Dissection of the Sequence Specificity of the Holliday Junction Endonuclease CCE1†

Mark J. Schofield, David M. J. Lilley, and Malcolm F. White*

CRC Nucleic Acid Structure Group, Department of Biochemistry, University of Dundee, Dundee DD1 4HN, U.K.

Received February 19, 1998; Revised Manuscript Received March 26, 1998

ABSTRACT: CCE1 is a Holliday (four-way DNA) junction-specific endonuclease which resolves mitochondrial DNA recombination intermediates in *Saccharomyces cerevisiae*. The junction-resolving enzymes are a diverse class, widely distributed in nature from viruses to higher eukaryotes. In common with most other junction-resolving enzymes, the cleavage activity of CCE1 is nucleotide sequence-dependent. We have undertaken a systematic study of the sequence specificity of CCE1, using a single-turnover kinetic assay and a panel of synthetic four-way DNA junction substrates. A tetranucleotide consensus cleavage sequence 5′-ACT↓A has been identified, with specificity residing mainly at the central CT dinucleotide. Equilibrium constants for CCE1 binding to four-way junctions are unaffected by sequence variations, suggesting that substrate discrimination occurs predominantly in the transition state complex. CCE1 cuts most efficiently at the junction center, but can also cleave the DNA backbone at positions one nucleotide 3′ or 5′ of the point of strand exchange, suggesting a significant degree of conformational flexibility in the CCE1:junction complex. Introduction of base analogues at single sites in four-way junctions has allowed investigation of the sequence specificity of CCE1 in finer detail. In particular, the N7 moiety of the guanine base-pairing with the cytosine of the consensus sequence appears to be crucial for catalysis. The functional significance of sequence specificity in junction-resolving enzymes is discussed.

Holliday junctions, or four-way DNA junctions, are a central intermediate of homologous recombination (*1*−*3*), a process that plays dual, fundamental roles both in the generation of genetic diversity and in the repair of certain types of DNA damage. The enzymology of homologous recombination is most clearly understood for *Escherichia coli*. A RecA-catalyzed strand exchange between homologous DNA sequences gives rise to a Holliday junction, which can branch migrate in a reaction catalyzed by RecG or the RuvAB complex to form heteroduplex DNA. Recombination is completed by the RuvC endonuclease, which resolves the Holliday junction into two recombinant DNA duplex molecules. Junction-resolving enzymes have been detected in a wide range of organisms from bacteriophage and eukaryotic viruses to eubacteria and eukaryotes. Sequence information predicts that they are a structurally diverse class, suggesting that enzymes have been recruited from other functions to fulfill this role several times during evolution. For example, solution of the crystal structure of RuvC revealed a common fold with the nuclease RNaseH1, while sequence comparisons suggest the zinc-binding domain of bacteriophage T4 endonuclease VII is present in proteins of unknown function from *E. coli* and HPV (*4*), and that the junction endonuclease CCE1 of *Saccharomyces cerevisiae* is homologous to a yeast group I intron splicing accessory protein, MRS1 (*5*). Despite the apparently varied pedigrees of the junction-resolving enzymes, all possess an exquisite specificity for the structure of the four-way DNA junction and manipulate that structure on binding (*6*).

Overlaying the structural specificity, several of the Holliday junction-resolving enzymes display a sequence specificity for DNA cleavage. This sequence specificity may be a remnant of the ancestral nuclease, or may be a functional requirement for homologous recombination. In eukaryotes, the sole junction endonuclease identified is CCE1 of *S. cerevisiae* (*7*, *8*) (together with its homologue in *Schizosaccharomyces pombe*; YDC2/spCCE1 (*9*, *10*)), which plays a role in the resolution of recombining mitochondrial DNA genomes (*11*). Previously, we overexpressed and purified recombinant CCE1 in *E. coli* and showed that the enzyme cleaves fixed and mobile four-way junctions, with a preference for 5′-CT↓ sequences (*5*). In this paper we report the results of a systematic effort to examine the sequence requirements of the enzyme, using a matrix of fixed four-way junction substrates in a single-turnover kinetic assay. Four-way junctions fold in the presence of divalent metal ions by pairwise coaxial stacking of helical arms, adopting an antiparallel stacked X-structure comprising two continuous and two exchanging strands (reviewed in (*12*, *13*)). Stacked X junctions can exist in two different

† This work was supported by the Cancer Research Campaign. M.F.W. is a Royal Society University Research Fellow.

* Author to whom correspondence should be addressed.

S0006-2960(98)00399-7 CCC: $15.00
Published on Web 05/06/1998

conformers, depending on the choice of stacking partners, and this depends on the local base sequence (reviewed in ref *14*). As resolving enzymes cleave at or near the junction center, any study of enzyme sequence preferences neccessitates the alteration of nucleotides flanking the point of strand exchange, which may alter the stacking preference of the junction. For CCE1, this complication is avoided, as comparative gel electrophoretic analysis and permanganate probing have shown that CCE1 unfolds the stacked X-structure of the four-way junction on binding, producing an open, 4-fold symmetric junction structure similar to that seen in the presence of EDTA (*15*). We define a tetranucleotide recognition sequence 5′-ACT↓A for CCE1, with cleavage favored at the point of strand exchange. Specificity resides principally at the G-C base pair and T nucleotide in the center of the recognition sequence; changes in any of these three nucleotides result in reductions in cleavage rate akin to those seen for restriction enzymes cutting noncanonical sites. We conclude that CCE1 has a high degree of sequence specificity for junction cleavage (though not binding), and that this may be a requirement of its biological role.

EXPERIMENTAL PROCEDURES

Oligonucleotides. Junctions Z1, Z1−3′, and Z1−5′. These are fixed four-way junctions with arms of 15 bp, made from the following sets of four 30 nt oligonucleotides. In the h strand, the point of strand exchange is indicated by a | symbol, and the tetranucleotide recognition sequence for cleavage by CCE1 is shown in bold.

Junction Z1

b 5′ TCCGTCCTAGCAAGGAGTCTGCTACCGGAA

h 5′ TTCCGGTAGCAG**ACT**|**A**AGCGGTGGTTGAAT

r 5′ ATTCAACCACCGCTTTCTCAACTGCAGCAG

x 5′ CTGCTGCAGTTGAGACCTTGCTAGGACGGA

Junction Z1-3'

b 5' AGGGACCTAGCAAGGGTCTGCTACCCTTCT

h 5' AGAAGGGTAGCAG**AC**|**TA**AGCGGTGGTTCCT

r 5' AGGAACCACCGCTTAACTCAACTGCAGACT

x 5' AGTCTGCAGTTGAGTCCTTGCTAGGTCCCT

Junction Z1-5'

b 5' AGGGACCTAGCAAGGTAGTCTGCTACGCTT

h 5' AAGCGTAGCAG**ACTA**|GGCGGTGGTTCCTAG

r 5' CTAGGAACCACCGCCACTCAACTGCTCTAG

x 5' CTAGAGCAGTTGAGTCCTTGCTAGGTCCCT

Junctions Z2-Z15 were created by changing one or more bases in the tetranucleotide recognition sequence of the h strand, together with the concomitant changes in the b or r strands to preserve base complementarity. All flanking sequences were unaltered from those for junction Z1.

	position			
Junction	−3	−2	−1	+1
Z1	A	C	T	A
Z2	T	C	T	A
Z3	C	C	T	A
Z4	G	C	T	A
Z5	A	G	T	A
Z6	A	A	T	A
Z7	A	T	T	A
Z8	A	C	A	A
Z9	A	C	C	A
Z10	A	C	G	A
Z11	A	C	T	C
Z12	A	C	T	T
Z13	A	C	T	G
Z14	T	C	T	T
Z15	C	C	T	T

Junctions Z16-Z21 contain single base changes in the h strand of junction Z1, incorporating a base analogue or mismatch at position −1 or −2. All flanking sequences, together with the sequences of the other three arms, are identical to junction Z1.

Junction	−3	−2	−1	+1
Z16	A	5meC	T	A
Z17	A	dU	T	A
Z18	A	T	T	A
Z19	A	C	dU	A
Z20	A	C	5meC	A
Z21	A	C	C	A

Junctions Z22-Z27 contain single base changes in the b strand of junction Z1, incorporating a base analogue or mismatch at position −1′ or −2′. All flanking sequences, together with the sequences of the other three arms, are identical to junction Z1.

Junction	−1'	−2'
(Z1	A	G)
Z22	G	G
Z23	7deazaA	G
Z24	C	G
Z25	T	G
Z26	A	7deazaG
Z27	A	A

Oligonucleotide Synthesis. Oligonucleotides were synthesized on a 394 DNA/RNA synthesizer (Applied Biosystems), purified by gel electrophoresis in 12% polyacrylamide containing 7 M urea, the bands excised and DNA eluted and recovered by ethanol precipitation. Oligonucleotides were radioactively labeled at their 5′-termini using [γ-^{32}P]ATP and T4 polynucleotide kinase.

Assembly of four-way DNA Junctions. Stoichiometric quantities of three unlabeled and one 5′-^{32}P-labeled strands were annealed by incubation in 50 mM Tris·HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol, 0.1 mM spermidine, and 0.1 mM EDTA for 3 min at 85 °C followed by slow cooling. Assembled junctions were purified by gel electrophoresis in 5% polyacrylamide gels. Bands were excised and DNA recovered by electroelution.

Protein Purification. Native recombinant CCE1 was expressed and purified to near-homogeneity as described previously (*15*).

Single-Turnover Kinetic Analysis. Determination of the first-order rate constants of junction cleavage were carried out by first incubating 8×10^{-8} M 5′-^{32}P-labeled junction with 5×10^{-7} M CCE1 dimer in binding buffer (20 mM Tris, pH 8.0, 200 mM NaCl, 0.2 mM DTT, 1 mM EDTA, 0.1 mg/mL bovine serum albumin). Under these conditions, the dissociation constant for junctions varies from 0.5 to 5×10^{-9} M, so all junction will be bound. Samples were preequilibrated at 37 °C, and reactions were initiated by the addition of $MgCl_2$ to a final concentration of 15 mM. At set time points an aliquot of the reaction mix was removed and the reaction was stopped by the addition of an equal volume of formamide loading buffer (95% v/v formamide, 50 mM EDTA, pH 8.0, 0.1% bromophenol blue, 0.1% xylene cyanol FF) and heating at 80 °C for 2 min, followed by storage on ice. Reaction products were analyzed by denaturing gel electrophoresis on 0.4 mm thick polyacrylamide gels in TBE buffer containing 7 M urea, run at 50 °C. After electrophoresis, gels were dried on Whatmann 3MM paper and exposed to X-ray film for documentation or phosphor screens for quantification on a Fuji BAS-1500 phosphorimager.

Measurement of Dissociation Constants. A range of concentrations of purified CCE1 protein was incubated with a known concentration of four-way DNA junction radioactively 5′-^{32}P-labeled on each of the four strands in binding buffer in 10 μL total volume. After 15 min at 20 °C, $^1/_6$ volume loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 35% Ficoll type 400) was added, and samples were loaded onto 5% polyacrylamide gels (29:1 acrylamide to bisacrylamide) and electrophoresed in 1× TBE buffer. After electrophoresis, gels were dried and exposed to X-ray film for documentation or phosphor screens for quantification as described above. Data were analyzed as fraction DNA bound (f_b) versus protein concentration and were fitted by nonlinear regression analysis to the equation:

$$f_b = \frac{(1 + K_A P_T + K_A D_T) - \sqrt{(1 + K_A P_T + K_A D_T)^2 - (4 D_T K_A{}^2 P_T)}}{2 D_T K_A}$$

where P_T is the total protein concentration (calculated for a dimer) and D_T is the total DNA concentration. At higher concentrations of CCE1, where a discrete further-retarded band was visible, it was counted as bound DNA. For the purposes of the calculation, CCE1 was considered to be a dimer in solution. The dissociation constant (K_D) is the reciprocal of K_A.

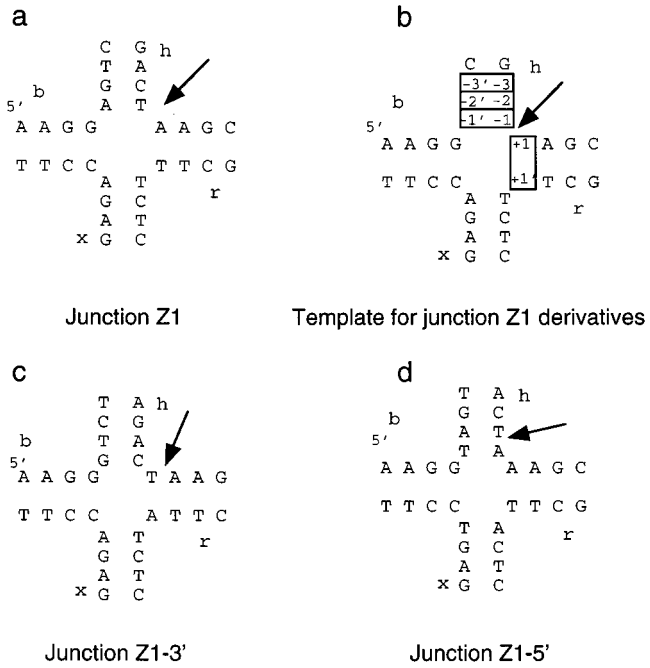

FIGURE 1: Central sequences of junctions used in this study. The central eight nucleotides of each strand (b, h, r, x) are shown, with the position of cleavage by CCE1 indicated by an arrow. (a) The central core of the fixed junction Z1, the reference junction for this study, which is cleaved by CCE1 on the h strand at the point of strand exchange, as indicated by the arrow. (b) Template for derivatives Z2-Z27 of junction Z1, indicating the nucleotide positions −3, −2, −1, and +1 where sequence changes were introduced. (c) The central core of junction Z1-5′, showing the displacement of the 5′-ACTA cleavage recognition sequence one nucleotide in the 5′ direction with respect to junction Z1. (d) The central core of junction Z1−3′, showing the displacement of the 5′-ACT↓A cleavage recognition sequence one nucleotide in the 3′ direction with respect to junction Z1.

RESULTS

Systematic Investigation of Sequence Specificity of CCE1. We have shown previously that CCE1 cuts a number of fixed and branch-migrating four-way junctions after 5′-CT sequences, and that alterations in the C to the three other possible nucleotides resulted in a shift in the cleavage site with a concomitant drop in reaction rate (*5*). Subsequent investigation of further junction sequences revealed that the recognition sequence 5′-ACT↓A, where the arrow represents the cleavage site at the point of strand exchange, was cut particularly quickly by CCE1. To investigate the sequence requirement of the enzyme in more detail, junction Z1 was synthesized, with arms of 15 nt, containing the 5′-ACT↓A sequence in the h strand (Figure 1). Junction Z1 was used as a basis for the production of a series of derivatives, each with one nucleotide changed at one of the four positions (together with the appropriate alteration in the sequence of the complementary strand to preserve Watson−Crick base-pairing). The four nucleotide positions of the recognition sequence are designated −3, −2, −1, and +1 to indicate their position in the h strand relative to the point of cleavage.

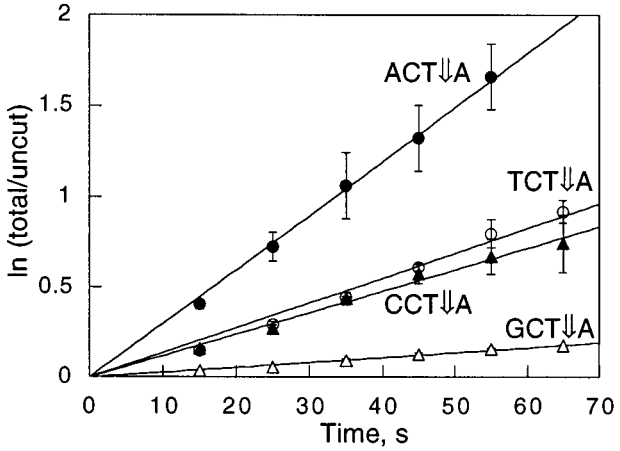

FIGURE 2: Variation in the rate of junction cleavage by CCE1 as a function of nucleotide sequence at position −3. Single-turnover kinetic analysis was carried out as described in Experimental Procedures. The natural logarithm of the ratio (total junction/uncleaved junction) is plotted against time of cleavage. Data points are the means of triplicate measurements, with error bars showing standard deviations. The first-order rate constant (k_c) was calculated from the gradient of a line fitted to the data by linear regression through the origin. Rate constants are summarized in Table 1. (●) cleavage sequence 5′-ACT↓A; (○) cleavage sequence 5′-TCT↓A; (▲) cleavage sequence 5′-CCT↓A; (△) cleavage sequence 5′-GCT↓A.

Table 1: First-Order Rate Constants (k_c) for Junction Z1 and Derivatives[a]

nucleotide position	A	C	T	G
−3	3000	1200 (0.56)	1400 (0.47)	270 (1.5)
−2	1.5 (4.7)	3000	4.5 (4.0)	2.3 (4.4)
−1	0.43 (5.4)	≤0.1 (≥6)	3000	≤0.1 (≥6)
+1	3000	610 (0.98)	950 (0.71)	1100 (0.62)

[a] Single-turnover rate constants were determined in triplicate, as described in Experimental Procedures, and are expressed as $s^{-1} \times 10^5$. Rates which were below the limits of detection of the assay are marked ≤0.1. The numbers in parentheses are values for $\Delta\Delta G^{\circ\ddagger}$ (kcal/mol) compared to the fastest cutting sequence. Where reaction rates were too slow to be measured, a lower limit of +6 kcal/mol can be estimated for the value of $\Delta\Delta G^{\circ\ddagger}$.

The complementary nucleotides in the b and r strands are designated −3′, −2′, −1′, and +1′. In total, twelve derivative junctions were synthesized, covering every possible single nucleotide substitution in the 5′-ACT↓A sequence. Kinetic analysis was performed under single-turnover conditions by incubating CCE1 (500 nM) with radioactively 5′-^{32}P-labeled four-way junction (80 nM), prior to initiation of the reaction by the addition of magnesium chloride to 15 mM, as described in Experimental Procedures. The assay has an upper limit of rate measurement of approximately 5×10^{-2} s^{-1}, and a lower limit of approximately 1×10^{-6} s^{-1}, allowing determination of reaction rates over a range of 4.5 orders of magnitude. A representative set of data, for junctions with the four possible nucleotides at position −3, are shown in Figure 2.

The data in Table 1 show that the preferred tetranucleotide cleavage sequence for CCE1 is 5′-ACT↓A. All other sequences are cleaved at least 2-fold more slowly. Rates can be expressed in terms of a free energy of activation ($\Delta G^{\circ\ddagger}$), i.e.,

$$k = A_o\, e^{-\Delta G^{\circ\ddagger}/RT}$$

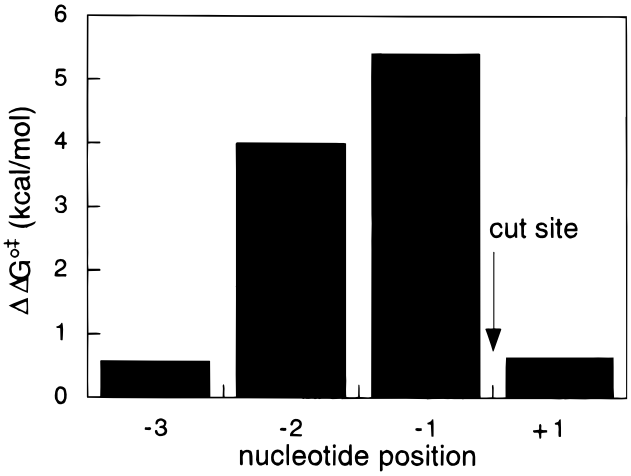

FIGURE 3: Sequence discrimination by CCE1 as a function of nucleotide position in the tetranucleotide cleavage sequence. The difference in free energy (kcal/mol) between the fastest- and second fastest-cutting sequence variant is plotted as a function of nucleotide position in the tetranucleotide cleavage site of junction Z1.

where R is the gas constant, T is the absolute temperature, and A_o is the preexponential factor. It follows that for two different substrates (with rates k_1 and k_2) we can determine a difference in activation free energy $\Delta\Delta G^{\circ\ddagger}$, calculated from:

$$\Delta\Delta G^{\circ\ddagger} = \ln(k_1/k_2)$$

The sequence discrimination at each nucleotide position can be demonstrated by plotting the $\Delta\Delta G^{\circ\ddagger}$ values resulting from the ratio of the fastest and second fastest rates (Figure 3). Specificity resides mainly at the −2 and −1 positions, where, in agreement with previous studies, a CT dinucleotide is highly favored. The next best sequence, TT, is cut over six hundred-fold more slowly. At the outer positions, the degree of sequence discrimination is reduced markedly, with only a 2-fold and a 3-fold discrimination at the −3 and +1 positions, respectively.

Double Substitutions in the Cleavage Sequence Have Additive Effects. To examine whether multiple substitutions in the cleavage sequence have cooperative effects on the cleavage rate, two junctions (Z14 and Z15) were synthesized with alterations in the 5′-ACT↓A reference sequence at both outer positions (positions −3 and +1). In junction Z14, the change in free energy ($\Delta\Delta G^{\circ\ddagger}$) for the double change to 5′-TCT↓T of 1.32 kcal/mol was close to the sum of the values of the equivalent two single changes Z2 and Z12, 1.18 kcal/mol (Table 1). Similarly, the $\Delta\Delta G^{\circ\ddagger}$ of 1.48 kcal/mol for junction Z15, sequence 5′-CCT↓T, is comparable to that of the sum of the two relevant single changes, where $\Delta\Delta G^{\circ\ddagger}$ = 1.27 kcal/mol (Z12 and Z3, Table 1). The additivity observed for double changes suggests that the interaction between protein and DNA at the outer positions of the tetranucleotide sequence can be disrupted independently, with a change at one site having little effect on the other. Equivalent experiments could not be carried out at either of the two inner sites due to the dramatic reduction in cleavage rate brought about by changes at those positions.

CCE1 Cuts Preferentially at the Point of Strand Exchange. Previously, CCE1 has been shown to cut junctions either at,

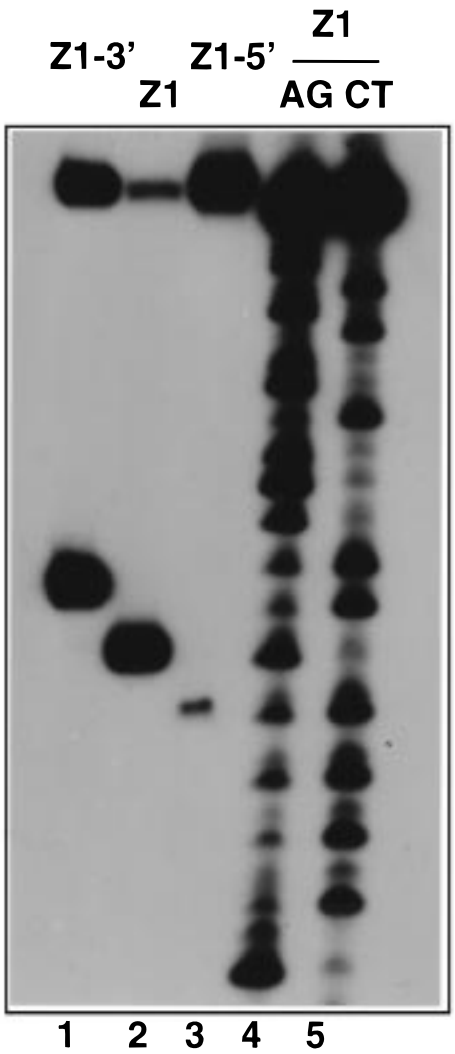

FIGURE 4: Position of cleavage relative to the junction center can be altered by the local sequence. Junctions were radioactively 5′-^{32}P-labeled on the h strand, and were incubated with CCE1 in cleavage buffer, followed by gel electrophoresis and autoradiography. In the reference junction (Z1), cleavage occurs in the 5′-ACT↓A sequence at the point of strand exchange. When the 5′-ACT↓A sequence is moved one nucleotide 3′ (Z1-3′) or one nucleotide 5′ (Z1-5′) with respect to junction Z1, cleavage still occurs at the 5′-ACT↓A site, and these junctions are therefore cleaved one nucleotide 3′ or 5′ of the junction center, respectively. Lane 1, junction Z1-3′ digested by CCE1; lane 2, junction Z1 digested by CCE1; lane 3, junction Z1-5′ digested by CCE1; lane 4 A+G sequence markers for junction Z1; lane 5, C+T sequence markers for junction Z1.

or one nucleotide 3′ of, the point of strand exchange (*5*). To obtain quantitative data on the importance of cleavage sequence position with respect to the point of strand exchange, two junctions were synthesized with the favored cleavage sequence 5′-ACT↓A shifted either one nucleotide 3′ (junction Z1-3′) or one nucleotide 5′ (junction Z1-5′) compared to junction Z1. Sequence mapping revealed that in all three cases CCE1 cut the junction at the same 5′-ACT↓A sequence (Figure 4). The enzyme can therefore cleave junctions at the center or at one nucleotide removed from the point of strand exchange. The cleavage rates (k_c), estimated from single-turnover kinetic analysis, were 3000 $\times$ 10^{-5} s^{-1} for Z1, 300 $\times$ 10^{-5} s^{-1} for Z1-3′, and 4 $\times$ 10^{-5} s^{-1} for Z1-5′. CCE1 therefore cuts junctions preferentially at the point of strand exchange, with cleavage at 1 nt 3′ and 5′ reduced by factors of 10 and 750, respectively.

Table 2: First-Order Rate Constants (k_c) and $\Delta\Delta G^{\circ\ddagger}$ Values for Derivatives of Junction Z1 with Mismatches and Base Analogues[a]

junction	Seq −2′- −2 −1′- −1	k_c ($\times$ 10^5 s^{-1})	$\Delta\Delta G^{\circ\ddagger}$ (kcal/mol)
Z1	G-C A-T	3000	
Z16	G-^{m}C A-T	810	0.81
Z17	G-dU A-T	12	3.40
Z18	G-T A-T	7.6	3.68
Z19	G-C A-dU	240	1.55
Z20	G-C A-^{m}C	4.2	4.04
Z21	G-C A-C	≤0.1	≥6
Z22	G-C G-T	520	1.08
Z23	G-C ^{7C}A-T	270	1.48
Z24	G-C C-T	131	1.93
Z25	G-C T-T	54	2.47
Z26	^{7C}G-C A-T	1.2	4.81
Z27	A-C A-T	0.6	5.24

[a] Single-turnover rate constants were determined in triplicate, as described in Experimental Procedures. Values for $\Delta\Delta G^{\circ\ddagger}$ are calculated by comparison to the cleavage rate for the fastest cutting junction, Z1. Rates below the limits of detection of the assay are marked ≤0.1, giving a lower limit of +6 kcal/mol for $\Delta\Delta G^{\circ\ddagger}$ compared to the fastest cutting sequence.

Introduction of Mismatches and Base Analogues. As the sequence specificity of CCE1 for junction cleavage resides primarily at positions −2 and −1, immediately 5′ of the site of cleavage, we designed a further series of four-way junctions to examine the effect of incorporation of base mismatches and analogues. In these experiments, single nucleotides were altered without changing their base-pairing partner, thus avoiding the potential ambiguity caused by simultaneous alteration of both partners of a base pair in prior experiments. In some cases, this resulted in the generation of mismatches or non-Watson−Crick base-pairing. All of the junctions appeared to fold correctly, as assessed by gel mobility during purification and the unaltered binding affinity displayed by CCE1 (see below). The kinetic data are summarized in Table 2.

Guanine at Position −2′. Substitution of the guanine residue in the b strand at position −2′ to 7-deazaguanine (Z26) resulted in a reduction in cleavage rate of over 3 orders of magnitude, representing an increase in $\Delta\Delta G^{\circ\ddagger}$ of approximately 5 kcal/mol. Substitution of the N7 ring nitrogen with carbon represents a small defined change in the major groove that is unlikely to affect the local structure of the junction or the base-pairing, and yet this has a dramatic effect on cleavage rate. CCE1 may therefore make a specific interaction with the N7 nitrogen which is integral to the catalytic mechanism. An adenine at this position (Z27) reduced the rate to approximately the same level as that seen for 7-deazaguanine.

Cytosine at Position −2. Cytosine was substituted by 5-methylcytosine, deoxyuracil, and thymine. Introduction of a 5-methyl group (Z16) had a relatively minor effect on the reaction rate, $\Delta\Delta G^{\circ\ddagger} = 0.8$ kcal/mol. Substitutions of cytosine by deoxyuracil (Z17) and thymine (Z18) result in a mismatched base pair, reducing the reaction rate by 3.4 and 3.7 kcal/mol, respectively.

Thymine at Position −1. In experiments mirroring those carried out for C^{-2}, T^{-1} was substituted with deoxyuracil, 5-methylcytosine, and cytosine. Introduction of deoxyuracil (Z19), which removes a methyl group and preserves base-pairing, resulted in an 11-fold drop in reaction rate, $\Delta\Delta G^{\circ\ddagger} = 1.6$ kcal/mol. Substitution with 5-methylcytosine (Z20), disrupting base-pairing but preserving the 5-methyl group in the major groove, resulted in an increase in $\Delta\Delta G^{\circ\ddagger}$ of 4 kcal/mol. Introduction of cytosine at this position, which both disrupts base-pairing and removes the 5-methyl group (Z21), resulted in a junction that is not cut to a detectable extent by CCE1 ($\Delta\Delta G^{\circ\ddagger} \geq 6$ kcal/mol).

Adenine at Position −1′. Substitution of $A^{-1'}$ results in much smaller changes in reaction rate than those seen for the other three bases of the core recognition site. A guanine at this position, resulting in a G-T mismatch (Z22), gave only a 5-fold drop in rate, while C-T (Z24) and T-T (Z25) mismatches had slightly larger effects. Clearly a mismatch at the −1′ position does not inevitably lead to a large drop in reaction rate, suggesting that the large changes seen in rate due to substitution of T^{-1} with 5-methylcytosine and cytosine are not predominantly due to disruption of base-pairing. Substitution with 7-deazaadenine (Z23) resulted in a 10-fold rate decrease ($\Delta\Delta G^{\circ\ddagger} = 1.5$ kcal/mol), in contrast with the much larger effect produced by substitution of $G^{-2'}$ with 7-deazaguanine (Z26) ($\Delta\Delta G^{\circ\ddagger} = 4.8$ kcal/mol).

Effect of Junction Sequence on CCE1 Binding Affinity. Gel electrophoretic retardation experiments have shown that CCE1 binds four-way junctions with dissociation constants of around 1 nM. A four-way junction that could not be cleaved by CCE1 was bound with a similar affinity to one that was cleaved readily, and the dissociation constant was not significantly affected by the presence or absence of magnesium (*15*). Electrophoretic retardation analysis was used to investigate the binding affinity of CCE1 to a cross-section of the junctions used in this study, including the fastest-cutting junction Z1, two junctions with altered sequences, two containing base analogues, and two with base analogues which result in a base mismatch. The results are summarized in Table 3. The dissociation constants range over less than 1 order of magnitude, from 0.6 to 5.6 nM, and show no correlation with the rate of cleavage, which ranges over 3 orders of magnitude for these junctions. Thus it appears that although CCE1 exhibits exquisite sequence specificity for junction cleavage, this specificity is not reflected in junction binding. CCE1 may rely on interactions with the phosphate backbone and carbohydrate moieties of DNA for binding, rather than making base-specific contacts.

Table 3: Binding Affinity of CCE1 for a Selection of Four-Way Junctions[a]

junction	nature of alteration	$\Delta\Delta G^{\circ\ddagger}$ (kcal/mol)	K_D (nM)	$\Delta\Delta G^{\circ}_{bind}$ (kcal/mol)
Z1	reference		5.6	
Z16	base analogue	0.81	1.9	−0.21
Z14	sequence	1.32	1.5	−0.16
Z15	sequence	1.48	2.4	−0.26
Z19	base analogue	1.55	0.6	−0.066
Z17	base analogue/mismatch	3.40	3.2	−0.35
Z20	base analogue/mismatch	4.04	3.2	−0.35

[a] K_D values were calculated as described in Experimental Procedures, and $\Delta\Delta G^{\circ}_{bind}$ calculated by comparison to the dissociation constant for the reference junction Z1. $\Delta\Delta G^{\circ\ddagger}$ values are taken from Tables 1 and 2.

DISCUSSION

Junction Binding Is Not Sequence-Dependent. We have demonstrated by gel retardation analysis that CCE1 binds four-way junctions containing sequence differences, mismatches, and base analogues with approximately equal affinity. Though a detailed understanding of the molecular basis for recognition necessitates a cocrystal structure, we can speculate that the protein interacts primarily with invariant parts of the DNA: the phosphate backbone and the deoxyribose moieties. The lack of sequence specificity suggests that CCE1 binds Holliday junctions, and then scans for favorably positioned cleavage sequences. It is unknown at present whether the enzyme works in concert with a protein or proteins that catalyze branch migration, as is the case with the eubacterial RuvABC system. It is conceivable that CCE1 promotes branch migration by binding and unfolding Holliday junctions, as formation of the folded, stacked X-structure in magnesium is a significant barrier to spontaneous branch migration (*16*). This possibility is currently being investigated.

Junction Cleavage by CCE1 is Highly Sequence-Specific. A single-turnover kinetic assay has been used to measure the rate of the chemical step of junction resolution as a function of local DNA sequence, giving a consensus cleavage sequence 5′-ACT↓A. By comparison, the consensus cleavage site 5′-($^{A}/_{T}$)TT↓($^{G}/_{C}$) was identified for RuvC using a different experimental approach which involved mapping cut sites in an α-structure containing a mobile Holliday junction (*17*). Sequence discrimination resides primarily at positions −2 and −1, where a CT dinucleotide sequence is highly favored over all other sequences. The two nucleotides flanking the CT sequence also influence the cleavage rate, but are much less important, and any influence by more distant nucleotides is likely to be minor. Although we have discussed changes in rate constants with respect to the nucleotide sequence of the h strand, all of the derivative junctions in Table 1 have changes in both partners of a base pair. Using this approach we could not therefore discriminate between the effect on cleavage rate of changes in the h strand from the effect of the simultaneous nucleotide change in the b or r strands. This limitation was partially addressed by the introduction of base analogues and mismatches (see below). Restriction enzymes show a similar level of sequence discrimination at the cleavage step. Single nucleotide changes in the canonical restriction site result in reduction in first-order cleavage rate constants from 70 to 1.7×10^6-fold for *Eco*RI (*18, 19*) and from 10^4 to 1.3×10^6-fold for *Eco*RV (*20*). However, the overall discrimination of canonical versus "star" sites by these enzymes, taking both binding and cleavage into

account, is considerably higher: from 7×10^4 to 3×10^9 fold (or 6.6 to 13 kcal/mol) in the case of *Eco*RI (*18*). By contrast, sequence discrimination by CCE1 appears to be exerted solely at the chemical step of catalysis. Incorporation of base analogues in the cleavage sequence has allowed the importance of specific functional groups to be investigated. The smallest changes introduced at the consensus 5′-CT sequence, addition of a 5-methyl group in C^{-2}, or deletion of a 5-methyl group at T^{-1}, result in modest changes in reaction rate ($\Delta\Delta G^{\circ\ddagger} = 0.8$ and 1.6 kcal/mol, respectively). These effects are consistent with small changes in nonpolar interactions in the transition state. $G^{-2'}$, which is paired with C^{-2} in the cleavage consensus sequence, also appears to be a major determinant of sequence specificity. Replacement with 7-deazaguanine results in near abolition of cleavage activity, suggesting either that N7 is a vital hydrogen bond acceptor in the transition state, or alternatively that the C-H group replacing it introduces unfavorable steric interactions. In contrast to $G^{-2'}$, $A^{-1'}$ plays a minor role in sequence discrimination, as base analogues and mismatches at this position are relatively well tolerated. Together, the data presented in Tables 1 and 2 emphasize the importance of the GC base pair at position −2, and the T nucleotide at position −1, for cleavage activity.

Conformational Flexibility in the CCE1:Junction Complex. By moving the consensus cleavage sequence with respect to the point of strand exchange, while preserving all base-pairing, we were able to investigate the influence of positional effects on enzyme activity. Surprisingly, though CCE1 displayed a clear preference for cleavage at the point of strand exchange, it was capable of catalyzing cleavage of the phosphodiester backbone at one nucleotide 3′ or 5′ of the junction center. These three cleavage sites span a distance of approximately 12 Å in a model of the extended, square form of the four-way junction which is thought to represent the bound junction structure. Either CCE1 is capable of remarkable flexibility in orienting its active site with respect to the junction center, or the junction is manipulated to a more significant extent on enzyme binding than thought previously. Prediction of a square, extended structure for the four-way junction bound by CCE1 is based on information from comparative gel electrophoresis experiments (*15*), which tell us nothing about more local changes such as unstacking of bases and disruption of base-pairing near the point of strand exchange. Such conformational changes should be amenable to study by chemical probing. Dramatic changes in DNA conformation are seen for many protein:DNA complexes, including *Eco*RI and *Eco*RV (*21*, *22*), and may serve to stabilize the transition state as an aid to catalysis.

The Significance of Sequence-Dependent Junction Cleavage. Sequence-dependent DNA cleavage is a general feature of the Holliday junction endonucleases, with the exception of bacteriophage T4 endonuclease VII and T7 endonuclease I. Branch migration of Holliday junctions ensures that favorable cleavage sequences can be positioned correctly with respect to the point of strand exchange. However, in addition to their role in recombination, the bacteriophage enzymes play a role in phage DNA packaging which requires the removal of several types of aberrant DNA structures, including DNA bulges and three-way junctions (reviewed in ref *23*). As these structures are incapable of branch migration, the bacteriophage enzymes cannot afford to discriminate at the level of sequence. It follows that imposition of a sequence requirement for endonuclease activity will largely prevent junction-resolving enzymes from cleaving other branched and distorted DNA species. This may be a significant advantage to the cell: it is noteworthy that both the T4 and T7 bacteriophage enzymes are highly toxic when overexpressed in *E. coli*, while the sequence-dependent enzymes CCE1, YDC2, RuvC, and RusA display little toxicity. Two other potential benefits of sequence-specific junction cleavage are apparent. First, sequence-dependent cleavage could allow Holliday junctions to branch migrate further before being resolved, facilitating recombination. Second, branch migration can be driven through regions of sequence heterology when catalyzed by the RuvAB complex (*24*, *25*). Junction resolution would be inhibited in these heterologous regions as a sequence-specific endonuclease would tend to introduce repairable nicks rather than paired cuts, thus allowing further branch migration to occur until homologous DNA was encountered.

ACKNOWLEDGMENT

We thank our colleague Richard Pöhler for helpful discussions.

REFERENCES

1. Holliday, R. (1964) *Genet. Res. 5*, 282−304.
2. Meselson, M. S., and Radding, C. M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 358−361.
3. Schwacha, A., and Kleckner, N. (1995) *Cell 83*, 783−791.
4. Giraud-Panis, M.-J. E., and Lilley, D. M. J. (1996) *J. Biol. Chem. 271*, 33148−33155.
5. White, M. F., and Lilley, D. M. J. (1996) *J. Mol. Biol. 257*, 330−341.
6. White, M. F., Giraud-Panis, M.-J. E., Pöhler, J. R. G., and Lilley, D. M. J. (1997) *J. Mol. Biol. 269*, 647−664.
7. Kleff, S., Kemper, B., and Sternglanz, R. (1992) *EMBO J. 11*, 699−704.
8. Ezekiel, U. R., and Zassenhaus, H. P. (1993) *Mol. Gen. Genet. 240*, 414−418.
9. Whitby, M. C., and Dixon, J. (1997) *J. Mol. Biol. 272*, 509−522.
10. White, M. F., and Lilley, D. M. J. (1997) *Mol. Cell. Biol. 17*, 6465−6471.
11. Lockshon, D., Zweifel, S. G., Freeman-Cook, L. L., Lorimer, H. E., Brewer, B. J., and Fangman, W. L. (1995) *Cell 81*, 947−955.
12. Lilley, D. M. J., and Clegg, R. M. (1993) *Annu. Rev. Biophys. Biomol. Struct. 22*, 299−328.
13. Seeman, N. C., and Kallenbach, N. R. (1994) *Annu. Rev. Biophys. Biomol. Struct. 23*, 53−86.
14. Lilley, D. M. J. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9513−9515.
15. White, M. F., and Lilley, D. M. J. (1997) *J. Mol. Biol. 266*, 122−134.
16. Panyutin, I. G., and Hsieh, P. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2021−2025.
17. Shah, R., Bennett, R. J., and West, S. C. (1994) *Cell 79*, 853−864.
18. Lesser, D. R., Kurpiewski, M. R., and Jen-Jacobson, L. (1990) *Science 250*, 776−786.

19. Thielking, V., Alves, J., Fliess, A., Maass, G., and Pingoud, A. (1990) *Biochemistry 29*, 4682−4690.
20. Alves, J., Selent, U., and Worfes, H. (1995) *Biochemistry 34*, 11191−11197.
21. Winkler, F. K., et al. (1993) *EMBO J. 12*, 1781−1795.
22. Kim, Y., Choi, J., Grable, J. C., Greene, P., Hager, P., and Rosenberg, J. M. (1994) in *Structural Biology: The State of the Art* (Sarma, R. H., and Sarma, M. H., Eds.), Adenine Press, New York.
23. Duckett, D. R., Murchie, A. I. H., Bhattacharyya, A., Clegg, R. M., Diekmann, S., von Kitzing, E., and Lilley, D. M. J. (1992) *Eur. J. Biochem. 207*, 285−295.
24. Iype, L. E., Wood, E. A., Inman, R. B., and Cox, M. M. (1994) *J. Biol. Chem. 269*, 24967−24978.
25. Parsons, C. A., Stasiak, A., and West, S. C. (1995) *EMBO J. 14*, 5736−5744.

BI980399S