- Upstream of the flanking sequence is the TATA box, which is found in the promoter region and binds RNA Polymerase II.
- Upstream of the TATA box is an octamer motif, which binds regulatory transcription factors.
- Association with enhancers will increase the rate of transcription and the promoters are relatively inactive

V(D)J recombinase

- The actual process of V(D)J recombination is mediated by V(D)J recombinase, a complex comprised of lymphocyte-specific enzymes and some enzymes that are ubiquitously expressed.
- RAG-1 (recombination-activating genes) and RAG-2, lymphocyte-specific enzymes, recognise nucleotide sequences known as recombination signal sequences (RSS), which flank the V, D and J gene segments.
- RAG-1 and RAG-2 are encoded by recombination-activating genes and are referred to as recombinases.

RSS

- The RSS is a noncoding sequence made up of conserved heptamer and nonamer sequence, which themselves are separated by an unconserved (in length, but not sequence) 12- or 23-nucleotide spacer.
- Efficient recombination occurs between segments with a 12-nucleotide spacer and a 23-nucleotide spacer. This '12/23' rule helps make certain that appropriate gene segments are joined together and simply states that you can only recombine a gene segments that has a 12bp spacer next to a 23bp spacer.
- On the heavy chain gene locus, RSS are found downstream of each V_H gene segment, on both sides of the D_H gene segment and then upstream of each J gene segment.
- On the light chain gene locus, they are found downstream of each V gene segment and then upstream of each J gene segment.

Mechanism

- RAG-1 and RAG-2 bind to the RSS and the high-mobility group (HPG) proteins bend the double-stranded DNA to bring together the two RSSs.
- RAG-1 and RAG-2 cleave the DNA between the coding sequence and the 5'-heptameric end of the RSS, creating a single-stranded break that attacks the opposite anti-parallel DNA strand in a transesterification reaction.
- This leads to the creation of a hairpin loop containing intervening sequences, which is repaired by non-homologous end-joining machinery. The coding sequences are then also religated.
- This happens when a D gene segment is joined to a J gene segment and then subsequently when V joins to DJ.

<u>Recombinatorial inaccuracies increase diversity</u>

When the DNA is cut between D and J gene segments and subsequently V and DJ, the exact junction can vary, as the cutting does not occur at a precise location, and therefore the coding ends are joined imprecisely. These are referred to as recombinatorial inaccuracies or junctional inaccuracies, a mechanism that creates further diversity.

N-nucleotide addition

This is another mechanism that creates further diversity. Following the cut, sometimes the enzyme terminal deoxynucleotidyl transferase (TdT) picks up a small number of nucleotides at random and adds them to the ends of the cut DNA before it is religated.

Somatic hypermutation

This is a mechanism specific to only immunoglobulin genes and occurs in rapidly proliferating B-cells in the lymphoid follicles after antigen activation of B-cells. Following activation, the immunoglobulin gene undergoes somatic mutation that is far greater than the normal rate of mutation. Mutations are usually single-base substitutions and insertions and deletions are less common. The area of this mutation is very precisely targeted to the hypervariable regions, but can also occur in framework regions. This mechanism can result in an increased affinity of the BCR to its cognate ligand in what is known as affinity maturation. The enzyme responsible for somatic hypermutation is known as activation-induced cytidine deaminase (AID), which is also very important for class-switch recombination.

Class-switch recombination

This is a mechanism by which the heavy chain constant region exon of a given antibody is exchanged for another exon downstream of that. This gives rise to antibodies of the same antigen specificity but different effector functions.
Class switch recombination has been studied extensively in mice. It is known that the heavy chain exons for IgD, IgG, IgE and IgA are located downstream of the heavy chain exon for IgM.

- CSR occurs between switch (S) regions, which are repetitive sequences often G-rich, that are found upstream of each C_H exon, with the exception of the constant exon for IgD. This is because IgM and IgD are always produced together.
- Breaks are introduced into the DNA of two S regions and fusion of the S regions leads to a rearranged locus in which the variable region exon is now joined to a new constant exon.
- The DNA between the two S regions is removed.
- Alternative splicing of the primary RNA transcript then gives rise to a new membrane-bound or soluble immunoglobulin. Once class-