Minireview

Caution! DNA Crossing: Crystal Structures of Holliday Junctions*

Published, JBC Papers in Press, October 16, 2003,
DOI 10.1074/jbc.R300033200

Franklin A. Hays, Jeffrey Watson,
and P. Shing Ho‡

From the Department of Biochemistry and Biophysics,
Oregon State University, Corvallis, Oregon 97331-7305

A four-stranded DNA junction (Fig. 1a) was first proposed by Robin Holliday in 1964 as a structural intermediate in a mechanistic model to account for the means by which genetic information is exchanged in yeast (1). This mechanism for genetic exchange is now generally known as homologous recombination and the four-stranded intermediate as the Holliday junction. The general mechanism for recombination has undergone a number of revisions in detail, but the Holliday junction remains a key component in the process and in a growing number of analogous cellular mechanisms (reviewed in Refs. 2 and 3), including site-specific recombination (4), resolution of stalled replication forks (5, 6), DNA repair (7, 8), and phage integration (9). Consequently, the structural and dynamic properties of the Holliday junction have been the focus of intense study since the mid-1960s. As is often the case in science, a multitude of single crystal structures of Holliday junctions in various forms have been solved over a relatively short period of time but only after decades of disappointment. The first structures of junctions in complexes with recombination and DNA repair proteins were reported in 1997 and 1998 (10, 11), whereas junctions in RNA-DNA complexes (12, 13) and in DNA-only constructs (14, 15) emerged just as the twentieth century came to a close. In this review, we will focus on the structures of DNA-only junctions and their geometries, as defined by sequence and ion-dependent interactions.

BC (Before Crystal) Structures

Before any of the crystal structures of junctions were available, the general architecture and dynamic properties of the Holliday junction were already being defined through biochemical and biophysical studies on immobilized junctions assembled from asymmetric sequences (reviewed in Ref. 2). From comparative gel and solution studies, the junction was known to undergo ion-dependent conformational transitions. In low salt solutions, the junction adopts an extended "open-X" form to minimize the strong repulsion between negatively charged phosphates at the crossover strands that exchange between DNA double helices (Fig. 1b). The junction collapses to a stacked-X form when the phosphates are screened by cations, particularly by magnesium (Mg$^{2+}$) and other divalent ions. This more compact junction pairs the arms into coaxially stacked, semicontinuous duplexes, with each stacked duplex having an outside non-exchanging strand and an inside crossover strand. The crossed-over strands can be oriented either in parallel or antiparallel directions.

Although the parallel configuration of the stacked-X junction (Fig. 1a) has been suggested to be formed by extrusion of cruciform structures in negatively supercoiled closed circular DNAs (16, 17) and in synthetic DNA arrays (18), it is the antiparallel form that is seen by gel (19–21) and, subsequently, fluorescence resonance energy transfer studies (22–24) and by atomic force microscopy in interlinked sheets of DNA (25, 26). The basic model that emerged for the antiparallel stacked-X junction (20, 27) has the arms paired and stacked as near-continuous B-type DNA duplexes, interrupted only by the turn of the junction crossover. The two opposing stacked duplexes are related by a right-handed twist to give an interduplex angle of +60° across the junction. One of the problems with the antiparallel configuration, however, is that it does not allow migration of the junction along homologous DNA sequences. A resolution to this problem is that the stacked-X junction is not static but can undergo dynamic conformational rearrangements: a conformational isomerization can occur in which the stacked duplex arms switch partners in a sequence-dependent manner (28–30). This isomerization transition has recently been observed in single molecules (31) and has been proposed to involve the open-X structure (which does support junction migration) as a transient intermediate configuration (14). Additionally, stacked-X junction isomerization can affect the genetic outcome of homologous recombination through production of either patched or spliced gene products. The single crystal structures of DNA junctions and junctions in complex with proteins are now providing conformational details of the two junction forms at the atomic level and, particularly with the stacked-X form, an understanding of how sequence and ions affect these structural details.

Single-crystal Structures of DNA Holliday Junctions

The first single-crystal structure of a DNA junction was seen in the Cre recombination complex (10). There are now structures of at least three unique complexes with recombination proteins (11, 32, 33), and in all cases, the junctions adopt the extended open-X form, with the DNA geometries determined primarily by the topology of the binding surfaces of the proteins (reviewed in Ref. 3). The question of how DNA sequence and other factors affect the inherent structure of the Holliday junction remains and is being answered by single-crystal structures in the absence of proteins.

The first DNA Holliday junction crystal structures came not by design but by happenstance while other problems were being attacked. The sequence d(CCGGGACCGG) was the first structure of a DNA Holliday junction, but the initial intent was to study the effects of tandem GA mismatches on B-DNA (14). Shortly afterward, the structure of a junction with all standard Watson-Crick base pairs (15) was solved as a control study for the psoralen cross-linked structure of d(CCGGTACCGG) (34). In these and all subsequent DNA-only studies, the junctions adopt the antiparallel stacked-X conformation (Fig. 2) with general structural features that are remarkably similar to the model proposed in 1988 from solution and gel studies (20, 27).
Minireview: Crystal Structures of DNA Junctions

There were, however, some surprising observations that could not be predicted from earlier studies on asymmetric junctions, first being that these decanucleotide sequences could be crystallized as junctions at all. The single-crystal structures of many 10-base pair sequences have been reported since 1987 (reviewed in Ref. 35) as B-DNA, A-DNA, and left-handed Z-DNA. In fact, early models of junctions were constructed using structures of decanucleotides in crystal lattices that resemble the general geometry of the stacked-X junction (36, 37). What became apparent when the actual Holliday junction structures were solved was that a common A6-C7-C8 trinucleotide sequence and an associated set of unique intramolecular interactions define a common core in the two junctions (15, 38). Since these two early breakthroughs, we are learning that both sequence and ions at and near this ACC core affect these interactions and consequently the geometry of the stacked-X junction, in both subtle and profound ways.

Sequence and Ions Affect the Geometry of DNA Junctions

The mismatched and inverted repeat junctions are very similar in structure (38) but show minor structural differences that could be attributed to an unusual base interaction at the G-A mismatched base pairs (14). The first indication that the standard geometry of Holliday junctions could be perturbed in a systematic way was seen in the structure of the d(CCGGTACm5CGG) (39), where m5C is a 5-methylcytosine. This methylated structure demonstrated that, in addition to the interduplex angle, the stacked duplexes can be rotated and shifted relative to each other across the junction crossover (we label these $J_{\text{roll}}$ and $J_{\text{slide}}$, respectively, Fig. 1d) and that these two geometric parameters were dependent on the interactions at the junction core. The $J_{\text{slide}}$ parameter measures the extent to which the stacked helices translate along their respective helix axes relative to the center of the junction crossover, whereas $J_{\text{roll}}$ rotates these duplexes about the helix axes (a large $J_{\text{roll}}$ opens the major groove surfaces of the junction crossover, with a value of $J_{\text{roll}} = 180^\circ$ indicating that the major and minor grooves of the junction are equally accessible).

There are now 12 junction structures deposited in the Nucleic Acid Database (40) representing six unique sequence types. All of the structures are in the antiparallel stacked-X form and vary according to their interduplex angles, $J_{\text{roll}}$, and $J_{\text{slide}}$ (Table I). What have we learned from these structures? Our first lesson is that the sequence requirements at the junction core are not exclusive to ACC but can be expanded to R6-C7-Y8, where R6 can be either the purine bases adenine or guanine, and Y8 can be cytosine, 5-methylcytosine, or 5-bromouracil but not thymine. This may be expanded further as more sequences become crystallized as junctions. The terminal base pairs (C1-G10) are not critical for formation of junctions and can be replaced by T1-A10 base pairs, as long as the trinucleotide core sequence remains intact (41).

As the sequence is changed and/or base substituent groups are added at the trinucleotide core, it is becoming increasingly clear that the core intramolecular interactions not only dictate the formation of the junction but also its geometry as defined by the interduplex angle and, more significantly, $J_{\text{roll}}$ and $J_{\text{slide}}$. In this discussion, we will start by defining a standard set of interactions that are associated with the reference “unperturbed” d(CCGGTACCGG) (ACC-4Na, see Table I for the description of this nomenclature) (15) (Fig. 2). In this case, $J_{\text{slide}} \approx 0^\circ$ and $J_{\text{roll}} \approx 160^\circ$. The key interactions in this reference structure involve base substituents in the major grooves of the stacked duplex arms and the phosphate oxygens of the junction crossover and include 1) a direct hydrogen bond between the N-4 nitrogen of the C8 base and the C7 phosphate oxygen (C8-C7 interaction), 2) a solvent-mediated interaction from the complementary G3 nucleotide and the A6 phosphate oxygen (G3-A6 interaction), and 3) a direct hydrogen bond from the N-4 nitrogen of the C7 base to the...
A6 phosphate oxygen (C7-A6 interaction) (Fig. 2, insets). The first interaction helps explain the requirement for a cytosine at the C8 position, whereas the last interaction, for the first time, provides a rationale for the C7 cytosine of the original ACC core trinucleotide. The involvement of the A6 phosphate in two of these three interactions may explain the importance of having a purine base at the R6 position of the core.

As we change this reference sequence, we see losses in the core interactions that affect both the structure and stability of the junction. For the most part (with one exception) structures that maintain all three key interactions effectively adopt the unperturbed reference structure of the ACC-4Na junction. The GA mismatched (gACC-2Na) junction (14) has lost the C7-A6 interaction, and, consequently, we see a significant perturbation to J\text{roll} and J\text{slide} in the closely related ACC-2Na junction (45).

The other significant cation interactions involve aqueous calcium (39) and strontium clusters (41) that sit in the grooves of the stacked duplexes in several junctions. Again, the effects of these cation complexes on the structure of the junction depend on how they perturb the core interactions. For example, five strontium-water complexes were located in the minor grooves of each stacked duplex of the tACC-2Sr1 (41) and tACC-2Sr2 junctions. In both cases, additional waters intervene in the core G3-A6 and C7-A6 interactions, resulting in nearly identical perturbations to J\text{roll} and J\text{slide}.

**Conclusions and Perspectives**

We are now starting to understand how sequence, ions, and drugs (this latter effect, although not discussed here, is covered previously in Refs. 34 and 46) affect the detailed conformation of the Holliday junction in single crystals. As with any set of crystallographic structures, however, we need to address the question of whether the conclusions that we draw from these studies correspond to how DNA behaves in solution and in the cell. In most respects, the crystal structures do indeed conform...
to the general features of Holliday junctions as defined from previous studies, with one exception. In all cases, the interduplex angles of the structural crystals are about 20° more shallower than the ∼60° observed from earlier gel mobility, fluorescence resonance energy transfer, and atomic force microscopy studies (19–25). One reason for this discrepancy may be that the crystallographic studies all use symmetric sequences where the junction is fixed in place by specific intermolecular interactions. In contrast, the non-crystallographic studies on junctions were nearly all applied to immobilized junctions that are assembled from four unique asymmetric sequences and are thus locked in place by sequence design rather than by any defined set of molecular interactions, other than standard Watson-Crick base pairing. Consequently, we would expect these sequence-immobilized constructs to be missing the core and even solvent-mediated interactions between the stacked arms and the phosphates at the junction crossover seen in the crystal structures and therefore would show significant deviation from the reference ACC-4Na junction. This is supported by atomic force microscopy studies, which show that indeed the interduplex angles of asymmetrically locked junctions are ∼60°, whereas the angle in similar constructs that incorporate the symmetric ACC core trinucleotide is ∼43°, as seen in the crystal structures (26). In addition, these same ACC-containing constructs were shown by hydroxyl radical footprinting to cross over at the nucleotide positions seen in the crystal structures (26). Thus, we can assert that the sequence-dependent interactions at the junction core, as identified in the single-crystal structures, are relevant in other systems and that these interactions do indeed define the conformational details of the stacked-X junction even in solution.

Are these molecular details important for the participation of Holliday junctions in recombination-dependent cellular processes? As we know, junctions in the antiparallel stacked-X form cannot migrate along the DNA sequence but can do so if it adopts the open-X form. In solution, stacked-X stability and conformational switching rates have been shown to be dependent on the population of open-X junction available for branch migration (47, 48). We see that disrupting one or more of the core interactions allows the stacked-X junction to roll, slide, and twist. We suggest, therefore, that sequences in which this set of core interactions is disrupted or compromised by intervening solvent will be more conformationally malleable and thus be more amenable to switching to an open-X form and, consequently, promote branch migration.

How might these core interactions affect protein binding? As we have already seen, the junctions in protein complexes all adopt the open-X and not the stacked-X form. Thus, the barriers to conformational switching that we have discussed above should also contribute to the energetics of protein binding, any effect that facilitates formation of the open-X structure should also favor protein binding. In addition, we see that these core interactions, as they are affected by sequence and ions, specify the accessibility of the major and minor grooves (as measured by the parameter $J_{\text{roll}}$). Many junction-resolving enzymes, such as T4 endonuclease VII (49), are known to recognize the stacked-X structure from the minor groove face, and therefore we see that accessibility also plays an important role in protein recognition and binding. In short, it appears that a stacked-X junction that is structurally malleable, which we propose is determined by a set of core interactions, will also have a lower barrier to conformational switching and be more accessible, both of which are important for binding and recognition by junction-resolving and DNA repair enzymes. At this point, however, it is unclear exactly how specific proteins take advantage of the structural and dynamic features of the stacked-X junction to define their sequence-dependent functions.

REFERENCES