The EMBO Journal (2012) 31, 1568–1578 | © 2012 European Molecular Biology Organization | Some Rights Reserved 0261-4189/12
www.embojournal.org

THE EMBO JOURNAL

EMBO open

Insights into Chi recognition from the structure of an AddAB-type helicase–nuclease complex

Kayarat Saikrishnan[1,3,4], Joseph T Yeeles[2,4,5], Neville S Gilhooly[2,4], Wojciech W Krajewski[1,4], Mark S Dillingham[2,*] and Dale B Wigley[1,*]

[1]Division of Structural Biology, Institute of Cancer Research, Chester Beatty Laboratories, London, UK and [2]School of Biochemistry, University of Bristol, University Walk, Bristol, UK

In bacterial cells, processing of double-stranded DNA breaks for repair by homologous recombination is dependent upon the recombination hotspot sequence Chi and is catalysed by either an AddAB- or RecBCD-type helicase–nuclease. Here, we report the crystal structure of AddAB bound to DNA. The structure allows identification of a putative Chi-recognition site in an inactivated helicase domain of the AddB subunit. By generating mutant protein complexes that do not respond to Chi, we show that residues responsible for Chi recognition are located in positions equivalent to the signature motifs of a conventional helicase. Comparison with the related RecBCD complex, which recognizes a different Chi sequence, provides further insight into the structural basis for sequence-specific ssDNA recognition. The structure suggests a simple mechanism for DNA break processing, explains how AddAB and RecBCD can accomplish the same overall reaction with different sets of functional modules and reveals details of the role of an Fe–S cluster in protein stability and DNA binding.

The EMBO Journal (2012) **31,** 1568–1578. doi:10.1038/emboj.2012.9; Published online 3 February 2012
Subject Categories: genome stability & dynamics
Keywords: Chi; DNA recombination; DNA repair; Dna2; iron–sulphur cluster

Introduction

Double-stranded DNA breaks (DSBs) are a potentially lethal form of damage and cells have developed two contrasting mechanisms for their repair. In all eukaryotic and many prokaryotic organisms, the non-homologous end joining pathway can rejoin and ligate the broken ends in a process that does not require a homologous template DNA, but which is prone to errors (Pitcher *et al*, 2007; Shuman and Glickman, 2007; Lieber, 2008). Alternatively, the break can be salvaged by the ubiquitous homologous recombination (HR) pathway, in which case a homologous DNA molecule acts as a template thereby ensuring faithful repair (Kowalczykowski, 2000; Wyman and Kanaar, 2006). Different types of damaged DNA structures require processing by specific initiator proteins to enter the HR pathway. In all organisms, the gatekeepers for entry of DSBs into the HR pathway are proteins with the helicase and nuclease activities required to convert a DSB into a long 3′-terminated ssDNA overhang. This is a suitable substrate for RecA/Rad51 binding, DNA strand exchange and subsequent steps of HR, which eventually lead to the repair of the break (Kowalczykowski, 2000; Mimitou and Symington, 2009; Yeeles and Dillingham, 2010). In bacteria, the primary pathway for DSB repair is initiated by a stable helicase–nuclease complex of which there are two distinctive classes: the AddAB- and RecBCD-type enzymes (Dillingham and Kowalczykowski, 2008; Yeeles and Dillingham, 2010). In addition, the alternative RecF pathway can be initiated by the combined actions of the RecQ helicase and RecJ exonuclease (Handa *et al*, 2009). Very recent studies of DNA break processing in eukaryotes have suggested mechanistic parallels with the prokaryotic system, with numerous helicase and nuclease activities implicated in the resection reaction (Mimitou and Symington, 2009; Cejka *et al*, 2010; Niu *et al*, 2010; Nimonkar *et al*, 2011). These include the Dna2 helicase–nuclease, which shares certain similarities with the bacterial AddAB/RecBCD complexes at the level of primary structure (discussed in Cejka *et al*, 2010; Yeeles and Dillingham, 2010).

AddAB and RecBCD complexes catalyse the same net reaction, in which a DSB is converted into a 3′-terminated ssDNA overhang, but they operate by distinctive mechanisms and this reflects their different architectures (Yeeles and Dillingham, 2010). RecBCD complexes comprise two DNA helicase motors of opposite polarity (Dillingham *et al*, 2003; Taylor and Smith, 2003), a single nuclease domain (Wang *et al*, 2000), and a Chi-scanning domain that is structurally related to a helicase, but which contains none of the associated motifs (Singleton *et al*, 2004). In RecBCD enzymes, unwinding of DNA is powered by a bipolar DNA translocation mechanism, which feeds both nascent DNA single strands to one nuclease domain situated at the rear of the enzyme. Processive cleavage of the 3′-terminated strand stops at the Chi sequence ($Chi_{E.\ coli}$ = 5′-GCTGGTGG) despite continued translocation and unwinding. This is thought to occur because the Chi sequence is bound by the Chi-scanning domain, preventing all DNA on the 3′-terminated strand downstream of Chi from ever reaching the nuclease active site (Singleton *et al*, 2004). In contrast, AddAB complexes contain a single helicase motor and two nuclease domains, which are dedicated to the cleavage of each of the nascent strands of DNA (Quiberoni *et al*, 2001; Yeeles and Dillingham,

*Corresponding authors. MS Dillingham, School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, UK. Tel.: +44 (0)117 331 2159; Fax: +44 (0)117 331 2168; E-mail: Mark.Dillingham@bristol.ac.uk or
DB Wigley, Division of Structural Biology, Institute of Cancer Research, Chester Beatty Laboratories, 37 Fulham Road, London SW3 6JB, UK. Tel.: +44 (0)207 153 5521; Fax: +44 (0)207 153 5457; E-mail: Dale.Wigley@icr.ac.uk
[3]Present address: Division of Biology, Indian Institute of Science Education and Research, Pashan, Pune 411021, India
[4]These authors contributed equally to this work
[5]Present address: Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY, USA

Received: 6 September 2011; accepted: 4 January 2012; published online: 3 February 2012

2007; Yeeles *et al*, 2011a). *Bacillus subtilis* AddAB recognizes a short pentameric Chi sequence ($Chi_{B.\ subtilis}$ = 5′-AGCGG), but counter-intuitively, interacts with Chi much more strongly than does *Escherichia coli* RecBCD (Chedin *et al*, 2006). Site-directed mutagenesis experiments have suggested that ATP binding at a conserved Walker A motif near the N-terminus of AddB might stabilize the Chi complex (Yeeles *et al*, 2011a). This motif is not important for the helicase activity of AddAB and, interestingly, there is no equivalent site in the RecBCD-type enzymes. A further distinction between AddAB- and RecBCD-type enzymes is that many AddAB complexes contain a 4Fe–4S cluster which is, minimally, important for structural integrity and essential for DNA binding (Yeeles *et al*, 2009). The cluster is associated with the AddB C-terminal nuclease domain, that is the prototypical member of a new class of 'iron-staple' nuclease domain, apparently also present in the eukaryotic helicase–nuclease, Dna2 (Yeeles *et al*, 2009), the mitochondrial replication/recombination factor Exonuclease V (Burgers *et al*, 2010), and the Cas4 component of some prokaryotic CRISPR systems (Makarova *et al*, 2006). It is currently unknown if this Fe–S cluster is redox active or whether electron transfer plays any role in the mechanism of AddAB or the related enzymes.

We have solved two structures of *B. subtilis* AddAB complexed with DNA. In one structure, the Fe–S cluster is intact and this helps to rationalize its role in DNA binding. Comparison of the AddAB structure with that of RecBCD suggests a likely Chi-recognition site. Consistent with this hypothesis, mutation of residues in this site abolish Chi recognition and shed light on the structural basis for Chi binding in these enzymes.

Results

The crystal structure of AddAB with an intact 4Fe–4S cluster was solved to 3.2 Å resolution (Figure 1). The heterodimeric complex is bound to a 19-bp DNA duplex possessing a hairpin loop at the distal end and a 5 base 3′-ssDNA tail. This substrate mimics a DNA break, the physiological substrate for AddAB, with the hairpin loop forcing AddAB to interact in the desired orientation. The interactions between AddAB and DNA are summarized in Supplementary Figure S1. A second structure, in which the 4Fe–4S cluster is missing, was solved to a higher resolution of 2.8 Å. With the exception of the missing cluster, and greater disorder in the bound DNA at the ss–dsDNA junction, this structure is virtually identical to that solved at lower resolution. Below, we describe the structures of the AddA and AddB polypeptides individually before returning to the complex and its interactions with the DNA.

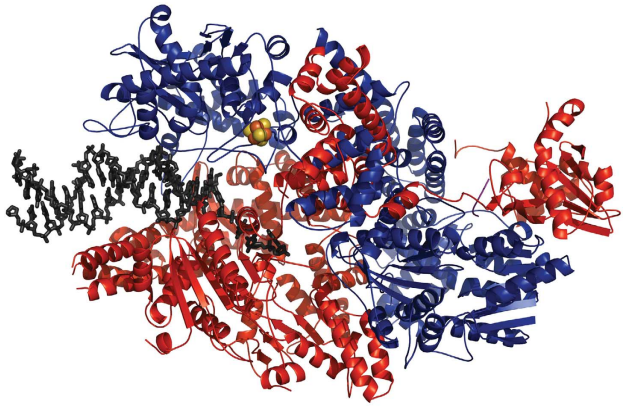

Figure 1 Overall structure of the AddAB–DNA complex. Ribbons representation of the AddAB–DNA complex with an intact 4Fe–4S cluster solved at 3.2 Å resolution. The AddA and AddB polypeptides are shown in red and blue, respectively. The DNA is shown in black sticks and the 4Fe–4S cluster is shown as yellow and orange spheres.

Structure of AddA

As predicted from the primary structure and biochemical analyses, the AddA protein comprises an N-terminal SF1A helicase domain and a C-terminal RecB-family nuclease domain joined by an extended linker (Figure 2A). In common with other UvrD-like helicases (Subramanya *et al*, 1996; Korolev *et al*, 1997; Singleton *et al*, 2004; Lee and Yang, 2006), the AddA helicase domain is divided into four subdomains, two of which (1A and 2A) are the helicase core domains found in all SF1 and SF2 nucleic acid motors, the other two (1B and 2B) being so-called accessory domains (Singleton *et al*, 2007). Helicase signature motifs involved in ATP binding and hydrolysis are found in their expected location at the interface of the core domains, but our structure does not contain a bound nucleotide (Figure 2B). Furthermore, several of the helicase motifs contact the 3′-terminated strand of the DNA substrate, which is bound in the canonical fashion across the top surface of both core domains (see below) (Singleton *et al*, 2007). The 1B accessory domain forms an extended 'arm' structure, the tip of which contacts the duplex portion of the substrate and the 2B accessory domain is extensively involved in protein–protein interactions with the AddB protein. The AddA nuclease domain displays a fold that is typical of RecB-family nucleases. The active site is formed by four conserved motifs, three of which are common to a very large family of nucleases and resolvases, whereas the fourth is unique to the RecB family (Figure 2C) (Aravind *et al*, 2000).

Structure of AddB

Like AddA, the fold of the N-terminal portion of AddB is divided into the four subdomains typical of UvrD-like helicases (Figure 2D). However, the conventional sequence motifs associated with a helicase are all missing from the core domains with the exception of helicase motif I (equivalent to the Walker A motif). Structural alignment of AddB with a conventional helicase shows that this motif is located at the interface of the core domains, consistent with a *bona fide* NTP-binding site (Figure 2E; Supplementary Figure S6). Furthermore, other conserved amino acids cluster around this site. These include an aspartate residue (D208), which occupies a position in the structure exactly equivalent to that of the Mg^{2+}-coordinating aspartate of helicase motif II (Soultanas *et al*, 1999), and an arginine (R283) equivalent to the ribose-interacting arginine in helicase motif IV (Soultanas *et al*, 1999). In DExx-box proteins (including helicases), the residue immediately following the aspartate in motif II is a conserved glutamate that acts as a catalytic base for activation of the water molecule required for ATP hydrolysis (Soultanas *et al*, 1999). However, in *B. subtilis* AddB, this residue is a glycine, which could not fulfil this catalytic role. Similarly, the 'arginine finger' from helicase

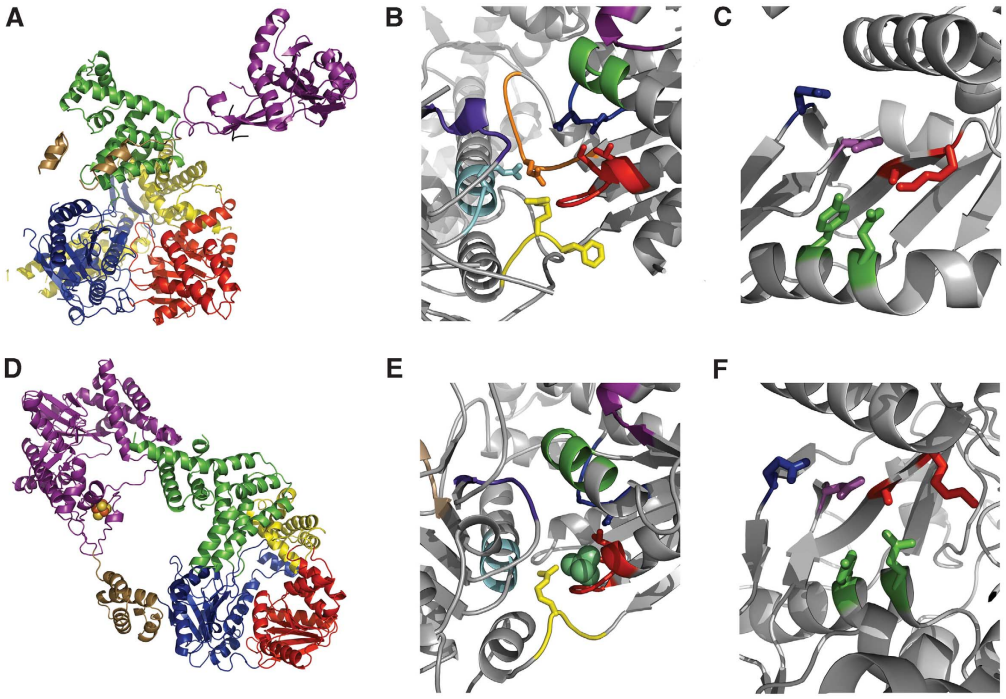

Figure 2 Structures of the AddA and AddB subunits. (**A**) Ribbons representation of AddA with the helicase core domains coloured in red (1A/N-core) and blue (2A/C-core). Accessory domains are yellow (1B) and green (2B) and the C-terminal RecB-family nuclease is magenta. A linker region between the helicase and nuclease domains is coloured brown. (**B**) Detail of the ATP-binding pocket in AddA. Helicase motifs are highlighted in different colours (I red, Ia green, Ib magenta, II blue, III orange, IV yellow, V purple, and VI cyan). Conserved residues involved in the binding and hydrolysis of ATP are shown in stick format (K36 and T37 in motif I, D407 and E408 in motif II, Q441 in motif III, F478 and R479 in motif IV, and R873 in motif VI). (**C**) Detail of nuclease active site of AddA. Nuclease motifs are highlighted in different colours (I blue, II magenta, III red, and green for the 'RecB-family'-specific motif). Conserved residues involved in phosphodiester hydrolysis are shown in stick format (E1129 in motif I, D1159 in motif II, D1172A and K1174 in motif III, as well as Q1200 and Y1204 in the RecB-family motif). (**D**) AddB structure coloured with the same scheme as (**A**). In addition, the 4Fe–4S cluster is shown as yellow and orange spheres. Note that the core helicase domains (red and blue) are inactivated (i.e., the helicase motifs are not conserved). (**E**) Detail of the putative ATP-binding pocket of AddB. Regions structurally equivalent to helicase motifs are coloured as in (**B**) and a sulphate ion is shown in turquoise spheres. Conserved residues that may be involved in ATP binding are shown in stick format (K14 and T15 in the motif I equivalent, D208 in the motif II equivalent, and R283 in the motif IV equivalent). (**F**) Detail of the nuclease active site of AddB coloured as in (**C**). The residues shown in stick format are E915 in motif I, D944 in motif II, D961A and K963 in motif III, as well as Q981 and Y985 in the RecB-family motif.

motif VI, which promotes NTP hydrolysis and coupling to DNA translocation, is replaced by a serine (S660) in AddB. Furthermore, the entire helicase motif III loop, which plays a crucial role in energy transduction and DNA translocation in UvrD-like helicases (Dillingham *et al*, 1999), is completely absent from AddB (Supplementary Figure S6). These observations raise the possibility that the AddB protein might bind, but not hydrolyse, NTP at the interface of the 'helicase-like' core domains. Interestingly, the 2.8-Å structure shows electron density in this putative nucleotide-binding pocket that we interpret as a bound sulphate ion (Figure 2E; Supplementary Figure S8).

The C-terminal region of AddB comprises the nuclease domain connected to the 'helicase-like' domains by a linker region. A portion of the nuclease domain is similar to the equivalent domain in AddA, being typical of a RecB-family nuclease domain (Figure 2F). However, there are a number of additions compared with the nuclease of AddA, including a part that contains a 4Fe–4S cluster and a region that makes extensive contacts with the DNA substrate (described below).

Overall structure of the AddAB complex and interactions with DNA

The AddAB complex displays an overall architecture that is reminiscent of the RecBC subcomplex of RecBCD (Singleton *et al*, 2004), which involves an intricate embrace of the two proteins around one another (Figure 1 and compare Figure 3A and B). There are extensive protein:protein interactions, particularly between the 2B subdomains of AddA and AddB, and the interface between the AddA and AddB monomers buries a total of 9060 Å^2, as calculated with the program PISA (Krissinel and Henrick, 2007). In addition, and in contrast to the RecBC complex, the linker peptides joining the helicase (or helicase-like) domains to the nuclease domain of each protein also interact to form a compact six helical domain on the side of the complex. Indeed, despite the superficial similarity, AddAB and RecBC display other striking differences. In both AddA and RecB, the 1B subdomain forms an 'arm' that contacts duplex DNA (Singleton *et al*, 2004), but the relative orientation of this structure with respect to the rest of the complex differs by about 90° and interactions with the duplex are quite different. In AddAB, the arm mainly contacts the phosphate backbone of the 3′-strand from a helix running above the major groove, extending out to a point around a dozen basepairs from the ss/dsDNA junction (Figure 4A). In addition to these contacts with the AddA subunit, the duplex portion of the substrate interacts extensively with the AddB nuclease domain (Figure 4B), which contacts a region covering about 10 bases on the 5′-strand, albeit not contiguously, and four or five bases on the

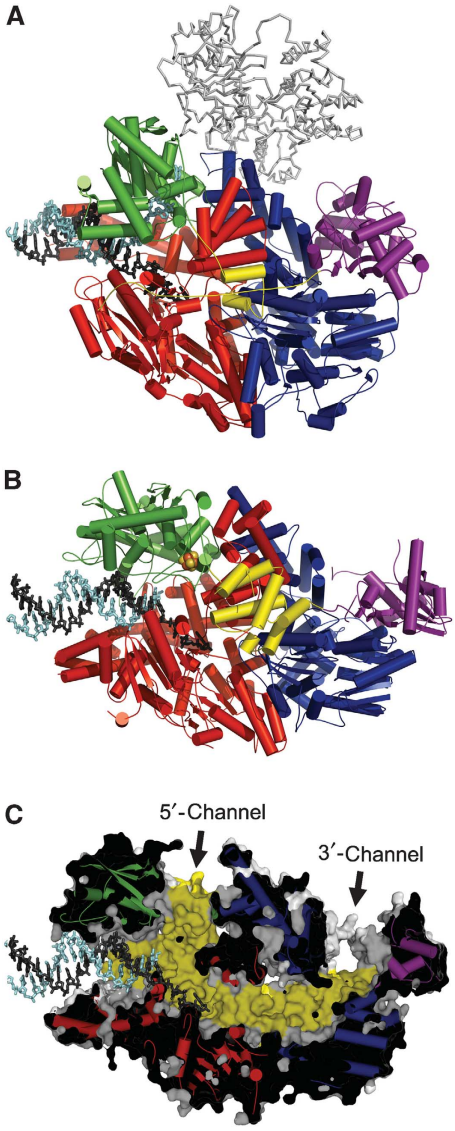

Figure 3 Architecture of the AddAB–DNA complex and comparison with RecBCD. (**A**) Ribbons representation of the RecBCD–DNA complex. The helicase and nuclease domains of RecB are shown in red and magenta, respectively. The helicase-like domain of RecC is shown in blue. The green domain displays a nuclease-like architecture but is *not* an active nuclease. The linker regions between the helicase and nuclease domains of RecB and RecC are both shown in yellow. The RecD subunit (for which there is no equivalent in AddAB) is shown as a grey backbone trace. The DNA is shown in a stick representation with the 3′-strand coloured black and the 5′-strand coloured pale blue. (**B**) Ribbons representation of the AddAB–DNA complex in a similar orientation to RecBCD. The helicase and nuclease domains of AddA are shown in red and magenta, respectively. The helicase-like and nuclease domains of AddB are shown in blue and green, respectively. Additionally, the linker regions between the helicase and nuclease domains of AddA and AddB are both shown in yellow and the 4Fe–4S cluster is shown as yellow and orange spheres. The DNA is shown in a stick representation with the 3′-strand coloured black and the 5′-strand coloured pale blue. (**C**) A cut-through view of AddAB indicating the channels for 3′- and 5′-terminated single-stranded DNA that extend right through the structure. The grey surface representation has been sliced to reveal the internal structure, which is coloured according to the scheme in (**B**). The back surfaces of the ssDNA channels are 'painted' yellow for clarity.

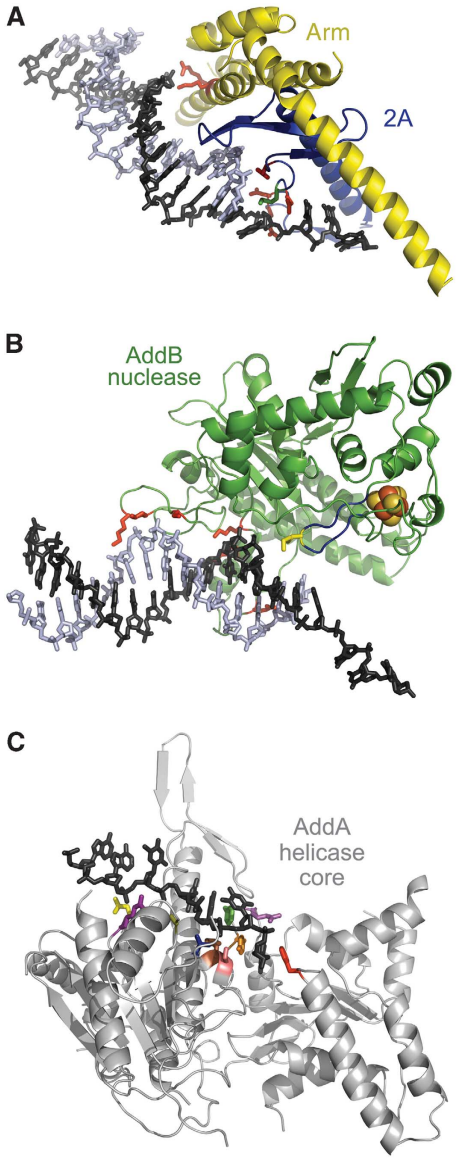

Figure 4 Interactions of AddA and AddB with DNA. (**A**) Ribbons representation of the tip of the AddA 1B 'arm' domain (yellow) and part of the 2A core domain (blue). Residues that interact with DNA (blue/black) are shown as red sticks (see Supplementary Figure S1 and text for details). A methionine residue that may act as a 'pin' at the ss–dsDNA junction is shown in green sticks. (**B**) Ribbons representation of the AddB nuclease domain (green) showing residues (red sticks) that interact with DNA (blue/black) (see Supplementary Figure S1 for details). The 4Fe–4S cluster is shown as orange and yellow spheres. The loop immediately preceding the second coordinating cysteine residue is coloured blue and includes a lysine residue (yellow sticks) involved in binding the 3′-strand (see text for discussion). (**C**) Ribbons representation of the core helicase domains of AddA (grey) bound to the 3′-ssDNA tail of the DNA substrate (black). Residues that form the ssDNA motor site are shown in sticks and coloured individually as follows: F65 red, Y444 orange, F446 green, R447 magenta, S591 olive, T792 blue, H794 brown, S795 pink, R811 purple, N812 yellow, N814 cyan.

3′-strand. The interactions are exclusively with the DNA backbone as would be expected for a protein that is not sequence specific and involve a number of conserved residues (including Q1017, K1033, K1036, K1068, K1069, and S1075). In the RecBC(D) complex, the equivalent region (the C-terminal domain of RecC) makes no contacts at all with the DNA duplex.

The AddAB complex was crystallized with a DNA hairpin substrate containing a 19-bp duplex and a five base

3′-terminated ssDNA tail. The DNA is fully base-paired right up to the end of the duplex where the 3′-tail extends, so it has not been unwound upon formation of this initiation complex. In RecBCD, several basepairs of duplex DNA are unwound in the absence of ATP (Farah and Smith, 1997; Wong *et al*, 2005; Saikrishnan *et al*, 2008) and this might be related to the differences in the orientation of the 'arm' in the two complexes. However, it should also be noted that the ability of RecBCD to unwind the DNA end is dependent upon calcium or magnesium ions (Wong *et al*, 2005), neither of which are present in the crystallization conditions. In common with many helicases including RecBCD (Singleton *et al*, 2007), AddAB appears to contain a 'pin' (M816) to split the duplex at the junction of single- and double-stranded DNA (Figure 4A). However, given that our substrate has not been unwound upon binding, and does not contain pre-formed ssDNA overhangs on both strands, we cannot exclude the possibility that other parts of the structure may be involved. Indeed, a DNA-binding loop supported by the Fe–S cluster may also help separate the strands (Supplementary Figure S5 and see below).

Inspection of the interior of the AddAB complex reveals two open channels that begin at the junction of single- and double-stranded DNA and extend right through the protein complex, finally exiting at different points at the top and rear of the structure (Figure 3C). The channels are large enough to accommodate single-, but not double-stranded DNA, and part of one channel is occupied by the 3′-tail of the DNA substrate, where it has been engaged by the AddA motor domain. The contacts between AddA and ssDNA are similar to those seen in related helicases such as PcrA, Rep, RecB and UvrD (Figure 4C) (Korolev *et al*, 1997; Velankar *et al*, 1999; Singleton *et al*, 2004; Lee and Yang, 2006). Therefore, it is likely that AddAB translocates along DNA in single base steps using the conserved inchworm mechanism for translocation first proposed for PcrA (Velankar *et al*, 1999). The 5′-end of the DNA at the junction is poised to enter the second channel. During active translocation along duplex DNA, each channel will accommodate one specific DNA strand thereby enforcing separation of the duplex (i.e., helicase activity), and they will be referred to hereafter as the 3′- and 5′-channels. This proposition is strongly supported by biochemical analyses of the strand specificity associated with each nuclease domain (Yeeles and Dillingham, 2007). The AddA and AddB nuclease domains were shown to cleave the 3′- and 5′-terminated strands, respectively, and the structure shows that the AddA and AddB nuclease active sites are located near the exit points of the 3′- and 5′-channels as would be expected. By contrast, in RecBCD, both DNA strands are degraded by a single nuclease active site situated at the rear of the complex.

Various functional modules in the AddAB complex appear at different points along each of the two channels (Figure 3B and C). From the entry to the exit of the 3′-channel, the translocating ssDNA would first encounter the AddA SF1A helicase motor, then the AddB inactivated SF1 helicase domain, and finally the AddA nuclease domain. The AddA SF1A motor domains were shown to catalyse ssDNA tracking in the 3′–5′ direction (Yeeles *et al*, 2011a), which is the appropriate polarity to move the enzyme into the duplex, feeding the single strands through the complex to the channel exit points. The 5′-channel is relatively short, containing only the AddB nuclease active site and the region of the structure that coordinates a 4Fe–4S cluster.

Role of a 4Fe–4S cluster in DNA binding and protein stability

Biochemical and electron paramagnetic resonance studies have shown that AddAB contains a cubane 4Fe–4S cluster and identified four conserved cysteine residues in AddB as the ligands (Yeeles *et al*, 2009). Furthermore, it was shown that disruption of this Fe–S cluster resulted in a complete loss of DNA-binding activity in the AddAB complex and a destabilization of the C-terminal nuclease domain of AddB (Yeeles *et al*, 2009). We have determined two structures of the AddAB complex: one with the Fe–S cluster intact (at 3.2 Å resolution) and another in which the cluster is absent (at 2.8 Å resolution). In the 3.2-Å structure, density consistent with a 4Fe–4S cluster is found near the linker region between the N- and C-terminal domains, and close to the interface with AddA protein (Figures 1 and 3). The cluster is deeply buried and surrounded by a protective shell of conserved aromatic residues (Supplementary Figure S5A and B). However, the cluster is partially exposed to the interior channel that accommodates the 5′-terminated DNA strand near the nuclease active site. As had been predicted (Yeeles *et al*, 2009), four conserved cysteine residues (C801, C1121, C1124, and C1130) are present at the four opposing vertices of the cubane cluster. The second, third, and fourth cysteine residues are presented on a short loop and an α-helix, an arrangement that is essentially identical to the equivalent region of EndoIII, an Fe–S containing DNA glycosylase (Supplementary Figure S5) (Thayer *et al*, 1995; Fromme and Verdine, 2003). In AddAB, the first cysteine (which is about 300 amino acids away from the second in the primary structure) is presented in a turn connecting two helices and there is no similarity in this respect with EndoIII.

In the EndoIII and MutY glycosylases, the amino acids that immediately precede the second cysteine ligand form a loop that is involved in binding DNA. This region of the protein is referred to as the iron–sulphur cluster loop and contains conserved residues, involved in binding a region of the DNA substrate, that are distant from the active site that cleaves the N-glycosidic bond (Thayer *et al*, 1995; Guan *et al*, 1998; Chepanoske *et al*, 2000). The position of a conserved lysine residue (AddB K1116) in the equivalent loop of AddAB is just close enough to the DNA to suggest that it might contact the phosphodiester backbone in the duplex region of the substrate (Figure 4B; Supplementary Figure S5). This loop might alternatively, or additionally, assist in the prising apart of the duplex because its position would block and divert the path of the 5′-terminated strand following translocation of the 3′-strand by the helicase domain. Indeed, an Fe–S cluster may play a similar role in the SF2 helicase XPD (Fan *et al*, 2008; Liu *et al*, 2008; Pugh *et al*, 2008, 2011; Wolski *et al*, 2008). The 2.8-Å structure does not contain density for the Fe–S cluster and, although the majority of the structure is indistinguishable from the Fe–S bound complex, the putative DNA-binding loop and the junction of the DNA itself are poorly ordered, helping to explain the importance of the Fe–S cluster in binding the substrate (Yeeles *et al*, 2009). In solution, the loss of the Fe–S cluster results in a local unfolding of the entire AddB nuclease domain and, consequently, would result in the loss of all associated DNA-binding contacts (Figure 4B). This is not the case in the crystal, and this implies that the Fe–S cluster was lost post-crystallization.

The inactivated helicase domain of AddB is the Chi-recognition locus

In RecBCD, an inactivated helicase domain in the RecC subunit is thought to be responsible for Chi recognition (Singleton *et al*, 2004). By analogy, the N-terminal region of AddB was the prime candidate for recombination hotspot recognition in AddAB. To test this possibility, we purified several AddAB complexes containing point mutations in conserved AddB residues that line the 3′-channel between the AddA motor and nuclease domains. The mutant complexes behaved similarly to wild-type AddAB during purification and retained the ability to bind tightly to dsDNA ends (Supplementary Figure S2). They were then tested for their ability to process DNA substrates that were either free of recombination hotspots or which contained a single Chi sequence at a defined position (Figure 5; Supplementary Figures S3 and S4). With substrates devoid of Chi, the wild-type enzyme produces a smear of different sized ssDNA products due to unwinding and stochastic cleavage of both nascent single strands (Chedin *et al*, 2000). When a Chi sequence is present, a prominent band appears in the smear of cleavage products, which reflects the downregulation of nuclease activity on the 3′-terminated strand and the protection of that strand between Chi and the distal DNA end. The yield of this 'Chi fragment' provides a simple quantitative test for Chi recognition in our mutant proteins. The wild-type and mutant complexes displayed similar DNA processing on substrates devoid of Chi sequences, suggesting that they retain comparable helicase and nuclease activities to wild-type AddAB. However, on substrates containing a recombination hotspot, seven of the eight mutant proteins were specifically defective in their ability to produce the Chi fragment. In some cases (D41A, Q42A, T44A, R70A, F210A), the Chi fragment was undetectable or barely detectable above the background smear of ssDNA products, whereas for other mutants (F68A, W73A) the efficiency of Chi recognition was significantly reduced relative to wild type. One mutant protein (F213A) produced the Chi fragment at levels comparable to wild type. There was no evidence for the formation of novel bands within the smear of ssDNA products, which would have been indicative of altered or relaxed Chi-recognition specificity. These results show that the 'helicase-like' core domains of AddB are important for the recognition and response to Chi sequences. Intriguingly, most of these mutations map to positions exactly equivalent to the helicase signature motifs in a conventional SF1A helicase (Figure 6; Supplementary Figure S6). The Chi-recognition apparatus is at least partly formed by modified versions of 'helicase' motifs Ia and Ib, both of which form part of the ssDNA motor in a conventional helicase. This work also provides evidence for the involvement of a region equivalent to motif II in sequence recognition. Interestingly, residues equivalent to Q42 and T44 in motif Ia as well as W73 (just outside motif Ib) appear to be common to both AddAB and its functional analogue *E. coli* RecBCD (SC Kowalczykowski, personal communication), which recognizes a different Chi sequence. These residues could very well be responsible for contacting conserved elements of the Chi sequence that are shared across diverse bacterial species (Halpern *et al*, 2007). For example, all known Chi sequences are G-rich and, more specifically, the recombination hotspots of *B. subtilis* and *E. coli* share a core sequence (GxGG) at their 3′-end. This is exactly the region of Chi expected to be closest to the AddA nuclease domain, and therefore likely to be contacted by regions equivalent to motifs Ia and Ib, which are located on the N-terminal core domain (1A). Likewise, specific differences in these motifs (Figure 6; Supplementary Figure S6) are likely to identify residues responsible for contacting elements of the Chi sequence that are unique to each bacterial species. It is striking that RecBCD-type enzymes contain well-conserved residues in regions equivalent to helicase motifs IVa and V, whereas the same parts of AddAB-type enzymes are relatively less well conserved. This might reflect the recognition on the 2A core domain of the additional three residues at the 5′-end of the octameric RecBCD Chi sequence. Finally, in RecBCD, the recognition of Chi is thought to cause a conformational change triggered by the unlatching of a nearby ionic interaction between a 'latch' in the 2A core domain and the 1B accessory domain of RecC (SC Kowalczykowski, personal communication). A similar 'ionic latch' structure is found in AddB and site-directed

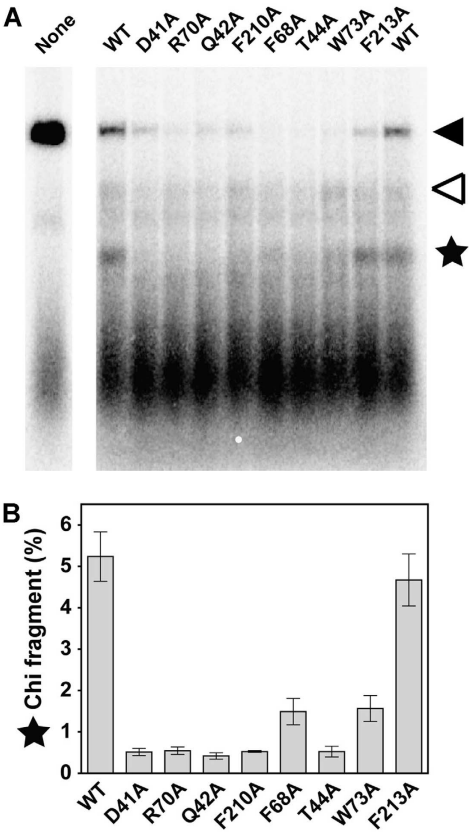

Figure 5 The inactivated helicase domain of AddB is the Chi-binding locus. (**A**) Comparison of wild-type and mutant AddAB proteins (2 nM) processing a linear DNA substrate (1.6 nM) containing a single *B. subtilis* Chi sequence. The major substrate and product species are highlighted with markers (filled triangle = substrate dsDNA, empty triangle = full-length ssDNA, star = Chi fragment). A smear of non-specific ssDNA products is also produced in all cases. Several of the mutant proteins display a reduced or undetectable yield of the Chi fragment (compare band marked with the star). A more extensive characterization of these mutant proteins is shown in the Supplementary data. (**B**) Quantification of the apparent yield of the Chi fragment for wild-type and mutant AddAB proteins (see Materials and methods for details). The error bars represent the standard error of the mean for three independent experiments.

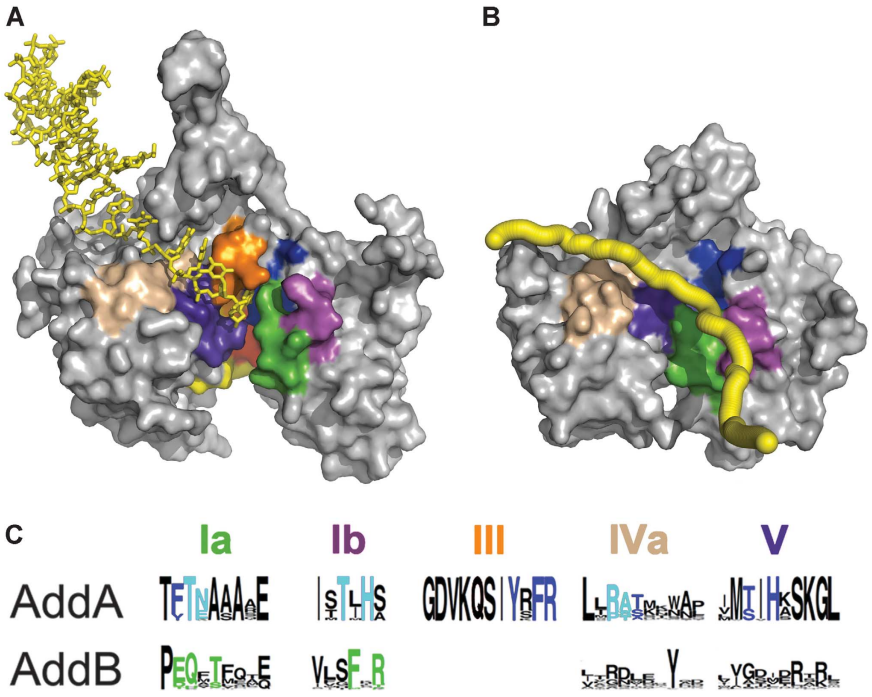

Figure 6 Comparison of the ssDNA motor site of AddA with the Chi-recognition locus of AddB. (**A**) The helicase core domains of AddA are shown as a grey surface representation with conserved helicase motifs highlighted in different colours (I red, Ia green, Ib magenta, II blue, III orange, IV yellow, IVa wheat, V purple, and VI cyan). The 3′-strand interacts directly with helicase motifs Ia, Ib, III, IVa, and V in related SF1A helicases such as PcrA and UvrD (Velankar *et al*, 1999; Lee and Yang, 2006) (see **C**). (**B**) The helicase-like core domains of AddB are shown in the same orientation as for AddA. AddB does not contain canonical helicase motifs but equivalent regions of the protein are coloured using the same scheme as in (**A**), although there is no equivalent of helicase motif III. The yellow 'worm' represents the centre of the 3′-channel calculated using MOLE (Petrek *et al*, 2007). (**C**) Conservation of the amino acids in some of the helicase or 'quasi-helicase' motifs of AddA and AddB. The motifs are shown in Weblogo format (Crooks *et al*, 2004) and were determined from ~200 AddAB sequences using COBALT (Papadopoulos and Agarwala, 2007). Motifs are colour coded as in (**A**, **B**). For AddA, the dark and light blue residues are actual or predicted contacts (respectively) with the 3′-strand of the DNA (Velankar *et al*, 1999; Lee and Yang, 2006). For AddB, the green residues are involved in Chi recognition (Figure 5). A more expansive version is available in the Supplementary data.

mutagenesis experiments support the idea that it is unlocked in response to Chi recognition (see Supplementary Figure S7 for discussion).

A mechanism for AddAB-type helicase–nucleases

The identification of the inactivated helicase domain of AddB as the Chi-recognition locus allows us to propose a simple mechanism for the DNA break processing reaction catalysed by AddAB (Figure 7). AddAB binds tightly to DNA ends, with extensive interactions between the duplex and the AddA arm, as well as the AddB nuclease domain including the Fe–S cluster loop. The 3′-terminated strand engages with the AddA SF1A helicase motor, which drives translocation at one base per ATP using an inchworm mechanism (Velankar *et al*, 1999; Lee and Yang, 2006). This forces the nascent ssDNA strands through the interior channels, where they each encounter a nuclease domain that makes occasional stochastic cuts. The DNA strands exit the complex in close proximity and in basepair register, explaining the high propensity for re-annealing to reform duplex DNA that is observed in the absence of SSB protein (Yeeles *et al*, 2011b). This 'single-motor, dual-nuclease mechanism' results in the processive translocation and cleavage of the DNA that occurs before Chi recognition (Figure 7A). When a Chi sequence passes through the 3′-channel, it is recognized and bound by the AddB Chi-recognition site. By sequestering the Chi sequence within AddAB, a growing ssDNA loop is formed as the complex continues to translocate (Figure 7B). This ssDNA loop, which has been detected by AFM, is both refractory to re-annealing and protected from cleavage by the AddA nuclease, allowing the formation of the 3′-terminated ssDNA overhang that will become the recombinogenic RecA nucleoprotein filament (Yeeles *et al*, 2011b). This intermediate participates in further steps in the HR repair pathway, but the details of how RecA is loaded onto the loop are not clear at this stage for AddAB. The recombinogenic ssDNA loop would require the presence of an alternative exit channel from the AddAB complex, similar to that proposed for RecBCD (Wong *et al*, 2006). Interestingly, a candidate exit channel is lined and partially occluded by the AddB latch structure described above (Supplementary Figure S7). It seems likely that unlocking of this latch structure upon Chi recognition fully opens the exit channel for extrusion of an ssDNA loop. However, a more complete picture of the effect of recombination hotspots on AddAB will await the crystal structure of AddAB in a *bona fide* Chi-recognition complex.

The overall mechanism presented here is similar in principle to that proposed for RecBCD (Singleton *et al*, 2004). In both cases, a combination of helicase, nuclease, and ssDNA-recognition domains are combined to create a machine capable of sequence-regulated processive nuclease activity. However, nature has developed at least two distinctive

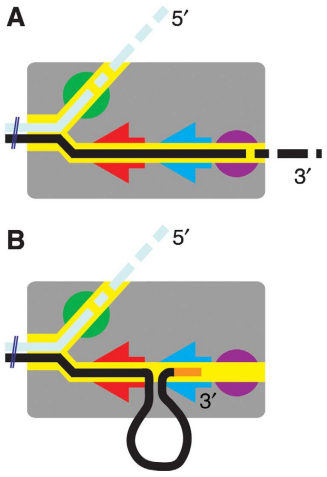

Figure 7 Proposed mechanism for the control of nuclease activity by recombination hotpots in an AddAB-type helicase–nuclease. (**A**) Schematic drawing of the AddAB structure (grey box) showing ssDNA channels (yellow) and various functional domains; the SF1A helicase motor is a red arrow, the Chi-binding locus is a blue arrow, and the AddA and AddB nuclease domains are magenta and green circles, respectively. The DNA is shown as black and pale blue lines representing the 3′- and 5′-terminated single strands, respectively. This part of the figure shows the pre-Chi translocation mode. The DNA strands are pumped through the complex by the activity of the SF1A motor domain. As they pass through the nuclease domains they are occasionally cleaved, resulting in the processive degradation of both strands of the duplex substrate. (**B**) Post-Chi translocation mode. Following the delivery of a correctly orientated Chi sequence (orange) to the recognition locus it is tightly bound within the enzyme complex, enforcing a final cleavage event on the 3′-terminated strand just upstream of Chi. The SF1A motor continues to translocate, resulting in continued cleavage of the 5′-strand by the AddB nuclease domain (green), and the formation of an ssDNA loop between the SF1A motor in AddA and the Chi-recognition locus in AddB. The net product is a 3′-ssDNA overhang that is the substrate for RecA-mediated strand exchange and subsequent steps of recombinational repair.

architectures that are equally adept at the job, but which differ significantly in the details of how the reaction is catalysed.

Discussion

In this work, we solved the crystal structure of an AddAB-type helicase–nuclease and identified the Chi-recognition locus using site-directed mutagenesis. This provided a structural rationale for the different Chi-recognition specificities of AddAB and RecBCD complexes and new insights into sequence-specific ssDNA–protein interactions. A remarkable feature of the Chi-recognition apparatus found in the AddAB and RecBCD systems is its apparent evolution from a structure that functioned originally as an ssDNA motor. This is evident not only at the level of tertiary structure similarity, but also in the Chi-recognition role played by amino acids found at locations exactly equivalent to the positions of key residues in conventional helicase motifs. This concept first became apparent through structural analysis of the RecC component of RecBCD (Singleton *et al*, 2004). The structure of AddB now extends this concept by providing a picture of gradual evolutionary morphogenesis from motor to sequence-recognition device. The core domains of AddB appear to retain some features of a *bona fide* helicase domain, including residues involved in the binding of NTP at the interface of the core domains. Moreover, it is apparent from sequence analyses that some AddAB complexes also retain amino acid residues that would be expected to promote NTP hydrolysis (Figure 6; Supplementary Figure S6). Previous work has shown that mutation of the conserved Walker A motif destabilizes the Chi-recognition complexes that are formed by AddAB (Yeeles *et al*, 2011a). This supports a model in which Chi recognition is allosterically stabilized by NTP binding in a mechanism that would presumably bear some structural analogy with the ATP-dependent conformational transitions of a conventional helicase. These ideas will be tested in future experiments.

Our structure suggests a simple mechanism for DNA break processing by AddAB, and helps to rationalize the distinctive architectures of AddAB and RecBCD enzymes. Both systems use ssDNA tracking motors (SF1α helicase domains) to unwind the duplex and feed the single strands along two channels to a ssDNA endonuclease. However, whereas RecBCD uses dual motors (in RecB and RecD) coupled to a single nuclease (in RecB), AddAB employs a single-motor (in AddA), dual-nuclease (in AddA and AddB) mechanism (Yeeles and Dillingham, 2007, 2010; Yeeles *et al*, 2011a). The architecture of *Bacillus* AddAB also provides a structural framework for the enigmatic AdnAB helicase–nucleases that are restricted to the actinomycete niche (Unciuleac and Shuman, 2010). It has been suggested that these enzymes might, uniquely, catalyse DNA end resection using a dual-motor, dual-nuclease mechanism (Sinha *et al*, 2009). In such a scenario, the second active motor is likely to be in a position equivalent to the Chi-scanning domain (i.e., the inactivated helicase domain) of AddB. In support of this idea, the primary structure of AdnA (equivalent to AddB) suggests that, in contrast to AddB and RecC, it retains the helicase motif III loop that forms a critical part of the ssDNA motor in PcrA/UvrD-like helicases (Supplementary Figure S6) (Sinha *et al*, 2009). Moreover, although its helicase motifs are by no means fully conserved, AdnA also retains Walker A and Walker B motifs for nucleotide binding and hydrolysis. One may speculate therefore, that the AdnAB system represents an intermediate or alternative step in the development of a novel activity from a SF1 helicase motor. In this respect, it will be of great interest to determine the function of AdnA; the AddB-like component of the AdnAB complex.

Finally, this work provides a structural basis for the role of an Fe–S cluster domain in the DNA-binding activity and stability of AddAB, and reveals an unexpected structural similarity with the DNA glycosylase EndoIII. Fe–S clusters have relatively recently been identified in a number of DNA-binding proteins including the SF2 helicase XPD (White, 2009; White and Dillingham, 2011). In both AddAB and XPD, the Fe–S cluster domain is located towards the front of the enzyme with respect to translocation and might provide a 'pin' or 'wedge' structure to assist separation of the DNA strands, thereby contributing to helicase activity (Pugh *et al*, 2008, 2011). In AddAB, the Fe–S cluster is associated with the C-terminal domain of AddB. This is the prototypical member of a class of Fe–S containing nucleases that have been termed 'iron-staple' nuclease domains and which are also apparent in Exonuclease V and Cas4 (Makarova *et al*, 2006; Yeeles *et al*, 2009; Burgers *et al*, 2010). Most intriguingly, an equivalent domain is present in the N-terminus of the Dna2 protein, a component of the eukaryotic Dna2–BLM–RPA–MRN complex, which like

RecBCD, contains nuclease and bipolar helicase domains. Therefore, as noted previously (Cejka *et al*, 2010; Yeeles and Dillingham, 2010), the bacterial AddAB and RecBCD complexes might provide a structural framework for interpreting the architecture of the Dna2–BLM–RPA–MRN complex, which promotes the equivalent DNA end processing reaction in eukaryotic cells, albeit without any apparent regulation by a recombination hotspot sequence.

Materials and methods

Protein expression and purification

The nuclease-dead AddAB mutant ($AddA^{D1172A}B^{D961A}$) was purified as described previously (Yeeles *et al*, 2009). Selenomethionine (SeMet) was incorporated into AddAB in place of methionine by expressing the protein in B834 (DE3) cells grown in LeMaster medium containing SeMet. The cells were grown at 37°C to an OD of 0.4, induced with 1 mM IPTG and grown at 25°C for 10 h. Purification of the SeMet protein was similar to the native protein. The DNA substrate (5′-TCTAATGCGAGCACTGCTATTCCCTAGCA GTGCTCGCATTAGATTTTG-3′) used for crystallization was prepared as described previously (Singleton *et al*, 2004).

Crystallization

Protein was concentrated to 10 mg/ml in 10 mM Tris–HCl pH 7.5, 100 mM NaCl and 1 mM dithiothreitol and mixed with the DNA substrate at 1:1.3 molar ratio for crystallization. Crystals of AddAB complexes were obtained by vapour diffusion in hanging drops at 12°C by mixing equal volumes of protein and mother liquor consisting of 15% polyethylene glycol 4000, 0.1 M Tris pH 7.5, 0.8 M sodium formate. Microseeding was used to improve crystal quality. Crystals were cryoprotected by transfer to the reservoir solution supplemented with 30% ethylene glycol, and flash cooled in liquid nitrogen. Crystals of SeMet-substituted AddAB–DNA complex were obtained in a similar way.

Structure determination and refinement

A 2.8-Å diffraction data set was collected from a native crystal at ESRF beamline ID23-1. The data were integrated with MOSFLM (Leslie, 2006), then scaled and merged using SCALA (Collaborative Computational Project, Number 4, 1994). The crystals belonged to space group $P2_1$ with one AddAB/DNA complex in the asymmetric unit. MAD data sets were collected from a selenomethionine-substituted protein crystal at peak, inflection point and remote wavelengths at the Diamond Light Source beamline IO2 to a maximum resolution of 3.2 Å. Diffraction data were integrated and scaled using XDS and XSCALE (Kabsch, 1993). Intensities were converted to structure factors using TRUNCATE (Collaborative Computational Project, Number 4, 1994). Positions of 58 selenium sites were either located using SHELXC and SHELXD (Schneider and Sheldrick, 2002) or manually from an anomalous difference Fourier map, and phases calculated using SHARP (DeLaBarre and Brunger, 2006). The initial phases were improved by density modification using DM (Cowtan and Main, 1998). The structure was built manually. The quality of the initial electron density map was sufficient to locate the various protein domains and the DNA. Homologous domains of RecBCD were placed in the unit cell based on the map to guide model building. Alternating rounds of model building and structure refinement were carried out using Turbo-Frodo (Roussel and Cambillau, 1991) or O (Jones *et al*, 1991), and CNS (Brunger, 2007), respectively. The model obtained was used to determine a structure solution for the 2.8-Å data. The model was rebuilt and refined. Final rounds of model building for both the structures were carried out using COOT (Emsley and Cowtan, 2004) and refined using PHENIX (Adams *et al*, 2010) involving restrained maximum likelihood refinement of atomic position and isotropic B-factors. Water molecules were identified using an Fo–Fc map at a contour level of 3 σ. The structures were validated using MOLPROBITY (Davis *et al*, 2007) as implemented in PHENIX. Crystallographic data statistics are shown in Table I. The 2.8 and 3.2 Å structures have 0.2 and 0.4% residues as outliers in the Ramachandran plots, respectively. The figures shown are all of the 3.2-Å structure with the intact 4Fe–4S cluster, with the exception of those showing the details of the ATPase and nuclease active sites of AddA and AddB (Figure 2B, C, E, and F) and the ssDNA motor site of AddA (Figures 4C and 6A), which are of the higher-resolution structure.

Site-directed mutagenesis and protein purification

Site-directed mutagenesis of the putative Chi-binding locus in the *addB* gene was performed using the QuikChange II XL Site-Directed Mutagenesis kit (Stratagene) and was validated by sequencing (University of Dundee DNA Sequencing Service). Mutant *addB* genes were subcloned from pET28a into the pCOLADuet expression vector (both Novagen), the latter of which already contained the wild-type *addA* gene in MCS2. Wild-type and mutant AddAB complexes were expressed and purified as described previously (Yeeles *et al*, 2009). The purities of the mutant AddAB complexes were comparable to wild type as judged by SDS–PAGE (data not shown) and their concentration was determined by Bradford assay using known concentrations of BSA as standards. The proportion of Fe–S cluster containing AddAB in each preparation is variable and was assessed using native polyacrylamide gels as described previously (Yeeles *et al*, 2009; Supplementary Figure S2).

Table I Crystallographic data and model building statistics

	Native (Fe–S lost)	Peak (Fe–S intact)	Inflection	Remote
(a) Data statistics				
Wavelength (Å)	1.1	0.9795	0.9796	0.9807
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$
a, *b*, *c* (Å)	100.6, 139.7, 103.1	100.2, 139.0, 103.0	100.4, 139.2, 103.1	100.5, 139.3, 103.2
β (deg)	105.3	105.3	105.4	105.4
Resolution (Å)	2.8	3.2	3.5	3.5
Completion (%)	95.7 (95.4)	92.8 (90.0)	95.4 (94.6)	95.3 (93.9)
R_{merge} (%)	6.8 (33.4)	6.0 (39.2)	7.9 (33.6)	9.0 (32.4)
I/σ (I)	7.0 (2.2)	14.0 (2.6)	11.5 (3.1)	11.9 (3.1)
		AddAB/Fe–S intact		AddAB/Fe–S lost
(b) Refinement and model statistics				
Resolution (Å)		30–3.2		30–2.8
R-factor (%)		23.3		23.3
R_{free} (%)		28.4		28.7
r.m.s.d bond length (Å)		0.003		0.002
r.m.s.d. bond angles (deg)		0.711		0.655

Figures in parentheses indicate highest-resolution shell.

Chi-recognition assays

DNA processing assays were performed essentially as described previously (Yeeles and Dillingham, 2007). Briefly, the pADGF6406 plasmid (Chedin *et al*, 2000), which contains a single *B. subtilis* Chi sequence, was linearized and de-phosphorylated using *Cla*I and Antarctic phosphatase (NEB) and then 5′-labelled with [γ32P]ATP (Perkin-Elmer) using T4 polynucleotide kinase (NEB). This DNA substrate (1.6 nM molecules) was pre-incubated with *E. coli* SSB (2 μM) at 37°C for 2 min in a buffer containing 20 mM Tris-acetate (pH 7.5), 1 mM ATP, 2 mM magnesium acetate and 1 mM DTT. Reactions were then initiated by the addition of AddAB (2 nM) and quenched after 4 min by adding an equal volume of 2 × Stop Buffer (1 mg/ml proteinase K, 100 mM EDTA, 10% (w/v) Ficoll 400, 5% (w/v) SDS, 0.125% (w/v) Bromophenol blue, 0.125% (w/v) Xylene Cyanol). The products were run on 1% (w/v) agarose gels at 40 V for 18 h. Gels were dried onto DE81 paper (Whatman) and exposed to a storage phosphor screen that was imaged using a Typhoon 9400 phosphorimager (Molecular Dynamics). TotalLab TL100 (non-linear dynamics) was used to analyse the gel images. The amount of Chi fragment was assessed by applying the rolling ball method to subtract the background non-specific ssDNA product smear from the Chi-specific peak. The amount of Chi fragment produced was normalized to account for the amount of the dsDNA that had been processed (>96% in all cases).

Accession codes

The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession numbers 3U44 and 3U4Q for the 3.2 and 2.8 Å structures, respectively.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We are grateful to Professor Steve Kowalczykowski and Dr Fernando Moreno Herrero for their comments on the manuscript. We thank the ESRF and Diamond synchrotrons for access to beamlines. This work was funded by the Royal Society, the Wellcome Trust and the European Research Council (MSD), by the BBSRC (JTY and NSG), by Cancer Research UK (DBW), and by EMBO (KS and WWK). This work was initiated when KS, WWK, and DBW were located at the CRUK London Research Institute, South Mimms, Herts, UK but completed at the Institute of Cancer of Research (WWK and DBW) and Indian Institute of Science Education & Research (KS).

Author contributions: KS, JTY, and MSD purified and crystallized the AddAB complex; KS, WWK, and DBW solved, built, and refined the AddAB crystal structures; NSG performed the biochemical characterization of the AddAB mutant proteins; MSD and DBW wrote the paper; all authors interpreted the data and commented on the final manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213–221

Aravind L, Makarova KS, Koonin EV (2000) SURVEY AND SUMMARY: holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res* **28:** 3417–3432

Brunger AT (2007) Version 1.2 of the crystallography and NMR system. *Nat Protoc* **2:** 2728–2733

Burgers PM, Stith CM, Yoder BL, Sparks JL (2010) Yeast exonuclease 5 is essential for mitochondrial genome maintenance. *Mol Cell Biol* **30:** 1457–1466

Cejka P, Cannavo E, Polaczek P, Masuda-Sasa T, Pokharel S, Campbell JL, Kowalczykowski SC (2010) DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. *Nature* **467:** 112–116

Chedin F, Ehrlich SD, Kowalczykowski SC (2000) The Bacillus subtilis AddAB helicase/nuclease is regulated by its cognate Chi sequence *in vitro*. *J Mol Biol* **298:** 7–20

Chedin F, Handa N, Dillingham MS, Kowalczykowski SC (2006) The AddAB helicase/nuclease forms a stable complex with its cognate chi sequence during translocation. *J Biol Chem* **281:** 18610–18617

Chepanoske CL, Golinelli MP, Williams SD, David SS (2000) Positively charged residues within the iron-sulfur cluster loop of E. coli MutY participate in damage recognition and removal. *Arch Biochem Biophys* **380:** 11–19

Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50** (Part 5)**:** 760–763

Cowtan K, Main P (1998) Miscellaneous algorithms for density modification. *Acta Crystallogr D Biol Crystallogr* **54:** 487–493

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* **14:** 1188–1190

Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall III WB, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35:** W375–W383

DeLaBarre B, Brunger AT (2006) Considerations for the refinement of low-resolution crystal structures. *Acta Crystallogr D Biol Crystallogr* **62:** 923–932

Dillingham MS, Kowalczykowski SC (2008) RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol Mol Biol Rev* **72:** 642–671, Table of Contents

Dillingham MS, Soultanas P, Wigley DB (1999) Site-directed mutagenesis of motif III in PcrA helicase reveals a role in coupling ATP hydrolysis to strand separation. *Nucleic Acids Res* **27:** 3310–3317

Dillingham MS, Spies M, Kowalczykowski SC (2003) RecBCD enzyme is a bipolar DNA helicase. *Nature* **423:** 893–897

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126–2132

Fan L, Fuss JO, Cheng QJ, Arvai AS, Hammel M, Roberts VA, Cooper PK, Tainer JA (2008) XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell* **133:** 789–800

Farah JA, Smith GR (1997) The RecBCD enzyme initiation complex for DNA unwinding: enzyme positioning and DNA opening. *J Mol Biol* **272:** 699–715

Fromme JC, Verdine GL (2003) Structure of a trapped endonuclease III-DNA covalent intermediate. *EMBO J* **22:** 3461–3471

Guan Y, Manuel RC, Arvai AS, Parikh SS, Mol CD, Miller JH, Lloyd S, Tainer JA (1998) MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nat Struct Biol* **5:** 1058–1064

Halpern D, Chiapello H, Schbath S, Robin S, Hennequet-Antier C, Gruss A, El Karoui M (2007) Identification of DNA motifs implicated in maintenance of bacterial core genomes by predictive modeling. *PLoS Genet* **3:** 1614–1621

Handa N, Morimatsu K, Lovett ST, Kowalczykowski SC (2009) Reconstitution of initial steps of dsDNA break repair by the RecF pathway of E. coli. *Genes Dev* **23:** 1234–1245

Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47** (Part 2)**:** 110–119

Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constraints. *J Appl Cryst* **26:** 795–800

Korolev S, Hsieh J, Gauss GH, Lohman TM, Waksman G (1997) Major domain swiveling revealed by the crystal structures of

complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. *Cell* **90:** 635–647
Kowalczykowski SC (2000) Initiation of genetic recombination and recombination-dependent replication. *Trends Biochem Sci* **25:** 156–165
Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372:** 774–797
Lee JY, Yang W (2006) UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* **127:** 1349–1360
Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62:** 48–57
Lieber MR (2008) The mechanism of human nonhomologous DNA end joining. *J Biol Chem* **283:** 1–5
Liu H, Rudolf J, Johnson KA, McMahon SA, Oke M, Carter L, McRobbie AM, Brown SE, Naismith JH, White MF (2008) Structure of the DNA repair helicase XPD. *Cell* **133:** 801–812
Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* **1:** 7
Mimitou EP, Symington LS (2009) Nucleases and helicases take center stage in homologous recombination. *Trends Biochem Sci* **34:** 264–272
Nimonkar AV, Genschel J, Kinoshita E, Polaczek P, Campbell JL, Wyman C, Modrich P, Kowalczykowski SC (2011) BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev* **25:** 350–362
Niu H, Chung WH, Zhu Z, Kwon Y, Zhao W, Chi P, Prakash R, Seong C, Liu D, Lu L, Ira G, Sung P (2010) Mechanism of the ATP-dependent DNA end-resection machinery from Saccharomyces cerevisiae. *Nature* **467:** 108–111
Papadopoulos JS, Agarwala R (2007) COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* **23:** 1073–1079
Petrek M, Kosinova P, Koca J, Otyepka M (2007) MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels. *Structure* **15:** 1357–1363
Pitcher RS, Brissett NC, Doherty AJ (2007) Nonhomologous end-joining in bacteria: a microbial perspective. *Annu Rev Microbiol* **61:** 259–282
Pugh RA, Honda M, Leesley H, Thomas A, Lin Y, Nilges MJ, Cann IK, Spies M (2008) The iron-containing domain is essential in Rad3 helicases for coupling of ATP hydrolysis to DNA translocation and for targeting the helicase to the single-stranded DNA-double-stranded DNA junction. *J Biol Chem* **283:** 1732–1743
Pugh RA, Wu CG, Spies M (2011) Regulation of translocation polarity by helicase domain 1 in SF2B helicases. *EMBO J* **31:** 503–514
Quiberoni A, Biswas I, El Karoui M, Rezaiki L, Tailliez P, Gruss A (2001) *In vivo* evidence for two active nuclease motifs in the double-strand break repair enzyme RexAB of Lactococcus lactis. *J Bacteriol* **183:** 4071–4078
Roussel A, Cambillau C (1991) *Silicon Graphics Geometry Partners Directory*. p 81 Mountain View, CA: Silicon Graphics
Saikrishnan K, Griffiths SP, Cook N, Court R, Wigley DB (2008) DNA binding to RecD: role of the 1B domain in SF1B helicase activity. *EMBO J* **27:** 2222–2229
Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* **58:** 1772–1779
Shuman S, Glickman MS (2007) Bacterial DNA repair by non-homologous end joining. *Nat Rev Microbiol* **5:** 852–861
Singleton MR, Dillingham MS, Gaudier M, Kowalczykowski SC, Wigley DB (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* **432:** 187–193
Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* **76:** 23–50
Sinha KM, Unciuleac MC, Glickman MS, Shuman S (2009) AdnAB: a new DSB-resecting motor-nuclease from mycobacteria. *Genes Dev* **23:** 1423–1437
Soultanas P, Dillingham MS, Velankar SS, Wigley DB (1999) DNA binding mediates conformational changes and metal ion coordination in the active site of PcrA helicase. *J Mol Biol* **290:** 137–148
Subramanya HS, Bird LE, Brannigan JA, Wigley DB (1996) Crystal structure of a DExx box DNA helicase. *Nature* **384:** 379–383
Taylor AF, Smith GR (2003) RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature* **423:** 889–893
Thayer MM, Ahern H, Xing D, Cunningham RP, Tainer JA (1995) Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO J* **14:** 4108–4120
Unciuleac MC, Shuman S (2010) Characterization of the mycobacterial AdnAB DNA motor provides insights into the evolution of bacterial motor-nuclease machines. *J Biol Chem* **285:** 2632–2641
Velankar SS, Soultanas P, Dillingham MS, Subramanya HS, Wigley DB (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97:** 75–84
Wang J, Chen R, Julin DA (2000) A single nuclease active site of the Escherichia coli RecBCD enzyme catalyzes single-stranded DNA degradation in both directions. *J Biol Chem* **275:** 507–513
White MF (2009) Structure, function and evolution of the XPD family of iron-sulfur-containing 5′→3′ DNA helicases. *Biochem Soc Trans* **37:** 547–551
White MF, Dillingham MS (2011) Iron-sulphur clusters in nucleic acid processing enzymes. *Curr Opin Struct Biol* (advance online publication, 12 December 2011; doi:10.1016/j.sbi.2011.11.004)
Wolski SC, Kuper J, Hanzelmann P, Truglio JJ, Croteau DL, Van Houten B, Kisker C (2008) Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD. *PLoS Biol* **6:** e149
Wong CJ, Lucius AL, Lohman TM (2005) Energetics of DNA end binding by E.coli RecBC and RecBCD helicases indicate loop formation in the 3′-single-stranded DNA tail. *J Mol Biol* **352:** 765–782
Wong CJ, Rice RL, Baker NA, Ju T, Lohman TM (2006) Probing 3′-ssDNA loop formation in E. coli RecBCD/RecBC-DNA complexes using non-natural DNA: a model for 'Chi' recognition complexes. *J Mol Biol* **362:** 26–43
Wyman C, Kanaar R (2006) DNA double-strand break repair: all's well that ends well. *Annu Rev Genet* **40:** 363–383
Yeeles JT, Cammack R, Dillingham MS (2009) An iron-sulfur cluster is essential for the binding of broken DNA by AddAB-type helicase-nucleases. *J Biol Chem* **284:** 7746–7755
Yeeles JT, Dillingham MS (2007) A dual-nuclease mechanism for DNA break processing by AddAB-type helicase-nucleases. *J Mol Biol* **371:** 66–78
Yeeles JT, Dillingham MS (2010) The processing of double-stranded DNA breaks for recombinational repair by helicase-nuclease complexes. *DNA Repair (Amst)* **9:** 276–285
Yeeles JT, Gwynn EJ, Webb MR, Dillingham MS (2011a) The AddAB helicase-nuclease catalyses rapid and processive DNA unwinding using a single Superfamily 1A motor domain. *Nucleic Acids Res* **39:** 2271–2285
Yeeles JT, Van Aelst K, Dillingham MS, Moreno-Herrero F (2011b) Recombination hotspots and single-stranded DNA binding proteins couple DNA translocation to DNA unwinding by the AddAB helicase-nuclease. *Mol Cell* **42:** 806–816

