MIC/BIO/BCH522

Spring 2006

Homologous genetic recombination
and
DNA DSB repair
Homologous genetic recombination

- Exchange of genetic material between 2 DNA molecules
- Essential to all organisms
- Generation of genetic diversity
- Maintenance of genomic integrity, e.g., DSB repair
- Proper segregation of chromosomes

1. Recombination
2. DNA strand exchange
3. Single strand binding proteins
4. RecA
5. DNA helicases
6. RecBCD
7. RuvABC
8. RecG
5' 3' 3' 5'

**Initiation**

1. Helicase and/or nuclease

2

**Homologous Pairing & DNA Strand Exchange**

3. RecA-like protein
   Accessory proteins

4

**DNA Heteroduplex Extension**

5. Branch migration proteins

6. Resolvase

Resolved

Spl iced

Patched
**E. coli**

- **Initiation**
  - RecB/C/SSB
- **Homologous Pairing & DNA Strand Exchange**
  - RecA DNA helicase
  - RecA DNA polymerase I
  - DNA topoisomerases
- **Resolution**
  - RuvC DNA ligase
  - DNA ligase

**Bacteriophage T4**

- **Initiation**
  - G46P
  - G47P
  - G32P
- **Homologous Pairing & DNA Strand Exchange**
  - UvsX
  - UvsW
  - G41P
  - G59P
  - G32P
- **Resolution**
  - Resolvase
  - DNA ligase

**Eukaryotes**

- **Initiation**
  - Endonuclease (Rad50 and Spo11)
- **Homologous Pairing & DNA Strand Exchange**
  - Rad51, Rad52, Rad54, Rad55, Rad57, RPA, and DNA polymerase
- **Resolution**
  - Resolvase
  - DNA ligase

**Spliced**

**Patched**
## Proteins involved in recombination

<table>
<thead>
<tr>
<th>General Function</th>
<th>Organism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>E. coli</strong></td>
</tr>
<tr>
<td><strong>Initiation</strong></td>
<td>RecBCD, RecQ</td>
</tr>
<tr>
<td>DNA strand exchange</td>
<td>RecA</td>
</tr>
<tr>
<td>ssDNA-binding</td>
<td>SSB</td>
</tr>
<tr>
<td>Accessory</td>
<td>RecF, RecO, RecR</td>
</tr>
<tr>
<td>Branch migration</td>
<td>RuvA, RuvB</td>
</tr>
<tr>
<td></td>
<td>RecG</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Holliday Junction cleavage</td>
<td>RuvC</td>
</tr>
<tr>
<td>Other proteins</td>
<td>DNA topoisomerase I and II</td>
</tr>
<tr>
<td></td>
<td>DNA ligase</td>
</tr>
<tr>
<td></td>
<td>DNA polymerase I</td>
</tr>
</tbody>
</table>
DNA strand exchange is used to study recombination *in vitro*.

**DNA substrates**

- **D-loop assay**: 
  - Reactant 1
  - Reactant 2
  - Result

- **3-strand reaction**: 
  - Reactant 1
  - Reactant 2
  - Result

- **Gapped duplex**: 
  - Reactant 1
  - Reactant 2
  - Result
DNA strand exchange can be visualized in agarose gels

![Diagram showing DNA strand exchange](image)

- **Homologous reaction**
- **Heterologous reaction**

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>M M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Higher order complexes
- Intermediates
- Product
- Linear dsDNA
- Displaced ssDNA

---

**DNA substrates**

**Intermediates**

**DNA substrates**
Stages of DNA strand exchange: Presynapsis (does not require ATP hydrolysis, although hydrolysis occurs)
Stages of DNA strand exchange:

2

2 and 3-synapsis (does not require ATP hydrolysis, although hydrolysis occurs)

3

DNA heteroduplex extension (requires ATP hydrolysis)
ssDNA binding proteins

Three of these to consider:
- Gene 32 protein or gp32 (T4); functions as a monomer
- Single-stranded DNA binding protein or SSB (E. coli), homotetramer (monomer is 18.5 kDa)
- Replication protein A or RPA (Eukaryotes) heterotrimer

- All contain the **OB-fold** (oligonucleotide-oligosaccharide fold), characteristic of SSB’s
- Generally, bind to ssDNA and remove secondary structure so that additional proteins can bind or protect the DNA from degradation by nucleases
- SSB is known as a helix-destabilizing protein and it binds preferentially and cooperatively to ssDNA
- Because of its preferential binding to ssDNA, it destabilizes dsDNA
- When SSB binds to ssDNA, the contour length of the DNA decreases and this is consistent with DNA wrapping around the protein
- The OB fold is a conserved ssDNA-binding motif observed in many ssDNA-binding proteins.
- A ssDNA-binding protein can have one OB fold (T4 SSB) or as many as four in the *E. coli* SSB tetramer.
- There are also examples of ssDNA-binding proteins that do not have any OB folds

A. T4 SSB; gp32
B. RPA;
C. human mitochondrial SSB;
D. (telomere end-binding protein.)
Surfaces involved in binding of ssDNA to SSB core

The molecular surface of the SSB tetramer calculated and displayed using GRASP. The surface is colored deep blue in the most positive regions and deep red in the most negative, with linear interpolation for values in between. Residues known to be involved in binding are shown.

Ribbon diagram of SSB (blue) with ssDNA (orange) wrapping around it. The key tryptophan residue side chains are highlighted in green. These side chains stack with bases.
Side-chains contacts between RPA and ssDNA.

A. Diagram showing the contacts of RPA with dC8. The bases are numbered 1-8 starting from the 5’ end, and the numbers have been placed in the corresponding bases. Hydrogen bonds between the amino-acid side chains and the DNA are marked by dashed lines. The aromatic residues that stack with the bases are marked in bold.

B. The aromatics (green) stacking with bases

C. Side chains forming hydrogen bonds (fuchsia)
Interaction of SSB with ssDNA

Electron micrograph of SSB bound to ssDNA

Crystal structure of SSB bound to ssDNA
Gp32 and RPA with ssDNA

SSB protein  Gene 32 protein  RPA
DNA strand exchange proteins

- *E. coli* RecA is the paradigm, UvsX in T4 and Rad51 in euks (also have meiosis specific Dmc1)
- RecA is a 38,742 Dalton, DNA-dependent ATPase
- Mutations in *recA* are pleiotropic (affect recombination, DNA repair, SOS mutagenesis, cell division, chromosome segregation)

**The details of the structure of the RecA-DNA Filament**

B. The filament with ATP
C. Cross section
D. Crystal structure of the filament in cross section
The RecA nucleoprotein filament that assembles on ssDNA is needed for recombination, SOS function and mutagenesis.
SOS genes in the presence of DNA damage
The nucleoprotein filament is the key structure in recombination

- The key structure in RecA-mediated reactions is the nucleoprotein filament that consists of RecA, ssDNA and ATP
- The formation of this structure is required for DNA repair, mutagenesis and recombination
- RecA binds ssDNA in the absence of ATP but the resulting filament is inactive for all RecA-promoted activities
- Binding of ATP to the RecA-ssDNA complex converts RecA into the active form
  - In the inactive form, RecA is in the low affinity ssDNA binding state
  - In the active form, RecA is in the high affinity ssDNA binding state, DNA is extended 1.5-fold
RecA protein is a complex allosteric enzyme

- RecA protein exists at about 1,000 molecules per cell
- Following DNA damage (e.g., MMC or UV irradiation) the level increases to >10,000 molecules per cell
- The basal level of RecA protein is sufficient to promote recombination
- For either the recombination or repair role of RecA protein, you need to keep RecA in an inactive state – keep it away from ssDNA!!
- RecA requires ATP and ssDNA for full activity; [ATP] in *E. coli* ~ 3mM, therefore [ssDNA] is rate limiting
- ATP is the allosteric effector molecule and ssDNA is the substrate
- ssDNA is produced in one of two ways:
  - By DNA damage followed by DNA replication
  - By the concerted action of DNA helicase/nuclease

The levels of *recA* as determined using a *recA-lacZ* fusion

![Graph](image)

- Induced levels (SOS induction and DNA repair)
- Basal levels (genetic recombination)
How is ssDNA produced for RecA?

- Method 1: By DNA replication following DNA damage
How is ssDNA produced for RecA?

- Method 2: By the combined action of DNA helicase/nuclease, e.g. RecBCD enzyme
How does ATP function as an allosteric effector for RecA protein?

- The key structure in RecA-mediated reactions is the nucleoprotein filament that consists of RecA, ssDNA and ATP.
- The formation of this structure is required for DNA repair, mutagenesis and recombination.
- RecA binds ssDNA in the absence of ATP but the resulting filament is inactive for all RecA-promoted activities.
- Binding of ATP to the RecA-ssDNA complex converts RecA into the active form:
  - In the inactive from RecA is in the low affinity ssDNA binding state.
  - In the active from, RecA is in the high affinity ssDNA binding state.
- RecA filaments rapidly and completely dissociate when the ratio of ADP/ATP reaches 0.4.

AN electron micrograph of a RecA nucleoprotein filament on ssDNA.
Extension of RecA filaments by ATP binding

The details of the structure of the RecA-DNA Filament
B. The filament with ATP
C. Cross section
D. Crystal structure of the filament in cross section

Three-Dimensional Reconstructions of RecA-DNA Filaments
The state of the RecA filament has been labeled below each surface. The pitch of filament 1 is 82 Å and for the others it is 91 Å. The nucleotide binding core of RecA is indicated with a black arrow, while the C-terminal domain is marked with an orange arrow.

Coordination of nucleoprotein filament activity by ATP binding
RecA is a potent DNA-dependent ATPase. So is ATP hydrolysis needed for DNA strand exchange?

- ATP binding converts RecA from the low affinity DNA binding state to the high affinity DNA binding state
- under typical conditions, ATP hydrolysis coincides with the pairing and exchange of DNA strands
- neither the hydrolysis of ATP nor the presence of a high energy phosphate bond is necessary for DNA strand exchange to occur
- ATP hydrolysis is **NOT** required for the exchange of DNA strands which is isoenergetic
- it is required at phases of DNA strand exchange that require the dissociation of RecA, because dissociation is induced by ADP, the product of ATP hydrolysis
- DNA heteroduplex extension phase of DNA strand exchange requires ATP hydrolysis, which renders this final phase of the strand change reaction unidirectional, with the branch moving in a 5’ to 3’-direction relative to the incoming single-stranded DNA (ssDNA)
- required for bypassing short heterologous sequences or other structural barriers in either of the DNA substrates
The RecA protein filament undergoes continuous cycles of polar assembly and disassembly that accompany ATP hydrolysis.

Figure 3  ATP hydrolytic cycle for RecA protein association with ssDNA. Association and dissociation are depicted as polar (5′ to 3′), resulting in net translocation of the protein filament. The inner cycle at step (C) represents the processive ATP hydrolysis pathway that occurs at high ATP concentrations; under these conditions, net polymerization ensues.
This leads to a 2-state model for RecA protein function with the hydrolysis of ATP being used as a switch.

- Low affinity DNA binding state
  - Low affinity for ssDNA
  - Low STMP
  - Low RFI
  - RecA can be displaced by SSB
  - No joint molecules
  - No coproteolytic activity
  - Filament has a compact structure induced by:
    - ADP (or NDP)
    - TTP
    - GTP

- High affinity DNA binding state
  - High affinity for DNA
  - High STMP
  - High RFI
  - RecA displaces SSB
  - Forms Joint Molecules
  - Coproteolytic activity
  - Filament has an extended structure induced by:
    - ATP
    - dATP
    - ATP-γ-S
    - ADP-ALF₄⁻
ATP concentration-dependent changes in RecA protein behaviour

Figure 5  ATP concentration–dependent changes of RecA protein–dependent behavior. Open circles represent nucleotide-free RecA protein; circles labeled T represent the ATP–RecA protein complex; and shaded circles represent different conformations of the ATP–RecA protein complex. Each step represents apparent equilibrium or steady-state properties. Apparent dissociation constants ($K_d$) are given for events that display a hyperbolic dependence on ATP concentration. For properties that display a sigmoid dependence on ATP concentration, the ATP concentration at the midpoint of the transition ($S_{0.5}$) is given. Though the numerical values are representative, they depend on reaction conditions, DNA, and nucleotide cofactor.
All three nucleoprotein filaments are similar

<table>
<thead>
<tr>
<th>Activity</th>
<th>RecA protein</th>
<th>UvsX protein</th>
<th>Rad51 protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{\text{cat}}$ (min$^{-1}$) on ssDNA</td>
<td>18 - 30</td>
<td>145 - 240</td>
<td>0.6 - 0.7</td>
</tr>
<tr>
<td>$k_{\text{cat}}$ (min$^{-1}$) on dsDNA</td>
<td>19 - 22</td>
<td>5</td>
<td>0.05 - 0.1</td>
</tr>
<tr>
<td>Length dependent</td>
<td>Yes</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Compositional preference</td>
<td>Yes</td>
<td>-</td>
<td>Yes</td>
</tr>
<tr>
<td>Cooperative</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Reaction Products</td>
<td>ADP+$P_i$</td>
<td>ADP+$P_i$, AMP+$PP_i$</td>
<td>ADP+$P_i$</td>
</tr>
<tr>
<td>Stimulated by</td>
<td>SSB</td>
<td>UvsY; gp32</td>
<td>RPA; Rad52</td>
</tr>
<tr>
<td>Amino Acid</td>
<td>Beta sheets</td>
<td>Alpha helices</td>
<td>Conservation Index</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>--------------</td>
<td>--------------------</td>
</tr>
<tr>
<td></td>
<td>N-terminus</td>
<td>Core domain</td>
<td>C-terminus</td>
</tr>
<tr>
<td>1</td>
<td>36-52</td>
<td>B1</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>54-75</td>
<td>B2</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>88-101</td>
<td>C1</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>104-110</td>
<td>C2</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>124-135</td>
<td>D</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>140-149</td>
<td>E1</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>160-195</td>
<td>E2</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>205-227</td>
<td>F</td>
<td>0.62</td>
</tr>
<tr>
<td>352</td>
<td>10</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

### Conservation Index:

<table>
<thead>
<tr>
<th></th>
<th>RDR</th>
<th>Eubacterial</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-terminus</td>
<td>0.47</td>
<td>0.66</td>
<td>0.43</td>
</tr>
<tr>
<td>Core domain</td>
<td>0.69</td>
<td>0.84</td>
<td>0.53</td>
</tr>
<tr>
<td>C-terminus</td>
<td>0.74</td>
<td>0.70</td>
<td>0.49</td>
</tr>
</tbody>
</table>

### Structures:

- **Conserved Domain:**
  - B1, B2, C1, C2, D, E1, E2, F

- **Amino Acid Index:**
  - 0-352

- **Beta Sheets:**
  - 0-6

- **Alpha Helices:**
  - A, B, C, D, E, F, G, H, I, J

- **Conserved Domain Diagram:**
  - N, B1, B2, C, C1, C2, D, E1, E2, F, a, b, c, d, e, f, g, h, i, j
What are DNA helicases?

- Specialized, essential group of linear motor proteins
- Translocate along a linear lattice - DNA
- Require an energy source - ATP
- Role - separate dsDNA into its components single strand of DNA

DNA replication
DNA recombination
DNA repair
Helicases are organized into families

Superfamily 1
-PcrA, Dda and RecB

Superfamily 2
-RecG and EcoR124I

“DNA helicase” motifs

I  Ia  II  III  Y  “IV”  V  VI

“engine”

“business end”
Common Features of SF1 & SF2

• PcrA and RecG are both monomeric inchworms

• Helicase activity results from the coupling of two separate activities:
  
  (i) DNA destabilisation which is coupled to,

  (ii) an ATP-dependent DNA translocating motor

• Helicases have a modular structure - translocase domains are combined with a variety of domains to recognise & split different DNA (or RNA) substrates
Modular Structure of Helicases

- RecG (Singleton et al., 2001)
- PcrA (Subramanya et al., 1996)
- UvrB (Theis et al., 1999)
- NS3 (Yao et al., 1997)
- Eif4a (Story et al., 2001)
Helicases have distinct oligomeric structures

- RecBCD heterotrimer
- RecB SF1
- RecD SF1
- RecC defunct helciase
- RecG monomer
- RuvB – two diametrically opposed homohexamers
Activities and organization of hexameric helicases

Activities of a hexameric helicase. The pyramid shows the hierarchy and the interdependence of the various activities of the hexameric helicases.

DNA exclusion model. A model of T7 gp4 helicase interaction with the unwinding junction. A fork DNA, with 30 bp of duplex DNA and 25 bases of ssDNA tails, is docked into the EM image of the T7 gp4 hexamer. The 5' tail is bound in the central channel and the 3' tail is excluded.
RecBCD enzyme is the recombination initiator

- Exonuclease V
- Responsible for initiating recombination in *E. coli*
- Responsible for destroying invading DNA in *E. coli*
- ~400 kDa holoenzyme consisting of 3 subunits:
  - RecB – 3’ to 5’ DNA helicase, ATPase, nuclease
  - RecC - chi recognition, dead (defunct) helicase
  - RecD – 5’ to 3’ DNA helicase, ATPase, processivity factor
- It is a combination nuclease - DNA helicase
- Activities are regulated by Chi
  - Effects of chi: modify nuclease activity, stimulate RecA loading, alter the enzymes translocation speed
Chi (χ) sequences in *E. coli*

- Originally discovered as a hotspot in phage lambda
- Chi = crossover hotspot instigator,
- **enhances recombination frequency**

- An 8 nucleotide sequence:

  5'- GCTGGTGG –3'

  - Controls the nuclease activity and recombinogenic activity of RecBCD
  - RecBCD must approach from the 3-side

- > 1000 chi sequences in *E. coli* genome (one every 5 kB)

- Recombination enhancers also exist in other organisms
A. \(\chi^0\) DNA

- 5' * 
- 3' 

- Full length ssDNA

B. \(\chi^+\) DNA

- 5' *  
- 3' 

- Top-strand, downstream \(\chi\)-specific fragment

- Bottom-strand, upstream \(\chi\)-specific fragment
What are the strand requirements for $\chi$-recognition?
The unwound, top, single-strand of DNA defines $\chi$

<table>
<thead>
<tr>
<th>$\chi$ is the top strand only</th>
<th>Recognition occurs after unwinding</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^0$ $\chi^+$</td>
<td>$\chi^+$ $\chi^+$ $\chi^+$</td>
</tr>
<tr>
<td>$\chi^0$ $\chi^+$</td>
<td>M-8  M-22  M-8  M-22</td>
</tr>
<tr>
<td>0  30  300</td>
<td>0  30  300 0  30  300</td>
</tr>
<tr>
<td>Top strand</td>
<td>0 300 0 300 0 300</td>
</tr>
<tr>
<td>Bottom strand</td>
<td>M-8 M-22 M-8 M-22</td>
</tr>
<tr>
<td>Time (sec)</td>
<td>0 300 0 300 0 300</td>
</tr>
</tbody>
</table>

$\chi$ is the top strand only.
Model for $\chi$-recognition

\[
\begin{align*}
 &5' \quad \chi \quad 3' \\
 &3' \quad 3' \quad 5'
\end{align*}
\]
Is the bottom strand necessary for $\chi$-recognition?

\[ \begin{array}{ccc}
\chi^+ & \chi^+ & \chi^o \\
0 & 1 & 5 \\
0 & 1 & 5 \\
0 & 1 & 5 \\
\end{array} \]

Substrates
Recombination initiation in *E. coli*
RecA promotes formation of the joint molecule but needs RecBC to load onto dsDNA. In the presence of chi, RecBCD cuts dsDNA.
Only the top strand is necessary for recombination

The coupled D-loop assay

SSB, RecA, RecBCD

Substrate construction

NruI SphI

Assay

555

Exo III

Coupled D-loop assay

Time (min) 0 1 2 5 10 15 20

D-loops

Full length ssDNA

CSF

Res. Enz. Fragment

Assay
How do you study DNA unwinding by a DNA helicase?

Quenching of the intrinsic fluorescence of SSB

Fluorescent dye displacement
Visualization of DNA helicase action on individual DNA molecules. a, Syringe pump and flow cell: the sample syringe contains helicase–DNA–bead complexes, and the reaction syringe contains ATP. 'X' indicates the laser trap position, and the red arrow indicates movement of the trapped DNA–bead complex across the boundary between solutions. Inset, the trapped DNA with bound helicase, and its unwinding after relocation into the reaction solution. b, Fluorescent DNA helicase assay. A trapped and stretched, fluorescent DNA molecule is shown. As RecBCD enzyme translocates, it both unwinds and degrades the DNA, simultaneously displacing dye molecules (black stars). B, biotinylated oligonucleotide.
Unwinding of dsDNA can be visualized in real time

23°C

37°C
Analysis of the time courses in unwinding of a DNA molecule by RecBCD enzyme
DNA unwinding by individual RecBCD enzyme molecules varies significantly.

\[ V_{\text{max}} = 521 \pm 60 \text{ bp/sec} \]

\[ K_{m}^{\text{ATP}} = 142 \pm 58 \mu\text{M} \]
The processivity of individual RecBCD enzyme molecules varies significantly.

- $N_{\text{max}} = 29,670 \pm 4,256 \text{ bp}$
- $K_{N^{\text{ATP}}} = 158 \pm 78 \mu\text{M}$
Recognition of chi alters RecBCD translocation velocity

Recognition of \( \chi \) Results in a Pause, Followed by a Reduction in the Rate of RecBCD Enzyme Translocation

(A) Analysis of a representative video visualizing RecBCD-mediated unwinding of a single DNA molecule (shown schematically at the right of the graph) containing \( \chi \) positioned 8.2 kb from the DNA end and an additional reverse-oriented \( \chi \) at 7 kb from the DNA end.

The length of the dsDNA molecule was measured for each frame and presented as a time trace (red squares). Zero time corresponds to the moment when the RecBCD-DNA-complex was transferred to the ATP-containing channel. The rates of RecBCD-mediated DNA unwinding and both the duration and the position of the pause were determined by fitting the data corresponding to phases II, III, and IV to a continuous three-segment line (blue line). (B) Schematic representation of the phases of RecBCD-mediated unwinding of \( \chi \)-containing dsDNA. (I) RecBCD-DNA-bead complex is trapped in the sample side of the flow cell. (II) dsDNA unwinding by RecBCD enzyme results in a linear decrease in the dsDNA length upon relocation of the bead-DNA-RecBCD complex to the reaction (ATP-containing) side of the flow cell. (III) interaction with \( \chi \) results in a pause in RecBCD enzyme translocation. (IV) the \( \chi \)-modified enzyme continues to unwind the DNA, but at significantly reduced rate. (V) RecBCD enzyme dissociates from the DNA.
How can the RecBCD translocation velocity change?

Model for the Control of the DNA Translocation Behavior of RecBCD Enzyme by $\chi$

RecBCD enzyme is shown as a bipolar helicase with its two motor subunits translocating on the opposite strands of the DNA substrate molecule. In panels IIa–IIIa, RecD subunit is assumed to be the leading motor subunit prior to $\chi$ recognition. RecD is proposed to be inactivated at $\chi$, and the pause reflects the time required for the slower (RecB) motor subunit to catch up with RecD and RecC stalled at $\gamma$. In panels IIb–IIIb, translocation of both motor subunits is assumed to occur at the same velocity. Pausing at $\chi$ in this model reflects the time required for the conformational change within RecBCD enzyme required to inactivate the RecD subunit. After $\chi$ recognition, the modified RecBCD enzyme is proposed to lack RecD function and, hence, translocates slowly. See text for more details.
How does RecBCD put all these activities together into one enzyme?

1. DNA helicase
2. Endonuclease
3. ATPase
4. Reads sequence of ssDNA, not duplex DNA
5. Responds to sequences of DNA while translocating at 1,000 bp/sec
RecB Subunit

1180 aa (134kDa)

SF1 3′-5′ helicase, sequence similarity to PcrA, UvrD

2 regions connected by long linker

Helicase domains at N-term

Nuclease domain at C-term
RecB comprises 5 domains, two interact with DNA in the structure.

4 base pairs of DNA duplex are melted, in absence of ATP (Initiation complex).
RecB Subunit

Nuclease domain at C-term

Same fold as λ-exonuclease

Ca²⁺ ion bound in active site
RecC Subunit

1122 aa (129kDa)
Less well characterised
Involved in Chi recognition

Large channels run through subunit
RecC Subunit

Unexpectedly, RecC is a defunct helicase

Same fold as PcrA (RMSD Ca ~3Å) but without detectable homology or helicase motifs
RecC Subunit

5 domains, one contacts DNA in the structure

Pin
RecD Subunit

608 aa (67kDa)  
SF1 5’-3’ helicase  
but missing some  
motifs (1A, 3 & 4)  

Contains 3 domains
RecD Subunit

Motor domains same as PcrA (but missing 1B & 2B),
additional domain at N-term that makes contacts with RecC
RecBCD - the complex

RecB contacts duplex ahead of the junction

ssDNA tails head towards each of the helicase motor subunits

RecB

RecC

RecD

Nuclease domain
Channels for ssDNA

Duplex is split across RecC pin and each tail of the DNA heads towards different helicase motor subunits along channels within the protein complex.
Chi recognition

3' tail of the DNA passes RecB subunit at one end of a long channel through RecC subunit, emerges adjacent to nuclease domain

Final cleavage event on 3' tail occurs 4-6 nucleotides before (3' side of) Chi

\[
\begin{align*}
5' & \quad G \ C \ T \ G \ G \ T \ G \ G \\
3' & \quad \chi
\end{align*}
\]
Chi recognition

RecC mutants that show altered Chi recognition map to lip of RecC at interface with nuclease domain which is located 4-6 nucleotides from active site.

Chi recognition could occur as ssDNA passes through RecC.

ssDNA footprint in SF1 helicases is 8 bases.
Summary of RecBCD

1) RecBCD has two helicase motors that split duplex DNA across RecC subunit

2) RecC is defunct helicase that is fed ssDNA by RecB and scans for Chi sites (helicase “footprint” and Chi are both 8 bases)

3) ssDNA tails are fed through channels in the complex that exit adjacent to (gated) nuclease domain
Initiation

Homologous Pairing & DNA Strand Exchange

DNA Heteroduplex Extension

Resolution

Initiation
RecBCD, SSB

Homologous Pairing & DNA Strand Exchange
RecA, RecFOR, SSB

DNA Heteroduplex Extension
RecA, RuvAB, RecG
DNA polymerase I, topoisomerases

Resolution
RuvC, DNA ligase
RuvC, DNA ligase

Spliced
Patched
Models of RuvAB-Holliday Structure Complex
RuvA is a stable tetramer and binds HJ
Directs loading of RuvB onto the HJ

RuvB is a homo-hexamer
dsDNA helicase, AAA^+-ATPase
Hydrolyzes ATP
Cleavage of HJ by RuvC

- RuvC is a small protein of about 20 kD. It requires and binds a magnesium ion.
- Involved in DNA repair and in the late step of RecE and RecF pathway recombination.
- RuvC protein cleaves cruciform junctions, which are formed by the extrusion of inverted repeat sequences from a super-coiled plasmid and which are structurally analogous to Holliday junctions, by introducing nicks into strands with the same polarity.
- The nicks leave a 5' terminal phosphate and a 3' terminal hydroxyl group which are ligated by E.coli or T4 DNA ligases.
- The active form of RuvC protein is a dimer.
- This is mechanistically suited for an endonuclease involved in swapping DNA strands at the crossover junctions.
Cleavage of Holliday Junctions by RuvC

(A) In the presence of divalent metal ions, protein-free Holliday junctions adopt a 2-fold symmetric stacked-X structure with the arms lying in an anti-parallel configuration.

(B) Binding by RuvC unfolds the junction into a 2-fold symmetric open complex in which equivalent sections of DNA pass through each active site in the RuvC dimer.

(C) In the presence of Mg$^{2+}$, resolution occurs by nicking of the two strands that constitute the wide angles in the open complex. In this diagram, the RuvC subunits are related by a dyad axis and the sites of incision are indicated by scissors.
RuvC cleavage of junctions displays a sequence specificity

**Effect of DNA sequence on RuvC-mediated Holliday junction resolution.** Reactions containing RuvC and 5'-32P-labelled Consensus (panel a), Mutant (panel b) or Hybrid (panel c) Holliday junctions were incubated in resolution buffer. Products were analysed by 6% neutral PAGE. The 11 bp homologous core of each junction is shown. The consensus (5'-ATTG-3') and mutant (5'-GTTG-3') sequences are highlighted by light and dark shading, respectively. Sites of cleavage are indicated by arrows. For simplicity, the DNA is drawn folded and parallel; within the RuvC–junction complex the DNA will lie in an open configuration as shown on previous page.
Model for BM and HJ resolution by RuvABC

A

B

Junction binding

Branch migration

Resolution

RuvA

RuvB

RuvC
In vitro Reconstitution of RecA & RuvABC Reactions
RecG and RuvAB are somewhat redundant
Coupling between replication and recombination

- Chromosomal replication in prokaryotes has long been described in terms of initiation from a single origin, followed by bidirectional replication of the chromosome until a termination sequence is reached.
- It has become apparent that the replication fork rarely progresses unimpeded from origin to terminator.
- It is likely that all replication forks stall at some point during the replicative process as a result of encountering various forms of damage in the DNA template.
- A pathway or pathways for repairing or avoiding the damaged DNA and restarting the stalled fork.
- One such mechanism for damage bypass involves the protein RecG.
- Biochemical and genetic studies have shown that the protein is able to convert stalled forks into so-called 'chicken-foot' structures in an ATP-dependent manner.

A. Repair via fork collapse or resetting

[Diagram showing the process of repair via fork collapse or resetting, involving proteins RuvABC and RecG, and processes like duplex DNA end and PriA-dependent replisome loading.]
RecG

- The ATP-dependent DNA helicase RecG plays a critical role in recombination and DNA repair.
- It helps to process Holliday junction intermediates to mature products by catalyzing branch migration.
- RecG has DNA unwinding activity characteristic of a DNA helicase with 3’ to 5’ polarity.
RecG bound to a forked substrate

- RecG protein bound to the DNA junction.
- The RecG is colour-coded as follows, domain 1-green, domain 2-blue, domain 3-gold.
- The DNA is shown in mauve, and the ADP molecule as a ball-and-stick representation.

- Molecular surface representation of RecG. Positive electrostatic potential is shown in blue, negative in red. The DNA is shown as a stick model. The junction can be clearly seen to sit in a cleft on the N-terminal of the protein; the leading and lagging strands are separated by a wedge-like protrusion.
- The way in which the DNA is bound, together with information obtained from other helicase structures, gives us some clue as to how the protein may catalyse chicken-foot and four-way junction formation.
Action of RecG on a 4-way junction

- Mechanism of formation of four-way junction from chicken-foot. If the newly synthesised lagging (orange) and leading (red) strands are allowed to reanneal after being stripped from the original template (blue and grey), a four-way junction may be formed.
- Thus the RecG protein contains domains specific to translocation of a nucleic acid substrate, and for specific targeting and unwinding of, a three-way junction. This activity allows it to initiate the template switching mechanism of DNA damage bypass.