

Background for March Literature Review: Structural Biology

1- What is Structural Biology

This month I can't come up with a better overview than the one posted on NIH_GMS website (<https://nigms.nih.gov/education/fact-sheets/Pages/structural-biology.aspx>). Here is my edited version of the content of that website.

Structural biology is the study of how biological molecules are built. Using a variety of imaging techniques, scientists view molecules in three dimensions to see how they are assembled, how they function, and how they interact. That has helped researchers understand how the thousands of different molecules in each of our cells work together to keep us healthy. Structural studies have also shown how misshapen molecules make us sick, and as a result, these studies have prompted new treatments for many diseases.

Structural biologists are particularly interested in proteins because they do so much of the work in the body. Increasingly, biologists are investigating large molecules made up of combinations of RNA and proteins, called RNA-protein complexes.

2- Protein and process

Proteins are molecules that contribute to virtually every activity in the body. They form hair and fingernails, carry oxygen in the blood, allow muscles to move, and much more. They are made up of amino acids hooked together like beads on a string. There are 20 amino acids found in nature. Each protein contains a unique combination of a few dozen to many thousands of amino acids. Some proteins consist of multiple amino acid strands wound together. To become active, proteins must twist and fold into their final, or "native," configuration (Figure 1).

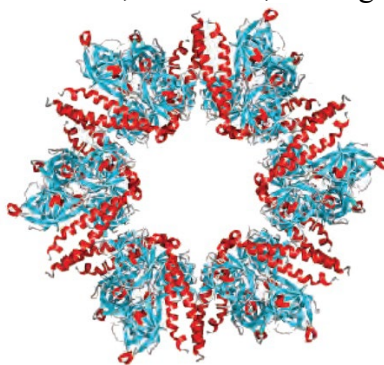


Figure 1: A mix of alpha helices (red; curled ribbon forming a six-sided star) and beta sheets (blue; thinner, tangled strands). Credit: RCSB Protein Data Bank.

2.1 How does a protein get its shape?

Even though proteins are strings of amino acids, they do not remain in a straight line. The strands twist, bend, and fold into specific shapes. The way they fold depends in part on the way the amino acids interact with each other. Some sections of proteins form standard "motifs": corkscrew-like coils called alpha helices and flat sections called beta sheets.

Researchers can easily determine a protein's amino acid sequence. The trick has been figuring out how and why the proteins fold. Scientists are beginning to solve that puzzle with research into how amino acids interact, and they are using powerful new computer programs that help predict protein motifs. Researchers can design their own brand-new proteins that perform specific jobs. This new work helps scientists understand not only how proteins fold but also how they misfold and malfunction in diseases such as Alzheimer's and cystic fibrosis. Knowing more about these processes might allow researchers to design new treatments.

2.2 Why does a protein's shape matter?

A protein's structure allows it to perform its job. For instance, antibodies are shaped like a Y. This helps these immune-system proteins bind to foreign molecules such as bacteria or viruses with one end while recruiting other immune-system proteins with the other. DNA polymerase III is donut-shaped. This helps it form a ring around DNA as it copies its genetic information. And proteins called enzymes have grooves and pockets that help them hold onto other molecules to speed chemical reactions. Misfolded, or misshapen, proteins can cause diseases. They often stop working properly and can build up in tissues. Alzheimer's disease, Parkinson's disease, and cystic fibrosis are examples of diseases caused by misfolded proteins.

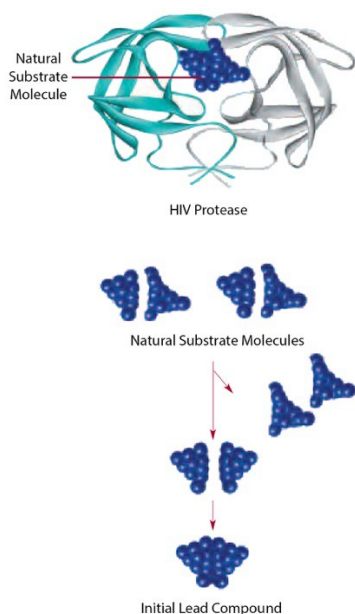


Figure 2: Knowing that HIV protease—an enzyme that breaks down HIV—has two symmetrical halves, pharmaceutical researchers initially attempted to block the enzyme with symmetrical, naturally occurring small molecules. They made these by chopping in half molecules of the natural substrate, then making a new molecule by fusing together two identical halves of the natural substrate. Credit: NIGMS.

2.3 How do scientists use protein structures to develop new drugs?

Drugs typically work by either blocking or supporting the activity of specific proteins in the body. Using an approach called structure-based drug design, scientists can make a template for a protein and use that blueprint for creating new medicines. They start with a computerized model of the protein structure they're interested in studying. For example, the computer model would allow researchers to examine how two proteins work together. Then, if scientists want to turn off one protein, they would try to design a molecule that would block or alter that interaction.

2.4 What's an example of a medicine developed using structure-based drug design?

Researchers used structure-based drug design to develop some anti-HIV drugs. HIV protease is an enzyme that keeps the virus alive. Knowing its structure allowed researchers to determine the

kinds of molecules that could stop HIV protease from working. Scientists used computer models to fine tune molecules that could halt virus production. This work led to medicines called protease inhibitors.

3- How do scientists determine protein structures?

Researchers use several imaging techniques to determine the structure of proteins and other complex molecules. Cryo-electron microscopy (cryo-EM) allows scientists to “see” individual proteins as well as larger structures such as molecular complexes (groups of proteins that combine and function as a unit), viruses, or organelles (specialized structures within the cell that perform specific functions). X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy also make it possible for researchers to view proteins. To date, researchers have used these techniques to unravel the structure of more than 122,000 proteins. The [Protein Data Bank](#) stores these structures and gives scientists access to them.

3.1 cryo-EM

In cryo-EM, researchers rapidly freeze a cell, virus, molecular complex, or other structure so that water molecules do not have time to form crystals. This preserves the sample in its natural state. Scientists use an electron microscope to blast the frozen sample with an electron beam. This creates a two-dimensional projection of the sample on a digital detector. By creating hundreds of projections of the sample from many different angles and then taking the average of these angles, scientists generate a three-dimensional model of its structure. Recent advances in cryo-EM provide highly detailed images of proteins and other biological structures, including larger structures such as RNA-protein complexes.

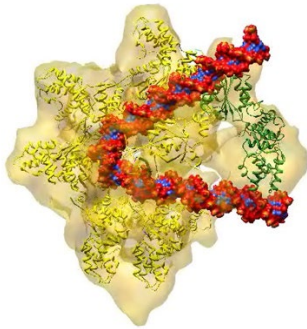


Figure 3: Cryo-EM can be used to determine the structures of large molecular complexes such as the origin recognition complex (ORC). The ORC recognizes and binds DNA to start the process that copies the cell's genetic material prior to cell division. Credit: Huilin Li, Brookhaven National Laboratory, and Bruce Stillman, Cold Spring Harbor Laboratory.

3.2 What is X-ray crystallography, and how does it work?

X-ray crystallography shoots a beam of X-rays through a tiny solid crystal made up of trillions of identical protein molecules. The crystal scatters the X-rays onto an electronic detector, similar to the way images are captured in a digital camera. A computer gauges the intensities of the scattered X-rays to assign a position to each atom in the crystallized molecule. The result is a three-dimensional digital image. This method has been used to determine more than 85 percent of known protein structures.

3.3 What is NMR spectroscopy, and how does it work?

NMR spectroscopy works using the natural magnets—the nuclei of certain atoms—inside proteins. Those natural cellular magnets interact with a big magnet inside the NMR machine. The big magnet forces the protein's magnets to line up. Researchers then blast the sample with a series of split-second radio-wave pulses and observe how the protein's magnets respond. Scientists use several sets of these NMR blasts and combine the data to get a more complete picture of the protein. Although X-ray crystallography can examine larger proteins than NMR is able to do, NMR technology can study proteins immersed in liquid solutions. In contrast, X-ray crystallography requires that proteins be organized into crystals.

4- Can scientists view how proteins act?

New technology is beginning to allow researchers to progress from creating static pictures of proteins and other molecules to making movies of their actions.

Images provide snapshots of what these cellular elements are doing at specific points in time. Although they supply valuable information, these still pictures don't capture how proteins and other molecules inside cells are constantly moving and changing, folding and unfolding as they interact. Understanding this dynamic system is critical to unlocking how life works. In addition, there is a whole class of proteins, called intrinsically disordered proteins, that do not hold a specific shape. Their shape adapts to what's going on inside the cell, making it nearly impossible to take a still picture of them. Scientists are now using powerful computer models to make molecular movies so they can see the full range of proteins in live action.

Researchers on the frontier of structural biology are merging all of the imaging techniques—X-ray crystallography, NMR, and cryo-EM. This allows them to create a more precise map of what proteins and other molecules look like and how they interact. Scientists can create a single image that zooms in to see specific proteins and also zooms out to see how they interact within the larger cellular structure. In addition to combining existing techniques, scientists are developing ever more powerful methods. For example, new X-ray lasers allow insights into processes that occur in less than one tenth of a trillion of a second, much faster timescales than that captured by other sources of X-rays.

Scientists use super-efficient methods to determine protein structures more quickly than ever before. They also use sophisticated techniques to predict three-dimensional structures of proteins. And they use high-powered computer models to design and create new proteins not found in nature that have useful functions, such as discovering and combating disease. This work will continue to increase our understanding of the diverse roles molecules play in biology and to spur advances in medicine.

5- Complementary information for selected article

5.1 Protein folding

Protein folding is the physical process by which a protein chain is translated into its native three-dimensional structure, typically a "folded" conformation, by which the protein becomes

biologically functional. Each protein exists first as an unfolded polypeptide or random coil after being translated from a sequence of mRNA into a linear chain of amino acids. At this stage, the polypeptide lacks any stable (i.e., long-lasting) three-dimensional structure (see the left side of the first figure). As the polypeptide chain is being synthesized by a ribosome, the linear chain begins to fold into its three-dimensional structure.

The primary structure of a protein, its linear amino-acid sequence, determines its native conformation.

The formation of a secondary structure is the first step in the folding process that a protein takes to assume its native structure. Characteristic of secondary structure are the structures known as **alpha helices** and **beta sheets** that fold rapidly because they are stabilized by intramolecular hydrogen bonds. The α -helices are formed by hydrogen bonding of the backbone to form a spiral shape. The β pleated sheet is a structure that forms with the backbone bending over itself to form the hydrogen bonds. The hydrogen bonds are between the amide hydrogen and carbonyl oxygen of the peptide bond.

The α -Helices and β -Sheets are commonly amphipathic, meaning they have a hydrophilic and a hydrophobic portion. This ability helps in forming tertiary structure of a protein in which folding occurs so that the hydrophilic sides are facing the aqueous environment surrounding the protein and the hydrophobic sides are facing the hydrophobic core of the protein. Once the protein's tertiary structure is formed and stabilized by the hydrophobic interactions, there may also be covalent bonding in the form of disulfide bridges formed between two cysteine residues. These non-covalent and covalent contacts take a specific topological arrangement in a native structure of a protein. Tertiary structure of a protein involves a single polypeptide chain.

Tertiary structure may give way to the formation of quaternary structure in some proteins, which usually involves the "assembly" or "coassembly" of subunits that have already folded; in other words, multiple polypeptide chains could interact to form a fully functional quaternary protein.

5.2 Protein-Protein interaction

Protein-protein interactions are physical contacts of high specificity established between two or more protein molecules as a result of biochemical events steered by interactions that include electrostatic forces, hydrogen bonding and the hydrophobic effect.

Hydrophobic surface is a surface that has the ability to repel water. It can be used to identify regions of a protein's surface most likely to **interact with a binding ligand**. This fast and simple procedure may be useful for identifying small sets of well-defined loci for possible ligand attachment.