preventing a buildup of aggregates. However, non-optimal conditions such as inflammation, overproduction of protein, aging/oxidation, and mechanical or chemical stresses may increase the amount of misfolded protein in vivo. These misfolded proteins may then form amyloid fibrils, which are associated with a large number of diseases, among them Alzheimer’s disease, Down syndrome, Parkinson’s disease, and catatons. A comprehensive understanding of protein folding and structure will contribute toward the logical design of therapies for these diseases.

Despite the stringent limits imposed on protein structure by function and the potential effects of aggregation in cells, amazingly regions within many proteins are unfolded or disordered in their native state. Many of these proteins fold upon binding other cellular components—for example, other proteins, nucleic acids, or small molecules—allowing regulation of protein function by a structural oscillation between active and inactive forms. Indeed, many intrinsically disordered proteins participate in inherently transient and tightly controlled processes such as transcription regulation and cell signaling.

Continuing challenges and potential benefits. Protein folding is the least well-characterized step in the process by which genetic information is stored, transmitted, and implemented (Fig. 1). Several significant challenges remain, including generating better algorithms or sufficient computational power for modeling large proteins in silico (via computer), monitoring kinetic folding events on very fast time scales, describing the unfolded state, identifying all possible folds (ternary structures), understanding the forces that drive proteins to misfolded states or aggregates and how these affects living cells, discovering general principles of protein folding despite varying folding pathways and diverse final structures, and elucidating how and why motion and disorder, like structure, can be encoded by amino acid sequence without eliciting protein destruction. Addressing these issues not only will illuminate a basic and critical aspect of biology but also will expand our ability to dissect uncharacterized biological pathways by predicting the structure and function of novel genes identified by genome sequencing projects. Finally, a thorough description of protein folding would allow exploitation of their functional specificity and efficiency by de novo design of artificial proteins as research reagents, industrial catalysts, and pharmaceutical products.


Protein kinase

One of a family of enzymes that exert regulatory effects on a variety of cellular functions, as well as malignant transformation, by adding a phosphate group to proteins according to the equation

\[
\text{ATP} + \text{protein} \rightarrow \text{protein kinase} + \text{protein-PO}_4^{2-} + \text{ADP}
\]

(ADP represents adenosine diphosphate, ADP represents adenosine triphosphate, and \(\text{OH}^-\) is a hydroxyl group attached to an amino acid residue.) Based upon the nature of the phosphorylated \(\text{OH}^-\) group, these enzymes are classified as serine/threonine protein kinases and tyrosine protein kinases, where serine, threonine, and tyrosine are amino acid residues found in proteins. Furthermore, there is a small group of dual-specificity kinases, which closely resemble serine/threonine kinases, that catalyze the phosphorylation of both threonine and tyrosine on target proteins. The ratio of phosphoserine/phosphothreonine/phosphotyrosine in proteins from animal cells is about 3000/300/1. Despite the scarcity of tyrosine protein phosphatase, it plays a paramount role in cell physiology.

**Actions.** Protein kinases are enzymes that play a role in nearly every aspect of cell biology. Work in the 1960s showed that protein kinases play a regulatory role in carbohydrate metabolism. Subsequent work indicates that protein kinases participate in nearly all cellular activities including regulation of gene expression, cell division, apoptosis (programmed cell death), differentiation, development, and antibody production. The brain, moreover, is an especially rich site of protein kinase activity. See CELL (BIOLOGY), NEUROBIOLOGY.

Protein phosphorylation represents the chief regulatory mechanism in animal cells. In some cases, phosphorylation leads to activation, and in other cases phosphorylation leads to inactivation of the phosphorylated protein. It is estimated that perhaps a quarter of all animal proteins can be phosphorylated by protein kinases. Moreover, many proteins can be phosphorylated at more than one residue by a given protein kinase, and many proteins can be phosphorylated by several protein kinases. See ENZYME, PROTEIN.

**Protein kinase genes.** There are more genes encoding protein kinases than there are genes encoding any other family of human enzymes except for proteases. Workers have identified 518 protein kinase genes in humans (478 typical and 40 atypical), about 2% of the total number of human genes (see table). All protein kinases have a similar overall three-dimensional structure and chemical mechanism. See GENE, GENE ACTION.
### Number of protein kinase genes found in humans

<table>
<thead>
<tr>
<th>Class</th>
<th>Number</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein-serine/threonine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receptor</td>
<td>429</td>
<td></td>
</tr>
<tr>
<td>Nonreceptor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dual specificity</td>
<td>12</td>
<td>Transforming growth factor-β</td>
</tr>
<tr>
<td></td>
<td>369</td>
<td>ERK, cyclic AMP-dependent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>kinase A, protein kinase C,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RAF protein kinase</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Protein-tyrosine</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Receptor</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>Nonreceptor</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>518</td>
<td></td>
</tr>
</tbody>
</table>

*ERK, extracellular regulated kinase; HER, human epidermal growth factor; MEK, mitogen-activated; ERK-activating kinase.*
*Possesses an extracellular domain that binds to a first messenger.*
*Contains an extracellular domain and is regulated by second messengers and by phosphorylation-dephosphorylation.*

### Serine/threonine protein kinases

There are many serine/threonine protein kinases, and three general classes are described here.

**Protein kinase A**

Protein kinase A, or cyclic adenylate monophosphate (AMP)-dependent protein kinase, was one of the first protein kinases to be described. Many hormones (first messengers) interact with their receptors on the cell membrane, and this interaction leads to the activation of adenyl cyclase. This enzyme catalyzes the intracellular production of cyclic AMP (a second messenger) from ATP. Cyclic AMP interacts with protein kinase A and activates it. Protein kinase A can lead to the phosphorylation of many different proteins, with many different effects. For example, protein kinase A-mediated phosphorylation of the enzyme that promotes triglyceride degradation leads to its activation, whereas protein kinase A-mediated phosphorylation of the enzyme that promotes glycerone synthesis is inhibitory.

**Protein kinase C**

Protein kinase C is another serine/threonine kinase that can lead to the phosphorylation of many different proteins. (The C refers to its requirement for ionic calcium.) It also requires phospholipids and diacylglycerol for full expression of activity. As a result of the interaction of a hormone or growth factor (first messengers) with a receptor on the exterior of a cell membrane, intracellular second messenger molecules such as diacylglycerol and calcium are liberated, and these activate protein kinase C.

Many protein kinases function in a series, or a cascade, where one enzyme catalyzes the phosphorylation of a second enzyme that in turn mediates the phosphorylation of a third enzyme. A generic form of this cascade is given in the illustration. A mitogen-activated protein kinase (MAPK) brings about a physiological result. This enzyme is phosphorylated by a kinase acting on a kinase (MAPK), and this is then activated by a kinase (MAP kinase kinase, or MAPKK). A stimulus initiates the process by activating the mitogen-activated protein kinase kinase (MAPKK). There are several examples of cascades that follow this paradigm, one of which (the Raf-MEK-ERK cascade) is shown in illustration. b. Protein kinase C-mediated phosphorylation of a serine/threonine protein kinase called Raf leads to its activation, and the activated Raf participates in a cascade that leads to growth and cell division (see illustration). Activation of Raf leads to the phosphorylation and activation of two different protein kinases, MEK1 and MEK2. MEK1 and MEK2 are examples of dual-specificity protein kinases. These lead to the phosphorylation of both a threonine and a tyrosine residue in each of the protein kinases ERK1 and ERK2, and this dual phosphorylation leads to their activation. ERK1 and ERK2 are serine/threonine protein kinases that mediate the phosphorylation of other protein kinases and regulators of gene transcription. As a result of this cascade of events, cell growth and cell division can result. See CELL DIVISION; GROWTH FACTOR; HORMONE.

**Cyclic-dependent protein kinases**

Cyclic-dependent protein kinases are another family of serine/threonine
protein kinases that are essential for cell division. Cyclins are proteins that are synthesized and activated during specific phases of the cell cycle, and they are degraded after they have acted. While the level of cyclin varies during the cell cycle, the level of their cognate protein kinases remains constant throughout the cell cycle. See CELL CYCLE.

Tyrosine protein kinases. Receptor tyrosine protein kinases possess an extracellular domain, a transmembrane domain, and an intracellular kinase domain. Binding of a growth factor to the extracellular domain results in receptor dimerization (a chemical reaction in which two identical molecular entities react to form a single molecule, a dimer) and protein kinase activation. The paired kinases phosphorylate each other and other substrate proteins. The phosphorylated tyrosines provide binding, or docking, sites for specific proteins. Binding of a target protein to a docking site results in its activation, and the activated protein initiates a series, or cascade, of downstream phosphorylation reactions. The binding of human epidermal growth factor to members of a family of four receptor tyrosine protein kinases attracts docking proteins that lead to the activation of the Ras-MEK-ERK cascade (see illustration) and growth and cell division. Variations of this cascade occur in yeast and insects in addition to higher animals. See SECOND MESSENGERS.

Protein phosphatases. Protein phosphorylation is a reversible process. Protein kinases add phosphorus groups to proteins, and protein phosphatases remove these groups during a hydrolysis reaction:

\[
\text{Protein} - \text{OPO}_4^{2-} + \text{H}_2\text{O} \rightarrow \text{Protein} - \text{OH} + \text{HoPO}_4^{2-}
\]

To achieve coordinate regulation of cellular processes, both protein kinase activity and protein phosphatase activity are carefully regulated. Much more is known about the regulation of kinases than phosphatases.

Human cells contain genes corresponding to 120 protein phosphatases: 32 are serine/threonine protein phosphatases, 42 are tyrosine protein phosphatases, and 46 are dual-specificity (threonine/tyrosine) phosphatases. There are three subclasses of tyrosine protein phosphatases. One group contains an extracellular receptor domain, a transmembrane segment, and an intracellular catalytic domain. A second class is tethered to the interior of the cell membrane; this class has the potential to be regulated by transmembrane receptors. A third class of tyrosine protein phosphatases is located within the cell, and this class functions in both the cytosol and the cell nucleus.

Serine/threonine protein phosphatases are generally located in the cell cytoplasm. Phosphoprotein phosphatase-1 is inhibited by a protein called phosphotyrosine phosphatase inhibitor-1. Phosphoprotein phosphatases-2a, -2b, and -2c are not inhibited by this protein. Phosphoprotein phosphatase-2b is activated by calcium, a regulatory ion that also activates some protein kinases. There is a balance in the regulation of kinases and phosphatases that leads to physiological responses. Substrates for dual-specificity phosphatases include ERK1, ERK2, and other MAP kinases. These enzymes target those lysine phosphorylations of both threonine and tyrosine residues for activation.

Drug targets. Owing to the numerous functions of protein kinases and phosphatases, these enzymes are recognized as important drug targets. The HER2 (human epidermal growth factor receptor) protein is overexpressed in about 30% of human breast cancers, and these tumors are treated effectively with trastuzumab (Herceptin), an antibody directed to the extracellular region of the HER2 tyrosine protein kinase. Imatinib (Gleevec) is an inhibitor of the Bcr-Abl tyrosine protein kinase that occurs in 95% of people with chronic myelogenous leukemia. Both of these medicinal agents are therapeutically effective. A large amount of research on identifying drug targets for kinases and phosphatases is being performed owing to the success of trastuzumab and imatinib. See CANCER (MEDICINE).

Robert Roskoski, Jr.


Protein metabolism

The transformation and fate of food proteins from their ingestion and assimilation to the elimination of their excretion products. Proteins are polymers of L-amino acids that are connected by peptide bonds. An average polypeptide chain in a protein contains about 500 amino acid residues. Insulin is a small protein with 51 amino acid residues. Titin, which is the largest known human protein, contains 34,530 residues. A dipeptide contains two amino acid residues, an oligopeptide contains several, and a polypeptide contains many amino acid residues. Proteins are the chief structural and functional components of all living organisms. Their importance was recognized by scientists in the midnineteenth century who coined the name from the Greek proteins, meaning first or primary. Proteins are the main building blocks of the cells, tissues, organs, and systems of the body. Proteins of one species differ from those of another species and, within a single animal, proteins of muscle differ from those of the brain, kidney, liver, and other organs. See AMINO ACIDS; METABOLISM; PEPTIDE; PROTEIN.

Functions of protein. A 70-kg (154-lb) person contains 10-12 kg (22-26 lb) of protein. After water, proteins are the most abundant component of humans, other animals, and bacteria. Muscle accounts