Nucleic Acids Research, 1999, Vol. 27, No. 4 **1039–1046**

DNA cleavage and degradation by the SbcCD protein complex from *Escherichia coli*

John C. Connelly, Erica S. de Leau and David R. F. Leach*

Institute of Cell and Molecular Biology, University of Edinburgh, Kings Buildings, Edinburgh EH9 3JR, UK

Received October 12, 1998; Revised and Accepted December 14, 1998

ABSTRACT

The SbcCD protein is a member of a group of nucleases found in bacteriophage T4 and T5, eubacteria, archaebacteria, yeast, *Drosophila*, mouse and man. Evidence from electron microscopy has revealed a distinctive structure consisting of two globular domains linked by a long region of coiled coil, similar to that predicted for the members of the SMC family. That a nuclease should have such an unusual structure suggests that its mode of action may be complex. Here we show that the protein degrades duplex DNA in a 3′→5′ direction. This degradation releases products half the length of the original duplex suggesting simultaneous degradation from two duplex ends. This may provide a link to the unusual structure of the protein since our data are consistent with recognition and cleavage of DNA ends followed by 3′→5′ nicking by two nucleolytic centres within a single nuclease molecule that releases a half length limit product. We also show that cleavage is not simply at the point of a single-strand/double-stand transition and that despite the dominant 3′→5′ polarity of degradation, a 5′ single-strand can be cleaved when attached to duplex DNA. The implications of this mechanism for the processing of hairpins formed during DNA replication are discussed.

INTRODUCTION

DNA sequences capable of forming hairpin secondary structures (e.g. long palindromes) are not propagated in wild-type strains of *Escherichia coli* (1). However, this problem is overcome in *sbcC* or *sbcD* mutants (2,3). This observation has led to the proposal that the primary function of the SbcC and SbcD polypeptides *in vivo* is to recognize and cleave hairpin structures that arise during DNA replication (1,4–6). Problems in the replication of long DNA inverted repeats have been observed in other bacteria (7–9), yeast (10–14), humans (15) and transgenic mice (16,17) suggesting that the need to deal with abnormally folded DNA structures is universal.

Hairpin DNA is also believed to arise as an intermediate in a variety of specialized recombination reactions. For example, excision of the *Tam3* and *Ac-Ds* transposon in plants (18,19), the *Ascot* transposons in fungi (20) and most notably mammalian V(D)J recombination which generates the variability found in immunoglobulin and T cell receptor genes (21,22). It has been suggested that the ability to cleave a hairpin substrate is a general feature of DNA metabolism as murine lymphoid and non-lymphoid cells both contain activities capable of nicking hairpin DNA (17,23).

SbcC and SbcD remain associated throughout purification and interact to form a 1.2 MDa protein complex. SbcC has 'Walker A' and 'Walker B' nucleotide binding motifs, separated by an α-helical region (several hundred amino acids long) predicted to form two coiled-coil domains (6,24,25). When visualized by electron microscopy SbcCD is seen to have the structure predicted for an SMC protein (26). SbcC belongs to a subgroup within the SMC family of proteins that show extensive sequence similarity in and around their ATP binding motifs. Members of this subgroup include the human and *Saccharomyces cerevisiae* Rad50 polypeptides (24), both of which are involved in pathways of double-strand break repair *in vivo* (27).

Human Rad50 (hRad50) and yeast Rad50 (scRad50) both interact with the (h and sc) Mre11 polypeptide (28,29). In mammalian cells Rad50 and Mre11 interact to form a large multiprotein complex (29). Mre11 shares sequence similarity with SbcD and both belong to a family of phosphoesterases that contain the conserved sequence $DXH(X)_{25}GDXXD(X)_{25}GNHD/E$ (24). Murine MRE11 is essential for normal cell proliferation in an embryonic cell line (30). hMre11 is a manganese-requiring double-strand exonuclease that can cleave hairpin DNA and facilitate the repair of DNA double-strand breaks *in vitro* (31,32).

In previous work we have demonstrated that the SbcCD protein complex possesses ATP-dependent double-strand DNA exonuclease and ATP-independent single-strand endonuclease activities (6,25). The protein can also recognize and cleave a DNA hairpin substrate and we have shown that hairpin loop cleavage can be dissociated from subsequent degradation by the use of suboptimal conditions (i.e. cleavage in the presence of the poorly hydrolysable analogue of ATP, ATPγS, at 16°C) (26). All of these activities have an essential requirement for manganese ions.

In this work we characterize in more detail the substrate specificity of the protein. We show that SbcCD is predominantly a 3′→5′ exonuclease and present evidence that degradation progresses to the point where half of the double-stranded DNA remains. This suggests that digestion occurs simultaneously from the two ends. SbcCD is also a sequence-independent hairpin endonuclease that cleaves a variety of hairpin substrates 5′ to the loop. In competition assays linear duplex DNA competes for hairpin endonuclease activity just as efficiently as hairpin DNA, whilst circular pUC19 and single-strand DNA do not compete. We also demonstrate that SbcCD possesses a single-strand endonuclease activity on a 5′ overhang of 10 nt. Therefore, SbcCD is a protein that can act at a variety of duplex ends, most notably to deprotect a DNA molecule sealed by a covalent linkage or to remove a 5′ overhang.

*To whom correspondence should be addressed. Tel: +44 131 650 5373; Fax: +44 131 650 8650; Email: d.leach@ed.ac.uk

MATERIALS AND METHODS

DNA substrates

The sequences of the synthetic oligonucleotides used in this study are shown in Table 1. To remove short chain termination products all oligonucleotides (synthesized by Oswel DNA Service, University of Southampton) were run on 10% denaturing polyacrylamide gels and visualized by UV shadowing. DNA was recovered from excised bands by eluting at 60°C for 16 h in TE (10 mM Tris–HCl, pH 7.5, 1 mM EDTA, pH 8.0) buffer, followed by ethanol precipitation. After 5′- or 3′-^{32}P-end-labelling oligonucleotides were once again gel purified and recovered as described above. All methods were according to standard protocols (33).

Oligonucleotides were labelled at the 5′-end with [γ-^{32}P]ATP (3000 Ci/mmol; Amersham) using T4 polynucleotide kinase (Boehringer Mannheim) to yield a substrate with a specific activity of ~6.7 × 10^6 c.p.m./μg total nucleic acid. 3′-End-labelled substrate was generated by filling in the 1 nt gap of 3′HP56-CGAG with [α-^{32}P]dCTP (3000 Ci/mmol; Amersham) using the Klenow fragment of DNA polymerase I (New England Biolabs) to yield a substrate with a specific activity of ~7.4 × 10^6 c.p.m./μg total nucleic acid.

Synthetic double-strand substrates were prepared in a PCR machine by adding a 3-fold excess of cold oligonucleotide to 5′-^{32}P-labelled DNA in a reaction containing 150 mM NaCl, 10 mM Tris–HCl, pH 7.5, 1 mM EDTA, pH 8.0. This mix was incubated at 98°C for 3 min, 65°C for 10 min and finally 37°C for 10 min. DNA was then run on a 10% native polyacrylamide gel and substrates were excised, eluted overnight at 37°C and resuspended in TE after ethanol precipitation.

Native polyacrylamide gel electrophoresis was used to confirm that double-strand and hairpin substrates exist as a single species (data not shown).

Circular pUC19 was purchased from New England Biolabs. Linear pUC19 was generated after treatment with *Eco*RI (Boehringer Mannheim).

SbcCD protein and nuclease assays

SbcCD protein was prepared as described (25). Protein concentrations were estimated using a protein assay kit (Bio-Rad) with BSA as a standard. The molar concentration of SbcCD was calculated assuming a stoichiometry of $SbcC_6$:$SbcD_{12}$. In SbcCD nuclease assays 2.5 nM protein and DNA (~1.5 nM) were incubated at the temperature and for the time indicated in a 20 μl reaction volume. The reactions were carried out in the presence of 5 mM Mn^{2+}, 1.25 mM DTT, 2% glycerol, 100 μg/ml BSA, 25 mM Tris, pH 7.5, and 1 mM ATPγS (or 1 mM ATP). Reactions were stopped and analysed by denaturing or native polyacrylamide gel electrophoresis prior to drying and autoradiography. After exposure to a PhosphorImager screen, ImageQuant software (Molecular Dynamics) was used to quantify the amount of product generated (data not shown).

Hairpin endonuclease competition assays

An aliquot of 1.5 nM hairpin (5′-HP56-CGAG) was incubated with 2.5 nM SbcCD at 16°C for 20 min in the presence of 1 mM ATPγS and SbcCD reaction buffer. The indicated amount of cold competitor DNA was added prior to the addition of protein. Reaction products were analysed by denaturing gel electrophoresis.

Denaturing gel electrophoresis

Reactions were stopped by adding an equal volume of DNA loading buffer (50 mM EDTA, pH 8.0, 95% formamide, 0.01% bromophenol blue). Samples were heated to 100°C for 4 min prior to electrophoresis and loaded as quickly as possible thereafter. To ensure conditions were as denaturing as possible polyacrylamide gels (0.4 mm thick) containing 7 M urea and 10% formamide were run on a Sequi-Gen® Nucleic Acid Sequencing Cell (Bio-Rad) at ~60°C for 1.5 h at 60 W, using a TBE (89 mM Tris–borate, pH 8.3, 2 mM EDTA, pH 8.0) buffer system. DNA was visualized by autoradiography.

Native gel electrophoresis

Reactions were stopped by adding an equal volume of DNA loading buffer (20 mM EDTA, pH 8.0, 30% w/v glycerol, 0.01% bromophenol blue). Polyacrylamide gels (0.4 mm thick) were run on a Sequi-Gen® Nucleic Acid Sequencing Cell (Bio-Rad) at ~25°C for 3 h at 20 W, using a TBE buffer system.

Thin layer chromatography

DNA was incubated with SbcCD for the times indicated. Reactions were treated as described for denaturing gel electrophoresis, except 1 μl samples were analysed on polyethyleneimine cellulose thin layer chromatography plates (Sigma) containing a 254 nm fluorescent indicator. The plates were developed in 1 M formic acid and 0.5 M LiCl. Products were visualized by autoradiography. In this system the R_f values for dATP and dCMP were 0.15 and 0.75, respectively.

RESULTS

Duplex stem length required for hairpin endonuclease activity

When SbcCD protein was incubated with HP78 (a 5′-^{32}P-labelled oligonucleotide capable of forming a hairpin structure with a 37 bp stem and 4 nt loop; 26; Table 1) at 16°C in the presence of ATPγS a major product of 37 nt and two minor products of 36 and 38 nt were obtained (Fig. 1, lanes 7 and 8). To test the minimum length of stem required for hairpin endonuclease activity, protein was incubated with HP78 and four other hairpin substrates; HP20, HP36, HP56 and HP94, that had a constant loop sequence but variable stem length (Table 1). Hairpin oligonucleotides with stem lengths of 16 (HP36), 26 (HP56), 37 (HP78) and 45 bp (HP94) were substrates for SbcCD (Fig. 1, lanes 3–10), whilst a substrate with a stem length of 8 bp was not (lanes 1 and 2). With substrates HP36 and HP56 products were seen that arise from cleavage on the 3′-strand of the blunt terminus (lanes 3–6). Such products were not seen with HP78 and HP94 (lanes 7–10), probably because they were difficult to detect on a 10% polyacrylamide gel. These data indicate that SbcCD can act at both the loop and blunt end of a hairpin substrate.

Direction of digestion

To determine in which direction SbcCD digests DNA, HP56-CGAG (Table 1) labelled at either the 3′- or 5′-terminus was treated with protein at 37°C in the presence of ATP and the reaction products analysed at various time intervals by denaturing polyacrylamide gel electrophoresis. Analysis of the products obtained when 5′-^{32}P-labelled HP56-CGAG was incubated with

Table 1. Oligonucleotide sequences

Name	Sequence	Length (bases)
	HAIRPIN FORMING OLIGONUCLEOTIDES	
	variable stem length, common 4T centre	
HP20	*5′-GCCCCGGG**TTTT**CCCGGGGC-3′	20
HP36	*5′-GTAATCCAGCCCCGGG**TTTT**CCCGGGGCTGGATTAC-3′	36
HP56	*5′-GCCCTTTGACGTAATCCAGCCCCGGG**TTTT**CCCGGGGCTGGATTACGTC-AAAGGGC-3′	56
HP78	*5′-GTTTCTATTCAGCCCTTTGACGTAATCCAGCCCCGGG**TTTT**CCCG-GGGCTGGATTACGTCAAAGGGCTGAATAGAAAC-3′	78
HP94	*5′-CCCGGGCACACAGATAGGTTTCTATTCAGCCCTTCCAGCCCCGGG**TTTT**C-CCGGGGCTGGAAGGGCTGAATAGAAACCTATCTGTGCCCGGG-3′	94
	common stem length, different centres	
HP56-TTTT	*5′-GCCCTTTGACGTAATCCAGCTTGGAA**TTTT**TTCCAAGCTGGATTACGTCA-AAGGGC-3′	56
HP56-CGCG	*5′-GCCCTTTGACGTAATCCAGCTTGGAA**CGCG**TTCCAAGCTGGATTACGTCA-AAGGGC-3′	56
HP56-AGCG	*5′-GCCCTTTGACGTAATCCAGCTTGGAA**AGCG**TTCCAAGCTGGATTACGTCA-AAGGGC-3′	56
5′-HP56-CGAG	*5′-GCCCTTTGACGTAATCCAGCTTGGAA**CGAG**TTCCAAGCTGGATTACGTCA-AAGGGC-3′	56
3′-HP56-CGAG	5′-GCCCTTTGACGTAATCCAGCTTGGAA**CGAG**TTCCAAGCTGGATTACGTCA-AAGGG-3′*	56
HP56-CTTG	*5′-GCCCTTTGACGTAATCCAGCTTGGAA**CTTG**TTCCAAGCTGGATTACGTCA-AAGGGC-3′	56
HP56-GTTC	*5′-GCCCTTTGACGTAATCCAGCTTGGAA**GTTC**TTCCAAGCTGGATTACGTCA-AAGGGC-3′	56
HP56-TTTA	*5′-GCCCTTTGACGTAATCCAGCTTGGAA**TTTA**TTCCAAGCTGGATTACGTCA-AAGGGC-3′	56
HP56-ATTT	*5′-GCCCTTTGACGTAATCCAGCTTGGAA**ATTT**TTCCAAGCTGGATTACGTCA-AAGGGC-3′	56
	10T centre, variable stem length	
HP62-T10	*5′-GCCCTTTGACGTAATCCAGCCCCGGG**TTTTTTTTTT**CCCGGGGCTGGATT-ACGTCAAAGGGC-3′	62
HP84-T10	*5′-GTTTCTATTCAGCCCTTTGACGTAATCCAGCCCCGGG**TTTTTTTTTT**CCC-GGGGCTGGATTACGTCAAAGGGCTGAATAGAAAC-3′	84
	Oligonucleotide containing trinucleotide repeats	
$(CTG)_{16}$GC clamp	*5′-GCCCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG-CTGCTGCCCGGGGC-3′	64
	OLIGONUCLEOTIDES USED TO FORM DUPLEX SUBSTRATES	
dup27(top)	5′-GTTCCAAGCTGGATTACGTCAAAGGGC-3′	27
dup27(bottom)	3′-CAAGGTTCGACCTAATGCAGTTTCCCG-5′	27
dup26(top)	5′-CCCGGGGCTGGATTACGTCAAAGGGC-3′	26
dup26(bottom)	3′-GGGCCCCGACCTAATGCAGTTTCCCG-5′*	26
3′-4T(top)	5′-CCCGGGGCTGGATTACGTCAAAGGGC-3′	26
3′-4T(bottom)	3′-**TTTT**GGGCCCCGACCTAATGCAGTTTCCCG-5′*	30
3′-10T(top)	5′-CCCGGGGCTGGATTACGTCAAAGGGC-3′	26
3′-10T(bottom)	3′-**TTTTTTTTTT**GGGCCCCGACCTAATGCAGTTTCCCG-5′*	36
5′-4T(top)	*5′-**TTTT**GCCCTTTGACGTAATCCAGCCCCGGG-3′	30
5′-4T(bottom)	3′-CGGGAAACTGCATTAGGTCGGGGCCC-5′	26
5′-10T(top)	*5′-**TTTTTTTTTT**GCCCTTTGACGTAATCCAGCCCCGGG-3′	36
5′-10T(bottom)	3′-CGGGAAACTGCATTAGGTCGGGGCCC-5′	26
	OLIGONUCLEOTIDE CAPABLE OF FORMING A BRANCHED SUBSTRATE	
56-br	*5′-**GCCCTTTGACGTAA**TCCAGCTTGGAA**CGAG**TTCCAAGCTGGACG**GTTGAC-CCTTCG**-3′	56

Single-strand extensions and hairpin loop sequences are shown in bold. *, position of ^{32}P label.

SbcCD suggested that DNA was digested mainly in a 3′→5′ direction from both the blunt and hairpin termini (Fig. 2A, lanes 1–7). However, small ^{32}P-labelled fragments were not detected by PAGE when 3′-labelled HP56-CGAG was used as a substrate, as might have been expected from a 3′→5′ exonuclease activity (Fig. 2A, lanes 8–14). This suggested that if these products existed, they were very small and were lost during processing of the polyacrylamide gel. To assay for the release of mononucleotides the same products were examined by thin layer chromatography. Single nucleotides were only visualized when 3′-labelled HP56-CGAG was used as a substrate, confirming that SbcCD digests DNA mainly in a 3′→5′ direction (Fig. 2B,

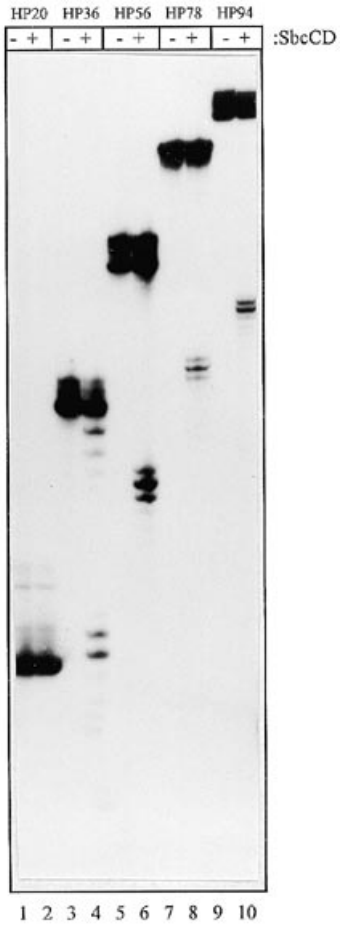

Figure 1. Duplex stem length required for hairpin endonuclease activity. The 5′-^{32}P-labelled substrate indicated was incubated with (lanes 2, 4, 6, 8 and 10) or without (lanes 1, 3, 5, 7 and 9) SbcCD in the presence of ATPγS at 16°C for 20 min. Reaction products were then resolved on a 10% denaturing polyacrylamide gel.

lane 8–14). Some prominent bands, at ~17 and 22 nt, are visible after gel electrophoresis of the 3′-labelled substrate (Fig. 2A, lanes 12–14). It is possible that these fragments could have been generated by a 5′→3′ nuclease activity of SbcCD. However, they could equally well have been produced by cleavage of a 5′ overhang generated after 3′→5′ digestion (see below for evidence of 5′ overhang cleavage).

When HP56-CGAG digestion was analysed by native gel electrophoresis the majority of 3′- and 5′-labelled products were not seen until the sample was boiled, indicating that SbcCD nicks but does not dissociate double-strand DNA (Fig. 2C, lanes 7 and 8, and lanes 15 and 16). However, prior to boiling substantial amounts of the 5′-labelled substrate are visible as a smear and limit digestion fragments are released. No smearing is visible with the 3′-labelled substrate and only a small amount of product is released. These observations suggest that a substantial amount of the 3′→5′ digestion from the hairpin end is coupled to 3′→5′ digestion from the blunt end. This simultaneous processing would lead to the release of the 5′-labelled limit product that would be held to its complement by a few hydrogen bonds. Again these data are consistent with a primary bias towards digestion in a 3′→5′ direction.

Effect of loop sequence on hairpin endonuclease activity

To test the effect of loop sequence on SbcCD action, hairpin endonuclease activity was assayed at 16°C in the presence of ATPγS using a variety of 56 base hairpin-forming oligonucleotides (Table 1). The nucleotide sequences at the centre of these hairpins were chosen as they have been investigated by NMR, molecular modelling and thermal transition studies and are believed to form a variety of two and four base loops (34 and references within). All of these hairpins were cleaved by SbcCD, preferentially at positions 5′ to the loop region (Fig. 3A, lanes 1–4, and B, lanes 1–14). Therefore, SbcCD acts as a sequence-independent hairpin endonuclease (i.e. it cleaves 5′ to the loop regardless of the nucleotide sequence). SbcCD also cleaves these substrates several nucleotides in from the 3′-end of the blunt terminus. HP56-CGCG, HP56-CGAG, HP56-CTTG and HP56-TTTA are predicted to form hairpins with 2 nt loops whilst HP56, HP56-TTTT HP56-AGCG, HP56-GTTC and HP56-ATTT are predicted to form 4 nt loops. It was not apparent from Figure 3A or B that there was any correlation between the site of SbcCD cleavage and the position of a stem/loop (or single-strand/double-strand) boundary.

In order to test whether a pseudo-hairpin containing T·T base pairs was cleaved in the region 5′ of the loop or at positions relating to the T·T pairs in the stem a substrate containing 16 CTG triplet repeats [$(CTG)_{16}$GC-clamp] was digested with SbcCD at 16°C in the presence of ATPγS. The preferred cleavage site for this substrate was 5′ of the loop. Very little cleavage was detected in the pseudo-hairpin stem (Fig. 3C).

To determine what effect loop size had on cleavage, HP62-T10 (a hairpin with a stem length of 26 bp and a predicted loop length of 10 nt; Table 1) was incubated with SbcCD, at 16°C in the presence of ATPγS. This substrate is identical to HP56 except it contains six extra thymidines in the loop region. To our surprise HP62-T10 yielded a variety of products ranging in size from 12 to 31 nt (Fig. 3D, lanes 1 and 2). This is in contrast to HP56 (and all the other hairpins tested) that yielded several discrete products (Fig. 3A–C). Three classes of product were seen with HP62-T10, those that result from cleavage at the predicted stem/loop boundary, those that arise from cleavage close to the tip of the hairpin and those entering the stem of the hairpin. The existence of the second class suggests that the thymidine residues in the predicted loop can form a quasi-duplex structure, perhaps being forced into this conformation by the hairpin arms whose tip is cleaved by SbcCD. These data indicate that a hairpin with a 10 nt loop is processed differently from one with a 2 or 4 nt loop. Particularly noticeable is the degradation of the stem that is not observed for the hairpins with tight loops under these conditions.

Degradation to half duplex stem length

We noted that the limit product generated using HP62-T10 was 12 nt, almost half the length of the double-strand stem (13 bp). To test if SbcCD could measure the length of half a duplex stem, it was incubated at 16°C in the presence of ATPγS with an oligonucleotide (HP84-T10; Table 1) capable of forming a hairpin with a longer (37 bp) stem and a 10 nt loop. Larger products were seen in size clusters around 37 and 42 nt. The limit product obtained with this substrate was 17 nt (Fig. 3E, lanes 1 and 2), again almost half the length of the stem (18.5 bp). Even after longer exposures no products smaller than half a stem length were seen. This agrees with the size of limit product seen when HP56-CGAG is digested at 37°C in the presence of ATP (Fig. 2A, lanes 1–7).

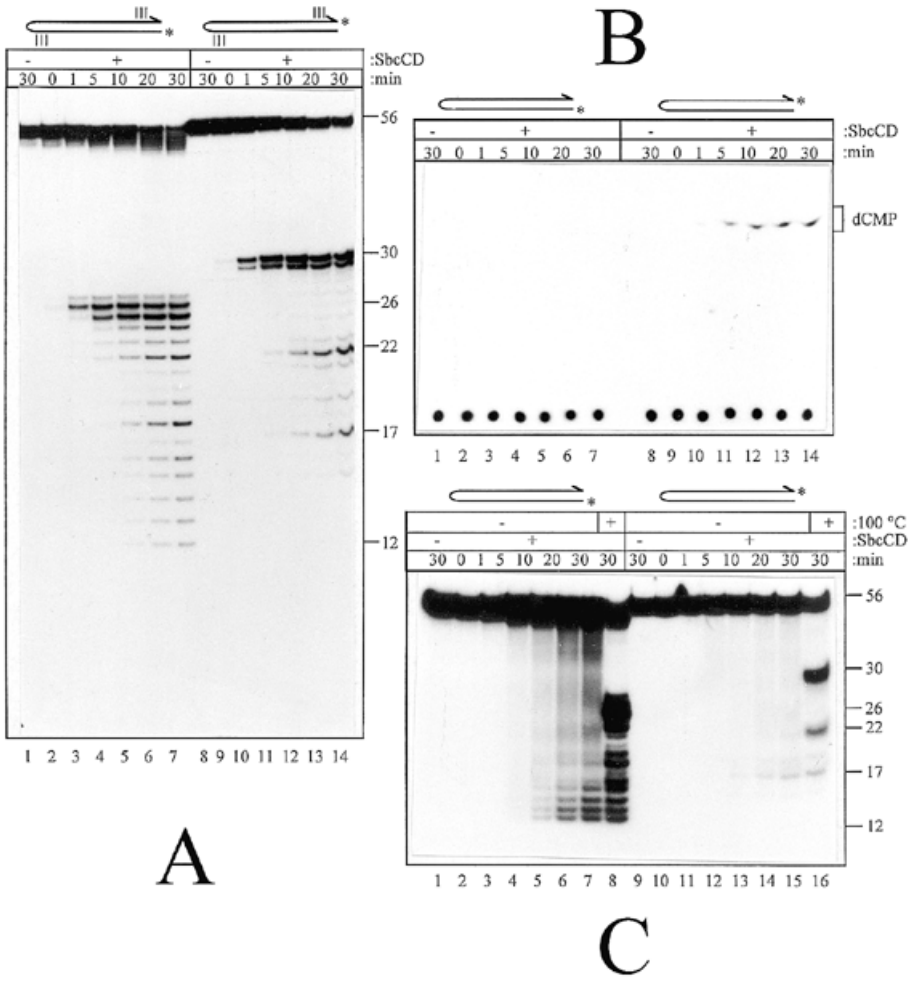

Figure 2. Direction of digestion. (**A**) Gel analysis of the fragments produced when SbcCD was incubated with 5′-^{32}P-labelled HP56-CGAG (lanes 1–7) or 3′-^{32}P-labelled HP56-CGAG (lanes 8–14) in the presence of ATP at 37°C for the time indicated. The products were resolved on a 10% denaturing polyacrylamide gel. Vertical lines indicate the sites of early cleavage by SbcCD. (**B**) A chromatogram showing the mononucleotide products released when the reaction mixtures obtained in (A) were analysed by thin layer chromatography. (**C**) Native gel analysis of the fragments produced when 5′-^{32}P-labelled HP56-CGAG (lanes 1–8) or 3′-^{32}P-labelled HP56-CGAG (lanes 9–16) was incubated with SbcCD, in the presence of ATP at 37°C for the times indicated. The sizes in nucleotides of various products are indicated. *, position of ^{32}P label.

Double-strand end preference of SbcCD

SbcCD appears to digest a hairpin molecule at its blunt or hairpin end equally well (Fig. 3A and B). To determine if SbcCD has a preference for binding either a blunt or hairpin DNA end, competition assays were performed in which hairpin endonuclease activity was assayed (at 16°C with ATPγS) in the presence of increasing amounts of cold competitor DNA (Fig. 4). A blunt-ended 27 bp duplex oligonucleotide (dup27) competes as efficiently for hairpin endonuclease activity as cold hairpin DNA (HP56-CGAG, compare lanes 3–6 with lanes 7–10). On the other hand, a 27 base single-strand DNA oligonucleotide (dup27-bottom) is a very poor competitor (lanes 11–14). Circular double-strand pUC19 DNA (lanes 19–22) is also a very poor competitor, in contrast to linear double-strand pUC19, which competes for the activity more efficiently than hairpin DNA or a small duplex (lanes 15–18). These observations imply that SbcCD prefers double-strand DNA that contains an end, but has no overall preference for a hairpin end over a blunt end.

SbcCD does not cleave at a single-strand/double-strand transition but does cleave within a 10 nt 5′ overhang and degrades to half of the duplex stem length

SbcCD can cleave a hairpin substrate in the vicinity of the stem/loop boundary. In an effort to elucidate what feature of a stem/loop SbcCD recognizes, cleavage activity was assayed at 16°C in the presence of ATPγS using a variety of DNA substrates. 5′-^{32}P-labelled duplex (dup26; Table 1), blunt at both ends, was digested in a 3′→5′ direction to form a limit product of ~12 nt (Fig. 5A, lanes 1 and 2). That this extended degradation occurred in ATPγS at 16°C signalled that duplex DNA behaves similarly to hairpins with 10 T loops and not tight looped hairpins. The same duplex substrate containing a 4 or 10 nt 3′ overhang (3′-4T and 3′-10T, respectively) also yielded a 12 nt limit product (Fig. 5A, lanes 3–6). Duplex DNA containing a 4 nt 5′ overhang (5′-4T) was digested in a similar manner to give a limit product of ~16 nt (Fig. 5B, lanes 4 and 5). When the 5′ extension was increased to 10 nt (5′-10T), the limit product arising from the

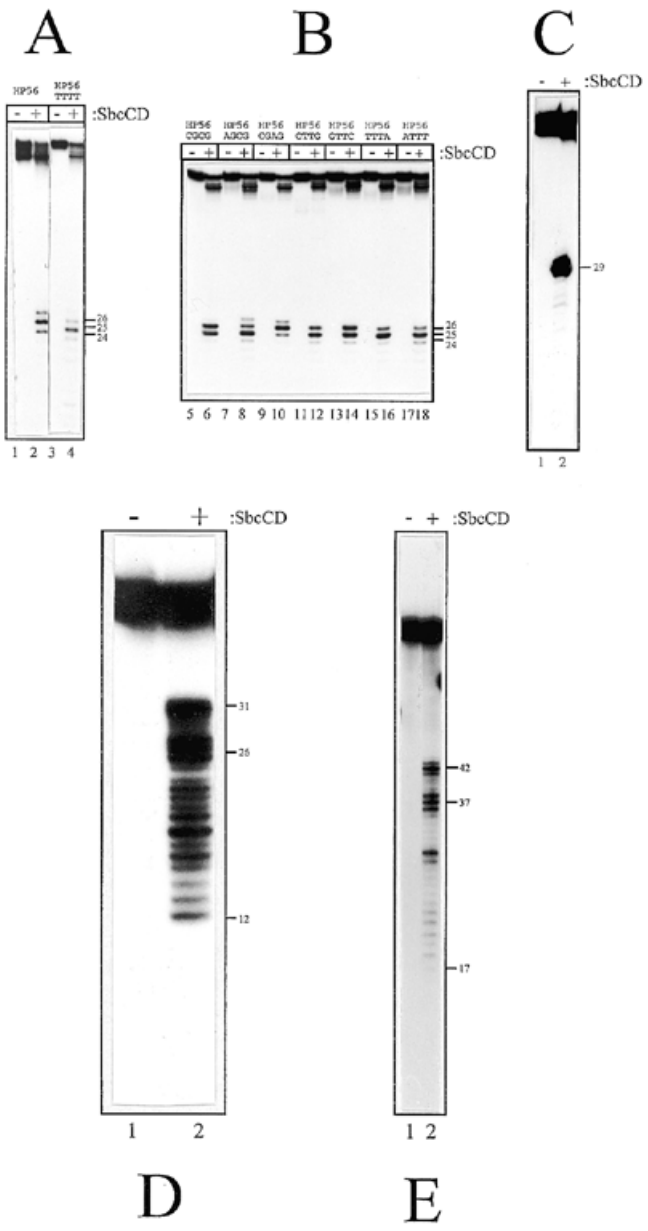

Figure 3. Effect of sequence on hairpin endonuclease activity. (**A** and **B**) Action of SbcCD on a variety of 56 base hairpin-forming oligonucleotides. The substrates indicated were incubated with (lanes 2, 4, 6, 8, 10, 12, 14, 16 and 18) or without (lanes 1, 3, 5, 7, 9, 11, 13, 15 and 17) SbcCD in the presence of ATPγS at 16°C for 20 min. Products were examined on a 10% denaturing polyacrylamide gel. (**C**) $(CTG)_{16}$ GC-clamp substrate incubated with (lane 2) or without (lane 1) SbcCD. The products were resolved on a 10% denaturing gel. (**D**) HP62-T10 was incubated with (lane 2) or without (lane 1) SbcCD and products were resolved on a 20% denaturing gel. (**E**) HP84-T10 treated with (lane 2) or without (lane 1) SbcCD and products analysed on a 10% denaturing gel. The size in nucleotides of various products are indicated. *, position of ^{32}P label.

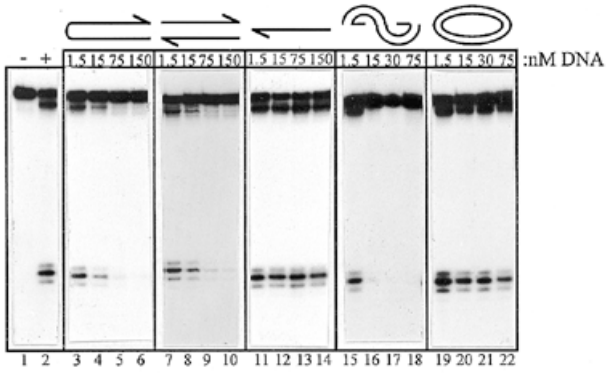

Figure 4. Competition assays. SbcCD was incubated with 1.5 nM 5′-^{32}P-labelled HP56-CGAG and challenged with the indicated concentration of cold competitor DNA. HP56-CGAG (lanes 3–6), 27 base duplex (dup27, lanes 7–10), 27 base single-strand oligonucleotide (dup27-bottom, lanes 11–14), linear double-strand pUC19 (lanes 15–18) and circular double-strand pUC19 (lanes 19–22). Lane 1 contains no protein or competitor DNA and lane 2 protein but no competitor. Samples were analysed by 10% denaturing polyacrylamide gel electrophoresis. *, position of ^{32}P label.

3′→5′ exonuclease activity of SbcCD was ~22 bases. In addition, a cluster of products migrating at ~3–8 nt was noted for the substrate with the 5′-10T overhang (Fig. 5B, lanes 6 and 7). This indicates that SbcCD possesses a single-strand nuclease activity on a 5′ overhang of sufficient length. When the 5′ overhang formed part of a branched structure (56-br) it was still acted upon by SbcCD (Fig. 5C, lanes 1 and 2). The activity seen with the 10 nt 5′ overhang is significantly greater than any activity on single-strand oligonucleotide substrates (5′-4T-top and 5′-10T-top; Table 1) (Fig. 5B, lanes 8–11). This argues that cleavage in the 5′ overhang is stimulated by the presence of adjacent duplex.

These data provide additional support for the hypothesis that SbcCD can degrade to half the length of a double-strand region of DNA. The 26 bp duplex (dup26) forms a limit product of ~12 nt, nearly half the stem length (Fig. 5A, lanes 1 and 2). When a 4 or 10 nt 3′ overhang (3′-4T and 3′-10T, respectively) was added the length of the limit product remained constant (Fig. 5A, lanes 3–6). This suggests that the protein senses the length of duplex and not the strand with the 3′ overhang. When a 10 nt 5′ overhang was added to the 26 bp duplex (5′-10T), the limit product obtained was ~22 nt (Fig. 5B, lanes 6 and 7). Again this suggests that SbcCD recognizes the size of the duplex and not the extended strand, in this case a 5′ overhang. In summary, these data show no evidence for cleavage at a single-strand/double-strand transition. However, they do argue that SbcCD recognizes the transition as if it were a double-strand end and degrades to half the length of the duplex region.

DISCUSSION

In this work we have investigated the activity of the SbcCD complex from *E.coli* on oligonucleotide substrates in order to reveal aspects of its mechanism of action. Initially we determined that the minimum duplex stem length necessary for cleavage of a hairpin substrate lies between 8 and 16 bp. This allowed us to concentrate the majority of our assays on substrates with 26 bp stems in the knowledge that this was well within the range accessible to the protein. The observation that SbcCD cleaves hairpins with stems as short as 16 bp suggests that the problems encountered *in vivo* in the replication of long palindromes (1) in *sbcCD*$^{+}$ cells (that are only noticeable above a length of 100–150 bp) relate primarily to the persistence of hairpin structure rather than a lower length limit for the activity of the protein.

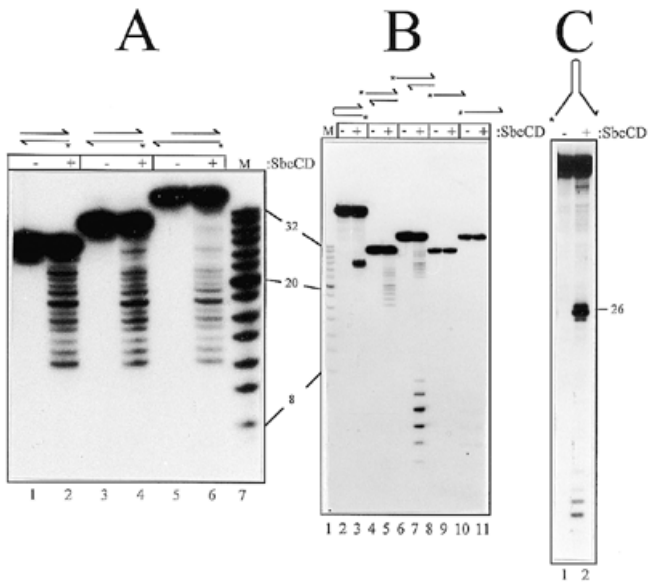

Figure 5. Action of SbcCD on a variety of synthetic substrates. (**A**) 20% denaturing gel showing the products obtained when SbcCD was incubated with various 5′-labelled oligonucleotides. Blunt-end 26 bp duplex (dup26, lanes 1 and 2), 4 nt 3′ overhang (3′-4T, lanes 3 and 4), 10 nt 3′ overhang (3′-10T, lanes 5 and 6). Lane 7, 8–32 nt marker. (**B**) 10% denaturing gel showing the products of SbcCD action on a 4 nt 5′ overhang (5′-4T, lanes 4 and 5), a 10 nt 5′ overhang (5′-10T, lanes 6 and 7) and two single-strand substrates of 30 (5′-4T-top, lanes 8 and 9) and 36 (5′-10T-top, lanes 10 and 11) nt. Lane 1, 8–32 nt marker. Lanes 2 and 3, HP56-CGAG. (**C**) 10% denaturing gel showing the action of SbcCD on a branched DNA substrate (56-br). Lane 1, no SbcCD; lane 2, with SbcCD. *, position of ^{32}P label.

SbcCD is a sequence-independent hairpin endonuclease

SbcCD is a double-strand exonuclease that can open a number of 56 nt hairpins containing tight loops, by nicking in the double-strand region 5′ to the loop. With these substrates the major sites of cleavage at the loop end of the molecule are 25 and 26 nt from the 5′-terminus. Although SbcCD acts as a sequence-independent hairpin endonuclease, it is sensitive to the exact sequence in the vicinity of the stem/loop. One might have expected that the single-strand endonuclease activity of SbcCD could attack the loop end of a hairpin. However, it prefers to open the hairpin by attacking the double-strand region adjacent to the loop. We have shown that single-strand DNA does not compete well for hairpin endonuclease activity nor is it a good substrate for SbcCD, unless it forms part of a duplex with a 10 nt 5′ overhang (see below). We suggest therefore that SbcCD does not cleave a hairpin loop at a single-strand/double-strand boundary or any other structural distortion, but in the first cleavable region of duplex next to the single-strand loop. SbcCD cleaves the blunt end of the hairpin and blunt end duplexes, therefore, it does not need a single-strand/double-strand transition to be active.

We could not find any direct correlation between the predicted size of the tight loops tested here (of 2 and 4 nt) and the site of SbcCD cleavage. The potential for structural distortion at a hairpin terminus is great and the relationship between loop geometry and the nucleotide sequence at a hairpin terminus is not a simple one. Important factors include; the sequence of the hairpin terminus; the particular nucleotides that form the terminal base pair that close the loop; stacking interactions. In addition loops undergo motions which hinder the analysis of structural data (35 has a discussion of these points). Despite these difficulties, there was no suggestion of any correlation between the site of SbcCD cleavage and any particular structural feature of a hairpin loop.

In competition assays hairpin DNA did not compete for hairpin endonuclease activity any better than the other linear duplexes tested. However, single-strand oligonucleotides and closed circular plasmids were poor competitors. These experiments suggest that SbcCD recognizes a double-strand end whether or not it is in a hairpin configuration.

SbcCD cleaves a CTG triplet repeat pseudo-hairpin adjacent to the loop

$d(CAG)_n \cdot d(CTG)_n$ triplet repeats are unstable DNA sequences that have been implicated in several human inherited neurological disorders (36). The d(CTG) strand has been shown to form a stable pseudo-hairpin structure containing T·T mispairs (37,38). Saltatory amplifications of $d(CAG)_n \cdot d(CTG)_n$ repeats can be detected in the SURE strain of *E.coli* (which carries a mutation in the *sbcC* gene) but not in SURE cells carrying a plasmid with the *sbcC* gene (39). Here we show that SbcCD can cleave a $d(CTG)_n$ pseudo-hairpin adjacent to the loop. No significant cleavages in the stem were observed suggesting that SbcCD does not recognize T·T mispairs.

SbcCD has a single-strand endonuclease activity on a 5′ overhang of 10 nt

We have shown previously that SbcCD has a single-strand endonuclease activity (6) and, here, that it possesses a single-strand nuclease activity on 5′ overhangs of 10 nt. This activity is strongly stimulated by the presence of adjacent duplex DNA since no cleavage was observed in single-strand oligonucleotides unpaired to their complements. Removal of a 5′ overhang *in vivo* by SbcCD may generate a blunt-ended substrate suitable for the entry of RecBCD.

Although SbcCD cleaves a 5′ overhang of 10 nt it does not act on one of 4 nt. The head regions of SbcC(D) are ~3.4 nm in diameter, a size which corresponds to ~10 nt of DNA. These head regions might be involved in end recognition and DNA binding. In support of this hypothesis, Ahkmedov *et al.* (40) have demonstrated that the C-terminal domains of yeast SMC1 and SMC2 (each presumed to form a head structure) bind DNA in an ATP-independent manner.

SbcCD is a 3′→5′ double-strand exonuclease that degrades duplex DNA to half its original length

We have suggested previously that SbcCD cleaves and removes hairpin secondary structures that arise during DNA replication, resulting in double-strand breaks which are then repaired by homologous recombination (6,41). We show here that SbcCD digests DNA in a 3′→5′ direction but does not dissociate the duplex strands until it has reached half the length of the duplex DNA. This pattern of degradation is observed on hairpin DNA in the presence of ATP at 37°C but is also seen in the presence of ATPγS at 16°C for duplex DNA without hairpin ends and hairpins with 10 T loops. Further experiments are required to determine why substrates with two free ends or with large loops

show more extensive degradation under the conditions that give only loop and end cleavage for the tight hairpins.

The observation of half length degradation is particularly interesting since it leaves behind two single strands whose 3′-ends lie across from each other. In the case of a hairpin formed during DNA replication, it is likely that only one of these 3′-ends is used directly for recombination since the other is not attached to chromosomal DNA. In the initiation of meiotic recombination, a similar bidirectional 3′→5′ degradation activity by Rad50/MRE11 could generate the 3′ overhangs observed (42,43). This is a theoretical means of resolving the apparent paradox that a 3′→5′ exonuclease is required for the generation of 3′ overhangs.

SbcCD has an unusual shape for a nuclease. Using electron microscopy a component of the SbcCD complex is seen as two head domains, of ~3.4 nm, linked by a rod region of ~80 nm (26). This is the organization predicted for the SMC family of proteins and a similar structure been observed for the SMC1 protein from *Bacillus subtilis* and MukB protein from *E.coli* (44,45). It is possible that each of the head domains of SbcCD recognizes a DNA terminus and digests a single DNA molecule from each end inwards until the two head domains meet and release half length single-strand products with 3′-ends remaining at the end points of digestion. Our experiments do not exclude the possibility that the release of half length products arises from the action of two separate SbcCD complexes acting at the two DNA ends. This situation has been observed for *E.coli* exonuclease III (46). However, our experiments differ from those carried out on exonuclease III in that we have used protein concentrations that result in <40% conversion of substrate to product within a time scale of 20 min, whereas the visualization of half length products with exonuclease III required concentrations of protein that gave complete degradation within 4 min. The nature of the products that we observe is also different. While we see DNA fragments varying in size between full length and half length and no fragments shorter than half length, the exonuclease III experiments reveal a statistical clustering of fragments about the half length end-point.

ACKNOWLEDGEMENTS

We would like to thank Gareth Cromie, Pan Xuefeng and Stuart Warner for their comments on the manuscript. This work was supported by a grant from the Wellcome Trust.

REFERENCES

1 Leach,D.R.F. (1994) *Bioessays*, **16**, 893–900.
2 Chalker,A.F., Leach,D.R.F. and Lloyd,R.G. (1988) *Gene*, **71**, 201–205.
3 Gibson,F.P., Leach,D.R.F. and Lloyd,R.G. (1992) *J. Bacteriol.*, **174**, 1222–1228.
4 Shurvinton,C.E., Stahl,M.M. and Stahl,F.W. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 1624–1628.
5 Lindsey,J.C. and Leach,D.R.F. (1989) *J. Mol. Biol.*, **206**, 7024–7027.
6 Connelly,J.C. and Leach,D.R.F. (1996) *Genes Cells*, **1**, 285–291.
7 Behnke,D., Malke,H., Hartmann,M. and Walter,F. (1979) *Plasmid*, **2**, 605–616.
8 Kieser,T. and Melton,R. (1988) *Gene*, **65**, 83–91.
9 Peeters,B.P.H., De Boer,J.H., Bron,S. and Venema,G. (1988) *Mol. Gen. Genet.*, **212**, 450–458.
10 Gordenin,D.A., Lobachev,K.S., Degtyareva,N.P., Malkova,A.L., Perkins,E. and Resnick,M.A. (1993) *Mol. Cell. Biol.*, **13**, 5315–5322.
11 Henderson,S.T. and Petes,T.D. (1993) *Genetics*, **113**, 57–62.
12 Ruskin,B. and Fink,G.R. (1993) *Genetics*, **133**, 43–56.
13 Nag,D.K. and Kurst,A. (1997) *Genetics*, **146**, 835–847.
14 Lobachev,K.S., Shor,B.M., Tran,H.T., Taylor,W., Keen,J.D., Resnick,M.A. and Gordenin,D.A. (1998) *Genetics*, **148**, 1507–1524.
15 Kramer,P.R., Stringer,J.R. and Sinden,R.R. (1996) *Nucleic Acids Res.*, **24**, 4234–4241.
16 Collick,A., Drew,J., Penberth,J., Bois,P., Luckett,J., Scaerou,F., Jeffreys,A. and Reik,W. (1996) *EMBO J.*, **15**, 1163–1171.
17 Akgun,E., Zahn,J., Baumes,S., Brown,G., Liang,F., Romanienko,P.J., Lewis,S. and Jasin,M. (1997) *Mol. Cell. Biol.*, **17**, 5559–5570.
18 Coen,E.S., Carpenter,R. and Martin,C. (1986) *Cell*, **47**, 285–296.
19 Peacock,W.J., Dennis,E.S., Gerlach,W.L., Sachs,M.M. and Schwartz,D. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 347–354.
20 Colot,E.S., Haedens,V. and Rossignol,J.L. (1998) *Mol. Cell. Biol.*, **18**, 4337–4346.
21 Roth,D.B., Menetski,J.P., Nakajima,P.B., Bosma,M.J. and Gellert,M. (1992) *Cell*, **70**, 983–991.
22 van Gent,D.C., McBlane,J.F., Ramsden,D.A., Sadofsky,M.J., Hesse,J.E. and Gellert,M. (1995) *Cell*, **81**, 925–934.
23 Lewis,S.M. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 1332–1336.
24 Sharples,G.J. and Leach,D.R.F. (1995) *Mol. Microbiol.*, **17**, 1215–1220.
25 Connelly,J.C., de Leau,E.S., Okely,E.A. and Leach,D.R.F. (1997) *J. Biol. Chem.*, **272**, 19819–19826.
26 Connelly,J.C., Kirkham,L.C. and Leach,D.R.F. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 7969–7974.
27 Tsakumoto,Y. and Ikeda,H. (1998) *Genes Cells*, **3**, 135–144.
28 Johzuka,K. and Ogawa,H. (1995) *Genetics*, **139**, 1521–1532.
29 Dolganov,G.M., Maser,R.S., Novikov,A., Tosto,L., Chong,S., Bressan,D.A. and Petrini,J.H.J. (1996) *Mol. Cell. Biol.*, **16**, 4832–4841.
30 Xiao,Y. and Weaver,D.T. (1997) *Nucleic Acids Res.*, **25**, 2985–2991.
31 Paull,T.T. and Gellert,M. (1998) *Mol. Cell*, **1**, 969–979.
32 Trujillo,K.M., Yuan,S.S.F., Lee,E.Y.H.P. and Sung,P. (1998) *J. Biol. Chem.*, **273**, 21447–21450.
33 Sambrook,J., Frisch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
34 Davison,A. and Leach,D.R.F. (1994) *Nucleic Acids Res.*, **22**, 4361–4363.
35 Kabotyanski,E.B., Zhu,C., Kallick,D.A. and Roth,D.B. (1995) *Nucleic Acids Res.*, **23**, 3872–3881.
36 Ashley,C.T. and Warren,S.T. (1995) *Annu. Rev. Genet.*, **29**, 703–728.
37 Smith,G.K., Jie,J., Fox,G.E. and Gao,X. (1995) *Nucleic Acids Res.*, **23**, 4303–4311.
38 Mariappan,S.V.S., Garcia,A.E. and Gupta,G. (1996) *Nucleic Acids Res.*, **24**, 775–783.
39 Sankar,P.S., Chang,H.-C., Boudi,F.B. and Reddy,S. (1998) *Cell*, **95**, 531–540.
40 Ahkmedov,A.T., Frei,C., Plugfelder,M.T., Kemper,B., Gasser,S.M. and Jessberger,R. (1998) *J. Biol. Chem.*, **273**, 24088–24094.
41 Leach,D.R.F., Okely,E.A. and Pinder,D.J. (1997) *Mol. Microbiol.*, **26**, 597–606.
42 Sun,H., Treco,D. and Szostak,J.W. (1991) *Cell*, **64**, 1155–1161.
43 Bishop,D.K., Park,D., Xu,L. and Kleckner,N. (1992) *Cell*, **69**, 439–456.
44 Niki,H., Imamura,R., Kitaoka,M., Yamanaka,K., Ogura,T. and Hiraga,S. (1992) *EMBO J.*, **13**, 5101–5109.
45 Melby,T.E., Ciampaglio,C.N., Briscoe,G. and Erickson,H.P. (1998) *J. Cell Biol.*, **142**, 1595–1604.
46 Shalloway,D., Kleinberger,T. and Livingston,D.M. (1980) *Cell*, **20**, 411–422.