	17.6	17.4	18.0	18.3	17.8
G_4·C_7/Py_5·Pu_6	17.9	17.9	18.1	18.1	17.8	18.5	18.3
Py_5·Pu_6/Pu_6·Py_5	17.8	18.5	17.3	17.4	17.0	17.2	18.0
Pu_6·Py_5/C_7·G_4[b]	**19.1**	**19.5**	**18.4**	**18.8**	17.8	18.5	18.2
C_7·G_4/Py_8·Pu_3	16.8	16.9	17.0	17.9	18.0	18.3	17.9
Py_8·Pu_3/G_9·C_2	16.8	17.1	17.2	16.7	16.7	17.0	18.1
G_9·C_2/G_{10}·C_1	16.8	16.4	16.8	17.8	18.1	18.3	17.1
average (SD)[c]	17.41 (0.76)	17.54 (0.96)	17.31 (0.65)	17.59 (0.64)	17.58 (0.60)	17.93 (0.66)	17.88 (0.39)

[a] The widths are compared for the junction structures of d(CCGGTACCGG) (ACC-J), d(CCGGCGCCGG) (GCC-J), d(CCGGTACm^{5}CGG) (ACm^{5}C-J), and d(CCAGTACbr^{5}UGG) (ACbr^{5}U-J) and the B-DNA duplex structures of d(CCAGTACTGG) in the monoclinic crystal form (ACT-B *C*2) and the hexagonal form (ACT-B *P*6$_1$22), and d(CCGGCGCCGG) (GCC-B). [b] Values for the dinucleotide Pu_6·Py_5/C_7·G_4 in bold represent the point where the phosphodeoxyribose backbone departs from the duplex to form the junction crossover in the junction structures. [c] SD = $\sqrt{((\sum(x - \langle x \rangle)^2)/(n-1))}$ where n = number of observations, x = calculated minor groove width for each dinucleotide, and $\langle x \rangle$ is the average of all observations.

Table 4: Comparison of Helical Twist, Measured Using a Global Axis in CURVES 5.1 (*28*) for the Junction Structures of d(CCGGTACCGG) (ACC-J), d(CCGGCGCCGG) (GCC-J), d(CCGGTACm^{5}CGG) (ACm^{5}C-J), and d(CCAGTACbr^{5}UGG) (ACbr^{5}U-J), and the B-DNA Duplex Structures of d(CCAGTACTGG) in the Monoclinic Crystal Form (ACT-B *C*2) and the Hexagonal Form (ACT-B *P6*$_1$22), and d(CCGGCGCCGG) (GCC-B)

dinucleotide	ACC-J (deg)	GCC-J (deg)	ACm^{5}C-J (deg)	ACbr^{5}U-J (deg)	ACT-B *C*2 (deg)	ACT-B *P*6$_1$22 (deg)	GCC-B (deg)
ref	*12*	this paper	*15*	this paper	*32*	this paper	*31*
C_1·G_{10}/C_2·G_9	38.0	38.0	37.4	45.1	30.1	26.5	36.3
C_2·G_9/Pu_3·Py_8	41.1	36.7	38.5	43.9	50.5	51.8	37.3
Pu_3·Py_8/G_4·C_7	39.0	41.8	38.1	34.6	22.1	26.9	36.4
G_4·C_7/Py_5·Pu_6	34.3	33.4	37.0	28.8	34.9	37.2	35.5
Py_5·Pu_6/Pu_6·Py_5	32.1	33.7	41.5	37.9	47.8	32.7	30.1
Pu_6·Py_5/C_7·G_4[a]	**33.6**	**30.6**	**26.5**	**32.6**	34.9	37.2	36.1
C_7·G_4/Py_8·Pu_3	40.9	43.0	40.0	32.9	22.1	26.9	42.6
Py_8·Pu_3/G_9·C_2	39.9	37.4	38.6	51.7	50.5	51.8	24.8
G_9·C_2/G_{10}·C_1	41.3	42.0	38.9	30.9	30.1	26.5	44.0
average (SD)[b]	37.80 (3.55)	37.40 (4.31)	37.39 (4.30)	37.6 (7.69)	35.89 (11.28)	35.28 (10.33)	35.90 (5.81)

[a] Values for the dinucleotide Pu_6·Py_5/C_7·G_4 in bold represent the point where the phosphodeoxyribose backbone departs from the duplex to form the junction crossover in the junction structures. [b] SD = $\sqrt{((\sum(x - \langle x \rangle)^2)/(n-1))}$ where n = number of observations, x = calculated helical twist for each dinucleotide, and $\langle x \rangle$ is the average of all observations.

(by 5.5°) as compared to the comparable steps in the B-DNA duplex of the same sequence. In fact, the duplexes in all of the junction structures are underwound by an average of ~2° (37.54°, SD = 4.99°) as compared to their B-DNA counterparts (35.69°, SD = 9.08°) (Table 4), and all of the dinucleotides that span the junction in these structures are underwound by ~8° relative to the identical or analogous sequences as standard B-DNA (Table 4).

Not surprisingly, many of the structural details seen in the GCC junction are similar to those in the ACC junction. For example, the direct hydrogen bonding interaction between the amino N4 nitrogen at the major groove surface of the cytosine C_8 base to the phosphate oxygen of C_7 that was first identified as helping to stabilize the ACC junction (Figure 4a) is also seen in the current GCC junction. The associated solvent-mediated hydrogen bonding interaction that bridges the keto oxygen of the complementary G_3 nucleotide to the crossover phosphate oxygen at A_6 in the ACC junction (S1 in Figure 4a), however, is now a network in the current GCC junction (Figure 4b), similar to that seen previously in the ACm^{5}C junction structure. Thus, the commonality of the direct C_8 amino to C_7 phosphate hydrogen bond and the similarities of the junction geometries across these structures suggest a correlation between this particular interaction at the atomic level and the general conformational features for this class of structures. However, the solvent that was assigned as a sodium ion in the center of the ACC junction is not present in the current structure, suggesting that this solvent interaction is not crucial to the overall geometry of this structural class.

Structure of d(CCAGTACTGG) as B-DNA Duplexes. Our experience with the d(CCGGCGCCGG) sequence indicated that, although originally reported to be a B-DNA duplex (*32*), we should study the structure of the sequence d(CCAGTACTGG) in our crystallization solutions. However, even under conditions that have yielded junctions, this sequence crystallized as double-stranded B-DNA, although in a different crystal form than was previously reported. The space group of the current crystals is hexagonal *P*6$_1$22 as compared to the earlier monoclinic *C*2 form (*32*). In both crystal forms, the asymmetric unit is one strand of the decanucleotide, with the second strand generated by the crystal symmetry. This precludes the sequence in either crystal form from being a Holliday junction since at least two unique strands are required to form the crossed four-stranded complex. The current and previous B-DNA structures of d(CCAGTACTGG) are very similar, RMSD = 1.03 Å for all nonhydrogen atoms, and with similar minor groove widths and helical twists at each dinucleotide step along the double-helix (Tables 3 and 4). The noticeable differences

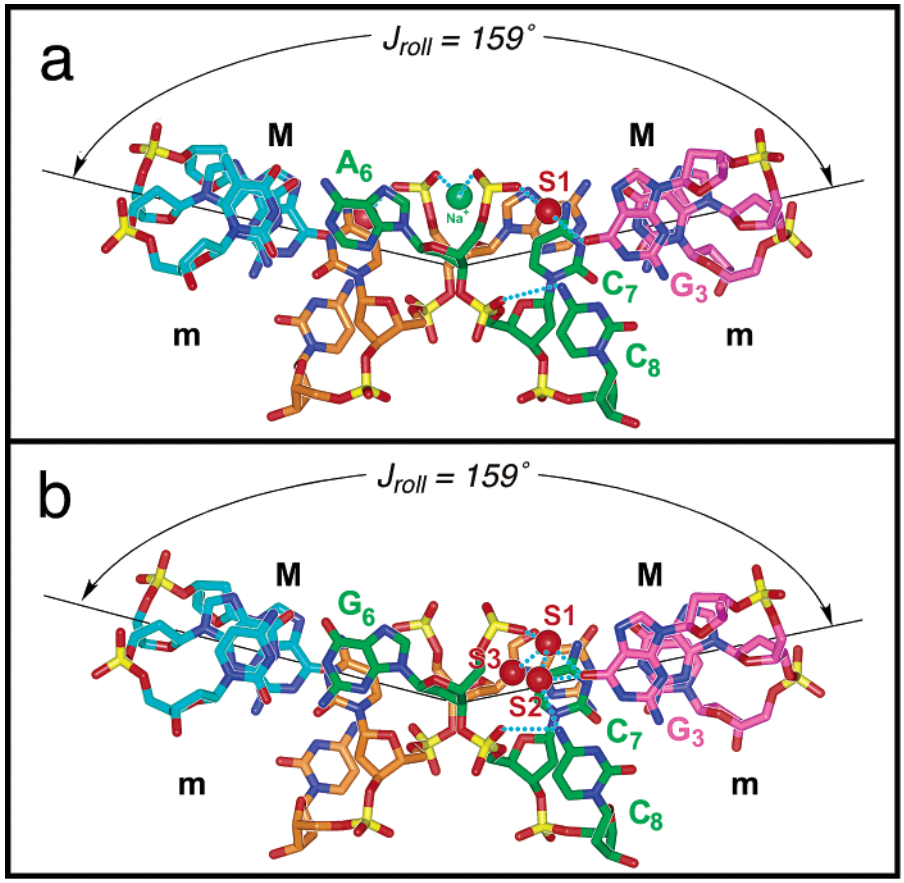

FIGURE 4: Atomic interactions within the trinucleotide core of the d(CCGGCGCCGG) junction viewed perpendicular to the junction dyad axis. Hydrogen bonds are shown as dotted lines, and solvent molecules are shown as spheres. The nucleotides in the junction crossover are labeled, along with the major (M) and minor (m) grooves. The J_{roll} angles are shown as the angle between the vectors extending from the center of the junction to the points that bisect the phosphates of the outside strands that complement the nucleotides of the junction gap in the inside strands. (a) The reference d(CCGGTACCGG) junction. The important interactions had been identified as the direct hydrogen bond between the amino N4-nitrogen of C_8 to the C_7 phosphate oxygen O2P, a water-mediated hydrogen bond between the keto O6 oxygen of G_3 and G_6 phosphate oxygen O1P and solvent (sodium ion) coordinated to the O2P phosphate oxygens of the G_6 nucleotides (*12*). (b) The d(CCGGCGCCGG) junction. This structure shows the direct hydrogen bond between nucleotide C_8 and the crossing phosphate and a set of interactions mediated by solvent (S1 to S3) from the major groove of the G_3 nucleotide to the opposing crossing phosphates. The solvents are labeled as follows: S1 is a water molecule that is hydrogen bonded to the Pu3 nucleotide that complements the Py8 of the junction trinucleotide core; S2 is a water that is hydrogen bonded directly to the Py8 position of the junction core; S3 is the solvent (proposed as the Na^+ ion displaced from the center of the ACC junction) that sits between the phosphates of the junction crossover strands.

result from differences in the duplex−duplex interactions of the two crystal lattices. For example, the duplexes in the monoclinic crystal form have the nonstandard BII conformations (as defined by a ζ dihedral angle ≈180°) at the phosphoribose spanning the C_2pA_3, G_4pT_5, and T_8pG_9 dinucleotide steps along the DNA chain. In contrast, only the C_2pA_3 dinucleotide step shows this BII conformation in the current hexagonal crystal form.

Structure of d(CCAGTACbr^{5}UGG) as a Holliday Junction. The sequence d(CCAGTACbr^{5}UGG) was initially designed as a means to introduce a bromine heavy atom at the nucleotide-8 position of d(CCAGTACTGG) to help solve this structure in the hexagonal crystal form. Although to date d(CCAGTACTGG) has only crystallized as resolved B-DNA duplexes, this brominated analogue is presented here as a four-stranded Holliday junction (Figure 5). The resulting structure of d(CCAGTACbr^{5}UGG) (ACbr^{5}U junction) is most closely related to that of the methylated d(CCGGTACm^{5}CGG) (ACm^{5}C junction) structure. Although the RMSD of 1.57 Å for common DNA atoms between the ACbr^{5}U and ACm^{5}C junctions is relatively high, both show similar structural perturbations relative to the ACC-type junctions in terms of the J_{roll} and J_{slide} (Table 2) around and along the axes of the stacked duplexes. The J_{slide} of ∼3.4 Å and J_{roll} of 170.4° rendered the ACm^{5}C junction more symmetric in terms of the relative positions of the ends and accessibility of the major and minor grooves of the stacked duplex arms. The current ACbr^{5}U junction is intermediate between this distorted ACm^{5}C and the parent ACC junctions. The interduplex angle of the ACbr^{5}U junction (38.8°) is the most shallow of any DNA-only junction (but not as shallow as the psoralenated junctions (*14*)), where this angle varies from 36.4−37.2°).

The detailed interactions at the ACbr^{5}U junction crossover (Figure 6b) show some similarities to the ACm^{5}C structure (Figure 6a). In both cases, the direct hydrogen bond from the Py_8 nucleotide to the crossing phosphate of C_7 has been disrupted. In the ACm^{5}C structure, this hydrogen bond from the amino N4 of cytosine C_8 to O2P oxygen of the phosphate was disrupted by addition of the methyl group at the C_8 nucleotide, but the two atoms are now bridged by a solvent molecule (S2 in Figure 6a). In replacing the cytosine C_8 with the substituted uracil base in the ACbr^{5}U junction, the hydrogen bond donating amino group of the cytosine has

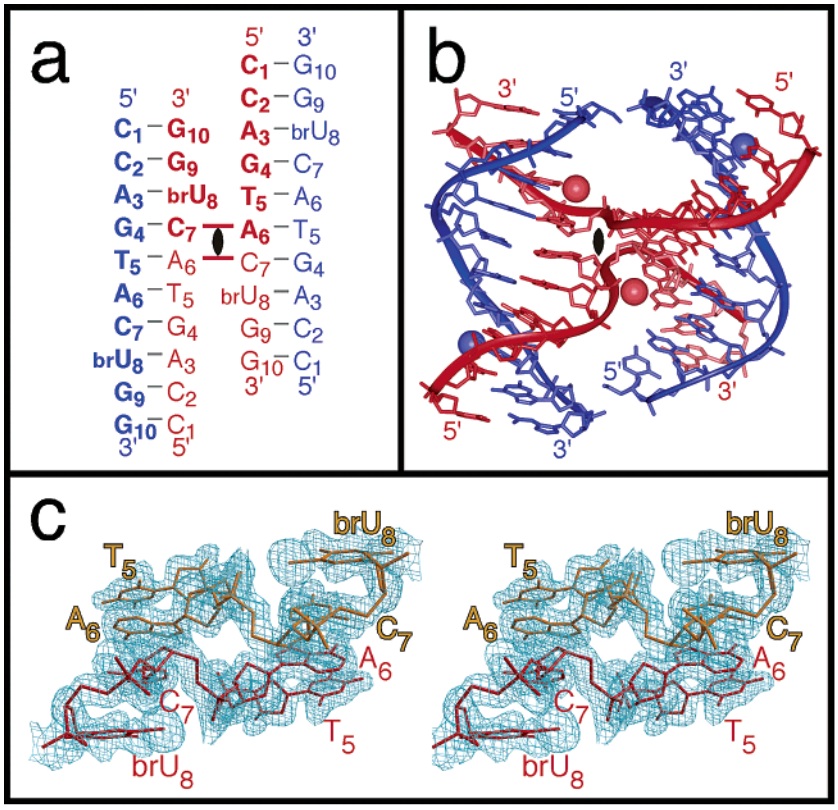

FIGURE 5: Crystal structure of d(CCAGTACBr^{5}UGG) as a DNA Holliday Junction. (a) Sequence topology of the d(CCAGTACBr^{5}UGG) junction. The four-stranded antiparallel stacked-X Holliday junction is generated by applying the crystallographic 2-fold symmetry to the two unique strands (bold). Strands are numbered from 1 to 10 in the 5′ to 3′ direction, with the inside crossing strands colored red and the outside noncrossing strands colored blue. (b) The atomic structure of d(CCAGTACBr^{5}UGG). Chemical bonds in the structure are rendered as sticks, the phosphodeoxyribose backbone is rendered as a solid ribbon, and bromines are rendered as spheres (colors and strand designations are as in panel a, figure rendered with InsightII from MSI/Biosym, Inc.). (c) Electron density map. The $2F_o - F_c$ map (contoured at 1σ) shows the discontinuity in electron density between nucleotides G_6 and C_7 in the stacked DNA duplexes, but bridging between adjacent stacked duplexes to form the junction crossovers (figure in panel created with Bobscript (*36*)).

been replaced by a hydrogen bond accepting keto oxygen. This precludes the formation of the direct hydrogen bonding interaction, regardless of the substituent group added to the C5 carbon of the base. A solvent (S2, Figure 6b) remains hydrogen bonded to this keto oxygen but does not span the gap to the phosphate oxygen of the junction crossover.

The ACC, GCC, and ACm^{5}C junctions all show a network of one or more solvent molecules that help bridge the complementary G_3 nucleotide base of the C_8•G_3 base pair to the crossing phosphate of A_6. The analogous A_3 nucleotide of the br$^5U_8$•A_3 base pair shows no such network of bridging solvent molecules. We had previously suggested that one of the bridging solvent molecules in the ACm^{5}C network (S3, Figure 6a) was the sodium ion in the center of the ACC junction but displaced by the collapse of the junction core. The loss of this series of solvent between the A_3 and A_6 phosphates would suggest that the ion could have buried itself back into the center of the ACBr^{5}U junction, and the electron density in the center of the junction supports that assertion. A solvent molecule was seen to be sandwiched between, and located very close to (<2.5 Å), the O1P phosphate oxygen atoms of the crossing A_6 nucleotides (Na^+, Figure 6b). It is interesting, however, that the volume estimated for this binding site corresponds to a sphere with radius 1.25 Å, which is now larger than that of the ACm^{5}C junction (at 0.9 Å radius) and even the unmodified ACC junction (at 1.13 Å radius) (*33*). Finally, a new solvent (S8, Figure 6b) is seen spanning the O1P oxygens of the C_7 cytosines in both crossing strands.

Thus, in the current structure of the ACBr^{5}U junction, there are no direct or solvent-mediated hydrogen bonds between the major groove surface of the br$^5U_8$•A_3 base pair and the phosphate oxygens of the junction crossover, even though many of the solvent molecules remain hydrogen bonded to the bases at the major groove surface. There is, however, an interaction at the other end of the $Pu_6C_7Py_8$ trinucleotide motif, with the N7 nitrogen of the A_6 nucleotide base now bridged through a solvent molecule (S9, Figure 6b) to the O2P oxygen of the A_6 nucleotide of the opposite crossing strand. This would seem to compensate for the loss of interactions between the stacked arms and the core of the junction that appears to be critical to the stability of the junction in this system.

DISCUSSION

In this study, we have determined the single-crystal structures of the sequences d(CCGGCGCCGG) and d(CCAGTACBr^{5}UGG) as four-stranded Holliday junctions in the antiparallel stacked-X conformation and the structure of d(CCAGTACTGG) as resolved B-DNA duplexes. The two new junctions are the first such structures from sequences that do not contain the previously identified ACC-core trinucleotide sequence at the $N_6N_7N_8$ positions of the general

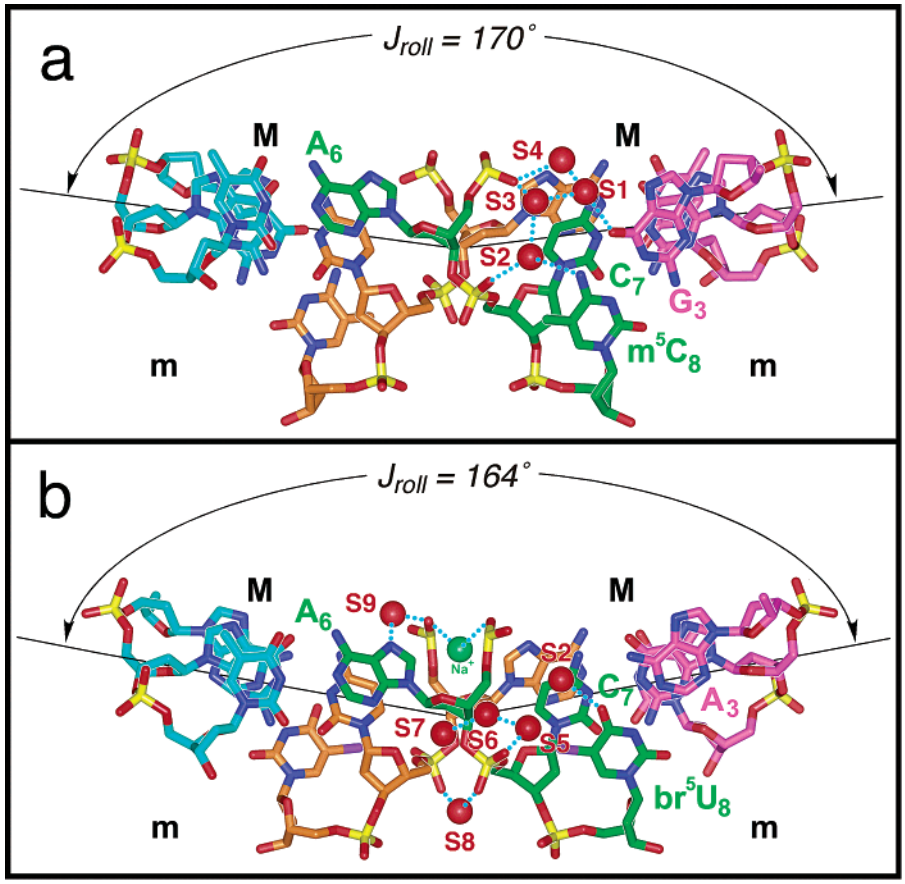

FIGURE 6: Atomic interactions within the trinucleotide core of the d(CCAGTACBr^{5}UGG) junction viewed perpendicular to the junction dyad axis. Hydrogen bonds are shown as dotted lines, and solvent molecules are shown as spheres. The nucleotides in the junction crossover are labeled, along with the major (M) and minor (m) grooves. The J_{roll} angles are shown as the angle between the vectors extending from the center of the junction to points that bisect the phosphates of the outside strands that complement the nucleotides of the junction gap in the inside strands. (a) The reference d(CCGGTACm^{5}CGG) junction. The important interactions had previously been identified as the network of solvent (S1 to S4) mediated hydrogen bonds from the C_8•G_3 base pair to the C_7 O2P and G_6 O1P phosphate oxygens (*15*). (b) The d(CCAGTACBr^{5}UGG) junction. The solvent (S2) at the br^{5}U$_8$•A$_3$ base pair is no longer continuous with the solvent at the phosphates of nucleotide C_7. A sodium ion (Na^+) is seen sandwiched between the phosphate oxygens of nucleotide A_6, and a solvent (S9) bridges the N7 nitrogen of A_6 to the phosphate of the A_6 nucleotide of the opposing crossover strand. The labels for solvent S1 to S3 occupy positions similar to those described for Figure 4. S4 is a water that bridges S1 to the phosphate of the crossover strand of the ACm^{5}C junction. The solvent molecules labeled S5 to S9 are new to the core of the ACBr^{5}U junction.

sequence d(CCN$_3$N$_4$N$_5$*N$_6$N$_7$N$_8$*GG). Thus, we have now expanded this central sequence motif for the crystallization of junctions to Pu$_6$C$_7$Py$_8$, where Pu is either a guanine or adenine purine-containing nucleotide, and Py is a cytosine, 5-methylcytosine, or 5-bromouracil pyrimidine nucleotide but not a thymine nucleotide. Our previous study with the methylated d(CCGGTACm^{5}CGG) sequence (*15*) showed that perturbations to the overall geometry of the Holliday junction can include a slide in the stacked duplexes across the junction (J_{slide}) and a roll of the duplexes along their helix axes (J_{roll}). With the current study, we see that sequence and substituent effects define distinct classes of stacked-X junctions: the ACC-type structure (the original junction structure) and ACm^{5}C class (which shows significant J_{slide} and J_{roll} relative to ACC). In addition, these effects are seen to perturb the interduplex angle relating the orientation of the stacked duplex arms across the junction.

A comparison of the GCC and ACC junctions show that the Pu$_6$ position of the Pu$_6$C$_7$Py$_8$ motif does not dramatically perturb the conformation of the Holliday junction. The only noticeable effect on the general geometry of the conformation in replacing the adenine with a guanine base at this position is to make the interduplex angle more shallow (from 41.4° in ACC to 40° in the GCC junction). It is not surprising that the two structures are so similar. The nucleotide at the N_6 position is the initial point of departure of the DNA strand that crosses over to form the junction. However, the purine base at this position remains paired with the complementary strand of the stacked duplex arms at this point and is not directly involved in the interactions of the crossover.

In contrast to Pu$_6$, the base and the substituent at the Py$_8$ position of the now expanded Pu$_6$C$_7$Py$_8$ trinucleotide motif have a significant affect on the conformation of the junction. The sliding (J_{slide}) and rolling of the duplexes along and around their respective helix axes (J_{roll}) previously defined in the d(CCGGTACm^{5}CGG) structure are also observed here with the br^{5}U deoxynucleotide at the Py$_8$ position but are less exaggerated. Thus, the J_{slide} distortions are primarily induced by steric interactions between the methyl or bromo substituent and the deoxyphosphoribose backbone of the crossover strands. In the current structure, the bromine is in direct contact with one of these phosphates. Thus, we can think of the junction as the pivot point not only to twist the stacked duplexes around (in defining the interduplex angle)

but to push against and thus slide these columns of base pairs relative to each other.

Rolling the stacked duplexes about their respective helical axes (J_{roll}) has the effect of opening the major groove surface of the junction. What we see by comparison of these structures is that the direct interaction between the C_8 cytosine and the crossover phosphates fixes the orientation of the opposing stacked duplexes of the ACC and GCC junctions. This is not surprising since hydrogen bonding interactions involving nucleotide bases have well-defined geometric constraints in terms of distances and angles between hydrogen bond donors and acceptors. Thus, such interactions predetermine the relationships between the major groove surface of the arms and the core phosphates of the junction. The addition of intervening solvent molecules, as in the ACm^{5}C and ACbr^{5}U junctions, can be thought of as lubricating the interaction surfaces, relieving the constraints imposed by the ACC and GCC interactions and thereby allowing more variations within the geometries of the junctions, allowing variability in the interduplex angle (as seen in the more shallow angle of the ACbr^{5}U junction), as well as perturbations to J_{slide} and J_{roll} (evident in both the substituted structures). This does not, however, suggest that the structure of the junction itself is more flexible. A comparison of all the junction crystal structures shows that the nucleotides directly at and those paired to nucleotides at the crossover have lower average temperature or *B* factors than those nucleotides that are distant from the junction core.

The other distortion seen in the d(CCAGTACbr^{5}UGG) structure, relative to all other junctions, is the shallower interduplex angle. This angle, which is approximately 40° (ranging from 40 to 41.4°) for all PuCC-type junction structures, becomes 38.4° with the ACbr^{5}U structure. The distortions to J_{slide} and J_{roll} means that the best comparison is between the two structures with substituents at the Py_8 position (ACm^{5}C and ACBr^{5}U), which places these two structures in their own class. In this comparison, we see that the br^{5}U base induces a 3° rotation about the junction relative to the ACm^{5}C junction, which results from substitution of an A_3·br$^5U_8$ for the G_3·m$^5C_8$ base pair. The rotation of the interduplex angle is thus not simply a substituent effect but more likely is associated with the A·br^{5}U base pair at this position.

Why does sequence have such an effect on the structure of the junctions? The answer to this question may be seen in the interactions at the ACbr^{5}U trinucleotide that compensate again for the loss of interactions between the major groove of the Py_8·Pu_3 base pair to junction phosphates seen in all of the other structures of the complex. There were no direct or even solvent-mediated bridges at this surface. However, there is a compensating solvent bridge from A_6 across to the phosphate of the opposite crossing strand. We can now imagine this as being a factor that both stabilizes and defines the interduplex angle of the ACbr^{5}U junction.

Alternatively or in addition, we had previously proposed that the interduplex angle is defined by a short phosphate−phosphate interaction that is distant (three base pairs away) from the actual junction crossover (*14*). We again see this very short oxygen to oxygen distance (3.69 Å), and in the GCC junction structure, observe two solvent molecules that are ∼2.4 Å from one set of the phosphate oxygen atoms. If these two closely interacting solvents are assigned as sodium ions rather than waters, then the net charge at the phosphate of the N_{10} nucleotide would be +1, which would provide a rationale for this short interarm phosphate−phosphate distance. This short phosphate−phosphate distance remains even in the ACm^{5}C (at 3.5 Å) and ACbr^{5}U (at 3.6 Å) structures. In both junctions, the direct amino nitrogen to the crossover phosphate of C_7 observed in the ACC junction has been disrupted (by the methyl group of ACm^{5}C, and additionally, by the replacement of the amino nitrogen with a keto oxygen in ACbr^{5}U). This supports a model that the components of the helical structure of the duplex arms (twist, rise, tilt, etc.) help to define this angle. For example, in the current ACPy structures, we see that both the average helical twist overall and across the unique trinucleotide steps at C_2-Pu_3-G_4 and C_7-Py_8-G_9 are significantly more shallow in the ACbr^{5}U junction as compared to the ACm^{5}C junction. The orientation of the phosphates in close approach would in fact result in a more shallow interduplex angle when the double-helical arms are unwound in this class of shifted and rolled junctions (Figure 7).

In the original ACC junction, we had observed a solvent molecule (which was assigned as a sodium cation) imbedded in a central cavity formed by the phosphates of the crossing strands. We now observe this same solvent molecule in the center of the ACbr^{5}U junction but not the ACm^{5}C or GCC structures. We can thus conclude that neither the J_{roll} nor the J_{slide} distortions alone account for the accessibility of the junction to the intruding ion.

Why does the ACbr^{5}U trinucleotide favor formation of crystals of the junction while ACT does not? The difference is obviously the bromine versus the methyl group at the C5 carbon of the Py_8 nucleotide base. Since the two substituents are nearly identical in size (both approximately 2 Å radius) and hydrophobicity, the difference in behavior likely arises from differences in their electronic properties—bromine is electron withdrawing, making it slightly electron rich, while methyl groups are electron donating to an aromatic ring system and thus are slightly electron poor. One can suggest then that the electrostatic interactions between the substituent and the close phosphate group from the junction crossover might be the distinguishing factor; however, this would predict that the methyl rather than the bromo group should stabilize the junction, which is the opposite of what is observed here. We suggest, therefore, that the electrostatic effect is indirect, affecting for example the base pair stacking. The helical twist at the C·G/br^{5}U·A dinucleotide step in the ACbr^{5}U junction is 8−10° larger than the comparable C·G/T·A steps of the B-DNA structures (Table 4). This is true regardless of whether the base pairs are at the junction core (C_7·G_4/br$^5U_8$·A_3) or not (A_3·br$^5U_8$/G_4·C_7), indicating again that the effects of nucleotide sequence and substituents on the structure of the stacked duplex arms of the junction mirror those seen in standard B-DNA duplexes (*14*). This difference in the helical twist at the junction core would dramatically affect the position of the substituent groups relative to the crossing phosphates. In the ACbr^{5}U junction, the bromine sits against the phosphate oxygen and above the deoxyribose ring of the preceding C_7 nucleotide. A rotation of this base pair by −10° in the stack (as would be expected in a theoretical ACT junction) would place the analogous methyl group of a ACT junction core nearly equidistant from the two crossing phosphates and in contact with neither. One

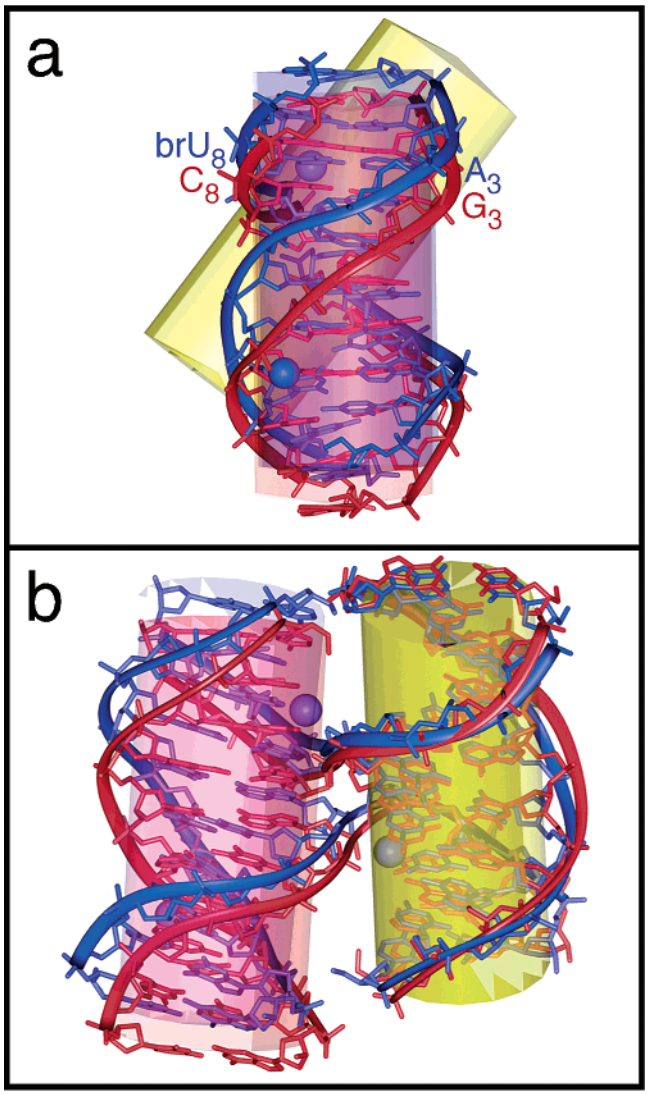

FIGURE 7: Superposition of the d(CCAGTACbr^{5}UGG) junction (blue) on the d(CCGGTACCGG) junction (red). Atoms are rendered as stick models with the phosphodiester backbone traced by ribbons. One set of stacked duplex arms (yellow cylinder) from each structure was used for superposition. (a) View along the junctions with the nonsuperimposed stacked duplex arms in front. The ACbr^{5}U junction shows ~2.3 Å J_{slide} along the helix axis. (b) View into the junctions with the superimposed stacked arms to the right. The nonsuperimposed stacked duplexes show the J_{slide} and J_{roll} effects imposed by the br^{5}U·A base pair of the ACbr^{5}U core trinucleotide on the overall geometry of the Holliday junction.

possible explanation therefore for why the thymine does not facilitate crystallization of the junction is that the effects of the methyl group on the electron distribution of the thymine base results in a base stacking geometry that does not force the substituent to push against the phosphates of the junction, and consequently, would not induce a J_{slide} of the stacked helices. The result may then be that the solvent bridged interaction between adenine A_6 (across the gap of the inside strands of the complex) could not be formed. Other factors, including solvent effects, may contribute to or may be more important than the electronic effect on base stacking described here.

Thus, we see from this series of structures that the overall geometry of the junctions vary according to the nucleotides within the trinucleotide core sequence motif. This is consistent with the understanding that the nucleotides immediately flanking the junction determine how helical arms pair to form the stacked duplexes (*20*, *21*) and the rates of conformational isomerization to switch these pairing partners (*22*). The current structures therefore provide an atomic level view of how these sequence variations would affect the stability of the four-way stacked-X junction, which also defines the ability of such junctions to undergo conformation-dependent dynamic processes, including branch migration and protein binding.

ACKNOWLEDGMENT

We thank Drs. Z. Wood and R. Faber for helpful discussion during data refinement and Dr. J. Watson for critical reading of the manuscript.

REFERENCES

1. Lilley, D. M., and White, M. F. (2001) *Nat. Rev. Mol. Cell Biol. 2*, 433−443.
2. Holliday, R. (1964) *Genet. Res. 5*, 282−304.
3. Ho, P. S., and Eichman, B. F. (2001) *Curr. Opin. Struct. Biol. 11*, 302−308.
4. Cooper, J. P., and Hagerman, P. J. (1987) *J. Mol. Biol. 198*, 711−719.
5. Cooper, J. P., and Hagerman, P. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7336−7340.
6. Clegg, R. M., Murchie, A. I. H., Zechel, A., Carlberg, C., Diekmann, S., and Lilley, D. M. J. (1992) *Biochemistry 31*, 4846−4856.
7. Clegg, R. M., Murchie, A. I. H., Zechel, A., and Lilley, D. M. J. (1994) *Biophys. J. 66*, 99−109.
8. Murchie, A. I. H., Clegg, R. M., von Kitzing, E., Duckett, D. R., Diekmann, S., and Lilley, D. M. J. (1989) *Nature 341*, 763−766.
9. Mao, C., Sun, W., and Seeman, N. C. (1999) *J. Am. Chem. Soc. 121*, 5437−5443.
10. Nowakowski, J., Shim, P. J., Prasad, G. S., Stout, C. D., and Joyce, G. F. (1999) *Nat. Struct. Biol. 6*, 151−156.
11. Nowakowski, J., Shim, P. J., Stout, C. D., and Joyce, G. F. (2000) *J. Mol. Biol. 300*, 93−102.
12. Eichman, B. F., Vargason, J. M., Mooers, B. H. M., and Ho, P. S. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 3971−3976.
13. Ortiz-Lombardía, M., González, A., Eritja, R., Aymamí, J., Azorín, F., and Coll, M. (1999) *Nat. Struct. Biol. 6*, 913−917.
14. Eichman, B. F., Mooers, B. H. M., Alberti, M., Hearst, J. E., and Ho, P. S. (2001) *J. Mol. Biol. 301*, 15−26.
15. Vargason, J. M., and Ho, P. S. (2002) *J. Biol. Chem. 277*, 21041−21049.
16. Thorpe, J. H., Gale, B. C., Teixeira, S. C., and Cardin, C. J. (2003) *J. Mol. Biol. 327*, 97−109.
17. Lilley, D. M. J. (2000) *Q. Rev. Biochem. 33*, 109−159.
18. Sha, R., Liu, F., and Seeman, N. C. (2002) *Biochemistry 41*, 5950−5955.
19. Miick, S. M., Fee, R. S., Millar, D. P., and Chazin, W. J. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9080−9084.
20. Grainger, R. J., Murchie, A. I. H., and Lilley, D. M. J. (1998) *Biochemistry 37*, 23−32.
21. Carlström, G., and Chazin, W. J. (1996) *Biochemistry 35*, 3534−3544.
22. McKinney, S. A., Declais, A. C., Lilley, D. M., and Ha, T. (2003) *Nat. Struct. Biol. 10*, 93−97.
23. Aymami, J., Pous, J., Lisgarten, J. N., and Coll, M. (2002) *Acta Crystallogr. D 58*, 310−311.
24. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) *Acta Crystallogr. D55* (Pt 2), 484−491.
25. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54* (Pt 5), 905−921.
26. Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R., and Schneider, B. (1992) *Biophys. J. 63*, 751−759.
27. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
28. McLachlan, A. D. (1982) *Acta Crystallogr. A 38*, 871−873.

29. Lavery, R., and Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 655−667.
30. Lu, X.-J., Shakked, Z., and Olson, W. K. (2000) *J. Mol. Biol. 300*, 819−840.
31. Heinemann, U., Alings, C., and Bansal, M. (1992) *EMBO J. 11*, 1931−1939.
32. Kielkopf, C. L., Ding, S., Kuhn, P., and Rees, D. C. (2000) *J. Mol. Biol. 296*, 787−801.
33. Kleywegt, G. J., and Jones, T. A. (1994) *Acta Crystallogr D50*, 178−185.
34. Brünger, A. T. (1992) *Nature (London) 355*, 472−475.
35. Gopaul, D. N., Guo, F., and Van Duyne, G. D. (1998) *EMBO J. 17*, 4175−4187.
36. Esnouf, R. M. (1999) *Acta Crystallogr D55 (Pt 4)*, 938−940.

BI0346603