Dr. Z Consulting Services[R]

P.O. Box 44
Old Greenwich, CT 06870
Phone: 203 698 0429
Fax: 203 698 0312
E-Mail: techserv@zenitech.com

Dr. Z on Biochemistry

Part 2

Proteins, Amino Acids, Peptides, and Polypeptides - Chemical Structure

Proteins consist of amino acids which are characterized by the -CH(NH_2)COOH substructure. Nitrogen and two hydrogens comprise the amino group, -NH_2, and the acid entity is the carboxyl group, -COOH. Amino acids link to each when the carboxyl group of one molecule reacts with the amino group of another molecule, creating a peptide bond -C(=O)NH- and releasing a molecule of water (H_2O). Amino acids are the basic building blocks of enzymes, hormones, proteins, and body tissues. A peptide is a compound consisting of 2 or more amino acids. Oligopeptides have 10 or fewer amino acids. Polypeptides and proteins are chains of 10 or more amino acids, but peptides consisting of more than 50 amino acids are classified as proteins.

In the animal kingdom, peptides and proteins regulate metabolism and provide structural support. The cells and the organs of our body are controlled by peptide hormones (see table below). Insufficient protein in the diet may prevent the body from producing adequate levels of peptide hormones and structural proteins to sustain normal bodily functions. Deficiency of good quality protein in the diet may contribute to seemingly unrelated symptoms such as sexual dysfunction, blood pressure problems, fatigue, obesity, diabetes, frequent infections, digestive problems, and bone mass loss leading to osteoporosis. Severe restriction of dietary protein causes kwashiorkor which is a form of malnutrition characterized by loss of muscle mass, growth failure, and decreased immunity.

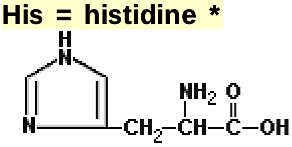
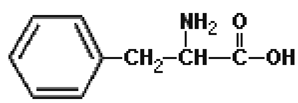
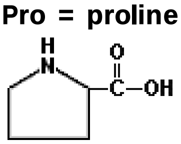
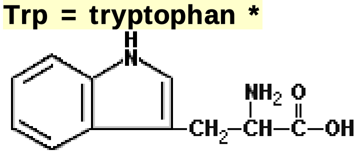
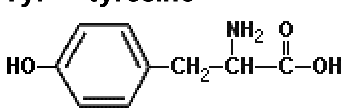
Allergies are generally caused by the effect of foreign proteins on our body. Proteins that are ingested are broken down into smaller peptides and amino acids by digestive enzymes called "proteases". Allergies to foods may be caused by the inability of the body to digest specific proteins. Cooking denatures (inactivates) dietary proteins and facilitates their digestion. Allergies or poisoning may also be caused by exposure to proteins that bypass the digestive system by inhalation, absorption through mucous tissues, or injection by bites or stings. Spider and snake venoms contain proteins that have a variety of neurotoxic, proteolytic, and hemolytic effects.

Many structures of the body are formed from protein. Hair and nails are made of keratins which are long protein chains containing a high percentage (15%-17%) of the amino acid cysteine. Keratins are also components of animal claws, horns, feathers, scales, and hooves. Collagen is the most common protein in the body and comprises approximately 20-30% of all body proteins. It is found in tendons, ligaments, and many tissues that serve structural or mechanical functions. Collagen consists of amino acid sequences that coil into a triple helical structure to form very strong fibers. Glycine and proline account for about 50% of the amino acids in collagen. Gelatin is produced by boiling collagen for a long time until it becomes water soluble and gummy. Tooth enamel and bones consist of a protein matrix (mostly collagen) with dispersed crystals of minerals such as apatite, which is a phosphate of calcium. Muscle tissue consists of approximately 65% actin and myosin, which are the contractile proteins that enable muscle movement. Casein is a nutritive phosphorus-containing protein present in milk. It makes up approximately 80% of the protein in milk and contains all the common amino acids.

Amino Acids

Naturally occurring amino acids, their abbreviations, and structural formulas

*** Essential amino acids**

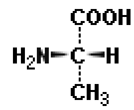
Ala = alanine $CH_3CH(NH_2)COOH$	**Arg = arginine** $H_2N\text{-}C(=NH)NHCH_2CH_2CH_2CH(NH_2)COOH$
Asn = asparagine $H_2N\text{-}C(=O)CH_2CH(NH_2)COOH$	**Asp = aspartic acid** $HOOC\text{-}CH_2CH(NH_2)COOH$
Cys = cysteine $HS\text{-}CH_2CH(NH_2)COOH$	**Gln = glutamine** $H_2N\text{-}C(=O)CH_2CH_2CH(NH_2)COOH$
Glu = glutamic acid $HOOC\text{-}CH_2CH_2CH(NH_2)COOH$	**Gly = glycine** $H_2N\text{-}CH_2COOH$
His = histidine * 	**Ile = isoleucine *** $CH_3CH_2CH(CH_3)CH(NH_2)COOH$
Leu = leucine * $CH_3CH(CH_3)CH_2CH(NH_2)COOH$	**Lys = lysine *** $H_2N\text{-}CH_2CH_2CH_2CH_2CH(NH_2)COOH$
Met = methionine * $CH_3\text{-}S\text{-}CH_2CH_2CH(NH_2)COOH$	**Phe = phenylalanine *** 
Pro = proline 	**Ser = serine** $HOCH_2CH(NH_2)COOH$
Thr = threonine * $CH_3CH(OH)CH(NH_2)COOH$	**Trp = tryptophan *** 
Tyr = tyrosine 	**Val = valine *** $CH_3CH(CH_3)CH(NH_2)COOH$

The term "essential amino acid" refers to an amino acid that is required to meet physiological needs and must be supplied in the diet. Arginine is synthesized by the body, but at a rate that is insufficient to meet growth needs. Methionine is required in large amounts to produce cysteine if the latter amino acid is not adequately supplied in the diet. Similarly, phenylalanine can be converted to tyrosine, but is required in large quantities when the diet is deficient in tyrosine. Tyrosine is essential for people with the disease phenylketonuria (PKU) whose metabolism cannot convert

phenylalanine to tyrosine. Isoleucine, leucine, and valine are sometimes called "branched-chain amino acids" because their carbon chains are branched.

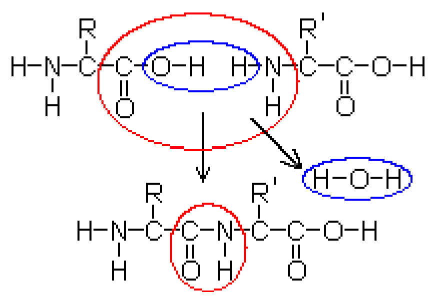
Stereochemistry

In all twenty amino acids, except glycine, the carbon atom with the amino group is attached to four different substituents. The tetrahedral bond angles of carbon and the asymmetry of the attachments make it possible for amino acids to have two non-superimposable structures, the L and R forms, which are mirror images of each other. Only L-amino acids are found in proteins. L-amino acids have the amino group to the left when the carboxyl group is the top, as illustrated here. The wedge bonds are above the display plane and the dotted bonds are below the display plane.



L-Alanine

Formation of a peptide from two amino acids



This illustration shows the reaction of two amino acids, where R and R' are any functional groups from the table above. The blue circle shows the water (H_2O) that is released, and the red circle shows the resulting peptide bond (-C(=O)NH-).

The reverse reaction, i.e., the breakdown of peptide bonds into the component amino acids, is achieved by ***hydrolysis***. Many commercial food products use hydrolyzed vegetable proteins as flavoring agents. Soy sauce is produced by hydrolyzing soybean and wheat protein by fungal fermentation or by boiling with acid solutions. Monosodium glutamate (MSG), a flavor enhancer, is a sodium salt of glutamic acid that is found naturally in seaweed and fermented soy products.

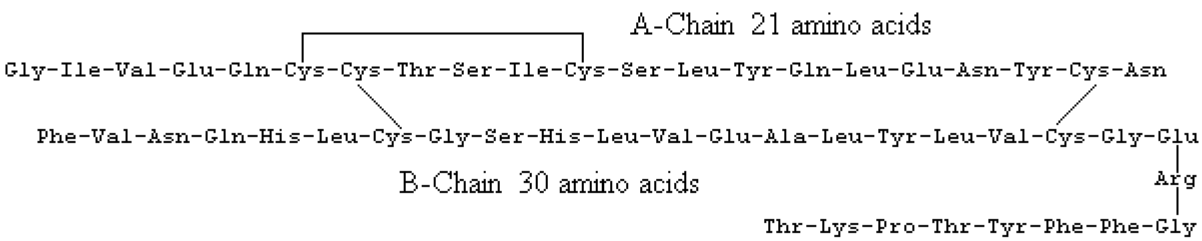
Peptides and Proteins

Peptides consist of two or more amino acids. Polypeptides and proteins both contain ten or more amino acids, but peptides consisting of more than fifty amino acids are classified as proteins.

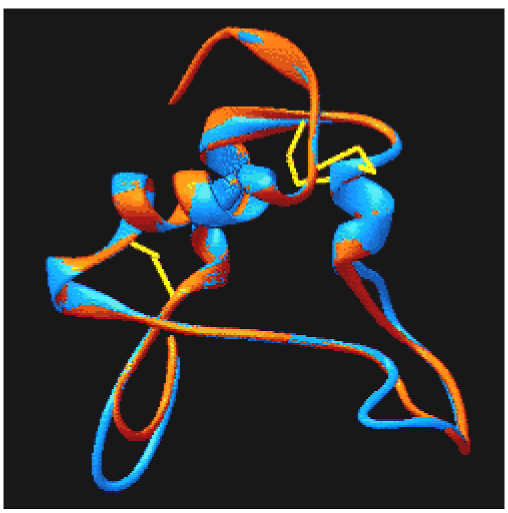
Some Important Peptide Hormones

Hormone	Number of amino acids	Function
Insulin	51	Lowers blood glucose level, promotes glucose storage as glycogen and fat. Fasting decreases insulin production.
Glucagon	29	Increases blood glucose level. Fasting increases glucagon production.
Ghrelin	28	Stimulates release of Growth Hormone, increases feeling of hunger.
Leptin	167	Its presence suppresses the feeling of hunger. Fasting decreases leptin levels.
Growth Hormone	191	Promotes amino acid uptake by cells and regulates development of the body. Growth hormone levels increase during fasting.
Prolactin	198	Initiates and maintains lactation in mammals
Human Placental Lactogen	191	Produced by the placenta late in gestation
Luteinizing Hormone	204	Induces the secretion of testosterone
Follicle Stimulating Hormone	204	Induces the secretion of testosterone and dihydrotestosterone
Chorionic Gonadotropin	237	Produced after implantation of an egg in the placenta
Thyroid Stimulating Hormone	201	Stimulates secretion of thyroxin and triiodothyronine
Adrenocorticotropic Hormone	39	Stimulates production of adrenal cortex steroids (cortisol and costicosterone)
Vasopressin	9	Increases the reabsorption rate of water in kidney tubule cells (antidiuretic hormone)
Oxytocin	9	Causes contraction of mammary gland cells to produce milk and stimulation of uterine muscles during childbirth
Angiotensin II	8	Regulates blood pressure through vasoconstriction
Parathyroid Hormone	84	Increases calcium ion levels in extracellular fluids
Gastrin	14	Regulates secretion of gastric acid and pepsin, a digestive enzyme consisting of 326 amino acids

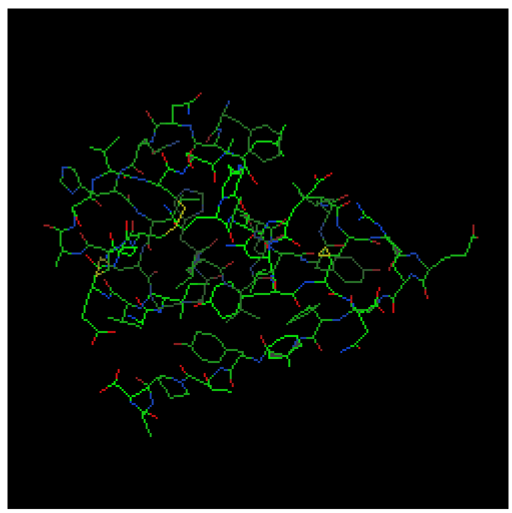
Peptide hormones are produced by the endocrine glands (pituitary, thyroid, pineal, adrenal, pancreas) or by various organs such as the kidney, stomach, intestine, placenta, or liver. Peptide hormones can have complex, convoluted structures with hundreds of amino acids. The following graphics illustrate the chemical structure of human insulin and its three-dimensional shape. Insulin is made of two amino acid sequences. The A-Chain has 21-amino acids, and the B-Chain has 30-amino acids. The chains are linked together through the sulfur atoms of cysteine (Cys). Peptide hormones are generally different for every species, but they may have similarities. Human insulin is identical to pig insulin, except that the last amino acid of the B-Chain for the pig is alanine (Ala) instead of threonine (Thr).



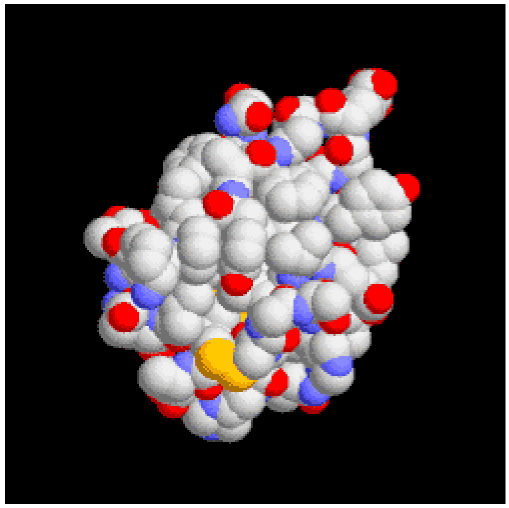
Chemical Structure of Human Insulin



Ribbon representation shows shape of peptide links



Stick representation shows all the atoms

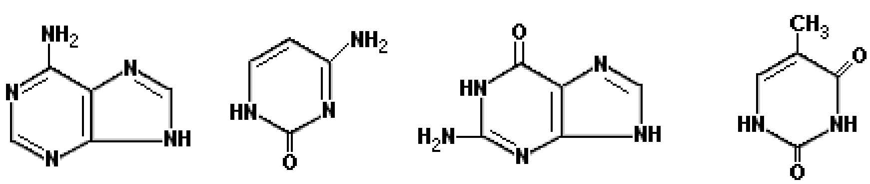


Space-filling representation shows external shape

How are proteins created?

The genetic code in DNA (deoxyribonucleic acid) provides the instructions for building proteins. In the 1960s, Marshal Nirenberg at the National Institutes of Health (NIH) deduced how DNA is mapped into proteins. DNA consists of long molecular sequences containing four nucleotide bases: Adenine (A), Cytosine (C), Guanine (G), and Thymine (T). Each combination of three bases, a DNA codon, corresponds to one specific amino acid. Since there are 64 different 3-base combinations and only 20 amino acids, some combinations do not have unique mappings. The genetic code applies to the vast majority of genes in animals, plants, and microorganisms. The same codons correspond to the same amino acids and to the same START and STOP signals, but in some rare cases one or two of the three STOP codons are assigned to an amino acid instead.

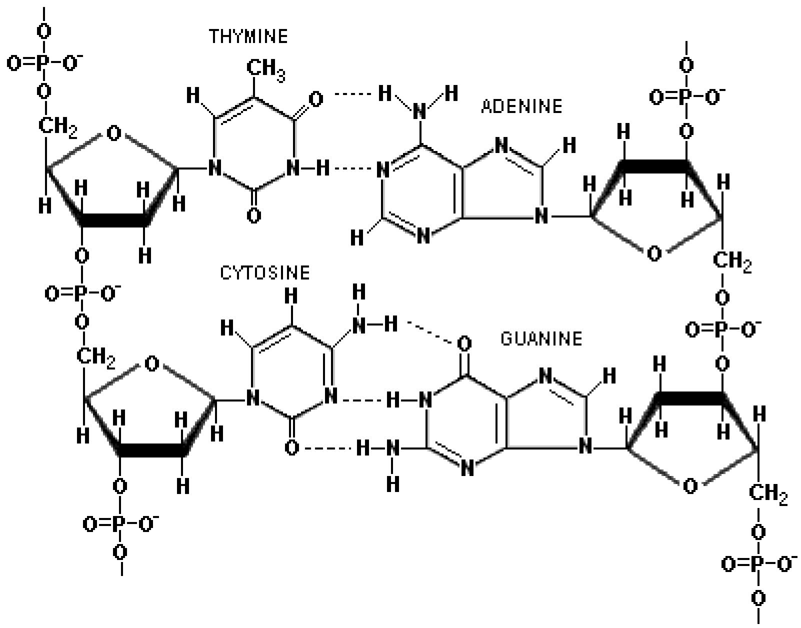
Nucleotide Bases

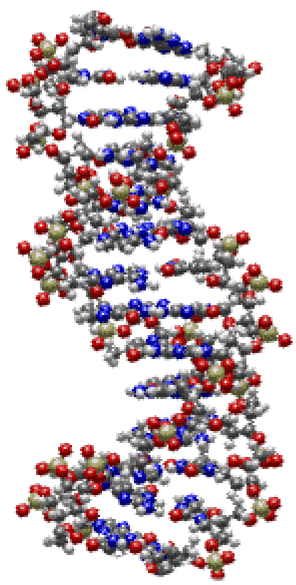


Adenine (A) Cytosine (C) Guanine (G) Thymine (T)

Chemical Structure of DNA

DNA forms a double helix in which the nucleotide bases are attached to deoxyribose units linked through phosphate groups. The bases in the center of the DNA helix always occur in complementary matched pairs, with cytosine linking to guanine and thymine linking to adenine through hydrogen bonding (shown as dotted lines). James Watson and Francis Crick described the structure of DNA in 1953, and received the Nobel Prize in 1962 for this work.

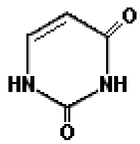




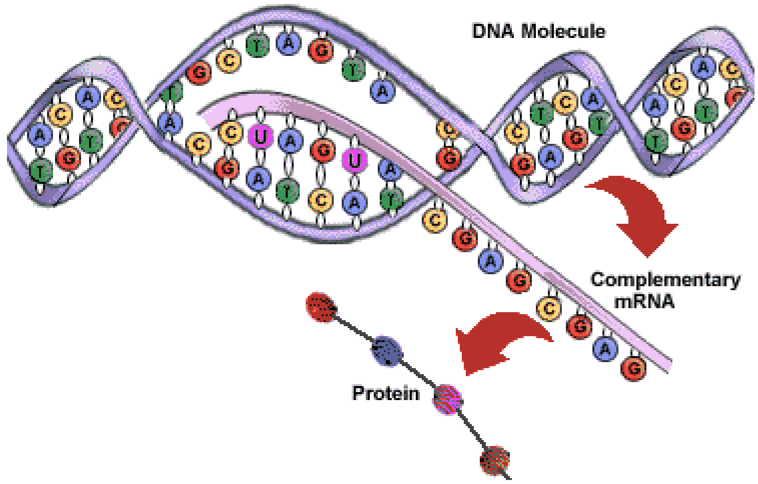
The nucleotide bases in the center of the DNA helix are flanked by deoxyribose units linked by phosphate groups. The figure on the right represents oxygen as red, nitrogen as blue, and phosphorus as olive green.

Transcription of DNA to mRNA, and of mRNA to proteins

The mechanism for producing proteins is analogous to *offset printing* where the image on a printing plate is covered with ink and transferred as a mirror image to a rubber blanket which is then pressed against a sheet of paper to produce the final image. The nucleotide sequence of DNA is not used directly in protein synthesis. Instead, the DNA molecule is transcribed into a complementary sequence of bases, called messenger ribonucleic acid (mRNA), which is then used for protein synthesis. Transcription begins when the DNA double helix hydrogen bonds break and each DNA nucleotide base links with a complementary matching base to build the mRNA molecule. Guanine links with cytosine and cytosine with guanine. Thymine links with adenine, but adenine, which would normally link to thymine, links with Uracil (U) during transcription. As an example, the DNA sequence GATACC is transcribed into the complementary mRNA sequence CUAUGG which builds the amino acid sequence Leu-Trp. The table below shows the correspondence of the amino acids and the mRNA codons.



Uracil (U)



The transcription process

The Genetic Code

Amino Acid	Abb.	SLC	mRNA codons
Alanine	Ala	A	GCA GCC GCG GCU
Arginine	Arg	R	AGA AGG CGA CGC CGG CGU
Asparagine	Asn	N	AAC AAU
Aspartic acid	Asp	D	GAC GAU
Cysteine	Cys	C	UGC UGU
Glutamic acid	Glu	E	GAA GAG
Glutamine	Gln	Q	CAA CAG
Glycine	Gly	G	GGA GGC GGG GGU
Histidine	His	H	CAC CAU
Isoleucine	Ile	I	AUA AUC AUU
Leucine	Leu	L	CUA CUC CUG CUU UUA UUG
Lysine	Lys	K	AAA AAG
Methionine*	Met	M	AUG
Phenylalanine	Phe	F	UUC UUU
Proline	Pro	P	CCA CCC CCG CCU
Serine	Ser	S	AGC AGU UCA UCC UCG UCU
Threonine	Thr	T	ACA ACC ACG ACU
Tryptophan	Trp	W	UGG
Tyrosine	Tyr	Y	UAC UAU
Valine	Val	V	GUA GUC GUG GUU
Stop codons			UAA UAG UGA

SLC is the single-letter code used to represent the amino acids in protein data bases.
Example: The single-letter code for human glucagon is:
HSQGTFTSDYSKYLDSRRAQDFVQWLMNT

Codon letters: A = Adenine, C = Cytosine, G = Guanine, U = Uracil
*** AUG signals "start" of translation when it occurs at the beginning of a gene.**

Amino acid profiles of food proteins

The following table shows representative amino acid profiles of some common foods and dietary protein supplements. The percentages are averages of several commercial products. Casein and whey are milk proteins. Casein is the protein that precipitates from milk when curdled with rennet; it is the basis for making cheese. Whey is the watery part of milk that remains after the casein is separated.

Percentage (%) by weight of amino acid

Amino Acid	protein						
	egg white	beef	chicken	whey	casein	soy	yeast
alanine	6.6	6.1	5.5	5.2	2.9	4.2	8.3
arginine	5.6	6.5	6.0	2.5	3.7	7.5	6.5
aspartic acid	8.9	9.1	8.9	10.9	6.6	11.5	9.8
cysteine	2.5	1.3	1.3	2.2	0.3	1.3	1.4
glutamic acid	13.5	15.0	15.0	16.8	21.5	19.0	13.5
glycine	3.6	6.1	4.9	2.2	2.1	4.1	4.8
histidine *	2.2	3.2	3.1	2.0	3.0	2.6	2.6
isoleucine *	6.0	4.5	5.3	6.0	5.1	4.8	5.0
leucine *	8.5	8.0	7.5	9.5	9.0	8.1	7.1
lysine *	6.2	8.4	8.5	8.8	3.8	6.2	6.9
methionine *	3.6	2.6	2.8	1.9	2.7	1.3	1.5
phenylalanine *	6.0	3.9	4.0	2.3	5.1	5.2	4.7
proline	3.8	4.8	4.1	6.6	10.7	5.1	4.0
serine	7.3	3.9	3.4	5.4	5.6	5.2	5.1
threonine *	4.4	4.0	4.2	6.9	4.3	3.8	5.8
tryptophan *	1.4	0.7	1.2	2.2	1.3	1.3	1.6
tyrosine	2.7	3.2	3.4	2.7	5.6	3.8	5.0
valine *	7.0	5.0	5.0	6.0	6.6	5.0	6.2

*** Essential amino acids**

Egg white protein is considered to have one of the best amino acids profiles for human nutrition. Plant proteins generally have lower content of some essential amino acids such as lysine and methionine. Soy protein is one of the best plant proteins, but nevertheless, the most prominent difference in this chart is the proportion of the essential sulfur-containing amino acid methionine. Egg white protein has approximately three times more methionine than is found in soy protein. The yeast information is for "brewer's yeast" (*Saccharomyces Cervisiae*).