PROTEIN STRUCTURE

Why is protein structure relevant for Genomics?

Protein structure is more conserved than amino acid sequence, which is more conserved than DNA sequence. Common structures/folds can be found in very distantly related proteins.

Characteristics of Protein Structure

The amino acid sequence of the polypeptide chain of a protein is called the **primary structure**. Local regions of the polypeptide fold into regular elements of **secondary structure**, including alpha helices, beta strands, or turns. The three-dimensional arrangement of these structural elements form the **tertiary structure** of the polypeptide chain. The complete protein may consist of several polypeptide chains arranged in the **quaternary structure**.

Most proteins fold into **globular domains** with many hydrophobic amino acid side chains in the interior, and more polar and charged residues on the protein surface and exposed to the mostly aqueous environment. One protein subunit may fold into two or more compact **domains**.

Primary structure = amino acid sequence of polypeptide chain

Secondary structure = regular arrangements of the backbone
alpha helices, beta sheets (parallel and antiparallel), and turns (mostly at protein surface).

Tertiary structure = 3D folding of polypeptide chain,
topological arrangement of secondary structural elements.
Common types of protein fold
All alpha helix, all beta sheet, mixed alpha and beta, and small disulfide rich structures.
The protein tertiary structure can be stabilized by **disulfide bonds** between two cysteines that connect different elements of secondary structure, or by **ligands**, including metal ions (Ca, Fe) or heme.
Domains or independent globular regions with hydrophobic core
Active sites or binding sites are often found between domains.

Quaternary structure = arrangement of subunits in oligomeric protein.
Can be formed by identical subunits or different polypeptides.
Dimers are the most common, then tetramers.
Many oligomeric proteins show allosteric effects for ligand binding, both cooperative and anti-cooperative effects which allow delicate regulation of the function.