Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural characterization of a new crystal form of the four-way Holliday junction formed by the DNA sequence d(CCGGTACCGG)$_2$: sequence *versus* lattice?

James H. Thorpe, Susana C. M. Teixeira, Benjamin C. Gale and Christine J. Cardin*

The University of Reading, Chemistry Department, Whiteknights, Reading RG6 6AD, England

Correspondence e-mail: c.j.cardin@rdg.ac.uk

Received 6 December 2001
Accepted 24 January 2002

NDB Reference: UD0015.

DNA-strand exchange is a vital step in the recombination process, of which a key intermediate is the four-way DNA Holliday junction formed transiently in most living organisms. Here, the single-crystal structure at a resolution of 2.35 Å of such a DNA junction formed by d(CCGGTACCGG)$_2$, which has crystallized in a more highly symmetrical packing mode to that previously observed for the same sequence, is presented. In this case, the structure is isomorphous to the mismatch sequence d(CCGGGACCGG)$_2$, which reveals the roles of both lattice and DNA sequence in determining the junction geometry. The helices cross at the larger angle of 43.0° (the previously observed angle for this sequence was 41.4°) as a right-handed X. No metal cations were observed; the crystals were grown in the presence of only group I counter-cations.

1. Introduction

DNA junctions play an important role in the physiology of cells, where the Holliday junction is a central recombination intermediate formed through the strand exchange of two homologous DNA molecules to give a four-way junction at the crossover (Holliday, 1964). Such systems are the target of many recombination enzymes (Ceschini *et al.*, 2001; Déclais *et al.*, 2001; Nishino *et al.*, 2001; Ristriani *et al.*, 2001; West & Austin, 1999; Sekiguchi *et al.*, 1996) that promote branch migration or catalyse its resolution, allowing the exchange of genetic information between DNA duplexes. The four-way junction is, therefore, a highly distinctive structural motif, which the junction-resolving enzymes must recognize at the molecular level through a currently unidentified mechanism (Lilley & White, 2001).

The recent structural characterization of several Holliday junctions (Eichman *et al.*, 2000, 2001; Ortiz-Lombardia *et al.*, 1999) formed by DNA decamer sequences have revealed a remarkable similarity in the crystal packing orientations of both the stacked-X structure Holliday junctions and the decanucleotide B-DNA structures, with near-identical unit-cell parameters in both cases for the monoclinic space group *C*2. In the present work, it was impossible to deduce the presence or absence of the junction from the lattice parameters alone, which may hint at a reason why the junction structure was not demonstrated for so many years. The two native junctions currently reported are the inverted repeat sequence d(CCGGTACCGG)$_2$, with all B-DNA arms, Watson–Crick base pairing and no internal symmetry, and the mismatched sequence d(CCGGGACCGG)$_2$, where two G·A mismatched base pairs sit adjacent to the crossover on a twofold symmetry axis. The inverted repeat structure therefore has an asymmetric unit of four nucleotide strands, whereas the mismatched structure has only two (see Table 1). Here, we report the inverted repeat structure but, unexpectedly, with the structural unit already reported for the mismatched sequence, each in the space group *C*2. The overall packing is the same in both cases, but our observation offered the possibility of defining to what extent the differences found between the two already published structures are a consequence of the presence or absence of symmetry constraints, set alongside the distorting effect of base-pair mismatching.

2. Materials and methods

Crystals were grown by sitting-drop vapour diffusion from conditions optimized from the Hampton Research Nucleic Acid Mini Screen yielding 40 m*M* sodium cacodylate pH 7.0, 12 m*M* spermine, 80 m*M* KCl, 10%(*v*/*v*) 2-methyl-2,4-pentanediol (MPD) with 1 m*M* d(CCGGTACCGG)$_2$ and equilibrated against 1 ml 35% MPD. Thin diamond-shaped crystals of approximate dimensions 300 × 200 × 50 µm were grown at 290 K. A single crystal was mounted in a loop with cryoprotectant perfluoropolyether (RS 3000) oil and flash-cooled to 100 K under a nitrogen stream. Data were collected using a MAR345 image plate in small mode with synchrotron radiation at λ = 1.073 Å from the EMBL X31 beamline, Hamburg and were further processed and reduced with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1995). The initial

Table 1
Summary of related native decamer sequences.

Sequence	Unit-cell parameters (Å, °)	Space group	Helix crossover angle† (°)	Motif type	P—P separation (Å)	Reference
d(CCGGTACCGG)	a = 64.9, b = 25.4, c = 37.4, β = 110.6	$C2$	43.0	Quadruplex	6.28	This work
d(CCGGTACCGG)	a = 66.5, b = 23.5, c = 76.9, β = 114.8	$C2$	41.4	Quadruplex	6.17	Eichman *et al.* (2000)
d(CCGGGACCGG)	a = 64.2, b = 23.7, c = 38.3, β = 112	$C2$	40	Quadruplex	6.23	Ortiz-Lombardia *et al.* (1999)
d(CCGCTAGCGG)	a = 64.1, b = 25.9, c = 39.9, β = 122.0	$C2$	44.1	Duplex	6.2	Eichman *et al.* (2000)
d(CGCAATTGCG)	a = 63.6, b = 24.7, c = 38.3, β = 110.3	$C2$	36	Duplex	6.14	Wood *et al.* (1997)

† Each value has been recalculated using *CURVES* 5.2 to ensure consistent results.

starting model was taken from the B-DNA decamer duplex sequence with the most closely correlated cell parameters, $d(CGCAATTGCG)_2$ (Wood *et al.*, 1997), with the appropriate bases changed. The orientation of this initial duplex model was optimized with the *CCP*4 (Collaborative Computational Project, Number 4, 1994) version of the program *MOLREP* (Vagin & Teplyakov, 1997), yielding a correlation coefficient of 60.3% and an R factor of 61.4%. The resulting model showed an overlap with a symmetry-related strand at the centre of the decamer sequence, suggesting the presence of the Holliday junction. The mismatched DNA stacked-X structure coordinates, $d(CCGGGACCGG)_2$, with the mismatched guanine replaced by thymine were then used as a starting model, rather than the $d(CCGGTACCGG)_2$ coordinates, because of the closely related unit-cell parameters (see Table 1) and the model was again optimized with *MOLREP*. This yielded a correlation coefficient of 44.6% and an R factor of 69.1%, which was less favourable than before. Refinement was carried out with *REFMAC* (Murshudov *et al.*, 1997) through maximum-likelihood methods with an overall isotropic B factor refined and model fitting with *XTALVIEW* (McRee, 1999). Both models showed clear continuous backbone density in the strand-crossover region of the map, with the structure in the form of the Holliday junction (Fig. 1*a*). Solvent sites were modelled by hand through σ_A and difference maps checked for stabilizing cations, but none could be definitively modelled and all were finally set to be oxygen. Refinement converged to give a final conventional R factor of 22.64% and an $R_{\rm free}$ of 26.78%. A summary of the refinement statistics is given in Table 2. Comparative helical analysis was carried out with *CURVES* (Lavery & Sklenar, 1988) and both coordinates and structure factors have been deposited with the NDB.[1]

Table 2
Summary of data processing and refinement statistics.

Values in parentheses refer to the outer resolution shell.

Data quality	
Resolution (Å)	30.00–2.35
$R_{\rm merge}$	0.056 (0.230)
Mean $I/\sigma(I)$	10.50 (3.37)
Completeness (%)	96.41 (97.3)
No. of observations	4466
No. of unique reflections	2390
Refinement	
No. of DNA atoms	404
No. of solvent atoms	43
R factor† (%)	22.64 (31.0)
$R_{\rm free}$ (%)	26.78 (44.0)
R.m.s. deviation, 1–2 bonds (Å)	0.067
R.m.s. deviation, 1–3 angles (Å)	2.195
Average DNA B (Å^2)	46.73
Average solvent B (Å^2)	33.15

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_{hkli} - \langle I_{hkl}\rangle|/\sum_{hkl}\sum_i \langle I\rangle$.

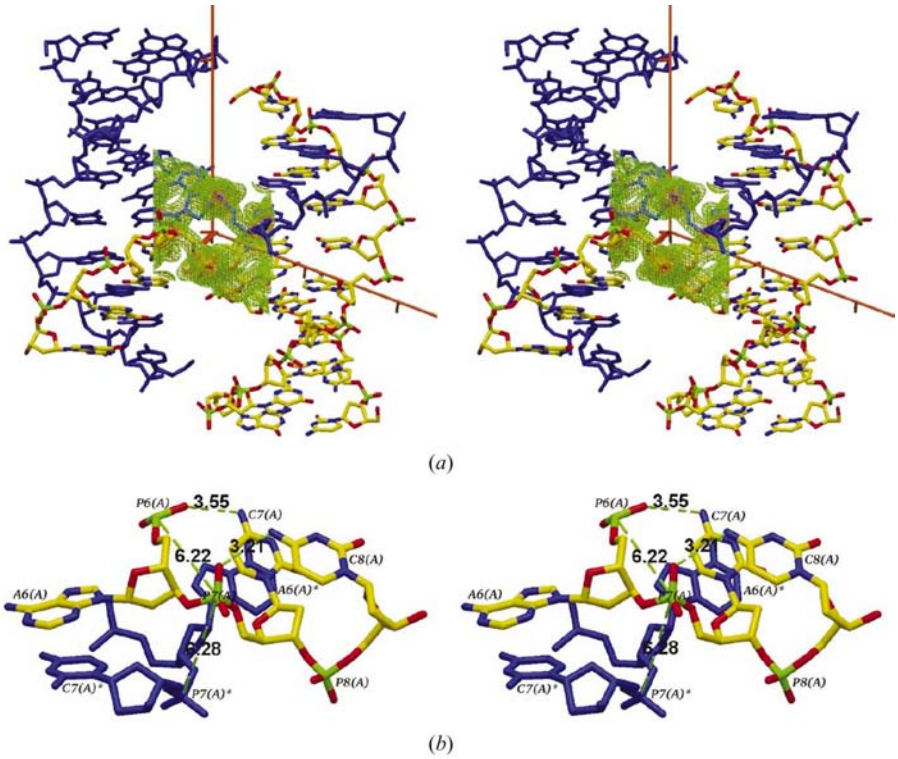

Figure 1
(*a*) Stereo diagram showing the minor-groove face (viewed down the b axis) of the stacked-X junction, illustrating the symmetry-related half (blue) generated by the twofold axis at the cell (orange) origin and the σ_A map covering the crossover (green = 1σ and orange = 2σ). (*b*) Stereo diagram illustrating the crossover fragment of the structure in an off-axis projection, with an inter-phosphate distance across the junction of 6.28 Å [P7(A)–P7(A)*] and long-range hydrogen bonds between cytosine N4 and phosphate O atoms at [P7(A)–C8(A)] and prior to the crossover [P6(A)–C7(A)]. Bond distances are given in Å and symmetry-related residues are coloured blue for clarity.

3. Results and discussion

The relationship between the $C2$ lattice adopted and the formation of duplex or quadruplex has been very clearly defined by Eichman *et al.* (2000). In their work, the asymmetric unit is either a quadruplex or a duplex, with the c-axial length doubled when the sequence crystallizes as a quadruplex, where the four identical DNA strands have an approximate twofold symmetry. The overall packing is remarkably similar whether a duplex or a quadruplex has been

[1] Supplementary data have also been deposited in the IUCr electronic achive (Reference: fw0019). Services for accessing these data are described at the back of the journal.

formed, with duplex helices crossing at approximately the same angle in either case, shown in Table 1. It is striking that the reported crossover angles for the two duplexes show a larger variability than all the reported quadruplexes. Table 1 also makes clear that it would be very unwise to use the lattice constants as a guide to whether a particular sequence has crystallized as a quadruplex or duplex; any potentially quadruplex-forming sequence in this packing should be fully refined to reveal the direction of the connected backbone density. Prior to the crystallization of any authentic junction, the close resemblance to a junction-like packing was reported for the DNA sequence d(CGCAATTGCG)$_2$, as a potential model for the Holliday junction (Wood *et al.*, 1997), prior to its final solution. In the present work, the four-way junction formed shows a twofold crystallographic axis running through the centre of the helix crossover (Fig. 1*a*), which is now isomorphous with the structure obtained using the mismatched sequence d(CCGGGACCGG) (Ortiz-Lombardia *et al.*, 1999). From these structural results, we can begin to build an overall picture of the consistency of the Holliday stacked-X conformation. Table 1 illustrates each structure to be remarkably similar at the low resolutions thus far obtained, with helix-crossover angles of $\sim$40° for these crystallographic studies, an angle not as steep as the 60° reported in solution studies (Duckett *et al.*, 1988). Also consistent, but in this case not differentiated from the two-duplex model, is the separation of the phosphates at the junction crossover (see Fig. 1*b*), where we observe a distance of 6.28 Å between them [P7(A)–P7(A)*]. These strained phosphates are stabilized through hydrogen-bonded interactions between the cytosine N4 sites on either side of the crossover, with the cytosine prior to the crossover hydrogen bonding to the phosphate O atom at the centre of the stacked-X motif [C8(A)–P7(A)] and the cytosine at the crossover hydrogen bonding to the phosphate O atom prior to the crossover [C7(A)–P6(A)] in the major groove. These observed hydrogen bonds in the current model are fairly long, with the shortest distance being 3.21 Å. However, they also go some way towards explaining why such a junction has not previously been observed for the numerous decanucleotide sequences which have been previously studied, as it would appear these two cytosine residues are essential for the stabilization of the strained phosphate crossover, as proposed by Eichman *et al.* (2000).

The role of metal ions in the structure and function of nucleic acids is central and has long been studied (*e.g.* Schneider *et al.*, 1996). Metal ions can bind to several donor sites on polynucleotides, but chiefly bind to phosphate links along the backbone and nitrogen donors on bases. They have been assumed to have key roles in junction stabilization, particularly of the junction phosphates. In the previously reported junction structure of Eichman *et al.* (2000), a cation site, identified as Na^+ from its short distances to adjacent O atoms (<2.5 Å), bridges the strained phosphates at the junction crossover. Analysis of the crystallization conditions employed for the other Holliday sequences shows each of them to have been crystallized in the presence of a group II metal halide (Mg^{2+} or Ca^{2+}) in addition to other Na^+ salts, but no divalent cations have been located in any of the quadruplex structures. The crystals grown here were grown in the presence of only the group I halide KCl (plus other Na^+ salts), with no ions definitively placed; more importantly, none were found on the twofold axis at the centre of the junction, although 43 other distinct water sites were located during the refinement. None of these have a coordination environment characteristic of a group I metal. The role of metal ions in junction stabilization clearly warrants further investigation.

The structural characterization of the Holliday junctions has provided a major extension of our insight into the conformational flexibility of DNA and its higher order structure. From these structures, we can begin to understand the enzyme-recognition processes which must occur, allowing the design of DNA-binding agents exhibiting a specificity towards such junctions. Such compounds providing a potential means of blocking the recombination process at the cellular level, a huge step forward in the drug-discovery program.

We acknowledge access to the EMBL Hamburg synchrotron facility at DESY supported by the TMR/LSF and HPRI programmes of the European Union. For financial support we thank the University of Reading (to JHT and SCMT) and the Association for International Cancer Research (to BCG).

References

Ceschini, S., Keeley, A., McAlister, M. S. B., Oram, M., Phelan, J., Pearl, L. H., Tsaneva, I. R. & Barrett, T. E. (2001). *EMBO J.* **20**, 6601–6611.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Déclais, A.-C., Hadden, J., Phillips, S. E. V. & Lilley, D. M. J. (2001). *J. Mol. Biol.* **307**, 1145–1158.
Duckett, D. R., Murchie, A. I. H., Diekmann, S., von Kitzing, E., Kemper, B. & Lilley, D. M. J. (1988). *Cell*, **55**, 79–89.
Eichman, B. F., Mooers, B. H. M., Alberti, M., Hearst, J. E. & Shing-Ho, P. (2001). *J. Mol. Biol.* **308**, 15–26.
Eichman, B. F., Vargason, J. M., Mooers, B. H. M. & Shing-Ho, P. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3971–3976.
Holliday, R. (1964). *Genet. Res.* **5**, 283–304.
Lavery, R. & Sklenar, H. J. (1988). *J. Biomol. Struct. Dyn.* **6**, 63–91.
Lilley, D. M. & White, M. F. (2001). *Nature Mol. Cell Biol.* **2**, 433–443.
McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nishino, T., Komori, K., Tsuchiya, D., Ishino, Y. & Morikawa, K. (2001). *Structure*, **9**, 197–204.
Ortiz-Lombardia, M., Gonzalez, A., Eritja, R., Aymani, J., Azorin, F. & Coll, M. (1999). *Nature Struct. Biol.* **6**, 913–917.
Otwinowski, Z. & Minor, W. (1995). *The HKL Manual.* New Haven, CT, USA: Yale University Press.
Ristriani, T., Nominé, Y., Masson, M., Weiss, É. & Travé, G. (2001). *J. Mol. Biol.* **305**, 729–739.
Schneider, B., Kabelác, M. & Hobza, P. (1996). *J. Am. Chem. Soc.* **118**, 12207–12217.
Sekiguchi, J., Seeman, N. C. & Shuman, S. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 785–789.
Vagin, A. A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
West, K. L. & Austin, C. A. (1999). *Nucleic Acids Res.* **27**, 984–992.
Wood, A. A., Nunn, C. M., Trent, J. O. & Neidle, S. (1997). *J. Mol. Biol.* **269**, 827–841.