Biomolecular chemistry

7. Antibodies: structure and function

Suggested reading: **Sections 5.1 to 5.3** of
Mikkelsen and Cortón, *Bioanalytical Chemistry*

Primary Source Material
- Biochemistry Chapter 33: Berg, Jeremy M.; Tymoczko, John L.; and Stryer, Lubert (NCBI bookshelf).
- Many figures and the descriptions for the figures are from the educational resources provided at the Protein Data Bank (http://www.pdb.org/)
- Most of these figures and accompanying legends have been written by David S. Goodsell of the Scripps Research Institute and are being used with permission. I highly recommend browsing the Molecule of the Month series at the PDB (http://www.pdb.org/pdb/101/motm_archive.do)

Cellular components of blood

- red blood cells
- white blood cells
- platelets
- plasma

http://www.emc.maricopa.edu/faculty/farabee/BIOBK/
BioBookcircSYS.html - Blood
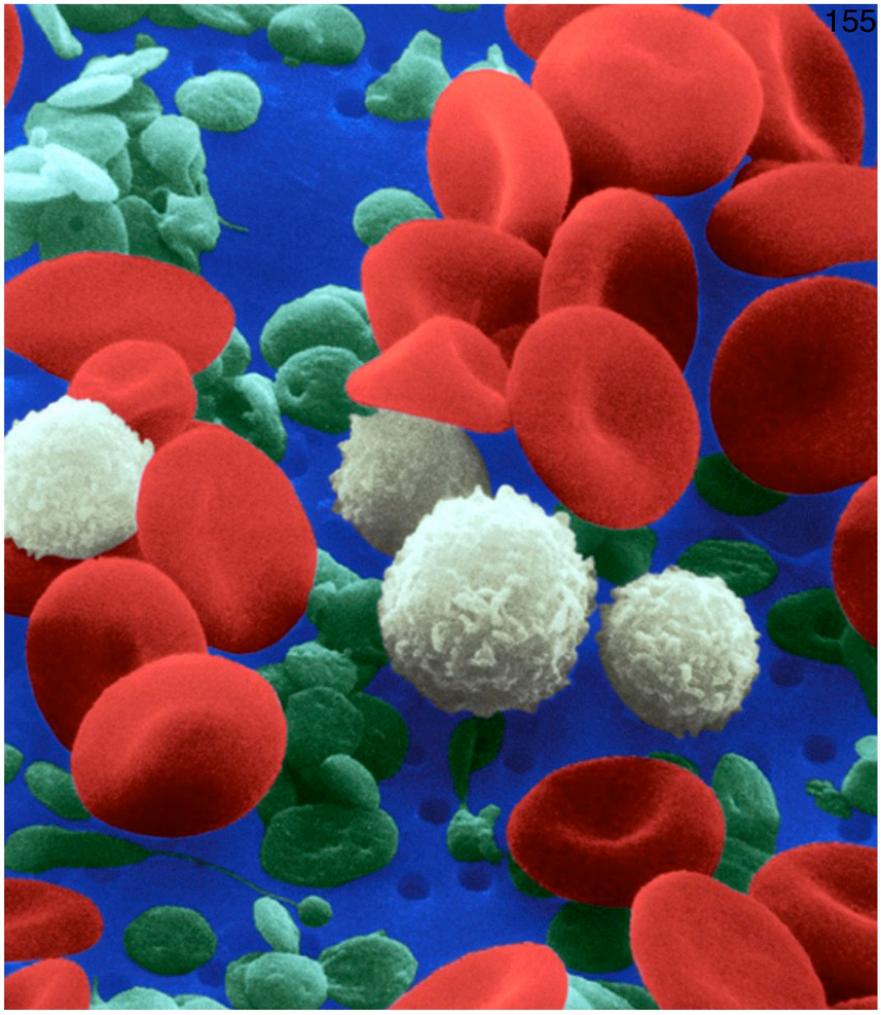
- Mammalian blood consists of plasma and a number of cellular and cell fragment components.
- Plasma: The liquid part of the blood, which makes up about half of its volume. Blood plasma contains antibodies and other proteins. It is taken from donors and made into medications for a variety of blood-related conditions. Plasma has 90% water and 10% dissolved materials including proteins, glucose, ions, hormones, and gases. It acts as a buffer, maintaining pH near 7.4. Note that serum is essentially similar in composition to plasma but lacks fibrinogen and other substances that are used in the coagulation (blood clotting) process.
- Red blood cells, also known as erythrocytes, are flattened, doubly concave cells about 7 μm in diameter that carry oxygen associated in the cell's hemoglobin. Mature erythrocytes lack a nucleus.
- White blood cells, also known as leukocytes, are larger than erythrocytes, have a nucleus, and lack hemoglobin. They function in the cellular immune response.
- Platelets result from cell fragmentation and are involved with clotting. Since they are fragments of a larger cell, they contain no DNA. However, they still contain a large number of enzymes that are capable of carrying out critical functions for the lifetime of the platelet (about a week). Interestingly, aspirin irreversibly inhibits a key enzyme in platelets (cyclooxygenase). This permanently inactivates the platelets and makes them unable to perform their normal clotting function. Since heart attacks are caused by blood clots that limit the supply of blood to the walls of the heart, aspirin can have a preventative effect.

The major serum proteins are albumin and immunoglobulin

Hemoglobin

Immunoglobulin G (IgG)
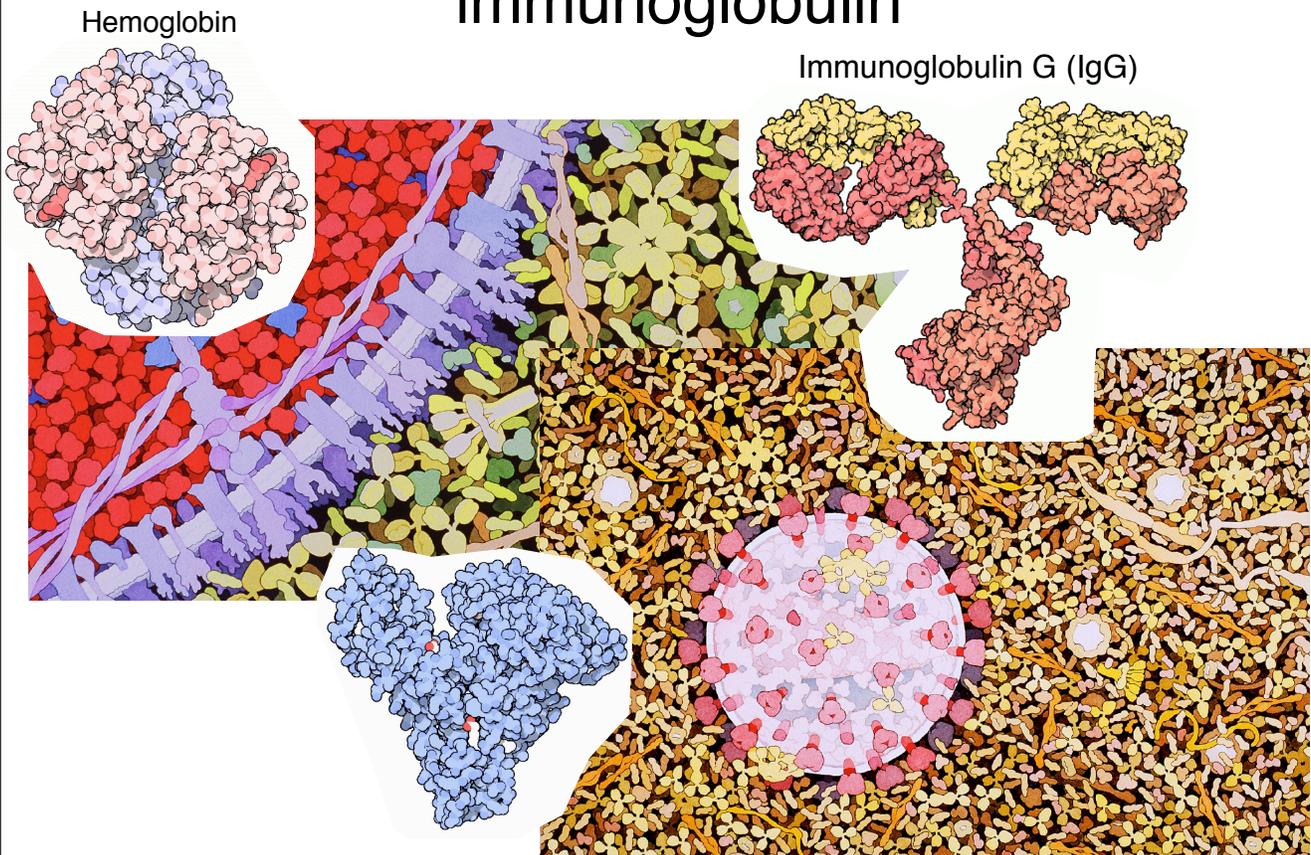
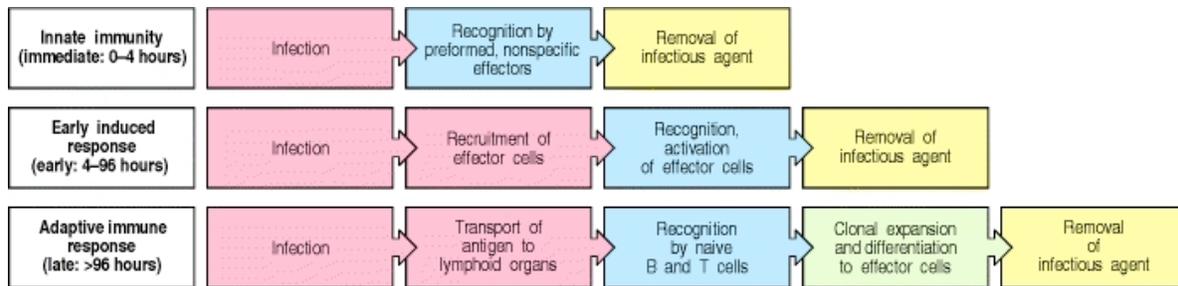David S. Goodsell: The Molecular Perspective

- An artistic/scientific rendition of blood by David S. Goodsell.
- This illustration shows a cross-section through the blood, with blood serum on the right hand side and a red blood cell (RBC) on the left hand side
- Blood serum is filled with antibodies, circulating and searching for foreign molecules. In this illustration, the antibodies are coloured yellow: look for Y-shaped IgG, IgA with two antibodies back-to-back, and IgM with five antibodies in a star.
- Other molecules in this portion of blood serum include stick-like fibrinogen molecules, snaky von Willebrand factor, low density lipoproteins (large circular molecules), and many small albumin proteins.
- The large UFO-shaped objects are low density lipoprotein and the six-armed protein is complement C1.
- The red blood cell is filled with hemoglobin which is shown in red.
- Hemoglobin is the main transporter of oxygen and carbon dioxide in the blood. It is composed of globin (a protein) and heme (a cofactor) which contains iron atoms and imparts the red color to hemoglobin. Hemoglobin is densely packaged into red blood cells.
- Albumin is the major constituent of serum protein (usually over 50%). It helps in osmotic pressure regulation, nutrient transport, and waste removal.
- Immunoglobulin G is a member of a class of blood plasma proteins known as globulins. Immunoglobulin is important in the immune response as we will see in the following slides.
- In the illustration on the right, an HIV surface protein is the antigen that is inducing the immune response. It is under attack by the immune system. An antigen is a substance capable of inducing a specific immune response. The term 'antigen' is derived from the generation of antibodies to such substances.
- Often antigens are foreign proteins (or parts of them) that enter the body via an infection. Sometimes, however, the body's own proteins, expressed in an inappropriate manner (where or when they are not usually seen), are treated like antigens by the immune system.
- It is important to recognize that bacteria or viruses are not themselves antigens but they contain antigens both on their surface and inside them. Such antigens can be isolated and used to safely vaccinate against infection with the whole organism.
- The immune response is a very complex process and we will be taking a greatly simplified view in this course.

- http://www.carbonbased.com/cbcblood.htm - Protein and
- David S. Goodsell: The Molecule of the Month appearing at the PDB
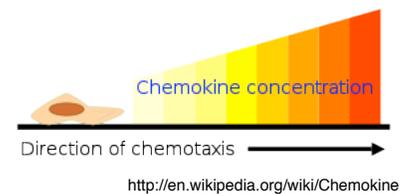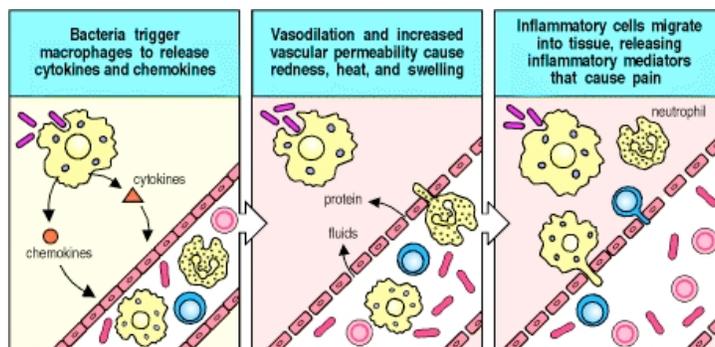
'Innate' and 'adaptive' immunity system'

- The human immune system recognizes and destroys foreign invaders, which could be molecules (generally proteins), viruses, bacteria, or other microorganisms
- The immune system is divided into two categories: innate and adaptive immunity

Innate immunity is the first response to an infection

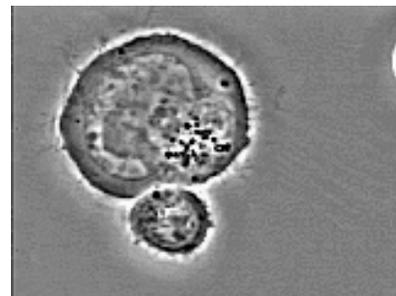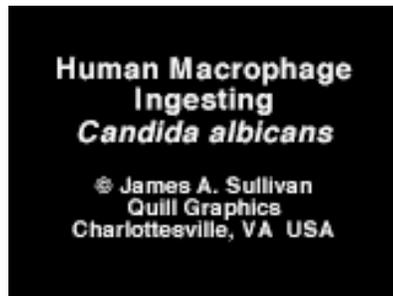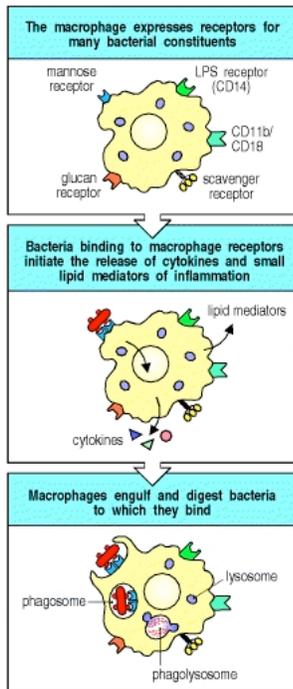| Innate immunity (immediate: 0–4 hours)	Infection		Recognition by preformed, nonspecific effectors		Removal of infectious agent		
Early induced response (early: 4–96 hours)	Infection	Recruitment of effector cells	Recognition, activation of effector cells		Removal of infectious agent		
Adaptive immune response (late: >96 hours)	Infection	Transport of antigen to lymphoid organs	Recognition by naive B and T cells	Clonal expansion and differentiation to effector cells	Removal of infectious agent		

The early induced response is better known as inflammation

Bacteria trigger macrophages to release cytokines and chemokines

cytokines

chemokines

Vasodilation and increased vascular permeability cause redness, heat, and swelling

protein

fluids

Inflammatory cells migrate into tissue, releasing inflammatory mediators that cause pain

neutrophil

Chemokine concentration

Direction of chemotaxis
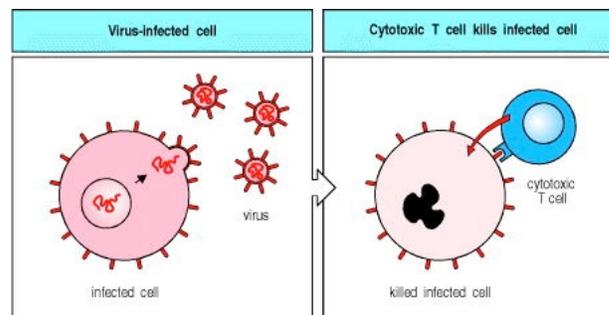
http://en.wikipedia.org/wiki/Chemokine

- The immediate innate immunity response is for macrophages (white blood cells) to attack and ingest the foreign invader
- The macrophages also release signalling molecules, termed cytokines and chemokines, that cause other 'defender' cells to come to the site of the body where the invaders have entered. This is the 'early induced response' and is better known as inflammation. As fluid and cells arrive at the site of the infection, the tissue swells, turns red and hot, and becomes painful.
- Cytokines are proteins that can be thought of much like hormones. Like hormones, they bind to a specfiic cell surface receptor on cells and induce a signalling cascade that lead to specific changes in the cell. For example, the stimulated cells could turn on the expression of genes involved in defending cells against viruses. Chemokines act to attract cells of the immune system to the site of bacterial or viral infection. Since they are being produced at the site of infection, other cells can follow the concentration gradient to travel through the body towards the infected site.

- Images from: Immunobiology: The Immune System in Health and Disease. 5th edition. Janeway CA Jr, Travers P, Walport M, et al. New York: Garland Science; 2001.
- Courtesy of NCBI bookshelf

Our **innate immune system** is the first line of defence again invaders

The macrophage expresses receptors for many bacterial constituents

mannose receptor — LPS receptor (CD14) — CD11b/CD18 — glucan receptor — scavenger receptor

Bacteria binding to macrophage receptors initiate the release of cytokines and small lipid mediators of inflammation

lipid mediators — cytokines

Macrophages engulf and digest bacteria to which they bind

phagosome — lysosome — phagolysosome

Human Macrophage Ingesting *Candida albicans*

® James A. Sullivan
Quill Graphics
Charlottesville, VA USA

Macrophages can recognize, ingest, and destroy antigens, viruses, bacteria, or other invaders.

If the macrophage is too late and the bodies own cells have been infected with virus, there are special lymphocytes that can attack and kill the infected cell.

Virus-infected cell — virus — infected cell

Cytotoxic T cell kills infected cell — cytotoxic T cell — killed infected cell

- The presence of foreign molecules on the surface of invaders mark these bacteria (or virus or whatever) for ingestion by phagocytic cells of the immune system. If the antigen happens to be on the surface of a virus or bacteria, the whole virus particle or bacterial cell will be ingested and destroyed
- There are special white blood cells (cytotoxic T lymphocytes) in the blood that will attack the bodies own cells if they have been infected with virus.
- The recognition of these foreign molecules can occur through 'general' receptors on the surface of the macrophage or cytotoxic T-cell. This is what would happen as part of the innate response.
- However, invaders can also be marked for destruction through the action of the adaptive immune response. The key to the adaptive immune response is that the antibody molecules bind with high specificity and affinity to the invader.
- Immunobiology: The Immune System in Health and Disease. 5th edition. Janeway CA Jr, Travers P, Walport M, et al. New York: Garland Science; 2001.
- Courtesy of NCBI bookshelf

Our **adaptive immune system** depends on having specific antibodies that bind the invader
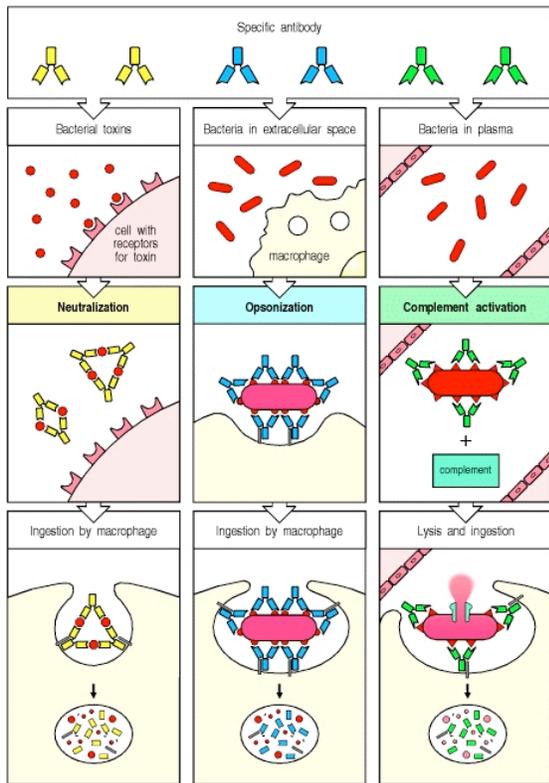
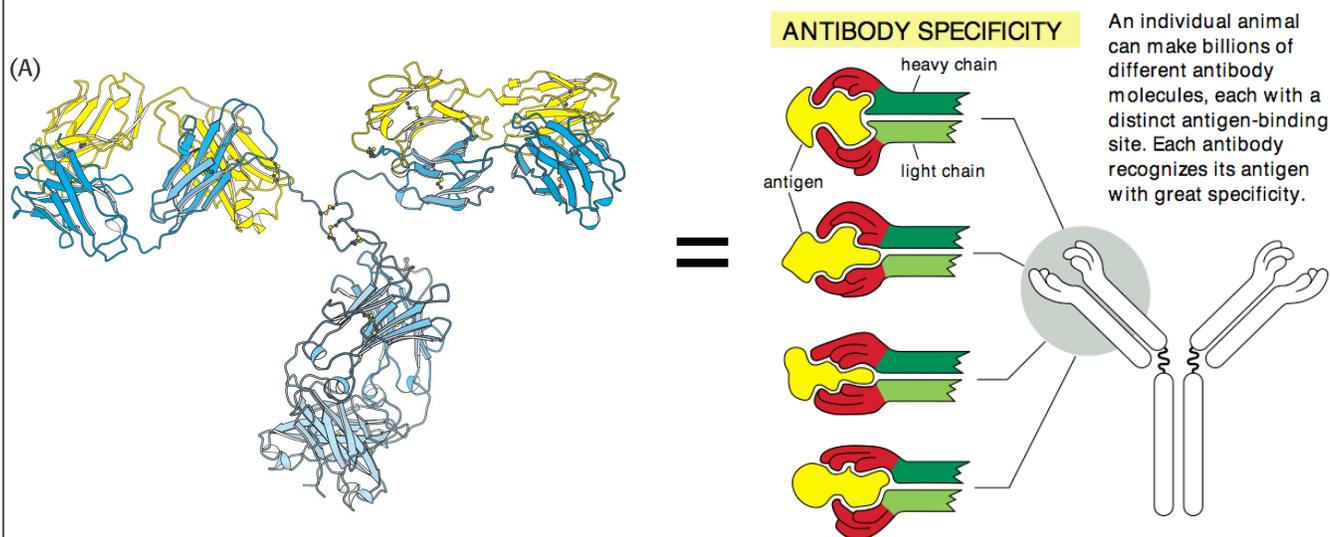**Neutralization:** by binding to a toxin, an antibody can block its function and make it non-toxic

Opsonization: by coating an invader, antibodies mark it as something to be destroyed, and it will be ingested and destroyed by a macrophage.

Complement activation: antibodies can activate the complement system, which is a set of blood proteins that can destroy an invader directly, and/or make it more likely to be eaten by a macrophage.

pathogen
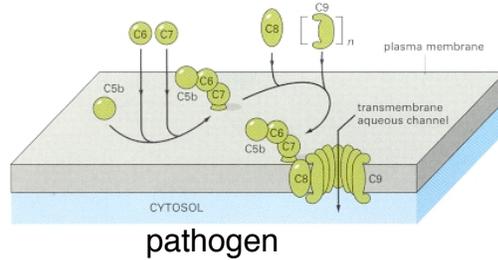
- The rest of this section will focus on antibodies.
- It is important to keep in mind that immunology is a huge subject that could be argued to rival all of chemistry in terms of its complexity and the size of the body of knowledge.
- Accordingly, we will be taking a highly simplified view of how the immune system works.

- Molecular Biology of the Cell. 4th edition. Alberts B, Johnson A, Lewis J, et al. New York: Garland Science; 2002.
- Immunobiology: The Immune System in Health and Disease. 5th edition. Janeway CA Jr, Travers P, Walport M, et al. New York: Garland Science; 2001.
- Courtesy of NCBI bookshelf

Humans have billions of different antibody molecules, each with a unique binding site


(A)

ANTIBODY SPECIFICITY

heavy chain

light chain

antigen

=

An individual animal can make billions of different antibody molecules, each with a distinct antigen-binding site. Each antibody recognizes its antigen with great specificity.

Why might flexible linkers between F$_{ab}$ domains be important for antibody function?

Access Excellence

- Immunoglobulin G consists of two kinds of polypeptide chains, a 25-kd light (L) chain and a 50-kd heavy (H) chain. The subunit composition is L2H2. Each L chain is linked to an H chain by a disulfide bond and non-covalent interactions, and the H chains are linked to each other by at least one disulfide bond plus non-covalent interactions.
- Each L chain comprises two homologous domains, termed immunoglobulin domains. Each H chain has four immunoglobulin domains. These domains have many sequence features in common and adopt a common structure, the immunoglobulin fold. The immunoglobulin fold is one of the most prevalent domains encoded by the human genome. More than 750 genes encode proteins with at least one immunoglobulin fold recognizable at the level of amino acid sequence.
- Overall, the molecule adopts a conformation that resembles the letter Y, in which the stem, corresponding to the Fc fragment obtained by cleavage with papain, consists of the two carboxyl-terminal immunoglobulin domains of each H chain and in which the two arms of the Y, corresponding to the two Fab fragments, are formed by the two amino-terminal domains of each H chain and the two amino-terminal domains of each L chain.
- The linkers between the stem and the two arms consist of relatively extended polypeptide regions within the H chains and are quite flexible.
- The most amazing property of IgG is its ability to recognize such an incredibly diverse range of molecular species ranging from small molecules to large proteins.
- The molecular basis for this versatility is the ability of antibodies to tolerate a wide variety of amino acid changes in its antigen recognition site at the two tips of the 'Y'.
- Each of the several billion antibodies circulating in your blood has a unique amino acid composition in this region of the antibody structure.
- But you only have ~3 billion base pairs of DNA…. How is it possible to encode such a large number of different gene products?
- *Q: You say that the heavy chains and light chains of IgG are homodimers. But I think the tips of the two arms, the CDR, are different. How can they be deemed as homodimer?*
- *A: What I meant by this is that the IgG can be though of as a homodimer of a heterodimer (made of one heavy chain plus one light chain). The CDR regions at the end of each arm are identical for a given antibody.*
- *Q: For antibodies in an organism, is it right to say they are only different in Fv, the remaining parts are all the same*
- *A: As far as this course is concerned, this is correct.*
- *Q: Why might flexible linkers between Fab domains be important for antibody function?*
- *A: Antibodies need to bind to viruses and bacteria that have many copies of the epitope on their surface. Since antibodies have two arms, they can engage in two binding interactions with the antigen, and thus they bind tighter due to avidity (multivalency). However, these epitopes on the surface could be at different distances and orientations for different viruses and bacteria. Flexible linkers allow antibodies to adopt a wide range of conformations so that they can 'reach' both epitopes.*

The VL and VH immunoglobulin domains each have 3 hypervariable loops

The 'hands' of antibodies are composed of two immunoglobulin domains which, together, provide a total of 6 hypervariable loops.
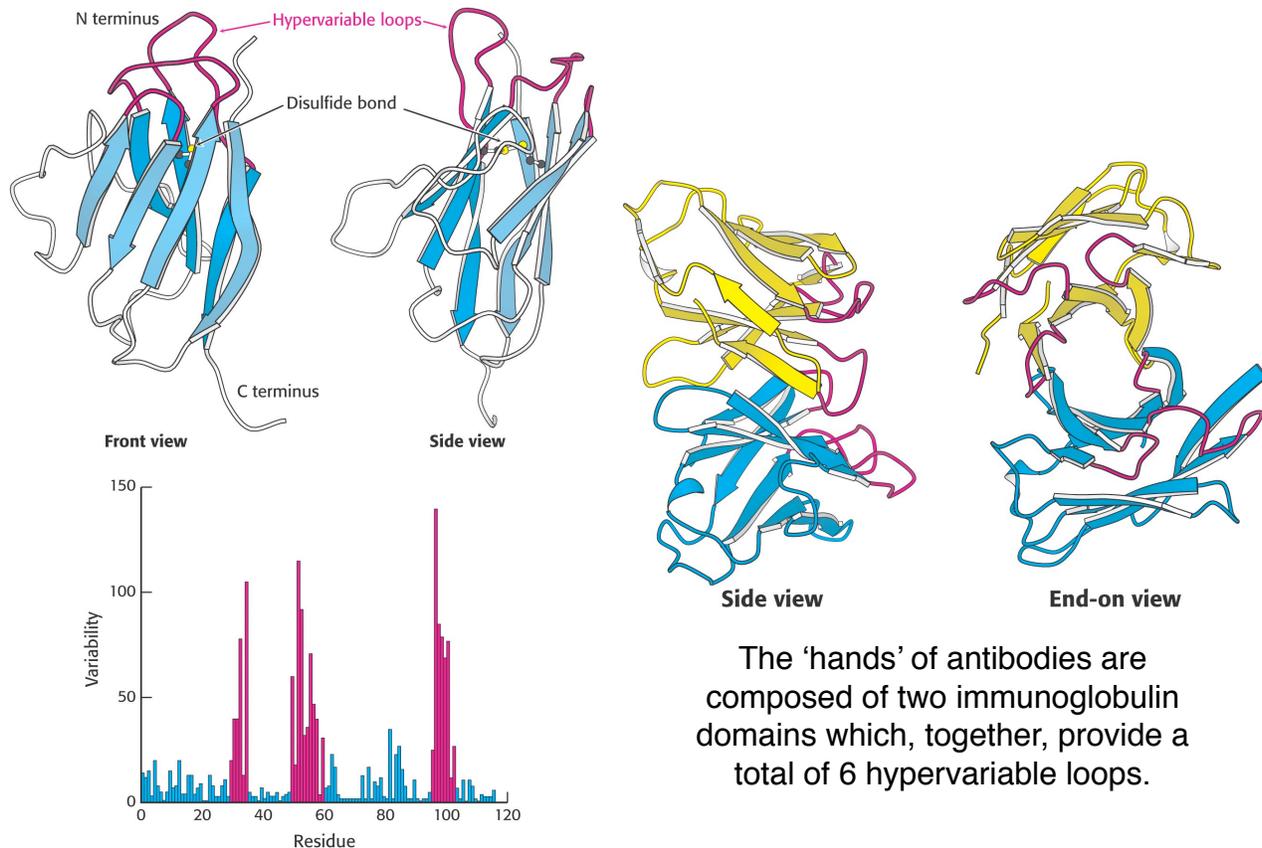
- The immunoglobulin fold consists of a pair of β-sheets, each built of antiparallel β-strands, that surround a central hydrophobic core. A disulfide bond bridges the two sheets.
- Two aspects of this structure are particularly important for its function.
- First, three loops present at one end of the structure form a potential binding surface. These loops contain the hypervariable sequences present in antibodies and in T-cell receptors. Variation of the amino acid sequences of these loops provides the major mechanism for the generation of the vastly diverse set of antibodies and T-cell receptors expressed by the immune system. These loops are referred to as hypervariable loops or complementarity determining regions (CDRs).
- Second, the amino terminus and the carboxyl terminus are at opposite ends of the structure, which allows structural domains to be strung together to form chains, as in the L and H chains of antibodies. If the termini werre on the same side of the domain, it is less likely that the domains could be strung together to make a chain since they would bump into each other.
- The amino-terminal immunoglobin domains of the L and H chains (the variable domains, designated VL and VH) come together at the ends of the arms extending from the structure.
- The positions of the complementarity-determining regions are striking. These hypervariable sequences, present in three loops of each domain, come together so that all six loops form a single surface at the end of each arm. Because virtually any VL can pair with any VH, a very large number of different binding sites can be constructed by their combinatorial association.

- *Q: Is somatic recombination a random process where some genes between V, D and J are being deleted (in two steps). I was wondering is there any governing factor which decides which gene would get deleted?*
- *A: For the sake of this course it is safe to assume that it is a completely random process. To the best of my knowledge this is essentially correct.*
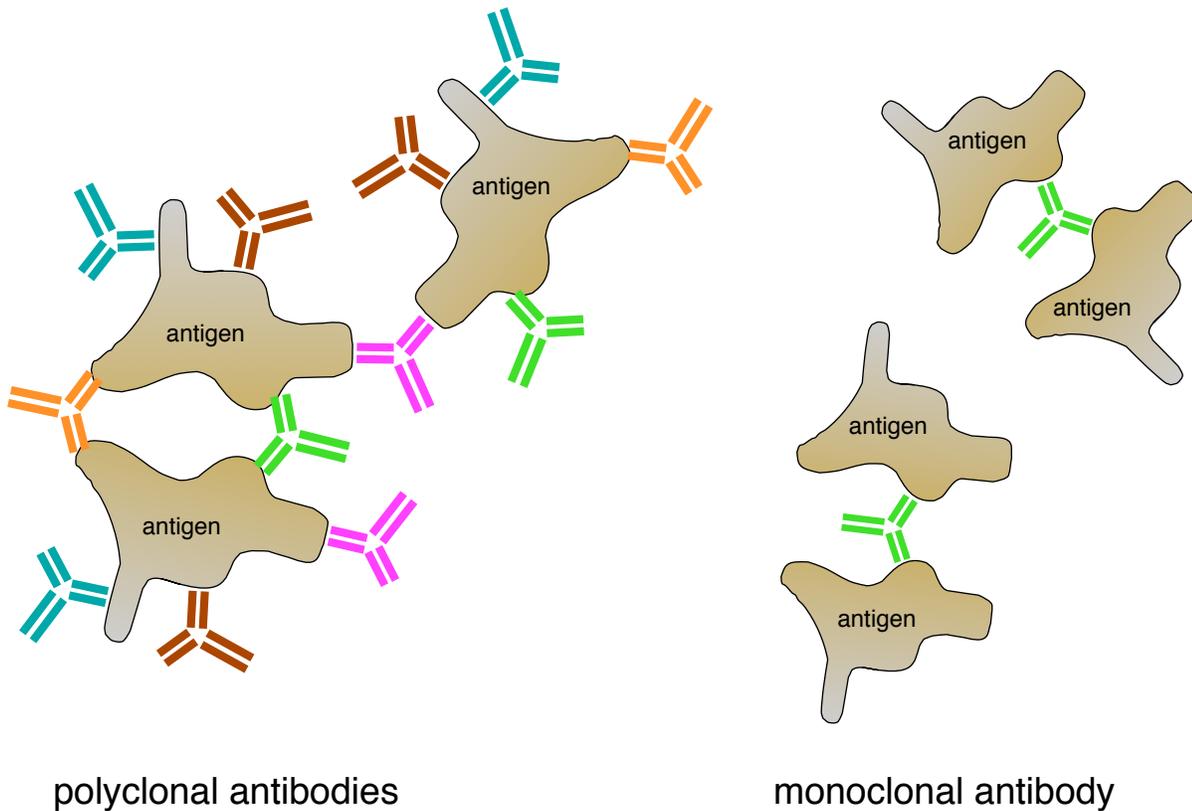
F$_{ab}$ fragments

(B)

Interchain disulfide bonds

Light chain

Heavy chain

Papain cleavage

F$_{ab}$ F$_{ab}$

F$_C$
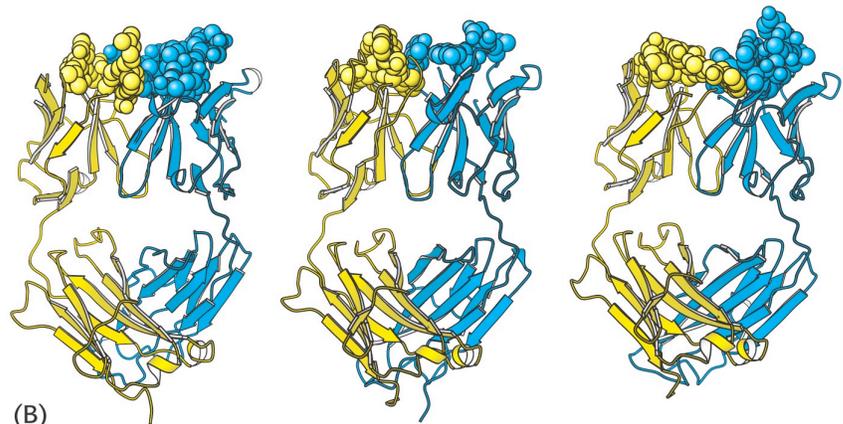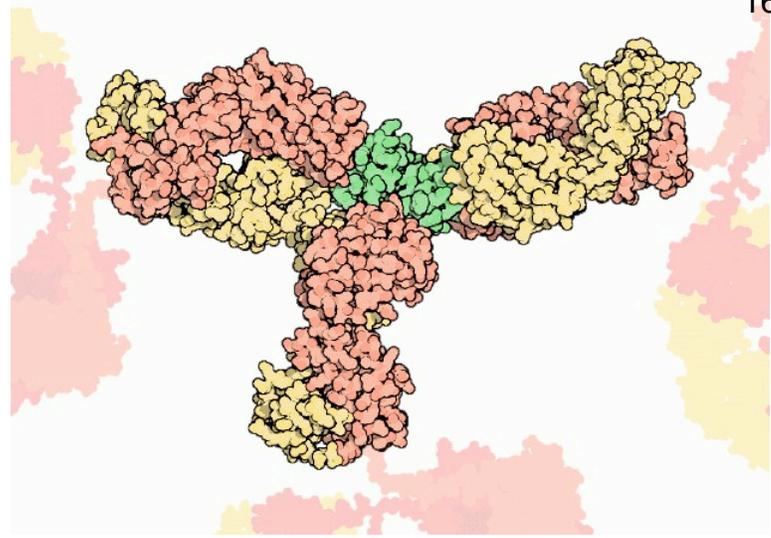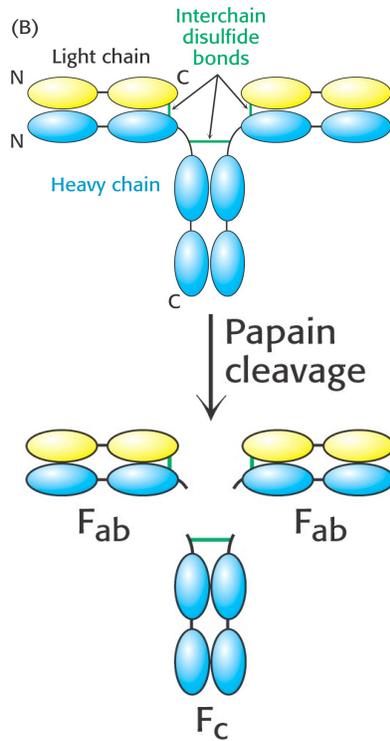
David S. Goodsell: The Molecule of the Month appearing at the PDB

(B)

- In 1959, Rodney Porter showed that immunoglobulin G (IgG), the major antibody in serum, can be cleaved into three 50-kd fragments by the limited proteolytic action of papain (an enzyme that cleaves specific peptide bonds). Two of these fragments bind antigen. They are called Fab (F stands for fragment, ab for antigen binding). The other fragment, called Fc because it crystallizes readily, does not bind antigen, but it has other important biological activities.
- How do these fragments relate to the three-dimensional structure of whole IgG molecules?

- Fab fragments are much easier to crystalize than whole antibodies, so most of what we know about antibody binding to antigens comes from Fab-antigen structures
- A large collection of antibodies raised against hen egg-white lysozyme has been structurally characterized in great detail. Each different antibody binds to a distinct surface of lysozyme.
- The specific part of the protein to which the antibody binds is known as the epitope.
- The models on this slide show how one antigen has potentially many different epitopes.
- Note that a mixture of polyclonal antibodies would contain individual antibodies that bind to all possible epitopes of a given antigen. A monoclonal antibody would bind to only one specific epitope.

- The results of x-ray crystallographic studies of many large and small antigens bound to Fab molecules have been sources of much insight into the structural basis of antibody specificity.
- The binding of antigens to antibodies is governed by the same principles that govern the binding of substrates to enzymes. The shape complementarity between the antigen and the binding site results in numerous contacts between amino acids at the binding surfaces of both molecules. Numerous hydrogen bonds, electrostatic interactions, and van der Waals interactions, reinforced by hydrophobic interactions, combine to give specific and strong binding.

Monoclonal vs. Polyclonal

antigen

antigen

antigen

antigen

antigen

antigen

antigen

antigen

polyclonal antibodies   monoclonal antibody

More on monoclonals: http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/M/Monoclonals.html
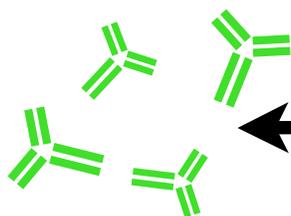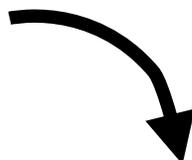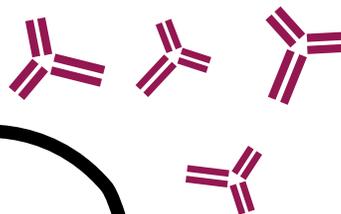
- Polyclonal antibodies are a mixture of many different antibodies with different affinities to different epitopes of the same antigen. It is technically incorrect to refer to a polyclonal antibody (that is, in the singular form) since 'polyclonal' implies many different molecular entities and so 'antibodies' is better than 'antibody'
  - Polyclonal antibodies can be made in the laboratory by injecting an animal (usually a mouse, rabbit, sheep, or goat) with antigen.
  - Repeated injections of the same antigen at intervals of several weeks stimulates specific B cells to secrete large amounts of anti-A antibodies into the bloodstream.
  - Because many different B cells are stimulated by antigen A, the blood will contain a variety of anti-A antibodies, each of which binds A in a slightly different way.
  - It is possible to prepare lots of antibodies (polyclonal) through affinity purification from the plasma of a previously immunized animal. The standard affinity purification would involve using a column on which either protein G or protein A (that is the name of an actual protein) is immobilized. These proteins are known to bind tightly to constant regions of IgG.
  - A better affinity purification is to use an immobilized version of the same protein as was used to immunize the animal. For affinity purification the antigen would be immobilized on a resin. The antibody would be bound to the antigen on the resin and proteins that don't bind would be washed away. The antibody could then be eluted by a dramatic change in pH. For example 100 mM Glycine pH 2.5 for acid elution or 100 mM ethanolamine pH 11.5 for base elution.
- A monoclonal antibody is a distinct antibody molecule that can only be prepared in a laboratory setting. An immune response always generates a mixture of antibodies. The trick is how do you isolate only one from this large population and then generate large quantities of it?
  - This problem was solved by Köhler and Milstein in 1975. They received the Nobel prize in Medicine in 1984 for this work (http://www.nobel.se/medicine/laureates/1984/).
  - A mouse is immunized by injection of an antigen to stimulate the production of antibodies targeted against it.
  - The antibody forming cells are isolated from the mouse's spleen.
  - Monoclonal antibodies are produced by fusing single antibody-forming cells to tumor cells grown in culture. The resulting cell is called a hybridoma.
  - Each hybridoma produces relatively large quantities of identical antibody molecules. By allowing the hybridoma to multiply in culture, it is possible to produce a population of cells, each of which produces identical antibody molecules.
  - These antibodies are called "monoclonal antibodies" because they are produced by the identical offspring of a single, cloned antibody producing cell. Once a monoclonal antibody is made, it can be continuously produced by the hybridoma cell line. Unlike polyclonal antibodies, monoclonal antibodies are thus reproducible.

- As we've seen, the whole purpose of antibodies is to recognize foreign proteins that happen to find their way into the blood of an animal
- Although antibodies from all mammals are practically identical in terms of overall structure and function, there are still enough minor differences for them to be recognized as foreign when introduced into another species. For example, cat antibodies are recognized as foreign when introduced into a goat. And cow antibodies are recognized as foreign when introduced into a rabbit.
- So, immunizing an animal with antibodies from another species leads to the generation of 'anti-antibodies'. In the first example mentioned above, the antibodies that are made (and could be purified from the blood of the animal) would be called 'goat anti-cat IgG' or 'anti-cat IgG produced in goat', or something like that.
- These anti-antibodies are known as secondary antibodies and incredibly useful in bioanalytical chemistry applications.
- However, this does one interesting question: where do we get antibodies for human therapeutic applications? Many of the most sophisticated cancer therapies rely on treating patients with antibodies that target the cancer cells (through a variety of mechanisms). Where could these antibodies come from?

The Fab fragment can be further minimized by protein engineering to give just the variable (VL and VH) domains

| Fab | Fv | scFv (V$_H$–linker–V$_L$) | scFv (V$_L$–linker–V$_H$) | dsFv (FRH2–SS–FRL4) | dsFv (FRH4–SS–FRL2) |

http://www.unizh.ch/%7epluckth/publications/pdf/APpub0108.pdf
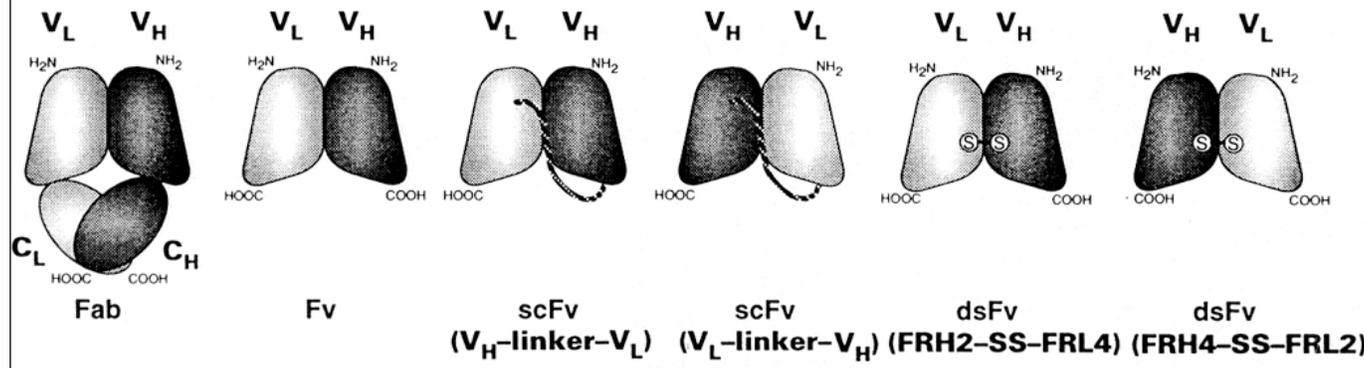
- A Fab fragment can be obtained by proteolysis of an antibody purified from hybridomas
- In contrast, all the Fv fragments must be generated by molecular biology techniques and expressed in bacteria.
- Note that the scFv fragments are single proteins, i.e. there is only one polypeptide chain.
- Fab, Fv, and dsFv are two polypeptide chains
- *Q: What does ScFv stand for?*
- *A: Single Chain Variable Fragment. This is an engineered protein that just contains the 1 domain from the heavy chain, and the 1 domain from the light chain, that contain the CDRs. The two domains are connected by a linker of about 20 amino acids or so.*

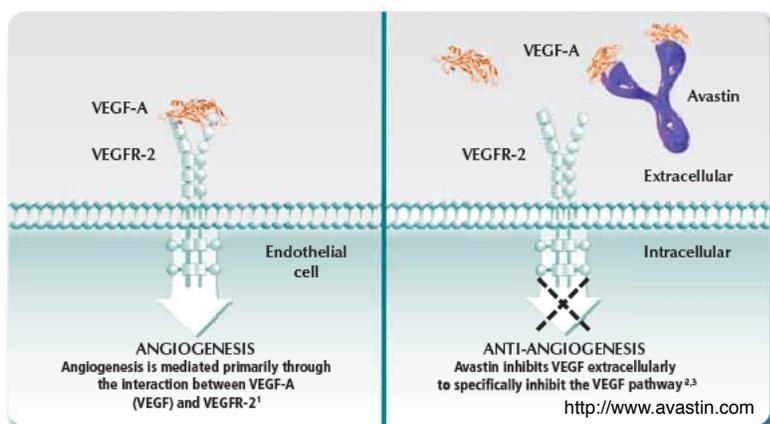# How to make human antibodies for therapeutics?

Mouse monoclonal
0% human

Chimeric
66% human

Humanized
90% human

HuCAL®
100% human

Avastin (bevacizumab) is a humanized monoclonal antibody against vascular endothelial growth factor (VEGF). Tumors release VEGF to stimulate the growth of blood vessels that will bring them nutrients needed for growth. It has been approved for multiple metastatic (actively spreading) cancers.

VEGF-A

VEGFR-2

Endothelial cell

ANGIOGENESIS
Angiogenesis is mediated primarily through the interaction between VEGF-A (VEGF) and VEGFR-2[1]

VEGF-A

Avastin

VEGFR-2

Extracellular

Intracellular

ANTI-ANGIOGENESIS
Avastin inhibits VEGF extracellularly to specifically inhibit the VEGF pathway[2,3]

http://www.avastin.com

Moroney, S., and Plückthun, A. (2005) in Modern Biopharmaceuticals: Modern Antibody Technology: The Impact on Drug Development (*Knäblein, J., ed*) Vol. 3, 1 Ed., pp. 1147-1186, 4 vols., Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

- Antibodies are now considered mainstream therapeutics and there are at least 21 different monoclonal antibodies now approved as human therapeutics. They are particularly useful for cancer therapy.
- The first antibodies tested for human therapeutic applications during the 1980s were of mouse origin. Immunogenicity was an obvious problem so researchers started looking for ways to create human antibodies. The most successful approach was to engineer a mouse antibody to look more "human"
- The first chimeric antibodies kept the mouse variable domains and the remainder was human. Several of these antibodies are now FDA approved. it seems that ~6-10% of patients will have an immune response to these antibodies.
- Another approach was to graft the CDRs from a mouse antibody onto a human antibody to produce something that is 90% human. This is a little bit trickier than it sounds because the framework influences the presentation of the loops.
- 'Most of the growing number of antibodies entering clinical trials are completely human and are derived from phage-display technology or transgenic mice that express human immunoglobulin genes.' (Paul J. Carter, Potent antibody therapeutics by design Nature Reviews Immunology 6, 343-357 (May 2006.