

Lecture 6: Genetic basis of Antibody Structure.

Chapter 6, Benjamini et al.

The following observations have to be accounted for:

1. Each B-cell produces antibodies that are homogenous in specificity.
2. At any given time there are about 10^8 different specificities present in the average human (or mouse)
3. Diversity in amino acid sequence is generally limited to the three hypervariable regions on the light and heavy chains.
4. The amino acid sequences in the third hypervariable regions on both the heavy and light chains are more diverse than the first two hypervariable region. The third hypervariable region on the heavy chain is more diverse than that on the light chain.
5. The specificity is maintained, even if the isotype (e.g. IgG and IgA) of the antibody changes.
6. The constant regions of the different isotypes are the same in all antibodies of the same isotype from a single individual.
7. Changes in isotype follow a regular sequence.
8. Once an isotype is no longer produced, it can never be produced by the same cell.
9. Membrane and soluble forms of the same isotype can exist, even from the same cell.
10. IgM and IgD are co-expressed on the surface of mature, no activated B-cells
11. The affinity is increased during the maturation of the B-cell after stimulation with the foreign antigen.

Early Models:

1. Each different polypeptide is produced by its own gene.
2. There are only a few immunoglobulin genes, and the diversity is generated by a high mutation rate in the B-cell during its development.
3. Two gene model (Dryer and Bennett, 1965). There are two genes that are used to construct a single functional immunoglobulin gene, one for the variable region and one for the constant region. Many different variable region 'cassettes' exist that can be combined with the constant region.

In 1976 Tonegawa showed that the genomic DNA that encodes light chain DNA changes its restriction map when stem cell DNA is compared to DNA from a mature B-cell.

Subsequent DNA sequencing produced the following structure of the kappa light chains in mice. Humans have a similar organization of DNA, some of which are discussed in Benjamini et al. (Figure 6.2).

A comparison of the protein sequences of κ -chains to the DNA sequence revealed that the κ -chain locus consists of a series of approximately 300 variable regions (with a leader exon), indicated by 'L V κ 1',

followed by 23 kb of DNA, five short segments of DNA called J segments that corresponded to CDR3, 2.5 kb of DNA and then DNA that encodes the constant region of the κ light chain. (The ψ indicates a pseudogene)

Based on the DNA sequence in mature B-cells it became clear that the following sequence of events occur:

1. One of the 300 L-V segments joins, at random, to one of the 4 J segments, leading to altered chromosomal DNA.
2. mRNA splicing removes the intro between the L-V segment and the J and C κ segment
3. The peptide encoded by the leader exon is removed by the leader peptidase as the growing light chain is synthesized on the rough endoplasmic reticulum.

The high diversity in CDR3 of the light chain is due to the fact that this sequence comes from the J-segment. This fact in itself does not generate diversity (there are only 4 J segments) but the DNA joining event is imprecise (see below) leading to considerable diversity.

The λ -chain DNA and the DNA encoding the heavy chain show a similar motif. In the mouse the λ -locus is less diverse, showing only two V segments and 3 functional J domains.

The heavy chain locus contains additional segments that are not seen in the light chain:

- D domains that are found between the V segments and the J domain. Both the D and J segments give rise to CDR 3 in the heavy chain.
- A series of segments that specify the different isotypes of the heavy chain appear to the 3' end of the J segments.

The sequence of events that occur to generate a viable heavy chain gene are as follows:

1. One of the 13 D segments joins to one of the four J regions, generating a DJ junction.
2. One of the 300-100 L-V segments joins to new DJ segment.
3. Introns are removed between the L segment and the V segment by mRNA splicing
4. Alternate splicing generates either an IgM (μ chain) or IgD (δ chain)

Combinatorial Diversity.

Number of possible λ -chains: $2 (V) \times 3 (J) = 6$ different λ light chains possible

Number of possible κ -chains: $300 (V) \times 4 (J) = 1200$ different κ light chains possible

Number of possible H-chains: $500 (V) \times 13 (D) \times 4 (J) = 26,000$ different V_H possible.

Total diversity, assuming all light chains can pair with all heavy chains:

$$26,000 \times 1206 = 3 \times 10^7 \text{ different antibodies}$$

Mechanism of Joining:

Need to insure that the right segments are joined. For example, you would not want to join two J segments together instead of a VJ joining event. There are two conserved recombination signal sequences (RSS) found directly adjacent to the V, J, and D regions in immunoglobulins:

CACAGTG - 23 bp - ACAAAAACC

GTGTCAC - 23 bp - TGTTTTTGG

Two-turn sequence

GGTTTTTGT - 12 bp - CACTGTG

CCAAAAACA - 12 bp - GTGACAC

One-turn sequence

When fragments are joined, a two-turn RSS can only join with a one-turn RSS. The position of these RSS in the heavy and light chain segments are as follows:

λ -chain: $V\lambda$ -2 RSS 1 RSS $J\lambda$

κ -chain: $V\kappa$ -1 RSS 2 RSS $J\kappa$

H-chain: V_H - 2 RSS 1 RSS - D - 1RSS 2 RSS - J

Diversity in Joining:

1. The end points of the RSS are the nominal points for joining heavy and light chains. However, there is some uncertainty (up to 5 bases) where the crossover point occurs. These leads to the loss of codons, and subsequent amino acids from CDR3.
2. Addition of up to 15 nucleotides, by the enzyme terminal transferase, adds codons to CDR3. The added nucleotides are referred to as N nucleotides.

The loss and gain of codons from the joining regions increases the diversity by 10-1000 fold.

Productive Rearrangements and Allelic Exclusion:

The mature B-cell:

1. must produce a functional antibody (productive rearrangement)
2. will only produce a single light and heavy chain, although the genes from both parents are present in the immature B-cell (Allelic exclusion).

The steps that insure that the above happens are as follows:

1. Cell attempts to make a functional heavy chain, if this fails the cell dies.
2. Once a single heavy chain is functional, rearrangement of the other heavy chain allele is inhibited.
3. Successful rearrangement of one κ -chain inhibits rearrangement of the other κ allele
4. Unsuccessful rearrangement of both κ -chains prompts rearrangement of λ -chains.
5. Unsuccessful rearrangements of λ -chains causes death.

Class Switching:

- There are recombinational signals in the DNA called switch regions.
- In mice these are located 5' to $C\mu$ ($S\mu$), between $C\delta$ and $C\gamma3$ ($S\gamma3$), between $C\gamma3$ and $C\gamma1$ ($S\gamma1$), and 5' to $C\epsilon$ ($S\epsilon$)
- Recombination between these signals results in deletion of the intervening H-chain gene and the use of a new H-chain constant region.

Membrane Bound or Secreted:

The 3' end of *every* constant region consists of the following structure (example shown for IgM, μ gene)

$\mu 4$ -S---polyA ----- M1 ----- M2 -polyA

1. $\mu 4$ is the last exon, coding for the 4th constant domain in IgM (the hinge region in IgM is replaced by another immunoglobulin domain).
2. S codes for the carboxy-terminal sequence found in soluble IgM
3. polyA: poly A cleavage and addition sites
4. M1, M2 exons that code for the membrane domain.

Alternate use of the two poly A sites control expression of soluble IgM ($\mu 4$ -S) or membrane bound ($\mu 4$ -M1-M2) IgM.

Co-expression of IgM and IgD.

These two isotypes are expressed at the same time in the cell. The two constant domain gene segments are adjacent to each other on the chromosome. The 3' end of the μ gene and the 3' end of the δ gene have two polyadenylation sites (as discussed above). Alternative use of these four sites (and alternative splicing) leads to expression of IgM or IgD.

Somatic Mutation:

After stimulation with antigen, affinity maturation occurs. The CDR1 and CDR2 regions of immunoglobulins acquire mutations. Since those B-cells with higher affinity antibodies on the surface are stimulated by antigen, the resultant B-cell population that is selected has higher affinities. The mechanism for this event is current unknown.