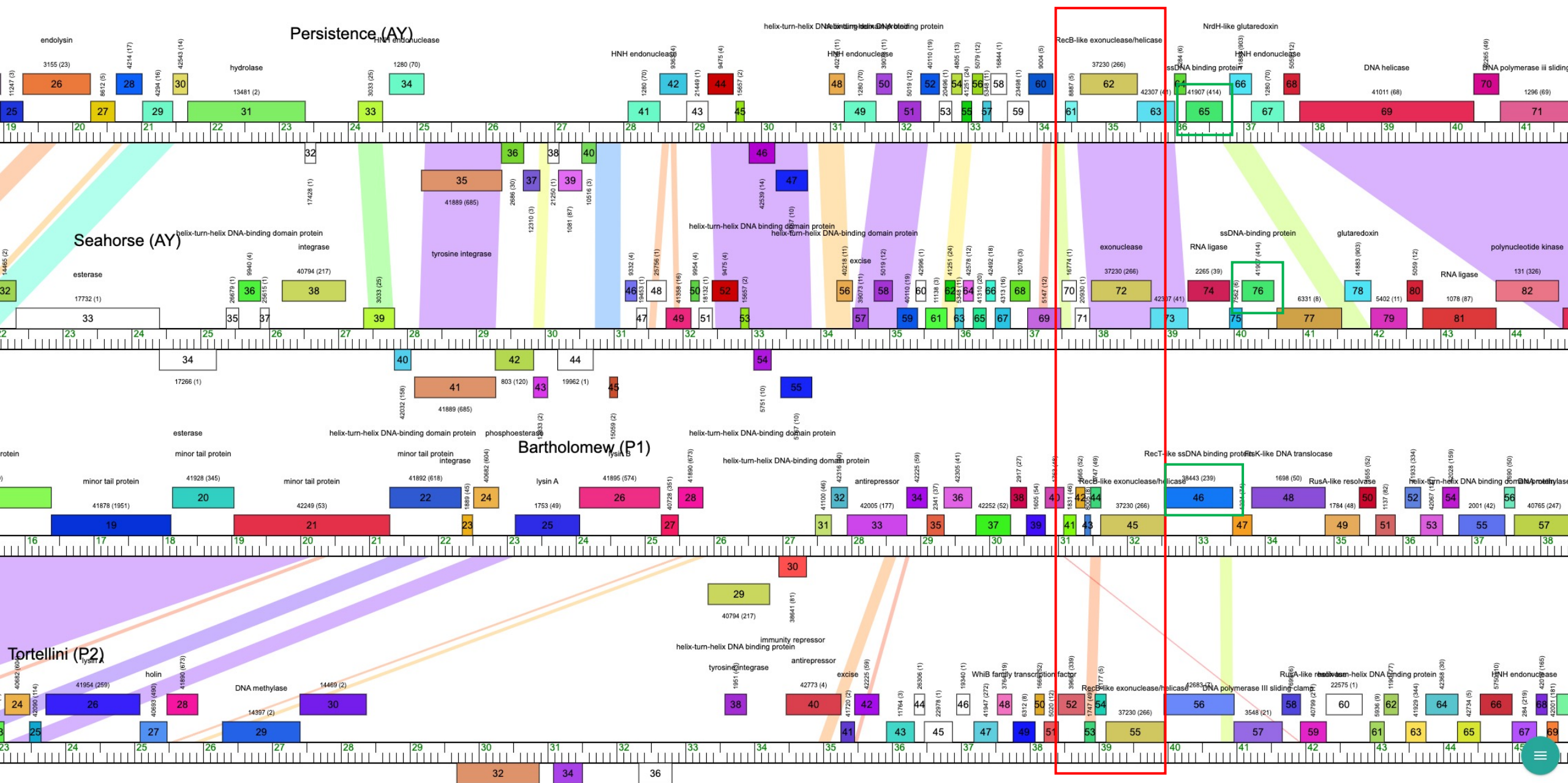




Langerak_46 RecE exonuclease (endonuclease 8) match

Template alignment | Template 3D Structure | PDBe

1. **3H4R_A** Exodeoxyribonuclease 8; Exonuclease, Recombination, Hydrolase, Nuclease; 2.8A {*Escherichia coli*}

Probability: 99.91%, E-value: 3.5e-21, Score: 163.81, Aligned cols: 250, Identities: 26%, Similarity: 0.38, Template Neff: 12

Q ss_pred		CCCCHHHHCcCCCCCHHHHHHHHCCCCCHHHHHHHHHc--CCCCCCHHHHHHHHHHHHHHHHCC--CcC E EeChHhcCcccC	
Q Langerak_46	14	FVDEDDYHADRGSLVS GAKLLPPSCPAKFRWEQDN--GRKP KKA WDFGHVAHKLVLGK---GAEF EVLDPSV HGLKAN	88 (309)
Q Consensus	14	++++. +++.+ + + .+. . ++ ++++. ++.+.++. +++ .+++. .+.+.+. . .	88 (309)
T Consensus	1	the the the the the the the the ~L S ~L the the the the Cp the the H the L the the the the the the the the the the the the	76 (265)
T 3H4R_A	1	GS HME N Y HAGP-GISKS QLD IA-- DTPALYLWRKNAPVDTTKT KL DGTA HF CRVLEEF S NR FIVAF FNRR T-NA	76 (265)
T ss_dssp		-----C-CBCHHHHHHHHH--HCHHHHHHHHTTSCCCCCCSCCTHHHHHHHHHHSSHHHHHTCC-----	
T ss_pred		CCcHHHccCC-CCCCHHHHHHhh--CCHHHHHHHhhCCCCCCCCcc <hhhh< h="">hhhhHHHHChHhhccEeCCccCCCC-cC</hhhh<>	

[illegible][illegible][illegible]

Langerak_46 Cas4 exonuclease match

Template alignment | Template 3D Structure | PDBe

27. **4R5Q_A** CRISPR-associated exonuclease, Cas4 family; MCSG, STRUCTURAL GENOMICS, PSI-BIOLOGY, EXONUCLEASE, HYDROLASE, Midwest Center for Structural Genomics; 2.65Å {Pyrobaculum calidifontis JCM 11548}

Probability: 99.04%, E-value: 2.3e-8, Score: 83.83, Aligned cols: 146, Identities: 15%, Similarity: 0.024, Template Neff: 10.7

Q ss_pred		CCCCHHHHHHHCCCC <hhhhhhhh>--hccc---<td></td></hhhhhhhh>	
Q Langerak_46	25	GSLSVSGAKLLPPSCPAKFRWEQ--DNGRK---PKAWDFGHVAHLVLGKGAEFEVLDPSVHGKLKANGEPSEKPTATA	99 (309)
Q Consensus	25	--ls~s~l~n~CP~~~~~G~~H~l~n~ ..+ +.+.+. . ++ +.+. .+. . . .+ +.+.+.+	99 (309)
T Consensus	54	--ls~s~l~n~-----Cpr~~~~~G~~H~l~n~-----	102 (216)
T 4RSQ_A	54	RRVSPSINDFE--YCPRLWQHKLGLKLLSEKSVVISIRGTLHERVER-----	102 (216)
T ss_dssp		TBCCCCHHHHHH--HCHHHHHHHHHTEEEEETTTTHHHHHHHHHHHHHHHH	
T ss_pred		ceehhhhhhhh--hcchhhhhhhhhcccccccccchhhhhhhhhhhhhhhhh	

Q ss_pred		HHHHHHHHHHCCCCcCHHHHHHHHHHHHHHHhCcchHHHhhCCCCcceeEEEEEEeCCCCeEEEEccccCC--C--c		
Q Langerak_46	100	MWRNAEAEARKGKGVPTIADFTRAYDMAERVVRQHTAGPIFADAEVQAEVALYHTDPTGVVLRGRIDWLTN--D--	174	(309)
Q Consensus	100	E.....~G~T~D~T~--g--i	174	(309)
T Consensus	103	-----G--D~i~n~g~n~i	137	(216)
T 4R5Q_A	103	-----LLSQYE--NVVAEYKVEIG--DLVGVDLVIKRGGEYTP	137	(216)
T ss_dssp		-----HHTTST--TEEEEEEEET-----TEEEEEEEEEETTEEEE		
T ss_pred		-----HHHHhh--cceeEEEEEEc--eeEEEEEEEEcCEEEEE		

[illegible]

Q ss_pred		HHHHHHHCCCCCCCcEeCCCHH	
Q Langerak_46	255	LYAE C METGRWPGYSDDVVTLSPAW	280 (309)
Q Consensus	255	P~	280 (309)
T Consensus	188	.+.+.+.+.+ . +.	
T 4R5Q_A	188	~~~~~C~~C~~	209 (216)
T ss_dssp		KIREILSAREAPP---AKCN S CIFK	209 (216)
T ss_pred		HHHHHHHCSSCCC---CCSSCTTT	
		HHHHHHHcCCCC--ccCCCCCCCh	

Langerak_46 RecB-like helicase/exonuclease match

Template alignment | Template 3D Structure | PDB

16. **6PPU_B** UvrD/REP helicase; DNA, DNA BINDING PROTEIN, DNA BINDING PROTEIN-DNA complex; 3.5A {Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155)}

Probability: 99.38%, E-value: 2e-11, Score: 125.61, Aligned cols: 195, Identities: 18%, Similarity: 0.037, Template Neff: 10.8

[illegible]

Q ss_pred		HHHHHHHCCCCcCHHHHHHHHHHHHHHHHHHCchHHHHhhCCCCceEEEEEEeCCCCeEEEEcccccC---C---cee		
Q Langerak_46	104	AEEAARKGKVPITIAIDFTRAYDMAERVQRHTAGPTIFADAEQAQEAVALYHTDPETGVRLRGRIDWLTN---D--IDDY	177	(309)
Q Consensus	104	+...+..-.....+...+.++..+.++..+ + . +	177	(309)
T Consensus	948	-----GAVSDSGRAVEESLAEIQDAFVKSPAARTPVEEVPDMV---LGETVVRGRIDAVFAEPDGTMMVLWD	1015	(1095)
T 6PPU_B	948	-----CCCCcchhhHHHHHHHHHHHHHHhhChhCCCEEEEEEEE-eCEEEEEEEEEECCCCCEEEEEEE	1015	(1095)
T ss_pred				

Q ss_pred		ecCCCCCHHHHHHHHcChHHHHHHHHHHCCCCCC--cEEEEECCECcEEEEEcCHHHHHHHHHHHHHHH		
Q Langerak_46	178	KTSTTANPAELKVKFYKLGYFMAAWYDLLVALGAENP--RRFTIQEKEPPYVVTPIEYDDDAIAEGRRNRQAIRL	255	(309)
Q Consensus	178	KT.....+... . +++++.++....+.++++..+.....+.+.....+.+	255	(309)
T Consensus	1016	Ktg.....QLY.....	1086	(1095)
T 6PUU_B	1016	KTDGDPPEAKEH-----AAVLAVYLAWAAMRGCPPEVSRAFHVYRSG---QTVIPETLP-GAEELVKLLAAPTE	1086	(1095)
T ss_pred		CCCCCCCCCHHHHH-----HHHHHHHHHHHHHHCCChhhHEEEEEEcCC--cEEcCCCC-CHHHHHHHHHcchh		

Q ss_pred		HHHHHHH	
Q Langerak_46	256	YAECMET	262 (309)
Q Consensus	256	~~~~~	262 (309)
		+ . + . . .	
T Consensus	1087	~~~~~	1093 (1095)
T 6PPU_B	1087	TAEEADR	1093 (1095)
T ss_pred		HHHHhhh	

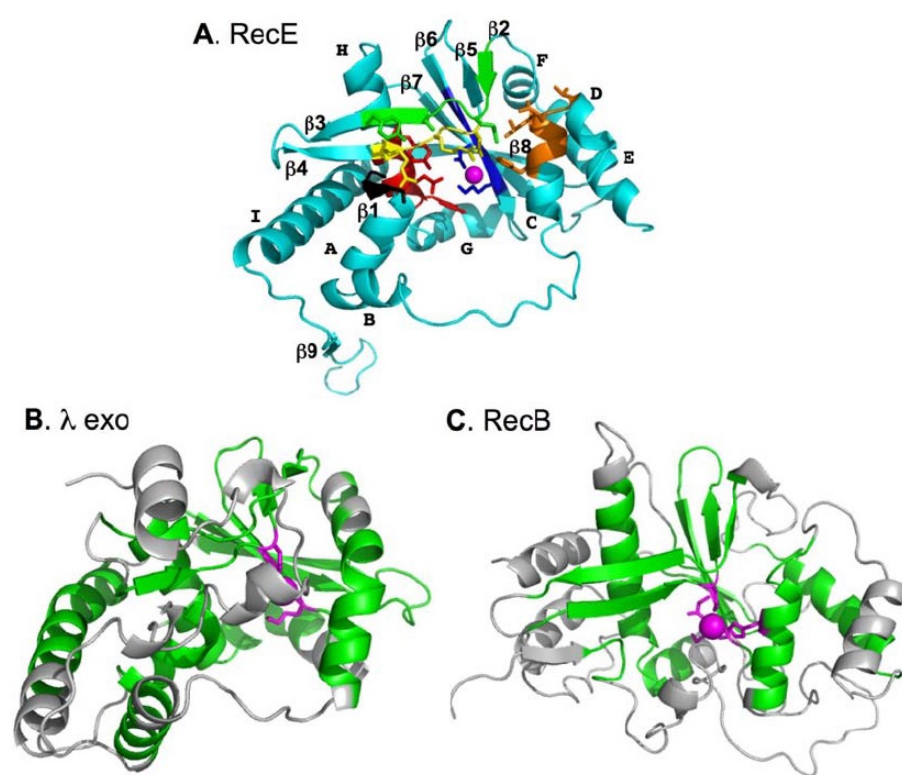


Figure S2. Structural Alignment of RecE, λ Cxonuclease, and RecB

(A) The RecE monomer, colored as in Figure 3.

(B) The λ exonuclease monomer, viewed in the same orientation as RecE, after structural superposition. Regions of λ exonuclease that align to within 4 Å of equivalent Cα atoms of RecE are shaded in green. Active site residues are shown in magenta.

(C) The RecB monomer, viewed in the same orientation as RecE, after structural superposition. Coloring is the same as for λ exonuclease in panel B.

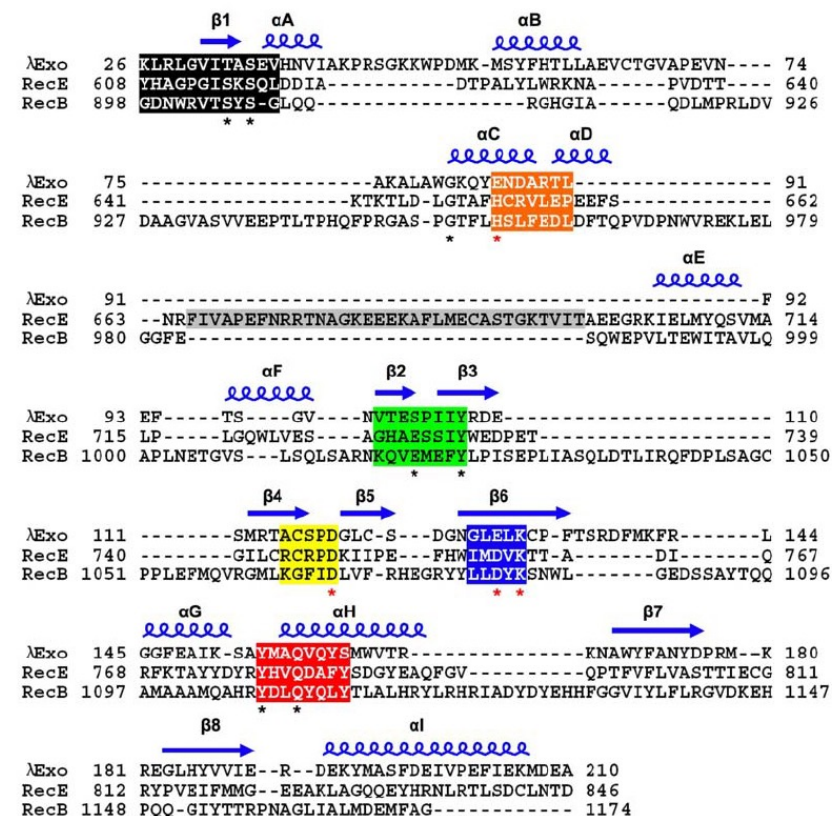
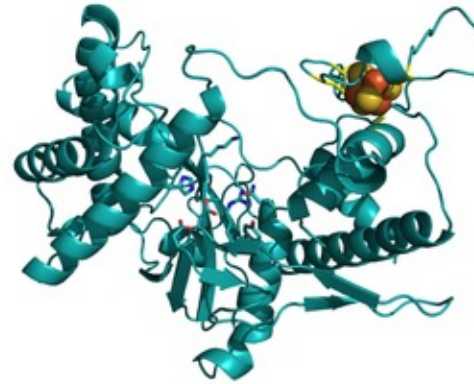
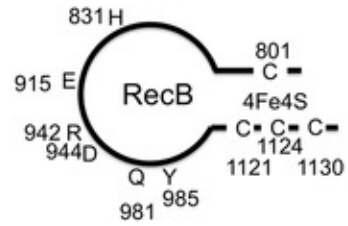


Figure S3. Structure-Based Sequence Alignment of RecE, RecB, and λ Exonuclease

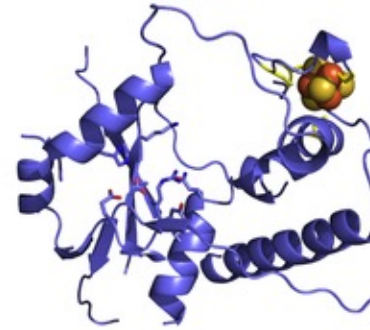
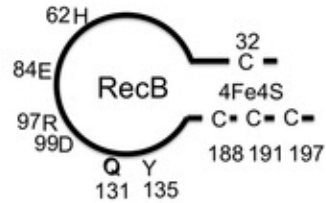
The structures of λ exonuclease and RecB were superimposed with the structure of RecE and the resulting sequence alignments are shown above and below the sequence of RecE, respectively. The conserved sequence motifs of RecE are colored onto the sequences, as in Figure 3 (Motif I, green, Motif II, yellow; Motif III, blue; Motif IV, red; Motif V, orange; Motif VI, black). The positions of secondary structures in RecE are indicated above the sequence alignment. Residues that are conserved in the three proteins are indicated below the sequence. Residues that coordinate the Mg^{2+} ion are shown with these symbols in red. The 34-residue insertion in RecE that forms an extended loop that is poorly resolved in the structure is shaded in gray.

Cas4 has a RecB domain with a C-terminal Fe-S cluster

AddB (3U44)



Cas4 (Sso0001 model)



DUF3799 (3L0A)

