# 3 Protein Structure and Function

## *Review the Concepts*

1. The primary structure of a protein is the linear arrangement or sequence of amino acids. The secondary structure of a protein is the various spatial arrangements that result from folding localized regions of the polypeptide chain. The tertiary structure of a protein is the overall conformation of the polypeptide chain, its three-dimensional structure. Secondary structures, which include examples such as the alpha (α) helix and the beta (β) sheet, are held together by hydrogen bonds. In contrast, the tertiary structure is primarily stabilized by hydrophobic interactions between non-polar side chains of the amino acids and hydrogen bonds between polar side chains. (The quaternary structure describes the number and relative positions of the subunits in a multimeric protein.)

2. Chaperones and chaperonins are proteins that play key roles in the proper folding of newly synthesized proteins. Chaperones bind to and stabilize unfolded or partially folded proteins and thereby prevent these proteins from aggregating and being degraded. The major chaperone protein in all organisms is the protein Hsp70 and its homologues. When bound to ATP, a hydrophobic region on Hsp70 binds to a hydrophobic region on the unfolded protein. Hydrolysis of the ATP induces a conformational change in Hsp70 that allows the bound protein to be folded. Chaperonins, which form large macromolecular assemblies, directly facilitate the folding of proteins. A partially folded or misfolded protein binds to the inner wall of the cylindrical chaperonin assembly and is folded into its native conformation. In an ATP-dependent step, the chaperonin assembly undergoes a conformational change that releases the folded protein.

3. An enzyme increases the rate of a reaction that is already energetically favorable by lowering the activation energy required for the reaction. The reaction occurs within the active site of the enzyme, which consists of a number of amino acid side chains brought into close proximity by the native folding of the enzyme. The active site serves two functions: to recognize and bind the substrate and to catalyze the reaction once the substrate is bound. For an enzyme-catalyzed reaction, the $V_{max}$ represents the maximal velocity of a reaction at saturating substrate concentrations and the $K_m$ represents the substrate concentration that yields a half-maximal velocity (1/2 $V_{max}$). For an enzyme-catalyzed reaction, the $K_m$ is a measure of the affinity of a substrate for its enzyme. A substrate with a low $K_m$ has a high affinity for the enzyme, and thus substrate B ($K_m$ = 0.01 mM) would have a higher affinity for enzyme X than substrate A ($K_m$ = 0.4 mM).

4. Motor proteins possess three general properties: 1) the ability to transduce a source of energy, either ATP or an ion gradient, into linear or rotary movement, 2) the ability to bind and translocate along a cytoskeletal filament, nucleic acid strand, or protein complex; and 3) net movement in a given direction. Linear motor proteins are responsible for moving cargo along microfilaments or microtubules. The revolution of rotary motors causes bacterial flagella to beat, they facilitate the packing of DNA into the capsid, and they are required for ATP synthesis in mitochondria.

5. Ubiquitin is a 76–amino acid protein that serves as a molecular tag for proteins destined for degradation. Ubiquitination of a protein involves an enzyme-catalyzed transfer of a single ubiquitin molecule to the lysine side chain of a target protein. This ubiquitination step is repeated many times, resulting in a long chain of ubiquitin molecules. The resulting polyubiquitin chain is recognized by the proteasome, which is a large,

cylindrical, multisubunit complex that proteolytically cleaves ubiquitin-tagged proteins into short peptides and free ubiquitin molecules. Proteasome inhibitors would be useful to treat cancers if they blocked the degradation of proteins e.g., tumor suppressors, required to halt the progression of uncontrolled cell growth. In the case of the proteasome inhibitor Velcade, which is used to treat patients with multiple myeloma, cells undergo apoptosis (programmed cell death), and because a protein serving as a pro-survival factor called NFκB cannot be activated when proteasome activity is blocked (reviewed in A. Fribley and C. Y. Wang, *Cancer Biol. Ther.*, 2006 July 1; 5(7):745-8.

6. Cooperativity, or allostery, refers to any change in the tertiary or quaternary structure of a protein induced by the binding of a ligand that affects the binding of subsequent ligand molecules. In this way, a multisubunit protein can respond more efficiently to small changes in ligand concentration compared to a protein that does not show cooperativity. The activity of many proteins is regulated by the reversible addition/removal of phosphate groups to specific serine, threonine, and tyrosine residues. Protein kinases catalyze phosphorylation (the addition of phosphate groups), while protein phosphatases catalyze dephosphorylation (the removal of phosphate groups). Phosphorylation/dephosphorylation changes the charge on a protein, which typically leads to a conformational change and a resulting increase or decrease in activity. Some proteins are synthesized as inactive propeptides, which must be enzymatically cleaved to release an active protein.

7. Proteins can be separated by mass by centrifuging them through a solution of increasing density, called a density gradient. In this separation technique, known as rate-zonal centrifugation, proteins of larger mass generally migrate faster than proteins of smaller mass. However, this is not always true because the shape of the protein also influences the migration rate. Gel electrophoresis can also separate proteins based on their mass. In this technique, proteins are separated through a polyacrylamide gel matrix in response to an electric field. Because the migration of proteins through a polyacrylamide gel is also influenced by shape of proteins, the ionic detergent sodium dodecyl sulfate is added to denature proteins and force proteins into similar conformations. During rate-zonal centrifugation, a protein of larger mass (transferrin) will sediment faster during centrifugation; whereas a protein of smaller mass (lysozyme) will migrate faster during electrophoresis.

8. Gel filtration, ion exchange, and affinity chromatography typically involve the use of a bead consisting of polyacrylamide, dextran or agarose packed into a column. In gel filtration chromatography, the protein solution flows around the spherical beads and interacts with depressions that cover the surface of the beads. Small proteins can penetrate these depressions more readily than larger proteins and thus spend more time in the column and elute later from the column; larger proteins do not interact with these depressions and elute first from the column. In ion-exchange chromatography, proteins are separated on the basis of their charge. The beads in the column are covered with amino or carboxyl groups that carry a positive or negative charge, respectively. Positively charged proteins will bind to negatively charged beads, and negatively charged proteins will bind to positively charged beads. In affinity chromatography, ligand molecules that bind to the protein of interest are covalently attached to beads in a column. The protein solution is passed over the beads and only those proteins that bind to the ligand attached to the beads will be retained, while other proteins are washed out. The bound protein can later be eluted from the column using an excess of ligand or by changing the salt concentration or pH.

9. Proteins can be made radioactive by the incorporation of radioactively-labeled amino acids during protein synthesis. Methionine or cysteine labeled with sulfur-35 are two commonly used radioactive amino acids, although many others have also been used. The radioactively labeled proteins can be detected by autoradiography. In one example of this technique, cells are labeled with a radioactive compound and then overlaid with a photographic emulsion sensitive to radiation. The presence of radioactive proteins

will be revealed as deposits of silver grains after the emulsion is developed. A Western blot is a method for detecting proteins that combines the resolving power of gel electrophoresis, the specificity of antibodies, and the sensitivity of enzyme assays. In this method, proteins are first separated by size using gel electrophoresis. The proteins are then transferred onto a nylon filter. A specific protein is then detected by use of an antibody specific for the protein of interest (primary antibody) and an enzyme-antibody conjugate (secondary antibody) that recognizes the primary antibody. The presence of this protein-primary antibody-enzyme-conjugated-secondary antibody complex is detected using an assay specific for the conjugated enzyme.

10. X-ray crystallography can be used to determine the three-dimensional structure of proteins. In this technique, x-rays are passed through a protein crystal. The diffraction pattern generated when atoms in the protein scatter the x-rays is a characteristic pattern that can be interpreted into defined structures. Cryoelectron microscopy involves the rapid freezing of a protein sample and examination with a cryoelectron microscope. A low dose of electrons is used to generate a scatter pattern that can be used to reconstruct the protein's structure. In nuclear magnetic resonance (NMR) spectroscopy, a protein solution is placed in a magnetic field and the effects of different radio frequencies on the spin of different atoms are measured. From the magnitude of the effect of one atom on an adjacent atom, the distances between residues can be calculated to generate a three-dimensional structure.

11. The four features of a mass spectrometer are 1) the ion source; 2) the mass analyzer; 3) the detector; and 4) a computerized data system. Basically, the investigator would collect protein samples from the cancerous cells and from the normal healthy cells, the latter serving as a control. Samples would be prepared for 2D PAGE and after electrophoresis the gels would be dyed and the profiles compared. If a protein "spot" were present in the sample from the cancer cell and not the control, it would be isolated out of the gel, protease-digested using trypsin to generate peptides that are mixed with a matrix, and applied to a metal target. A laser is used to ionize the peptides, which are vaporized them singly charged ions. In the case of a time of flight (TOF) mass analyzer, the time it takes the ions to pass through the analyzer before reaching the detector is inversely proportional to its mass and directly proportional to the charge they carry, generating a spectrum in which each molecule has a distinct signal, allowing the investigator to calculate each ion's mass. The fourth essential component is a computerized data system that acquires and stores the data, which are then compared to information in databases. The mass and charge signature, or fingerprint, of the unknown, is compared to that of peptides in a database and the best match protein is identified.

## *Analyze the Data*

a. Proteins 3, 5, 6, and 7 do not change in response to the drug. Protein 1 declines in response to the drug. Proteins 2 and 4 are induced in response to the drug.

b. Proteins 2, 3 and 6 are phosphoproteins; the others are not. Protein 2 is a phosphoprotein normally present in cells whose level increases in response to the drug. Protein 3 is not a phosphoprotein in control cells and is phosphorylated in response to the drug. Protein 6 is a phosphoprotein whose level does not change in response to the drug.

c. Proteins 1 and 6 are strictly nuclear proteins. Proteins 2, 3, and 4 are strictly cytoplasmic proteins. Protein 5 is present in the nucleus and cytoplasm. Protein 7 is a cytoplasmic protein that migrates to the nucleus in response to drug treatment.

d. Protein 1 is a nuclear protein whose level declines in response to drug treatment. Protein 2 is a cytoplasmic phosphoprotein whose level increases in response to drug treatment. Protein 3 is a cytoplasmic phosphoprotein whose level does not change in response to drug treatment. Protein 4 is a cytoplasmic protein whose level increases in response to drug treatment. Protein 5 is a nuclear and cytoplasmic protein whose level does not

change in response to drug treatment. Protein 6 is a nuclear phosphoprotein whose level does not change in response to drug treatment. Protein 7 is a cytoplasmic protein which migrates from the cytoplasm to the nucleus, in response to drug treatment.