Genome Editing and Engineering
Prof. Utpal Bora
Department of Bioscience and Bioengineering
Indian Institute of Technology, Guwahati

Module - 03
Recombination
Lecture - 06
Homologous and non-homologous recombination

Welcome to the course on Genome Editing and Engineering, module number 3 on recombination. Today we are going to learn about homologous and non-homologous recombination processes.

(Refer Slide Time: 00:52)

From our previous discussion you know that organisms are continuously exposed to a myriad of DNA damaging agents during its lifetime.

You also now know that DNA damage are broadly of two types:

(1) Endogenous damage caused by reactive oxygen species (ROS) that are derived from metabolic byproducts, replication errors, DNA base mismatches and topoisomerase-DNA complexes
(2) Exogenous damage caused by radiation (UV, X-ray, gamma), hydrolysis, plant toxins, and viruses.

If left unrepaired DNA damage impact health and modulate disease-states.

Luckily, robust DNA repair and damage-bypass mechanisms sincerely protect the DNA by either removing or tolerating the damage to ensure an overall survival.

Genetic or metabolic diseases occur when the repair mechanism fails for any reason which is an exception rather than a usual event.

From your previous classes you have come to know that organisms are continuously exposed to a myriad of DNA damaging agents during its lifetime. You also now know that DNA damage are broadly of two types, the endogenous damage which is caused by reactive oxygen species which are derived from metabolic byproducts, then replication errors, DNA base mismatches and topoisomerase DNA complexes.

The second type of DNA damage is the exogenous damage caused by radiation like UV X-ray gamma and hydrolysis plant toxins and viruses. You know that if these DNA which is damaged is not repaired, it will impact health and also modulate certain disease states in various organisms. Luckily, there are robust DNA repair and damage bypass mechanisms

which essentially protect the DNA by either removing or tolerating damage to ensure an overall survival of the organism.

Certain genetic or metabolic diseases occur when these repair mechanisms fail for any reason which is an exception rather than unusual event.

(Refer Slide Time: 02:05)

Organisms respond to DNA damage at the cellular and molecular level by activating robust DNA damage response (DDR) pathways, that physically remove or repair the damage in a substrate-dependent manner.

There are at least five major known DNA repair pathways active throughout different stages of the cell cycle, allowing the cells to cope up with the inflicted the DNA damage.

- base excision repair (BER),
- nucleotide excision repair (NER),
- mismatch repair (MMR),
- homologous recombination (HR) and
- non-homologous end joining (NHEJ)

A few other specific lesions can also be removed by direct chemical reversal and interstrand crosslink (ICL) repair.

Whenever there is a DNA damage the organisms respond to it at the cellular and molecular level by activating robust DNA damage response pathways or DDR pathways, which physically remove or repair the damage in a substrate dependent manner. There are at least five major known DNA repair pathways which are active throughout different stages of the cell cycle allowing the cells to cope up with the infliction of DNA damage by various agents.

Briefly these are the base excision repair nucleotide excision repair mismatch repair homologous recombination and non-homologous end joining. A few other specific lesions are also removed by direct chemical reversal and interstate cross link repair.

(Refer Slide Time: 02:49)

DSBs are considered as the most toxic lesion.

Double-strand DNA breaks (DSBs) are generated by various ways,

i. Endogenous stresses resulting from cellular metabolism, such as replication stress and reactive oxygen species (ROS),

ii. Exogenous factors, such as ionizing radiation and chemotherapy agents (e.g., topoisomerase inhibitors).

iii. DSBs may also arise through errors in DNA replication or as normal intermediates during programmed cellular processes such as meiosis or V(D)J recombination (which generates mature immunoglobulin or T-cell receptor genes from the separate fragments of the germline genome).

DSBs can also be programmed to trigger beneficial genomic rearrangements during meiotic differentiation or the establishment of the immune system.

Adapted from https://encyclopedia.pub/entry/9091

Double strand breaks are considered as the most toxic lesion. These DSBs are generated by various ways which is already known to you, the endogenous stresses or the exogenous factors. Besides these DSBs may also arise through errors in DNA replication or as normal intermediates during programmed cellular processes such as meiosis or VDJ recombination which generates mature immunoglobulin or T-cell receptor genes from the separate fragments of the germline genome.

These DSBs can be programmed to trigger beneficial genomic rearrangement during meiotic differentiation or the establishment of the immune system. So, DSBs are kind of double edged swords they are harmful on the one hand, but we can also exploit them in certain cases.

(Refer Slide Time: 03:39)

Strategies for repairing double-strand breaks.

Whenever there is a DNA breakage particularly DSB what are the strategies the cell adopts to repair those DSBs?

(Refer Slide Time: 03:56)

Whenever DSBs occur and by whatever process, these breaks must be repaired to ensure survival of the cell.

i. **Homologous recombination**. In this strategy, a double-strand break present in one chromatid is repaired using its intact sister chromatid as a template. The repair of breaks by homologous recombination is a high fidelity process as it ensures that all the genetic information at the break site is retained.

ii. **Non-homologous end joining.** NHEJ involves the simple rejoining of the broken DNA ends, regardless of the DNA sequence. This mechanism is error-prone, as small deletions may be introduced at the break site.

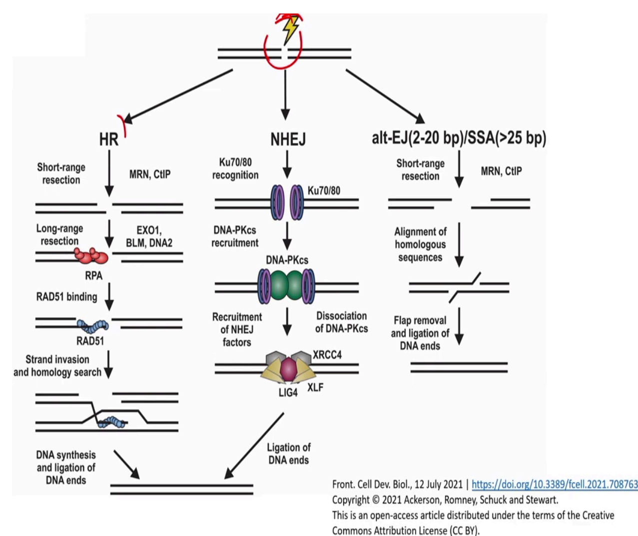
iii. **Alternative form of NHEJ** and is frequently abbreviated as A-NHEJ, or simply A-EJ

Without the repair of the DSBs we have emphasized time and again that the survival of the cell will be at stake and if it survives at all then it may lead to certain disease states. So, for the healthy existence of an organism the DNA is repaired whenever there is a occurrence of DSBs and this falls mainly into the following types, the one is the homologous recombination.

In this strategy the cell will undertake the repair of the DSB strand break whereby in one chromatid the repair process takes place using the sister chromatid as a template. So, this is kind of a template dependent repair. The repair breaks by homologous recombination therefore, is a high fidelity process as it ensures that all the genetic information at the break site is retained due to the presence of the template or the sister chromatid chromatid during its repair.

The other process by which DSBs are repaired are the non-homologous end joining method which involves the simple rejoining of the broken DNA ends regardless of the DNA sequence. This mechanism is error prone a small deletions may be introduced at the break site. There is a third form which is similar to the second form and this is therefore, known as the alternative form of NHEJ and is frequently abbreviated as A-NHEJ. So, A coming from the word alternative or simply as A-EJ.

(Refer Slide Time: 05:45)

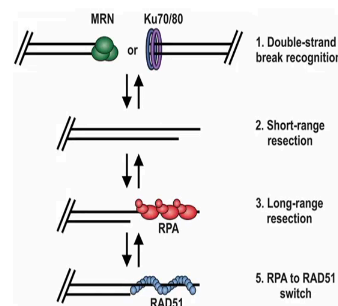


Front. Cell Dev. Biol., 12 July 2021 | https://doi.org/10.3389/fcell.2021.708763
Copyright © 2021 Ackerson, Romney, Schuck and Stewart.
This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

So, these are the three main DNA double strand brake repair pathways. The first one is the homologous recombination, the second one is the NHEJ, the third one is the A-EJ or A-NHEJ. So, these are various pathways through which the DSBs are repaired and you can see in each pathway there are many protein molecules involved in every step. We will be discussing about the protein molecules that help us in repairing the DSBs following a certain pathway whether it is homologous recombination or NHEJ or alt EJ.

Now, from this diagram you can see whenever there is a double strand break the DNA is either repaired by homologous recombination or any of the other two pathways. Now, a DNA break is a DNA break. How does the cell identify or take decisions, which pathway to follow for the repair of these particular break over here? Whether it will go the hr pathway or the other two pathways?

(Refer Slide Time: 07:03)



Key decision points in the repair of DSBs

(1) The DNA ends are bound by either MRN or the Ku70/80 heterodimer. Binding and retention of MRN will shift repair toward HR and the binding of Ku shifts repair toward to NHEJ.

(2) Short range resection by MRN.

(3) Long-range resection of the DNA and RPA binding. (4) RPA is exchanged for RAD51, which facilitates strand invasion, DNA synthesis and HR repair.

Front. Cell Dev. Biol., 12 July 2021 | https://doi.org/10.3389/fcell.2021.708763
Copyright © 2021 Ackerson, Romney, Schuck and Stewart.
This is an open-access article distributed under the terms of the Creative Commons Attribution License.

So, in this case the decision is taken by the involvement of certain players. For example, the DNA ends are bound by either MRN or Ku70/80 heterodimer as you can see in this picture. Binding and retention of MRN will shift repair towards HR and binding of the Ku shifts repair towards NHEJ.

So, these are the critical players, which decides which pathway the DNA repair will follow and there is a short range resection MRN. Long range resection of the DNA and RPA binding; the RPA is exchanged for RAD51 which facilitates strand invasion DNA synthesis and repair HR repair.

(Refer Slide Time: 08:02)

Homologous recombination promotes the pairing between identical or nearly identical DNA sequences and the subsequent exchange of genetic material between them.

Homologous recombination is termed as the **guardian of genome integrity**, as it acts to repair DNA damage.

Homologous recombination is found to be are involved in rescue replication forks that have stalled for various reasons, such as a missing factor (e.g., the helicase), or a particular difficulty upstream of the fork, such as supercoiling or intense traffic of proteins.

Homologous recombination is a highly conserved process from bacteria to humans, that serves to repair double-stranded breaks or single-stranded gaps in the DNA. For eg. in higher organisms, the Rad51 protein is a structural and functional homologue of the bacterial strand-exchange protein RecA.

Homologous recombination promotes pairing between identical or nearly identical DNA sequences and the subsequent exchange of genetic material between them. As the name homologous recombination suggests there must be some kind of homology present between the two strands of DNA that is why it involves nearly identical DNA sequences.

Homologous recombination is termed as the guardian of genome integrity, as it acts to repair DNA damage. Homologous recombination is found to be involved in rescue rescuing of replication forks that has stalled for various reasons, such as the missing factors or a particular difficulty upstream of the fork, such as super coiling or intense traffic of proteins.

Homologous recombination is a highly conserved process whether it is in bacteria or in humans, it serves to repair double-strand breaks or single stranded gaps in the DNA. For example, in higher organisms, there is a protein called RAD51 and its structure and functional homologue of the bacterial strand exchange protein is the RecA protein in bacteria.

(Refer Slide Time: 09:18)

HR is also a driving force for the evolution of multigene families.

Homologous recombination has played a major role in evolution and genome dynamics, by changing gene copy numbers through deletions, duplications, and amplification.

Intrachromosomal recombination between ribosomal operons or between mobile elements scattered into the genome leads to deletion or tandem duplications of large regions within the genome, up to several hundred kilobases.

The emerging technical application of HR constitutes the basis of targeted gene replacement for gene therapy as well as for the precise design of engineered organisms.

It is also a driving force for the evolution of multigene families. This is something very interesting. The HR is considered as a guardian of genome integrity. So, it guards the genome at all times, but it is also a driving force for evolution of multigene families at the same time. So, there is some kind of relaxation and escape somewhere in between which is allowed, without these the diversity would not have happened.

So, HR is a very very interesting phenomena, which not only guards the genome on the one side, but it also allows evolution to happen particularly in multigene families at the same time. HR plays a major role in evolution and genome dynamics by changing gene copy numbers through deletions, duplications and amplification.

Intrachromosomal recombination between ribosomal operands or between mobile elements, scattered into the genome leads to deletion or tandem duplications of large regions with within the genome of the several hundred kilo bases. The emerging technical application of homologous recombination today constitutes the basis of targeted gene replacement for gene therapy as well as for the precise design of engineered organism.

In our course we will be using this technique or depending on this technique to a large extent.

(Refer Slide Time: 11:00)

The basic model for homologous recombination was largely derived from genetic studies in fungi such as ***Ustilago maydis*** and ***Saccharomyces cerevisiae***

Studies with the bacterium ***Escherichia coli***, have provided us crucial biochemical insights into the mechanism of homologous recombination.

In *E. Coli* , about 20 genes are found to be involved in recombination. They produce, specific proteins which carry out each of the key steps in homologous recombination.

The basic model for homologous recombination was largely derived from genetic studies in fungi such as *Ustilago* and *Saccharomyces cerevisiae*. Studies with the bacterium *E. coli* have provided us crucial biochemical insights into the mechanism of homologous recombination.

In *E. coli* about 20 genes are found to be involved in homologous recombination. They produce specific proteins which carry out each of the key steps in the homologous recombination pathway.

(Refer Slide Time: 11:30)

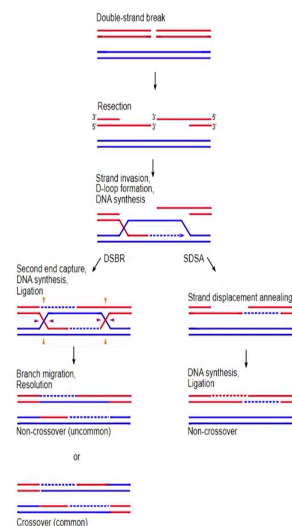
Let us examine the events following a DSB for **Resection.**

Most molecular models of homologous recombination describe the process in three key steps:

i. Strand exchange,
ii. Branch migration and
iii. Resolution.

Current Biology 11, Issue 7, 3 April 2001, Pages R278-R280.
https://doi.org/10.1016/S0960-9822(01)00138-5

Nature Reviews Molecular Cell Biology 7: 739-750. DOI:10.1038/nrm2008.
This file is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license.



So, you can see in this diagram the various steps involved in homologous recombination as well as alternative recombination repair pathways. So, let us examine the events that occur whenever a DSB happens. So, once a double strand break occurs one of the events that follows immediately is the resection.

Most molecular models of homologous recombination describe the process in three key steps. The first one is the strand exchange, you can see over here, then the branch migration and finally, the resolution.

(Refer Slide Time: 12:16)

Strand exchange, involves pairing of the broken DNA end with the homologous region of its sister chromatid, followed by strand invasion to form a DNA crossover or Holliday junction.

This process generates regions of heteroduplex DNA comprising DNA strands from different sister chromatids.

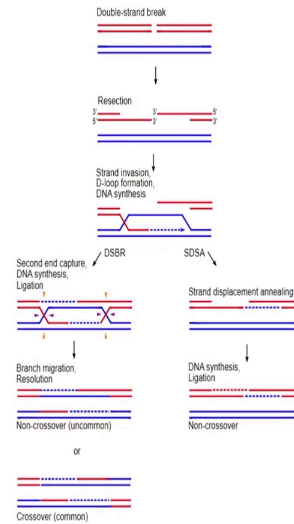
Strand exchange involves pairing of the broken DNA end with the homologous region of its sister chromatid followed by strand invasion to form a DNA crossover or Holliday junction. This process generates regions of hetero duplex DNA comprising DNA strands from different sister chromatids.

(Refer Slide Time: 12:40)

Resection: First there is 5′-to-3′ resection of the double-strand break (DSB), producing DNA ends with 3′ single-stranded DNA tails.

Strand Invasion: The free 3′ ends invade a homologous DNA duplex, forming a DNA crossover or Holliday junction and act as a primer to initiate new DNA synthesis.

Branch migration: Branch migration of the Holliday junction extends the region of heteroduplex away from the initial site of crossover.



So, what happens its resection? First there is a 3 5 to 3 prime resection of the double strand break, which produces DNA ends with 3 prime single stranded DNA tails. This is followed by strand invasion. The free 3 prime ends invade a homologous DNA duplex forming a DNA crossover or Holliday junction and act as a primer to initiate, new DNA synthesis. This is followed by branch migration where the Holliday junction extends the region of hetero duplex away from the initial site of the crossover.

(Refer Slide Time: 13:20)

Branch migration.
during branch migration, the Holliday junction is translocated along DNA, extending the region of heteroduplex away from the initial crossover site.

In the last step the Holliday junction intermediate is resolved by cleavage of the junction to form separate duplex DNA molecules again

Holliday Junction Resolution:
Holliday junctions are resolved by endonucleolytic cleavage of either the crossed strands or non-crossed strands of the junction.

Current Biology 11, Issue 7, 3 April 2001, Pages R278-R280.
https://doi.org/10.1016/S0960-9822(01)00138-5

During branch migration the Holliday junction is translocated along the DNA molecule extending the region of hetero duplex away from the initial crossover site. In the last step the Holliday junction intermediate is resolved by cleavage of the junction to form separate duplex DNA molecules again.

The Holliday junction resolution is therefore a very very important step in this entire process. This is resolved by endo nucleolytic cleavage of either the crossed strands or non-crossed strands of the junction.

(Refer Slide Time: 13:56)

Molecular Mechanisms of Repair

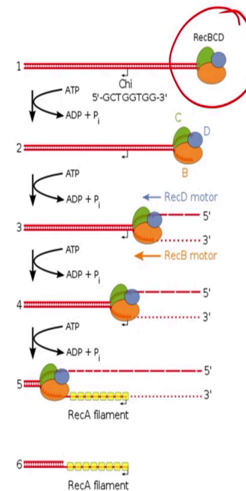
Let us now discuss the molecular mechanisms involved in the repair of DSBs.

(Refer Slide Time: 14:04)

DNA ends resulting from a double-strand break is processed by a multi-functional enzyme complex called RecBCD

RecBCD is a sequence-regulated **bipolar helicase-nuclease** that splits the duplex into its component strands and digests them until it encounters **Chi site which is** a recombinational hotspot.

The nuclease activity is then attenuated and RecBCD loads RecA onto the 3′ tail of the DNA.



Nature 432: 187-193. doi: 10.1038/nature02988 (CC BY 3.0)

You can see in this figure certain proteins called RecBCD and then RecA over here. We will discuss the role of these particular proteins in the repair mechanism one by one. DNA ends resulting from a double strand break is processed by a multifunctional enzyme complex called RecBCD. The RecBCD is a sequence regulated bipolar helicase nuclease that splits the duplex into its component strands and digest them until it encounters Chi site which is a recombinational hotspot.

The nuclease activity is then attenuated and RecBCD loads RecA into the 3 prime tail of the DNA. So, these are the critical steps. One thing we have to remember that this RecBCD has two functions. One is it opens the DNA with its helicase activity, the next is its nucleus activity.

(Refer Slide Time: 15:23)

A Chi (χ) site is a short stretch of DNA in the genome of a bacterium near which homologous recombination is more likely to occur than on average across the genome. For this reason Chi sites are also referred to as "recombination hot spots".

Chi sites serve as stimulators of DNA double-strand break repair in bacteria and its sequence is unique to each group of closely related organisms;

In enteric bacteria E. coli and Salmonella, the core sequence is

5'-GCTGGTGG-3'

In addition it includes about 4 to 7 nucleotides to the 3' side of the core sequence which are important.

The Chi site is a short stretch of DNA in the genome of a bacterium near which homologous recombination is more likely to occur than on average across the genome. For this region Chi sites are also referred to as recombination hotspots. Chi sites serve as stimulators of DNA double strand break repair in bacteria and its sequence is unique to each group of closely related organisms.

In enteric bacterial *E. coli* and *Salmonella,* the core sequence for example, is 5′-GCTGGTGG-3′. In addition, it includes about 4 to 7 nucleotides to the 3 prime side of the core sequence which play important role.

(Refer Slide Time: 16:13)

Primary structure of the RecBCD enzyme.
The total number of amino acids (aa) in each polypeptide is indicated in parentheses.

The RecB protein is modular. The N-terminal domain contains seven motifs characteristic of SF1 helicases. The C-terminal domain contains motifs characteristic of a diverse family of nucleases. The nuclease motif contains key catalytic aspartate and lysine residues.

Microbiol Mol Biol Rev. 2008 Dec; 72(4): 642–671.

Let us examine the primary structure of the RecBCD enzyme. The total number of amino acid in each polypeptide will be shown in the figure which we are going to study in the next slide. This RecB protein which is a component of the RecBCD enzyme is a modular protein or modular subunit.

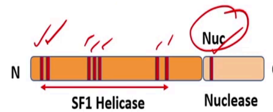
The N-terminal domain contains seven motifs which is characteristics of a SF1 helicases and these are involved in opening up of the DNA double strand molecule. The C-terminal domain contains motifs characteristics of a diverse family of nucleases. The nuclease motif contains key catalytic aspartate and lysine residue.

(Refer Slide Time: 17:01)

Primary structure of the RecBCD enzyme.

Let us study each component one at a time

The RecB protein contains 1180 aa residues and is modular.



The N-terminal domain contains Helicase function and has seven characteristic SF1 motifs 1, 1a, 2, 3, 4, 5, and 6.

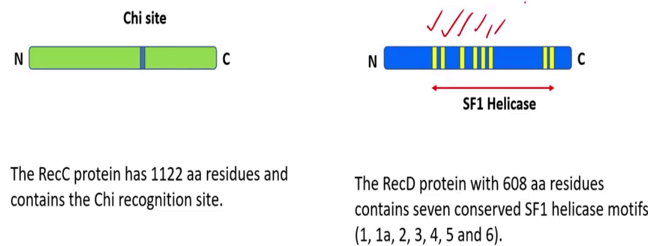
The C-terminal domain contains nucleases motifs. "Nuc" marks the position of nuclease motif 3, which contains key catalytic aspartate and lysine residues.

Microbiol Mol Biol Rev. 2008 Dec; 72(4): 642–671.

So, in this picture you can see the simplified diagram of the RecB protein which contains around 1180 amino acid residues and as already told to you it is modular having two modules, one having the helicase activity the other having the nucleus activity. The N-terminal domain contains the helicase function and it has seven characteristic SF1 motifs as you can see here 1a, 1, 1a, 2, 3, 4, 5, and 6 total seven motifs.

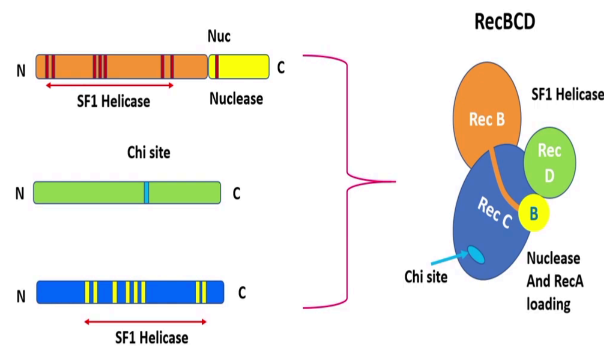
The nucleus activity lies in the C-terminal domain and you can see here Nuc motif which marks the position of the nucleus motif, which contains key catalytic aspartate and lysine residues.

(Refer Slide Time: 18:01)



The other two components of the RecBCD protein are RecC and RecD. So, in RecC there are around 1122 amino acid residues and these contains the important region called as the Chi recognition site. The RecBCD protein has around 608 amino acid residues and it contains seven conserved SF1 helicase motifs as in the case of RecB and you can see those in the yellow colour strips in this diagram.

(Refer Slide Time: 18:41)



So, we now know that RecB is a modular protein having helicase and nucleus activity, RecC is having the Chi site and RecD is again having the helicase activity. All these three join

together to form the RecBCD protein complex and in this you can clearly see that RecB and RecD which contains the helicase activity will be involved in opening the DNA double strands. This Chi site is the place where due to which the stoppage of this molecule will happen.

And a part of the RecB the nucleus domain you can see lying over here in this position. These are all schematic and not exact locations, but you can see in this particular protein complex the different functional domains with respect to one another. Now, what are the different functions of these different constituent proteins?

The RecC recognizes the Chi sequence as already told to you. Once it recognizes the Chi sequence, it signals RecD to stop not to go farther. RecD will then stop and signals RecB to cleave the DNA. RecB cleaves and continuous unwinding DNA and loads the RecA molecule.

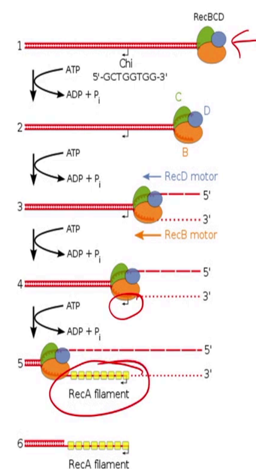
(Refer Slide Time: 20:33)

RecBCD-catalyzed DNA end-processing reaction.

(1) RecBCD binds tightly to a blunt (or nearly blunt) DNA end of a linear DNA duplex.

(2) RecBCD couples the hydrolysis of ATP to DNA translocation and unwinding (helicase activity). The ssDNA products are cleaved asymmetrically, with the degradation of the 3′-terminated ssDNA tail being much more vigorous than the degradation of the complementary tail.

(3) The enzyme continues to translocate until it pauses at a correctly oriented Chi sequence. At the Chi sequence the biochemical properties of the enzyme are altered dramatically. After Chi recognition, RecBCD facilitates the loading of the RecA protein onto the 3′ ssDNA tail.



Microbiol Mol Biol Rev. 2008; 72(4): 642–671.
doi: 10.1128/MMBR.00020-08

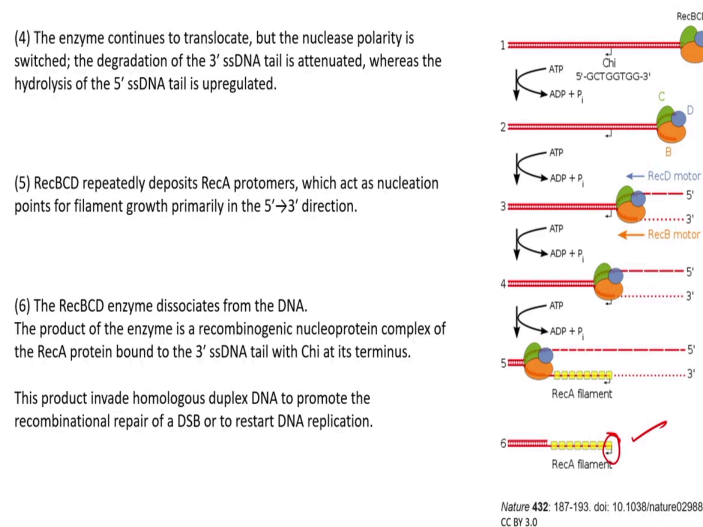
Nature **432**: 187-193. doi: 10.1038/nature02988
CC BY 3.0

Let us look into the various steps where RecBCD catalyzes DNA end processing resection. So, there are various steps laid over here 1 2 3 4 5 6, we will discuss these steps one by one. RecBCD binds tightly to a blunt or nearly blunt DNA end of a linear DNA duplex. The RecBCD couples the hydrolysis of ATP to DNA in the next step and the DNA translocation and unwinding due to the helicase activity takes place.

The single stranded DNA products are cleaved asymmetrically with the degradation of the three prime terminated ssDNA tail being much more vigorous than the degradation of the complementary tail. In the next stage the enzyme continues to translocate until it poses as a correctly oriented Chi sequence and the Chi sequence the biochemical properties of the enzyme are altered dramatically.

After Chi recognition, RecBCD facilitates the loading of the RecA protein on to the 3 prime ssDNA. So, you can see the Chi site here. So, once it arrives at the Chi site after that it will facilitate the loading of the RecA protein.

(Refer Slide Time: 22:05)



After these the enzyme continues to translocate, but the nucleus polarity is switched the degradation of 3 prime ssDNA tail is attenuated whereas, the hydrolysis of the 5 prime ssDNA tail is up regulated. RecBCD then repeatedly deposits RecA promoters which act as nucleation points for filament growth primarily in the 5 to 3 prime direction.

Following these the RecBCD enzyme dissociates from the DNA. The product of the enzyme is a recombinogenic nucleoprotein complex of RecA protein bound to the 3 prime ssDNA tail with Chi at its terminus. This product invades homologous duplex DNA to promote the recombinatorial repair of a DSB or to restart DNA replication.

(Refer Slide Time: 23:09)

It has been found that the permanent inactivation of RecBCD enzyme by Chi sites in duplex DNA occurs by the disassembly of the enzyme into its three constituent subunits.

It is hypothesized that this inactivation occurs in two distinct steps.

1. Upon encountering a Chi sequence, RecBCD enzyme undergoes its first change: it retains its ability to travel along the DNA and to cut a hairpin DNA structure at the distal end of the DNA but loses its ability to nick at subsequently encountered Chi sites on the same DNA molecule (Taylor and Smith 1992).

2. The second change, the disassembly of the enzyme into three inactive subunits, may occur either during continued unwinding beyond Chi or upon reaching the end of the DNA.

Andrew F. Taylor and Gerald R. Smith. Genes & Dev. 1999. 13: 890-900

It has been found that the permanent inactivation of RecBCD enzyme by Chi sites, in duplex DNA occurs by the disassembly of the enzyme into its three constituent subunits. It is hypothesized that this inactivation occurs in two distinct steps. Number 1 upon entering a Chi sequence RecBCD enzyme undergoes its first change.

It retains its ability to travel along the DNA and cut a hairpin DNA structure at the distal end of the DNA, but loses its ability to nick at subsequently encountered Chi sites on the same DNA molecule. The second change the disassembly of the enzyme into three inactive subunits may occur either during continued unwinding beyond Chi or upon reaching the end of the DNA molecule.

(Refer Slide Time: 24:03)

Protein: Protein RecA
Gene: recA
Organism: Escherichia coli (strain K12)

Length: 353
Mass (Da): 37,973

RecA is a protein of about 350 amino-acid residues. Its sequence is highly conserved among eubacterial species. It is also found in the chloroplast of plants.

RecA-like proteins are found in archaea and diverse eukaryotic organisms, like fission yeast, mouse or human.

RecA is involved in homologous recombination and bypass mutagenic DNA lesions by the SOS response.
It catalyzes the,

i. ATP-driven homologous pairing and strand exchange of DNA molecules necessary for DNA recombinational repair
ii. hydrolysis of ATP in the presence of single-stranded DNA,
iii. ATP-dependent uptake of single-stranded DNA by duplex DNA, and
iv. ATP-dependent hybridization of homologous single-stranded DNAs.

(PubMed:22412352).

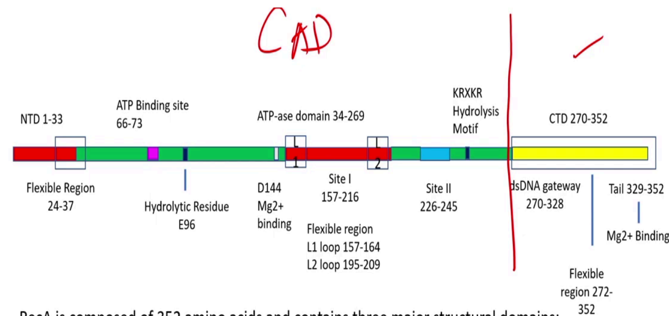
https://www.uniprot.org/uniprot/P0A7G6

Now, let us study a little bit about the RecA protein which binds to the single stranded DNA and helps in the strand invasion. RecA protein is about 350 amino-acid residue long. Its sequence is highly conserved among eubacterial species. It is also found in the chloroplast of plants.

RecA like proteins are found in RecA and diverse eukaryotic organisms like fission yeast mouse or humans. The RecA protein of E. coli has around 353 amino-acid residues. The RecA protein is involved in homologous recombination as already known to you and it bypasses mutagenic DNA lesions by the SOS response.

RecA protein catalyzes various reactions like the ATP driven homologous pairing and strand exchange of DNA molecules, necessary for DNA recombinatorial repair. Then the hydrolysis ATP in the presence of single stranded DNA molecules and ATP dependent uptake of single stranded DNA by duplex DNA. And finally, it catalyzes the ATP dependent hybridization of homologous single stranded DNAs.

(Refer Slide Time: 25:30)



RecA is composed of 352 amino acids and contains three major structural domains:

i. a central, core ATPase domain (CAD), which extends from the 34th to the 269th amino acid (in green), and two smaller

ii. NTD and CTD, which extend from the 1st to 33rd (in red) and from the 270th to 352nd amino acids, respectively (in yellow).

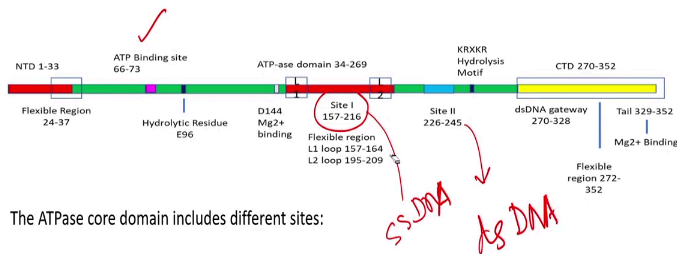
Figure Adapted from Biochemical Society Transactions (2019) 47 1511–1531

Let us now study about the various structural domains of RecA protein. You can see in this picture various domains which we have coloured and these colours represent certain domains which we will be discussing one by one. RecA is composed of around 350 amino acids and contains three major structural domains. It is a central core ATPase domain or CAD domain which extends from the 34th to the 269th amino acid residue and here we are depicting these in the green colour.

Apart from this core ATPase domain there are two other smaller domains the N-terminal domain NTD and the C-terminal domain. So, we can divide this into three parts the CAD the NTD and the CTD. The NTD and CTD extends from the first to 33rd residues and from the 270th to the 352nd amino acids respectively and shown here in the colour red and yellow.

Then there are various other important features in this particular protein. You can see certain flexible region, they are two flexible regions then there are two important sites called site I and site II. Then you have a dsDNA gateway we will study about these various flexible regions as well as these Mg2 binding sites and various site I and site II domains in our next slides.

(Refer Slide Time: 27:30)



The ATPase core domain includes different sites:

- the ATP binding site (residues 66–73);
- site I, which is the ssDNA binding site (residues 157–216); and
- site II, which is the dsDNA binding site, (residues 226–245).
- It also includes residues responsible for the ATP hydrolysis activity (E96 and the [KR] × [KR] motif in positions 248–250).

The CTD can be divided into two domains: the dsDNA gateway (residues 270– 328) and the CTD tail (residues 329–325) which modulates the protein activity.

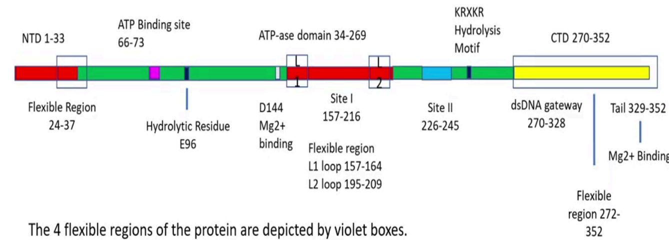
The Mg2+ binding is coordinated by D144 and CTD tail.

Let us first focus on the ATPase core domain. It includes different sites number one is the ATP binding site between the residue 66 to 73. Then there is site I as you can see over here having the residues 157 to 216. This is a single strand binding single strand DNA binding site.

Site II is double strand DNA binding site. So, this particular molecule can bind to single stranded DNA and double stranded DNA at the same time and for these it has specific sites, site I for ssDNA binding and site II for dsDNA binding. It also includes residues responsible for the ATP hydrolysis activity you can see this in the residues around 240 to 250 here into here.

This C-terminal domain can be divided into two domains the dsDNA gateway domain between the residues 270 to 328 and the CTD tail between the residues 329 to 325 which modulates the protein activity. The Mg2+ binding is coordinated by D144 and the CTD tail as shown in this picture

(Refer Slide Time: 29:27)



These regions are the regions between residues 24 and 37, at the end of the N-terminal domain (NTD) and the beginning of the core domain),

157 and 164 (loop L1),

195 and 209 (loop L2) and

270 and 352 (CTD).

Figure Adapted from Biochemical Society Transactions (2019) 47 1511–1531

Now let us focus on the flexible regions. There are total four flexible regions of the protein and these are depicted by these violet boxes. So, these regions are the regions between residues 24 and 37 at the end of the N-terminal domain and the beginning of the core binding domain. Then the next flexible region is L1, 157 and 164 amino acid residue, the third one is the L2 or the loop 2 and these lies between 195 and 209 amino acid residues, the last one is the CTD which lies between 270 and 352 amino acid residues.

(Refer Slide Time: 30:22)

RecA activity

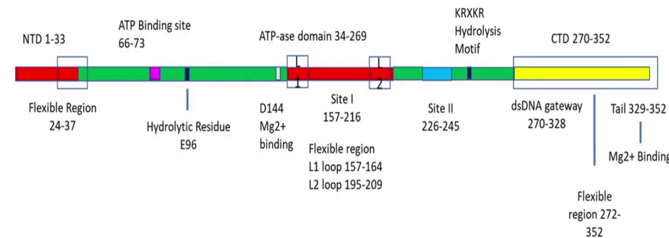
ssDNA and RecA along with its cofactor ATP form an active right-handed helical nucleoprotein filament with six RecA monomers per turn.

This active nucleoprotein filament searches for and captures a homologous dsDNA to produce a synaptic structure. Once a region of homology is found, the ssDNA strands on the homologous chromosomes are exchanged, producing heteroduplex DNA.

So, we now know the various structural domains in a RecA protein which binds to RecBCD which is being loaded into the single stranded DNA by RecBCD and then finally, which drags the Rec single stranded DNA into a duplex DNA structure. How does it act, what are the mode of its activity? ssDNA and RecA along with its co factor ATP form an active right-handed helical nucleo protein filament with six RecA monomers per turn.

This active nucleoprotein filament will search for and capture a homologous double stranded DNA to produce a synaptic structure. Once a region of homology is found the single stranded DNA on the homologous chromosomes are exchanged producing hetero duplex DNA structure.

(Refer Slide Time: 31:30)



The nucleoprotein filament has hydrolytic activity, and this hydrolysis is carried out by the [KR] × [KR] hydrolysis motif containing Lys248 and Lys250, which co-operate with Glu96 on the other monomer.

1. The single-stranded DNA (ssDNA) binds to site I of the protein, forming a nucleoprotein filament.
2. Then, ATP, a RecA cofactor, binds to the ATP binding site and activates the filament.
3. After that, RecA performs a homology search to find a homologous double-strand (dsDNA). In this process, dsDNA interacts first with the NTD and then
4. with the CTD, by which it can move to site II,
5. where it binds dsDNA.
6. If the bound dsDNA is homologous to the ssDNA, strand exchange is performed; if it is not, the dsDNA is released.

The nucleoprotein filament has hydrolytic activity and this hydrolysis is carried out by the KR into KR hydrolysis motif which contain lysine 248 and 250 which cooperates with Glu96 on the other monomer. The single stranded DNA binds to site I of the protein forming a nucleoprotein filament. Then ATP RecA co-factor binds to the ATP binding site and activates this filament.

After that RecA performs a homology search to find the homologous double strand DNA, in this process double stranded DNA interacts first with the NTD and then with the CTD by which it can move to site II where it will bind to double stranded DNA. If the bound double stranded DNA is homologous to the single stranded DNA, strand exchange is performed if it is not the double stranded DNA is released.

So, in many way this is some kind of a random interaction, where it tries to do some match making. If the match making happens the reaction will proceed forward, if the match making fails due to the non-availability of complementary sequences the double stranded DNA will be released and the system will keep on trying.

(Refer Slide Time: 33:06)

Homologous Recombination in eukaryotes

Let us now move to the homologous recombination in eukaryotes. Most of our studies discussions we had till now are based on prokaryotic research. The process is similar in eukaryotes, but there are certain players which are unique to the eukaryotic system. Here also resection takes place.

(Refer Slide Time: 33:40)

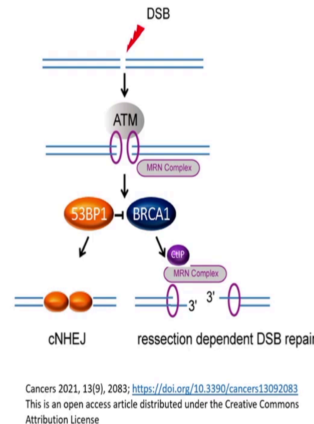
2.1. Resection

Sensor of DSBs:

Double-strand breaks are recognized by the Mre11–Rad50–Nbs1 (MRN) or Mre11 complex.

Capture of DNA ends by the MRN complex leads to the rapid activation of the ataxia-telangiectasia mutated (ATM) kinase which phosphorylates diverse substrates participating in DNA damage response.

BRCA1 and p53-binding protein 1 (53BP1)- has opposing roles and influences the choice of homologous recombination over non-homologous end joining and potentially other mutagenic pathways of DSB repair.



Cancers 2021, 13(9), 2083; https://doi.org/10.3390/cancers13092083
This is an open access article distributed under the Creative Commons Attribution License

Apart from i) Initiation of resection BRCA1 also helps in ii) Loading of RAD51.

And the double stranded breaks are recognized by Mre11-Rad50-Nbs or Mre11 complex. The capture of DNA ends by the MRN complex will lead to the rapid activation of the ATM kinase which phosphorylates diverse substrate participation in DNA damage response.

BRCA1 and p53-binding protein 1 (53 BP), here you can see this. They have opposing roles and influences the choice of homologous recombination over non-homologous end joining and potentially other mutagenic pathways of DSB repair. So, in the case of eukaryotic DSB repair the involvement of BRCA and p53 are very important in the decision making whether it will go HR recombination pathway or other recombination repair pathway.

Apart from number one initiating the resection BRCA1 also helps in the loading of the RAD51.

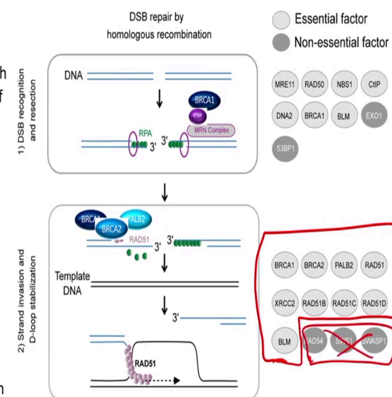
(Refer Slide Time: 34:53)

2.2. Loading RAD51 on ssDNA, Search for Homology and Strand Invasion

The 3′ ssDNA stretch created by resection is coated with replication protein A (RPA), protecting it. The loading of RAD51 onto ssDNA is performed by the BRCA2-PALB2 complex.

This protein complex interacts with BRCA1 and catalyzes the replacement of RPA by RAD51 on the stretch of 3′ ssDNA, creating the RAD51-ssDNA presynaptic complex.

The ssDNA-RAD51 filament scans the genome to search for homology and on success, the filament invades the duplex homologous DNA and initiates strand exchange, creating a displacement loop (D-loop).



Cancers 2021, 13(9), 2083; https://doi.org/10.3390/cancers13092083
This is an open access article distributed under the Creative Commons Attribution License

Let us discuss about the loading of the RAD51 on ssDNA. So, earlier we were discussing about the loading of the RecA on the single stranded DNA and you already now know in details how that loading and strand invasion takes place as a result of that.

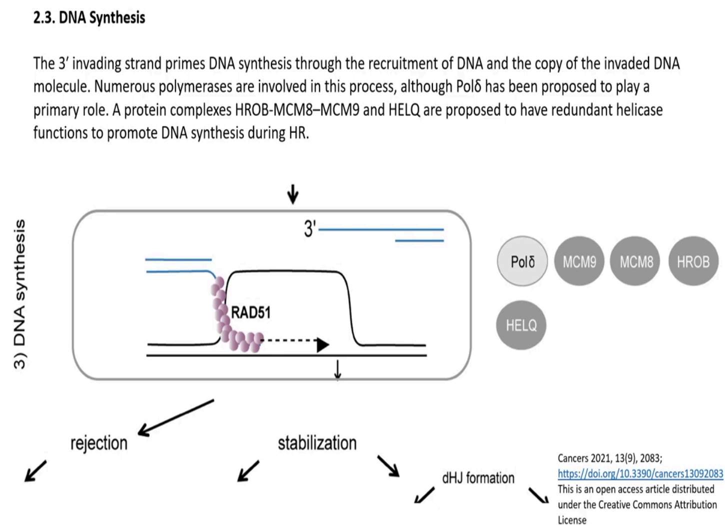
So, here also for there will be loading of the RAD51 on the single stranded DNA and this complex will search for homology and helps in strand invasion. The 3 prime ssDNA stretch created by resection is coated with replication protein A and it protects it from digestion. You can see here the involvement of the RPA proteins. The loading of RAD51 on the ssDNA is performed by the BRCA1-PALB2 complex ok.

So, you have the BRCA BRCA2-PALB2 complex over here sorry not this one here. So, this protein complex interacts with BRCA1 and catalyzes the replacement of RPA by RAD51 on the stretch of the 3 prime ssDNA creating the RAD51 ssDNA pre-synaptic complex. The ssDNA RAD51 filament scans the genome to search for homology and on success the filament in which the duplex homologous DNA and initiate strand exchanges creating a displacement loop.

So, you can see here the events taking place sequentially. First there is DSB recognition and resection and here are the essential factors for this particular step and these steps these are not essential in this stage. The second step is the strand invasion and D loop stabilization where BRCA1, BRCA2, PALB2, RAD51 and so many different players are critical for carrying the

process forward, while these three molecules at the end are not essential factors non-essential factors.

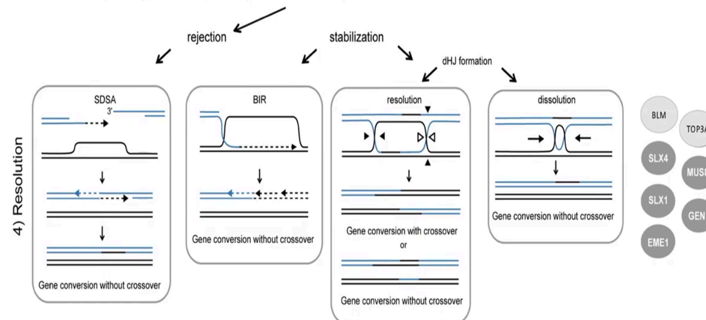
(Refer Slide Time: 37:39)



How DNA synthesis happens once the strand invasion is successful, because it found out a homology sequence and the structure is now stabilized and DNA synthesis can takes place. The DNA 3 prime invading strand will act like a primer. So, it will prime DNA synthesis through the recruitment of DNA and the copy of the invaded DNA molecule.

Numerous polymerases are involved in this process although delta polymerase has been proposed to play a primary role in this process. A protein complex HROB, MCM8, MCM9 and HELQ are proposed to have redundant helicase functions to promote DNA synthesis during this homologous recombination.

(Refer Slide Time: 38:40)



Once the DNA synthesis starts and completed the formation and resolution of the intermediates HR intermediates will have to take place. The strand invasion in DNA synthesis lead to the formation of different intermediates whose processing leads to gene conversion either associated with crossover products or not.

The invading strand can be disassembled, channelling DSB repair towards synthesis dependent strand annealing called SDSA. If stabilized the D loop can lead to DSB repair by break induced repair or to the formation of double Holliday junctions that can be either dissolved by the BLM-TOP3A-RMI1/2 complex or resolved by the structure specific resolvases.

(Refer Slide Time: 39:33)

BLM is a RecQ family DNA helicase. It is mutated in Bloom syndrome and plays several roles, sometimes contradictory.

For eg. It has been shown to suppress crossovers in mitotic cells while repair mitotic DNA gaps via HR*.

BLM is involved in different steps of HR, including end resection at HR initiation, D-loop rejection and double Holliday Junction (dHJ) resolution at HR termination.

At resection initiation, depending on the cell cycle phase that modifies its interacting partners, BLM either favors the loading of 53BP1 on the DSB in G1 phase, preventing the initiation of unscheduled resection, or, in contrast, favors resection in S phase when interacting with TOP3.

Cancers 2021, 13(9), 2083;https://doi.org/10.3390/cancers13092083

*Genetics. 2017 Nov; 207(3): 923–933.

This BLM is a RecQ family DNA helicase. It is mutated in Bloom syndrome and plays several roles sometimes which are contradictory. For example, it has been shown to suppress crossovers in mitotic cells while repair mitotic DNA gaps by a homologous recombination. BLM is involved in different steps of homologous recombination including end resection at HR initiation, D loop rejection and double Holliday junction resolution at the HR termination.

At restriction initiation depending on the cell cycle phase that modifies its interacting partners, BLM favors the loading of 53BP1 on the DSB in G1 phase, preventing the initiation of unscheduled resection or in contrast favors resection in S phase when interacting with TOP3. Besides these players there are several other accessory proteins, which are required in the process.

(Refer Slide Time: 40:40)

2.5. Accessory Proteins

RAD54, a member of the SWI2/SNF2 protein family (ATP-dependent chromatin remodelers), interacts with RAD51, and in vitro studies have proposed that it functions as a RAD51 cofactor. RAD54 catalyzes the extension of joint molecules and stabilizes the D-loop.

A family of six proteins (RAD51B, RAD51C, RAD51D, XRCC2, XRCC3, and RAD51AP1), known as the RAD51 paralogs (i.e., proteins that share sequence homology with RAD51 in a given species), has been identified in mammals. Two distinct complexes have been identified: RAD51B–RAD51C–RAD51D–XRCC2 (BCDX2) and RAD51C–XRCC3 (CX3).

RAD51 paralogs favor the recruitment of RAD51 to DNA damage sites and promote the formation and stabilization of the RAD51 nucleoprotein filament.

RAD51 paralogs also influence gene conversion tract length.

The SWSAP1 protein, a noncanonical paralog of RAD51, forms the so-called SHU complex when associated with SWS1 (SWSAP1-SWS1). SHU interacts with RAD51 and regulates its function.

Cancers 2021, 13(9), 2083; https://doi.org/10.3390/cancers13092083

For example, we need RAD54 which is a member of the SWI2/SNF2 protein family which are an ATP dependent chromatin remodels. It interacts with RAD51 and in vitro studies have proposed that it functions as a RAD51 co-factor. RAD54 catalyzes the extension of joint molecules and stabilizes the D loop.

A family of six proteins RAD51B, C, D, XRCC2, 3 and RAD51AP1 known as the RAD51 paralogs, which are proteins that share sequence homology with RAD51 in a given species have been identified in mammals. Two distinct complexes have been identified, RAD51B-C–D-XRCC2 and RAD51C-XRCC3. RAD51 paralogs favored the recruitment of RAD51 to DNA damage sites and promotes the formation and stabilization of the RAD51 nucleoprotein filament.

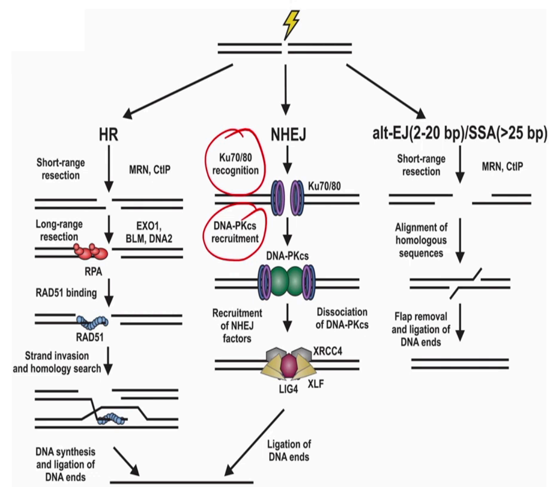
The RAD51 paralogs also influence gene conversion tract length. The SWSAP1 protein, a non-canonical paralog of RAD51 forms the so-called SHU complex when associated with SWS1. SHU interacts with the RAD51 and regulates its function.

(Refer Slide Time: 42:11)

NHEJ

NHEJ.

(Refer Slide Time: 42:17)



Let us discuss about the NHEJ pathway now. You will know how the homologous recombination pathway operates. Now, we are going to discuss about the NHEJ pathway and the various players involved in it Ku70/80 then DNA-PKcs and so on.

(Refer Slide Time: 42:40)

Canonical non-homologous end joining (C-NHEJ) pathway.

C-NHEJ depends on Ku heterodimer and DNA-PK catalytic subunit (DNA-PKcs), which together form the DNA-PK holoenzyme. Ku is a first responder at DSBs and provides a docking site for DNA-PKcs.

Unlike MRN, which can bind internally, Ku requires a free DNA end for binding and cannot associate with most blocked ends.

Several nucleases, including tyrosyl-DNA phosphodiesterase 1 and 2 (TDP1/2) and Artemis, can remove hairpins, damaged bases or protein-DNA adducts.

The DNA ends are processed by additional enzymes and rejoined by the LIG4/XRCC4/XLF complex

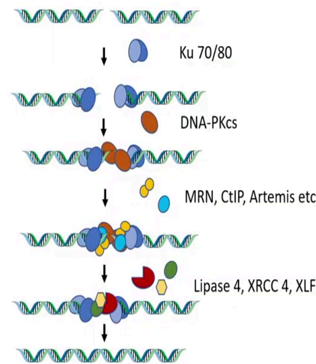


Figure Adapted from Transl Cancer Res 2013;2(3):163-177

Front. Cell Dev. Biol., 12 July 2021 | https://doi.org/10.3389/fcell.2021.708763

The canonical non-homologous end joining pathway depends on Ku heterodimer and DNA-PK catalytic subunit, which together form DNA-PK holoenzyme. Ku is the first responder at DSBs and provides a docking site for DNA-PKcs. Without the presence of Ku70 80 in the breakage site DNA-PKcs cannot be recruited to it.

Unlike MRN which can bind internally Ku requires a free DNA and for binding and cannot associate with most block dense. Several nucleases including tyrosyl-DNA, phosphodiester 1 and 2 and Artemis can remove hairpins damaged bases or protein-DNA adducts. The DNA ends are processed by additional enzymes and rejoined by the LIG4, XRCC4 and XLF complex. So, you can see here the various players like MRN, CtIP and Artemis etcetera.

(Refer Slide Time: 44:01)

Corresponding Enzymes in Prokaryotic and Eukaryotic NHEJ Components

Functional component	Prokaryotes	Eukaryotes	
		S. cerevisiae	Multicellular eukaryotes
Toolbelt protein	Ku (30-40kDa)	Ku 70/80	Ku 70/80
Polymerase	POL domain of LigD	Pol4	Polmu and lambda
Nuclease	?	Rad50:Mre11:Xrs2 (FEN-1)	Artemis:DNA-PKcs
Kinase/Phosphatase	PE domain of LigD	Tpp1 and others	PNK and others
Ligase	LIG domain of LigD	Nej1:Lif1:Dnl4	XLF:XRCC4:DNA ligase IV

Annu Rev Biochem. 2010; 79: 181–211.

Let us now compare the enzymes present in prokaryotes and eukaryotes which carries out the non-homologous end joining process. So, here you have the prokaryotes and for the toolbelt protein you have the Ku protein which is around 30 to 40 kilodalton and in eukaryotes you have the Ku 70 80 in *Saccharomyces cerevisiae*, while in multicellular eukaryotes also you have the similar kind of molecule. The polymerase action or the carried out by the POL domain of the LigD.

Here it is carried out by Pol4 *Saccharomyces cerevisiae* and in higher eukaryotes a Polmu and lambda are involved. The nucleus activity in prokaryotes it is still not understood although many reports have suggested certain candidate proteins. In eukaryotes it is the Rad50:Mre11:Xrs2 complex while in multicellular organisms it is the Artemis DNA-PKcs.

The kinase phosphatase activity lies in the PE domain of LigD in prokaryotes, in eukaryotes like *Saccharomyces cerevisiae* it lies in Tpp1 and others. And in multicellular organisms it is present in the PNK. The ligase activity in prokaryotes again lies in the Lig domain of Lig2. In eukaryotes like *Saccharomyces cerevisiae* it lies in the Nej1 Lif1and Dnl4, while multicellular organisms it lies in XLF:XSCC4:DNA ligase IV.

So, from this comparison we can see the diversity of the protein molecules and players which are involved in prokaryotes and also within eukaryotes, lower eukaryotes and higher eukaryotes have different proteins which carry out certain defined functions in the NHEJ pathway. The Ku is a critical player without which the NHEJ repair cannot begin.

(Refer Slide Time: 46:53)

Ku is an abundant, highly conserved DNA binding protein found in both prokaryotes and eukaryotes that plays essential roles in the maintenance of genome integrity.

In eukaryotes, Ku is a heterodimer comprised of two subunits, Ku70 and Ku80, that is best characterized for its central role as the initial DNA end binding factor in the "classical" non-homologous end joining (C-NHEJ) pathway, the main DNA double-strand break (DSB) repair pathway in mammals.

Ku binds double-stranded DNA ends with high affinity in a sequence-independent manner through a central ring formed by the intertwined strands of the Ku70 and Ku80 subunits.

At the break, Ku directly and indirectly interacts with several C-NHEJ factors and processing enzymes, serving as the scaffold for the entire DNA repair complex.

Clark et al. Molecular Biology 2019 Elsevier Inc.

This is abundant and highly conserved DNA binding protein found in both eukaryotes and prokaryotes and it plays essential role in the maintenance of the genome integrity. In eukaryotes Ku is a hetero dimmer comprised of two subunits Ku 70 and Ku 80. And it is best characterized for its central role as the initial DNA and binding factor in the classical non-homologous end joining pathway.

The main DNA double strand break repair pathway in mammals. Ku binds the double stranded DNA ends with high affinity in a sequence independent manner, through a central ring formed by the intertwined strands of the Ku 70 and Ku 80 subunits. At the break Ku directly and indirectly interacts with several C-NHEJ factors and processing enzymes serving as the scaffold for the entire DNA repair complex.

(Refer Slide Time: 47:44)

General Steps of Nonhomologous DNA End Joining

Ku binding to the DNA ends at a DSB improves binding by nuclease, polymerase and ligase components.

Flexibility in the loading of these enzymatic components, the option to load repeatedly (iteratively), and independent processing of the two DNA end all permit mechanistic flexibility for the NHEJ process.

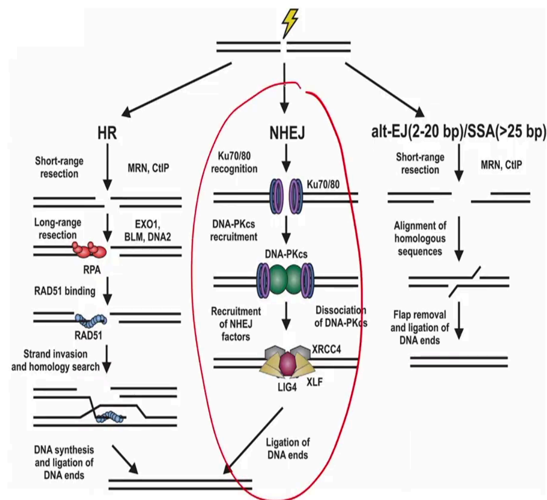
This mechanistic flexibility is essential to permit NHEJ to handle a very diverse array of DSB end configurations and to join them. In addition to the overall mechanistic flexibility, each component exhibits enzymatic flexibility and multifunctionality.

Annu Rev Biochem. 2010; 79: 181–211.

What are the general steps of non-homologous DNA end joining? It starts with the binding of Ku to the DNA ends at the DSB and it improves the binding by nucleases polymerases and ligase components there upon. Flexibility in the loading of these enzymatic components the option to load repeatedly and independently processing of the two DNA ends all permit mechanistic flexibility to the NHEJ process.

This mechanistic flexibility is essential to permit NHEJ to handle a very diverse area of DSB and configurations and to join them. In addition to the overall mechanistic flexibility each component exhibits enzymatic flexibility and multi functionality.

(Refer Slide Time: 48:36)



So, now you know about the homologous recombination and non-homologous recombination repair pathway. In addition to these conventional NHEJ and homologous recombination repair there is a third pathway of double strand break repair which we call as alt-EJ or A-NHEJ.

(Refer Slide Time: 48:43)

In addition to C-NHEJ and HRR, there is a third pathway of DSB processing, functioning on simple end-joining principles, but repairing DSBs in a slower speed than C-NHEJ.

This repair pathway is considered to be an alternative form of NHEJ and is frequently abbreviated as A-NHEJ, or simply A-EJ.

A-NHEJ or B-NHEJ

A-EJ is suppressed by C-NHEJ, and possibly also by HRR, and becomes active only when these standard repair processes fail, globally or locally. Therefor it is considered to be a backup pathway and has been abbreviated as B-NHEJ in many instances

Transl Cancer Res 2013;2(3):163-177

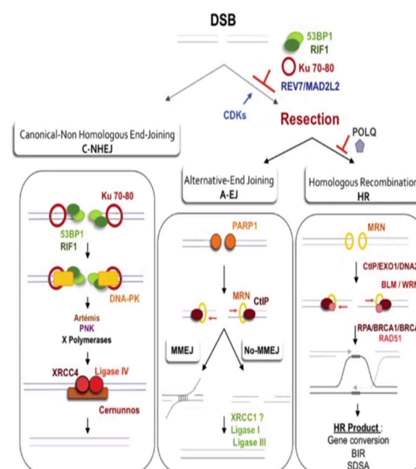
It is functioning on simple end joining principles, but repairing DSB in a slower speed than conventional NHEJ. This repair pathway is considered to be an alternative form of the NHEJ as already told to you and is abbreviated as A-NHEJ. A-NHEJ or B-NHEJ: A-NHEJ or A-EJ

is supposed suppressed by C-NHEJ. The conventional NHEJ will not allow the alternative NHEJ to operate and it is also possibly suppressed by homologous recombination repair pathway. It is the least preferred pathway because of the suppression by the other two pathways.

And it becomes active only when the standard repair processes fail. The C-NHEJ and HRR fails to take care of the genome integrity under that circumstances only alternative NHEJ will come into action. And this repair may take place at the local level to the global level. For this region this is considered as a backup pathway. If conventional NHEJ and HRR fails then alternative NHEJ comes into action as a backup mechanism.

So, for this region it is also sometimes referred to be called as B-NHEJ by certain groups of people.

(Refer Slide Time: 50:47)

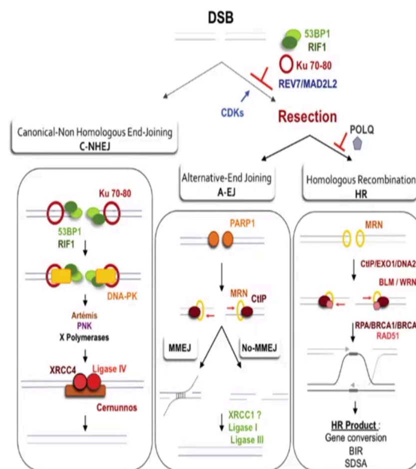


Several factors have been implicated in A-EJ and their functional diversity has led to the postulate that there are several sub-pathways in operation, engaging prospectively at each DSB on the basis of as of yet undefined parameters in competition with other repair pathways.

Gelot et al., Genes 2015, 6(2), 267-298; https://doi.org/10.3390/genes6020267, CC BY 4.0

So, several factors have been implicated in alternative NHEJ and the functional diversity has led to the postulation that. There are several sub pathways in this operation. It is not just one single pathway. And it engages prospectively at each double strand break on the basis of as yet undefined parameters in competition with other repair pathways. So, although it is the last one to operate when the first two fails, the entire landscape of its operation and mechanism is still largely unexplored.

(Refer Slide Time: 51:27)



A-EJ will engage at DSBs where either C-NHEJ or HRR have attempted processing but somehow failed. Thus, at each DSB where A-EJ engages, factors of either C-NHEJ or HRR, particularly those involved at early steps, will be present when A-EJ takes DSB processing over.

Also, it is possible, and even likely, that these factors have already operated at DNA ends and have carried out one or more of the initial steps of C-NHEJ or HRR, which of course alters the state of the substrate presented to A-EJ.

Furthermore, the presence of C-NHEJ and HRR factors at the DNA ends may either facilitate or compromise A-EJ.

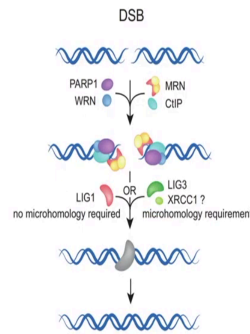
Gelot et al., Genes 2015, 6(2), 267-298; https://doi.org/10.3390/genes6020267. CC BY 4.0

So, this A-EJ will engage at DSBs as already told you when either of the two pathways have somehow failed to act. Does it each DSB where A-EJ engages factors of either C-NHEJ or HRA HRR particularly those involved in early steps will be present when this alternative NHEJ takes DSB processing over.

So, C-NHEJ and HRR try to repair the DNA, so those players are present, but they fail to do it then this process is taken over by alternative pathway. So, in the alternative pathway we will always find the presence of the initial players involved in the earlier two processes. Also it is possible and even likely that these factors have already operated at the DNA ends, so resection has taken place.

And it has carried out one or more of the initial steps of the C-NHEJ and HRR which of course, alters the state of the substrate presented to the alternative NHEJ. Furthermore the presence of the C-NHEJ and HRR factors is the DNA ends may either facilitate or compromise the operation of A in A alternative NHEJ.

(Refer Slide Time: 52:53)



When the engagement of A-EJ follows failure of C-NHEJ, several of the early C-NHEJ factors may be present at the junction, but the process must be abrogated prior to ligation by LIG4.

End ligation is carried out with either of the remaining ligases, LIG3 and LIG1.

PARP1 is a sensor for DNA discontinuities, originally shown to operate in base excision and single-strand break repair. Previous work implicated PARP1 also in repair by A-EJ. There is even evidence for competition between Ku and PARP1 for DSBs raising the possibility that pre-existing C-NHEJ factors at the DSB compromise A-EJ.

DNA end stabilization provided in C-NHEJ by Ku may be provided in A-EJ by histone H1. However, it should be emphasized that to date the evidence for a role of histone H1 in A-EJ is of purely biochemical nature.

Transl Cancer Res 2013;2(3):163-177

When the engagement of this alternative EJ follows failure of C-NHEJ several of the early C-NHEJ factors are present as already told you, but the process must be abrogated prior to the ligation of the LIG4. End ligation is carried out with either of the remaining ligases; ligase 3 and ligase 1.

PARP1 is a sensor for DNA discontinuities originally shown to operate in base excision and single strand brake repair. Previous works have implicated PARP1 in the repair of this alternative EJ there is even evidence for competition between Ku and PARP1 for DSBs raising the possibility that pre existing conventional NHEJ factors that the DSB compromise the alternative EJ pathway.

DNAs tend end stabilization provided by the conventional NHEJ by Ku may be provided in alternative EJ by histone H1. However, it should be emphasized that to date the evidence for a role of histone in alternative EJ is of purely biochemical nature.

(Refer Slide Time: 54:09)

A-EJ is considered to be a mechanistically distinct repair pathway, and has been shown to be active throughout the cell cycle.

It is markedly enhanced in the G2 as compared to G1 phase, and is compromised in stationary-phase cells tested either in the G1 or G2 phase of the cell cycle.

There are speculations that the latter response may be regulated by phosphorylation of BRCA1 at S988 through Chk2, where in its phosphorylated form BRCA1 promotes error-free NHEJ and suppresses mutagenic A-EJ.

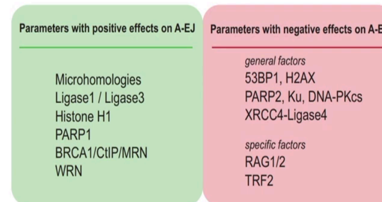
Therewith, it reduces the size of deletions at the breakpoint junction. However, this dependency is more likely in G2 than in G1 cells as BRCA1/CtIP/MRN initiates DSB resection during S/G2 phases, and therefore alternative mechanisms should be explored.

Transl Cancer Res 2013;2(3):163-177

These alternative EJ is considered to be mechanistically distinct pathway and it is shown to be active throughout the cell cycle. It is markedly enhanced in the G2 phase as compared to the G1 phase and is compromised in stationary phase cells tested either in the G1 or G2 phase of the cell cycle.

There are speculations that the later responses may be regulated by phosphorylation of BRCA1 at S988 through Chk2, where in its phosphorylated form BRCA1 promotes error free NHEJ and suppresses the mutagenic alternative EJ. There it reduces the size of deletions at the breakpoint junctions. However, this dependency is more likely in G2 then in G1 cells as BRCA1, CtIP and MRN initiates DSB resection during S/G2 phases and therefore, alternative mechanisms should be explored.

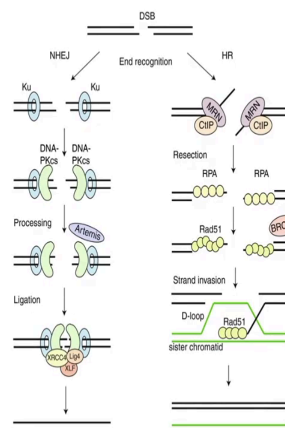
(Refer Slide Time: 55:06)



Transl Cancer Res 2013;2(3):163-177

There are certain parameters which have positive effects on the alternative EJ and there are certain others which have negative impacts on them. The positive ones are micro homologies, the presence of ligases 1 and 3, histone H1, PARP1, BRCA1/CtIP/MRN and WRN and we have discussed about the involvement or presence of these factors in the alternative EJ. There are other factors like the P53 or 53BP1, H2AX, PARP1, Ku, DNA ligase etcetera are reported to have negative effects.

(Refer Slide Time: 55:51)



Non Homologous End Joining. It starts with recognition of the DNA ends by the Ku70/80 heterodimer, which recruits DNA-PKcs. If the ends are incompatible, nucleases such as Artemis can trim the ends. The XRCC4-DNA Ligase IV-XLF ligation complex seals the break.

Homologous Recombination. The MRN-CtIP-complex starts resection on the breaks to generate single stranded DNA (ssDNA). After resection the break can no longer be repaired by NHEJ. The ssDNA is first coated by RPA, which is subsequently replaced by Rad51 with the help of BRCA2. These Rad51 nucleoprotein filaments mediate strand invasion on the homologous template. Extension of the D-loop and capture of the second end lead to repair.

Genome Integrity volume 3, Article number: 9 (2012). This is an Open Access article distributed under the terms of the Creative Commons Attribution License

So, finally, let us once again visit the two main DNA repair pathway especially for DSBs. We have the canonical or conventional, NHEJ, repair pathway and the homologous repair pathway. You now know what are the protein molecules involved in taking the decision which pathway it will follow and you know that Ku is an important player in NHEJ.

And in HR you have a different set of important critical players and the formation of certain invading strands takes place in the homologous recombination. In brief in the homologous end joining starts with the recognition of the DNA strand by the Ku 70/80 heterodimer, which recreates the DNA PKcs.

If the ends are incompatible, nucleases such as Artemis can trim the ends. The XRCC4-DNA ligase IV-XLF ligation complex will seal the break in this process. In homologous recombination the MRN-CtIP complex starts resection on the breaks to generate single stranded DNA. After resection of the breaks it can no longer be repaired by NHEJ and it has to follow the HR pathway.

The ssDNA is first coated by RPA, which is subsequently replaced by Rad51 with the help of BRCA2. These Rad51 nucleoprotein filaments mediate strand invasion on the homologous template. Extension of the D-loop in the capture of the second end lead will lead to and will lead to repair.

Thank you.