ANSWERS FOR THE BIOCHEMISTRY FINAL EXAM

by CHRISTOS CHINOPOULOS, M.D. Ph.D.

Semmelweis University of Medicine
Department of Medical Biochemistry

Budapest 1998

Dear students,

This manuscript is by no means intended to replace any textbook of Biochemistry, or your lecture notes. While you will start reading it, you will realize that you will not understand everything, simply because Biochemistry is a very coherent subject, and the information data that it contains cannot be just dissected away from each other. The purpose of this manuscript, is not more than to give you a brief reference to each question required to be known by the student at the exam, and also to provide some information on the field of Biochemistry that are not readily apparent after reading the book.
After having read Biochemistry, and being fed up with the hundreds of regulations, metabolic intermediates and enzymes, you must realize that the most important enzyme (which is not mentioned in any textbook of Biochemistry) is the biochemase, in order for you to digest this huge material, which can only be done by PPP (Pen, Paper, and Patience), by writing the pathways again and again.

Good Luck

Christos Chinopoulos, M.D. Ph.D.

BIOENERGETICS

For bioenergetics only the answers 16 and 17 will be provided, as the rest of the questions rely on material of the First Year, Second Semester under the title "Descriptive Biochemistry and Bioenergetics".

16) What is the mechanism of development of lactic acidosis in the absence of pyruvate dehydrogenase activity? What further symptoms can be observed if pyruvate dehydrogenase deficiency is caused by the lack of lipoamide dehydrogenase subunit?

Theoretical Background:

-The lactic acid that circulates in the human body is the product of anaerobic metabolism of glucose that takes place primarily in RBC`s, skin, kidney medulla, and white skeletal muscle. Some of it is oxidized by red muscle and kidney cortex, but the bulk of it is taken up by the liver and converted into glucose.

-Some tissues, such as blood, have either very few (white blood cells) or no (red blood cells) mitochondria, and thus a glycolytic mode of metabolism is obligatory. Other tissues have such a high glycolytic capacity that oxidative metabolism is suppressed, a phenomenon known as the Crabtree effect.

-In case of pyruvate dehydrogenase deficiency, lactic acidosis develops because of pile up of pyruvate, that it will convert to lactate, in order to regenerate NAD^+ as the citric acid cycle (TCA) is nonfunctional.

-Further symptoms if PDH deficiency is caused by the lack of lipoamide dehydrogenase subunit: As this subunit participates in other enzymes too, namely the α–Kg DHase, the Branched Chain Keto Acid DHase, and the glycine cleavage system, the symptoms will be the following: Slowing down of the TCA, inability of catabolism of BCAA which are preferred by the muscle and the brain, and hyperglycinemia.

17) What is the mechanism of brain damage detected in pyruvate carboxylase deficiency? Explain the development of hyperammonemia. Which metabolic pathways can be associated with the pyruvate carboxylase function?

-The reasons for brain damage are:

a) Absence of TCA.

b) Hyperammonemia resulting to cerebral edema, due to absence of OAA that would give Aspartate. Asp would give a Nitrogen to combine with ammonia and give urea (see Harper, Fig. 31-14).
c) No, Glutamate, No GABA formation.
d) No Acetylcholine formation.
e) No myelin synthesis, because of low citrate levels.

-Metabolic pathways associated with PC function: See question ***22***.

CARBOHYDRATES

1) Sequence of reactions in Glycolysis I

-Harper, Fig. 19-2.
-Glycolysis is found in the cytosol of the cell, along with the Pentose Phosphate Pathway (PPP), and the Fatty Acid synthesis (plus others). It was the first pathway discovered by researchers, and the first pathway on earth, simply because it does not require oxygen, which was not available in the atmosphere when life started to evolve.
-Although the enzymes of glycolysis are usually described as soluble components of the cytosol, it seems, that they form large multienzyme complexes, for the effective shuttling of the metabolic intermediates and for minimizing the diffusion factor that could limit the speed of the reactions. Furthermore, certain glycolytic enzymes form specific noncovalent complexes with structural components of the cell, which may serve to organize reaction sequences and assure efficient transfer of intermediates between cellular compartments; e.g. (though not a glycolytic enzyme) Glucose –6-Phosphatase is found in the luminal side of the ER.. Certain glycolytic enzymes bind to microtubules, or to actin microfilaments, bringing these enzymes into close association and holding them in a specific region of the cytoplasm; e.g. hexokinase (HK) is bound to the outer membrane of the mitochondria, allowing ATP produced in the mitochondrion to move directly to the catalytic site of the HK, without entering and being diluted by the cytosol.
In malignant cells, glycolysis proceeds at much higher rates than normal, more than the citric acid cycle can take. Thus, an excess of pyruvate is formed that it can only be metabolized to lactate (for the provision of NAD^+). Therefore, lactate builds up, favoring an acidic environment for the tumor. The pathway from the glycolysis until the formation of lactate is also known as Embden-Meyerhof pathway.
-Certain tissues and cell types (retina, brain, RBC`s), convert glucose to lactate even under aerobic conditions.
-Glycolysis is also the pathway for the oxidation of Fructose (bypassing the Phosphofructokinase [PFK]-1 step), Galactose (see question 6) and Mannose.
-Glucose enters cells through the action of glucose transporters. These are:
i) GLUT 1: (RBC`s, brain, muscle, adipocytes)
ii) GLUT 2: (liver, beta pancreatic cells, kidney, intestinal epithelial cells). It has a high K_M (low affinity-great capacity). It works both ways, in and out.
iii) GLUT 3: (brain)

iv) GLUT 4: (muscle, adipocytes). It is the only insulin-dependent glucose transporter.
v) GLUT 5: fructose transporter.
vi) GLUT 6: similar to GLUT 3.
vii) GLUT 7: found within the ER of liver and kidney cells. It translocates glucose molecules to the exterior of the cell after the action of Glc-6-Phosphatase.
-Hexokinase (the first step in glycolysis) is found in all extrahepatic tissues, except the beta cells of the pancreas. In liver and beta pancreatic cells, there is glucokinase (GK). Hexokinase and glucokinase are isoenzymes, they differ in their kinetic properties (Harper, Fig. 21-5), and in the way they are regulated (the other name of glucokinase is hexokinase D).
-Hexokinase (or glucokinase) phosphorylates glucose to Glc-6-P as it enters the cell for the following reasons:

a) To maintain the gradient between the extracellular and intracellular compartment, in order for glucose to have always a tendency towards the interior of the cell.
b) All phosphorylated molecules cannot exit the cell, therefore, Glc-6-P is trapped in the interior of the cell; of course, a gradient is formed for the Glc-6-P towards the extracellular milieu, but now Glc-6-P cannot even approach the inner side of the plasma membrane, because both Glc-6-P and the inner plasma membrane are negatively charged).
c) The phosphorylation of glucose by hexokinase (or glucokinase), serves as an "activation" for the provision of energy from glucose as a substrate, at the substrate level.
d) It serves also for the facilitation of the glycolytic reactions, due to the fact that, binding of phosphate groups to the active sites of enzymes provides binding energy that contributes to lowering the activation energy, and increasing the specificity of enzyme catalyzed reactions.

2) Sequence of reactions in Glycolysis II

-Harper, Fig. 19-2.
-In red blood cells, the 6th step of glycolysis which is the Phosphoglycerate kinase (PGK) step- the only reversible kinase of glycolysis- (Harper, Fig. 19-4), can be bypassed in order for the RBC`s to form 2.3 BPG. The latter molecule, binds to hemoglobin and decreases its` affinity for oxygen, with a concomitant displacement of the oxyhemoglobin dissociation curve to the right, therefore promoting the release of oxygen to the tissues. However, if it is 2,3 BPG to be formed, PGK is bypassed, and therefore there will be no net formation of 2 ATP`s, as PGK generates 1 ATP two times. This bypass happens only in the arterial blood when there is a need for oxygen release to the tissues; in venous blood 2,3 BPG is not formed, and PGK is not bypassed. The two enzymes participating in the bypass of PGK, are Bisphosphoglycerate mutase and 2,3 Bisphosphoglycerate phosphatase, which are both parts of a tandem enzyme (tandem enzyme means that both enzymes are found on the same polypeptide chain).
-Glycolysis in RBC`s always terminates in lactate formation (RBC`s posses no mitochondria).
-In mitochondria containing cells, as pyruvate is formed, it enters the mitochondria either with a cotransport of a proton, or the antiport of a hydroxyl group. Note that pyruvate possesses a negative charge, and this proton cotransport or hydroxyl antiport system, contributes to the

maintenance of the charge balance between the two sides of the mitochondria; moreover, the inner membrane space of mitochondria is very negatively charged, and it would be very unlikely for a negative molecule such as pyruvate to enter this space without the expenditure of energy, or the usage of a transport system.

3) Regulation and energetics of glycolysis. The Pasteur effect

-Harper, Fig.21-1, Table 19-1.
-"Nonequilibrium" reactions (monodromic) are potential targets for regulation. A nonequilibrium reaction is identified as a reaction in which the K_M of the enzyme is considerably lower than the normal substrate concentration. In glycolysis, there are three such targets:
a) At the level of the phosphorylation of glucose immediately as it enters the cell; namely the HK or GK step.
b) At the level of PFK-1.
c) At the level of Pyruvate Kinase (PK).
-a) Hexokinase is subject to allosteric inhibition by its` own product (negative feedback), Glc-6-P (and this is the only regulation of hexokinase).
-Glucokinase is subject to the following regulations: Induced by insulin, and repressed by glucagon (glucagon action mediated by cAMP), high carbohydrate feeding upregulates the GK gene expression while starvation and diabetes have opposite effects. GK is also inhibited by Frc-6-P. This partial inhibition of GK by Frc-6-P is mediated by an additional protein, the regulatory protein. This regulatory protein has also affinity for Frc-1-P, which competes with Frc-6-P and cancels its` inhibitory effect on GK; (Frc-1-P displaces Frc-6-P from the regulator protein, so the latter cannot inhibit GK). Because Frc-1-P is present in the liver only when there is fructose in the blood, this property of the regulator protein explains the observation that ingested fructose stimulates the phosphorylation of glucose in the liver. The beta pancreatic cells also contain this regulator protein.
-PFK-1 is the committed step of glycolysis. That means that x% inhibition of PFK-1 would give an x% inhibition of the final glycolytic product, that is pyruvate. PFK-1 is allosterically activated by:
a) Frc-6-P (positive feedforward)
b) AMP, ADP
c) Frc-2,6-BP_2 (Frc-2,6-BP_2 is formed by PFK-2 which is inhibited by glucagon, acting through cAMP, see Harper, Fig.21-3). Frc-2,6-BP_2 is the most potent allosteric effector on PFK-1 than any other signal.
-PFK-1 is allosterically inhibited by:
a) ATP
b) Citrate (it is a metabolic index in the mitochondria)
-PFK-1 gene is subject to induction by insulin and high carbohydrate feeding, while starvation and diabetes have opposite effects.
-PK is allosterically inhibited by:
a) Alanine (gluconeogenic signal)
b) ATP
c) Acetyl-CoA and Long Chain Fatty Acids (LCFA)
while PK is allosterically activated by:
a) Frc-1,6-BP_2 (it is a metabolic index in the cytosol)

b) Frc-1-P
-PK is very tightly regulated in the liver by the aforementioned mechanisms, and this serves to prevent the **"leak down"** of pyruvate formed by phosphoenolpyruvate carboxykinase [PEPCK] (in case of gluconeogenesis) that would convert to oxaloacetate (OAA) and then to phosphoenolpyruvate (PEP), and that would be a futile cycle.
-PK is also subject to covalent modification by phosphorylation/de phosphorylation mechanisms.
-The Pasteur effect: Originally, Pasteur observed that in the presence of oxygen, yeast consume more than 10 times less their carbohydrate supplies, than when they are incubated in the absence of oxygen. The explanation of this phenomenon is the following: The aerobic oxidation (via citric acid cycle, TCA) inhibits the anaerobic degradation of glucose. This is because in the presence of oxygen, the NADH/NAD ration would be low, and therefore, PDH would be uninhibited to form Acetyl-CoA, and this would give citrate, and a lot of ATP through the TCA. Citrate and ATP would inhibit PFK-1, and therefore, glycolysis would cease. In the absence of oxygen, the NADH/NAD ratio would be high, as there would be no final reducing equivalent acceptors, PDH would be inhibited to give Acetyl-CoA and therefore citrate and ATP. So, PFK-1 would not be inhibited, and glucose would be consumed at a much higher rate, in order to compensate for the "loss" of TCA and its` product, ATP.

4) Digestion and absorption of carbohydrates

Dietary carbohydrates provide a major portion of the daily caloric requirement. They consist of mono-, di-, and polysaccharides. Monosaccharides need not be hydrolyzed for absorption. Disaccharides require the intestinal surface enzymes for hydrolysis into monosaccharides, while polysaccharides depend on pancreatic amylase for degradation (amylase of the saliva plays a very minor role). Exceptionally, trisaccharides cannot be hydrolyzed by humans. Starch consists of amylose (1,4 bonds) and amylopectin (1,6 bonds). Hydrated starch and glycogen are attacked by the endosaccharidase α-amylase present in the pancreatic juice (and in saliva). Hydration of the polysaccharides occurs during heating (cooking) and is essential for efficient digestion. The products of α-amylase are maltose, maltotriose, and α-dextrins, containing on average 8 glucose units. Final hydrolysis of di- and oligosaccharides to monosaccharides is carried out by surface enzymes of the intestinal epithelial cells. Most of those are exoenzymes that cleave off one monosaccharide at a time from the nonreducing end. The capacity of the α-glucosidases is normally much higher than that needed for completion of the digestion of starch. Similarly, there is usually excess capacity for sucrose hydrolysis relative to dietary intake. In contrast, β-galactosidase (lactase) can be rate-limiting in humans for hydrolysis and utilization of lactose, the major milk carbohydrate.
The consequences of an inability to hydrolyze lactose in the upper small intestine are inability to absorb lactose and bacterial fermentation of ingested lactose in the lower small intestine. Bacterial fermentation results in the production of gas (distention of gut and flatulence) and osmotically active solutes that draw water into the intestinal lumen (diarrhea). The lactose in yogurt has already been partially hydrolyzed during the fermentation process of making yogurt. Thus, individuals with lactase deficiency can often tolerate yogurt better than

unfermented dairy products (e.g. milk). The enzyme lactase is commercially available to pretreat milk so that lactose is hydrolyzed.
-The well-known problem of flatulence after ingestion of leguminous seeds (beans, peas, soya) is caused by oligosaccharides, which cannot be hydrolyzed by human intestinal enzymes.
-At least 2 types of monosaccharide transporters catalyze monosaccharide uptake form the lumen into the cell:
i) a Na^+ monosaccharide cotransporter (SLGT)
ii) a Na^+ independent, facilitated diffusion type of monosaccharide transport system, with specificity for fructose (GLUT 5). In addition, a Na^+ independent monosaccharide transporter (GLUT 2) is present in the contraluminal plasma membrane.

5) Metabolism of fructose, fructose intolerance, essential fructosuria

-Harper, Fig. 22-5.
-Essential fructosuria: It is a benign condition, attributed to lack of hepatic fructokinase. Diets low in fructose and sucrose is beneficial.
-See also answer ***23***.

6) Metabolism of galactose, galactosemia, lactose synthesis

-Harper, Fig. 22-6.
-Galactose is derived from the intestinal hydrolysis of lactose.
-There cannot be direct conversion of galactose to glucose without to have glycogenesis first. Glucose is liberated from UDPGlc as Glc-1-P probably after incorporation into glycogen followed by phosphorolysis.
-Since the epimerase reaction is freely reversible, glucose can be converted to galactose, so that preformed galactose is not essential in the diet. Galactose is required in the body not only in the formation of lactose (in mammary gland), but also as a constituent of glycolipids (cerebrosides), proteoglycans, and glycoproteins.
-Galactosemias can be the cause of inherited defects of galactokinase, 4-epimerase, or uridyl transferase. Galactose, which increases in concentration in the blood, is reduced by aldose reductase in the eye to the corresponding polyol (galactitol), which accumulates, causing cataract. The general condition is more severe if it is due to a defect in the uridyl transferase, since galactose-1-P accumulates and depletes the liver of inorganic phosphate.

7) Reactions of the pentose phosphate pathway I. Oxidative branch

-Harper, Fig. 22-1, 22-2.

PPP does not generate ATP. It generates NADPH+ H^+ (oxidative phase), for reductive syntheses, such as FA synthesis and steroid biosynthesis. It is also used for the provision of ribose residues (non-oxidative phase) for nucleotide and nucleic acid biosynthesis.
-Deficiencies of PPP cause hemolytic anemias, e.g. G6PD deficiency.
-PPP is found in the cytosol.
-The oxidative phase is nonreversible, while the oxidative is reversible.
-Thiamin is a cofactor for the transketolase reaction.
-PPP is active in the liver (FA synthesis), adipose tissue (FA synthesis), adrenal cortex (steroid biosynthesis), RBC`s (maintenance of reduced glutathione, see Harper, Fig. 22-3).
-PPP has a low activity in skeletal muscle.
-PPP has a very high activity in cancer cells (ribose availability).
-NADPH+H^+ is also used for the synthesis of amino acids (aa) via glutamate dehydrogenase.

8) Reactions of the pentose phosphate pathway I. Non-oxidative branch

See question ***7, and 9***.

9) Metabolic significance, regulation and clinical aspects of the pentose phosphate pathway

-It is probable that the presence of active lipogenesis or of a system which utilizes NADPH+H^+ stimulates an active degradation of glucose via the PPP. This provides NADPH+H^+, which is normally very low in concentration because of the non-equilibrium nature of the first reactions in the pathway. The synthesis of Glucose-6-P Dehydrogenase and 6-Phosphogluconate Dehydrogenase may also be induced by insulin during conditions associated with the “fed state”.
-Even though skeletal muscle has very low activities of Glucose-6-P Dehydrogenase and 6-Phosphogluconate Dehydrogenase, it is still capable of forming ribose-5-P for nucleotide synthesis. This is accomplished by a reversal of the non-oxidative branch of PPP, utilizing fructose-6-P. Thus, it is not necessary to have a completely functioning PPP, for a tissue to synthesize ribose phosphates. Moreover, ribose is not a significant constituent of the blood, and therefore, tissues must satisfy their own requirements for this nucleotide precursor.
-**Harper, Fig. 22-3**
-PPP in the RBC`s provides NADPH for the reduction of oxidized glutathione to reduced glutathione catalyzed by glutathione reductase, a flavoprotein enzyme containing FAD. In turn, reduced glutathione removes hydrogen peroxide from the erythrocyte, in a reaction catalyzed by glutathione peroxidase, an enzyme that contains the trace element **selenium**. This reaction is important, since hydrogen peroxide may decrease the life span of the erythrocyte by increasing the rate of oxidation of hemoglobin to methemoglobin. Methemoglobin is also produced by the

ingestion of nitrate and/or nitrite contaminated water, and this is especially important for newborns, as their methemoglobin reductase activity is very low (methemoglobin reductase is the enzyme that by utilizing NADH [not NADPH!] would revert methemoglobin to hemoglobin). Navaho Indians also have very low activities of methemoglobin reductase, and therefore are prone to develop methemoglobinemias.

10) NADPH producing reactions in the cell

-Harper, Fig. 23-5.
-These are: i) extramitochondrial Isocitrate Dehydrogenase
ii) Malic enzyme
iii) PPP
iv) a proton-translocating transhydrogenase is a source of intramitochondrial $NADPH+H^+$; it is a protein in the inner mitochondrial membrane, coupling the passage of protons down the electrochemical gradient form outside to inside the mitochondrion, with the transfer of H^+ from intramitochondrial $NADH+ H^+$ to form $NADPH+ H^+$. It appears to function as an energy linked redox buffer and as a source of NADPH for intramitochondrial enzymes, such as glutamate dehydrogenase and hydroxylases, involved in steroid synthesis.

11) The synthesis and biological role of UDP-glucose

-Harper, Fig.20-1, Fig. 20-2.
-UDP-glucose is the glycosyl donor for biosynthesis of glycogen and glycosyl disaccharides e.g. lactose.
-UDP-glucose can exchange (4-epimerase) with other UDP-sugars serving as sugar donors for biosynthesis of the oligosaccharides of glycoproteins and proteoglycans.
-UDP-glucose is also the starting point (UDP Dehydrogenase) for the building block of glucuronides and proteoglycans, and also for the formation of UDP-glucuronic acid which is the glycosidic acid donor for the conjugation reactions that form the urinary glucuronide conjugates of bilirubin or drugs such as aspirin.

12) The pathway of glycogen biosynthesis. The biomedical importance of glycogen

-Harper, Fig.20-1, Fig. 20-3, Fig 20-4.

-Glycogen formation requires a primer: glycogenin. It is a protein in which 4 or more glucose units are attached on 4 or more tyrosine residues. 0.3 % of the protein part of glycogen in the liver is glycogenin.
-Glycogen function:
a) To reduce intracellular osmotic forces; 1 mmol of glycogen is equivalent to 400 mmol of glucose.
b) As an immediate storage (time is an important parameter; in an immediate storage place we need plenty of entrances and plenty of exits [branching]).
-Glycogenolysis is an energetically favorable reaction.
-The site of glycogen degradation is on the non-reducing end of glycogen.
-Glycogen phosphorylase requires PLP (in order to form a Schiff`s base during catalysis). It also possesses a calmodulin component.
-Glycogen synhtase has 9 phosphorylation sites.
-10 % of the liver mass is glycogen.
-Formation of UDPGlc is an energetically unfavorable reaction, but the subsequent hydrolysis of PP_i, pulls the overall reaction to the right.
-Phosphoglucomutase (PGM) can be inhibited by DIPF as all enzymes possessing a serine in their active site (however, PGM is not a serine protease!)

13) The pathway of glycogenolysis. Glycogen storage diseases

-See question ***12***. Harper, Table 20-2.
-Muscle does not have Glucose-6-Phosphatase, and therefore, it cannot contribute to the maintenance of blood glucose level, it can only consume its own glycogen.
-In glycogenolysis, the glycogen phosphorylase step is the rate limiting one.
-cAMP activates muscle phosphorylase
-Ca^{2+} synchronizes the activation of phosphorylase with muscle contraction.
-Glycogenolysis in liver can be cAMP independent.
-The debranching enzyme releases free glucose.

14) Regulation of glycogenesis in the liver and in the muscle

-Harper, Fig.20-7.
-See question ***13***.

15) Regulation of glycogenolysis in the liver and in the muscle

-Harper, Fig.20-6, Fig. 20-8.

-See question ***13***.
-In resting muscle nearly all the phosphorylase is in the β (inactive) form, because ATP is present at a much lower concentration than AMP. Vigorous muscular activity increases the AMP /ATP ratio, very rapidly activating (in msec) phosphorylase β by allosteric means. On a larger time scale (sec to min) hormonally induced phosphorylation of phosphorylase β converts it into phosphorylase α (active form), the activity of which is independent of the AMP/ ATP ratio.
-Liver glycogen phosphorylase, like that of muscle, is subject to allosteric regulation, but in this case the allosteric regulator is glucose, not AMP.

16) The reactions of gluconeogenesis

-Harper, Fig. 21-1, 21-2.
-The overall reaction of gluconeogenesis is: 2 pyr + 4ATP + 2GTP + 2NADH + $2H_2O \longrightarrow$ Glc + 4ADP + 2GDP + 2NAD + $2P_i$ ΔG = -9 kcal/mol
-Liver and kidney can convert noncarbohydrate metabolites such as lactate, glycerol and amino acids to glucose (gluconeogenesis).
-Gluconeogenesis is used for the maintenance of the basal requirement for glucose (4.5 to 5.5 mmol/L for the brain, RBC`s and the retina), and in addition to clear the products of the metabolism of other tissues from the blood, e.g. lactate produced by the muscle, brain and RBC`s, and glycerol which is continuously produced by adipose tissue.
-Pyruvate carboxylase (PC) which is considered to be an enzyme of gluconeogenesis requires biotin as a cofactor.
-PEPCK (in humans) is found both in the cytosol and inside the mitochondria.
-The presence of Fructose-1,6-Bphosphatase, determines whether or not a tissue is capable of synthesizing glycogen not only from pyruvate, but also from triosophosphates. It is present in the liver, kidney and striated skeletal muscle.
-Glycerol kinase has a very low activity in the muscle and in adipose tissue, so most of the glycerol-3-P must be derived from an intermediate of the glycolytic system, Dihydroxyacetone-P, which forms glycerol-3-P by reduction with $NADH+H^+$ catalyzed by Glycerol-3-P Dehydrogenase. Glycerol kinase activity is very high in the liver and in the kidney.
-The activation of PC and the reciprocal inhibition of PDH by Acetyl-CoA derived from the oxidation of FAs, helps to explain the action of FA oxidation in sparing the oxidation of pyruvate and in stimulating gluconeogenesis in the liver. The reciprocal relationship between the activity of PDH and PC in both liver and kidney alters the metabolic fate of pyruvate as the tissue changes from carbohydrate oxidation, via glycolysis, to gluconeogenesis during transition from a fed to a starved state.
-A major role of FA oxidation in promoting gluconeogenesis is to supply ATP required in the PC and PEPCK reactions, as well as reversing the PGK reaction of glycolysis. That is why, impairment of FA oxidation leads to hypoglycemia.

17) Regulations of gluconeogenesis

-Harper, Fig. 21-1, Table 21-1.
-Under conditions of glucose shortage, gluconeogenesis is stimulated by a decrease in the concentration of Frc-2,6-BP_2, which deactivates PFK-1, and deinhibits Fructose-1,6-Bphosphatase. **This mechanism also ensures that glucagon stimulation of glycogenolysis in the liver results in glucose release rather than glycolysis.**
-The glucose concentration in the blood is an important factor controlling the rate of uptake of glucose in both liver and extrahepatic tissues.
-Insulin has an immediate effect of increasing glucose uptake in tissues such as adipose tissue and muscle. This action is due to an enhancement of glucose transport through the cell membrane by recruitment of glucose transporters from the interior of the cell to the plasma membrane. In contrast, there is no direct effect of insulin on glucose penetration of hepatic cells; this finding agrees with the fact that the glucose metabolism by liver cells is not rate-limited by their permeability to glucose. However, insulin does -indirectly- enhance uptake of glucose by the liver as a result of its` actions on the enzymes controlling glycolysis and glycogenesis (positive inductive effect at the level of gene expression).
-PEPCK is regulated at the level of gene expression **only**. It is the committed step of gluconeogenesis.

18) Coordinated regulation of gluconeogenesis and glycolysis in the liver

-Harper, Fig. 21-3.

19) Metabolism and biological role of amino sugars

Harper, Fig. 22-7.
-Amino sugars are important components of glycoproteins, glycolipids, and of glycosaminoglycans. See also question ***20***.

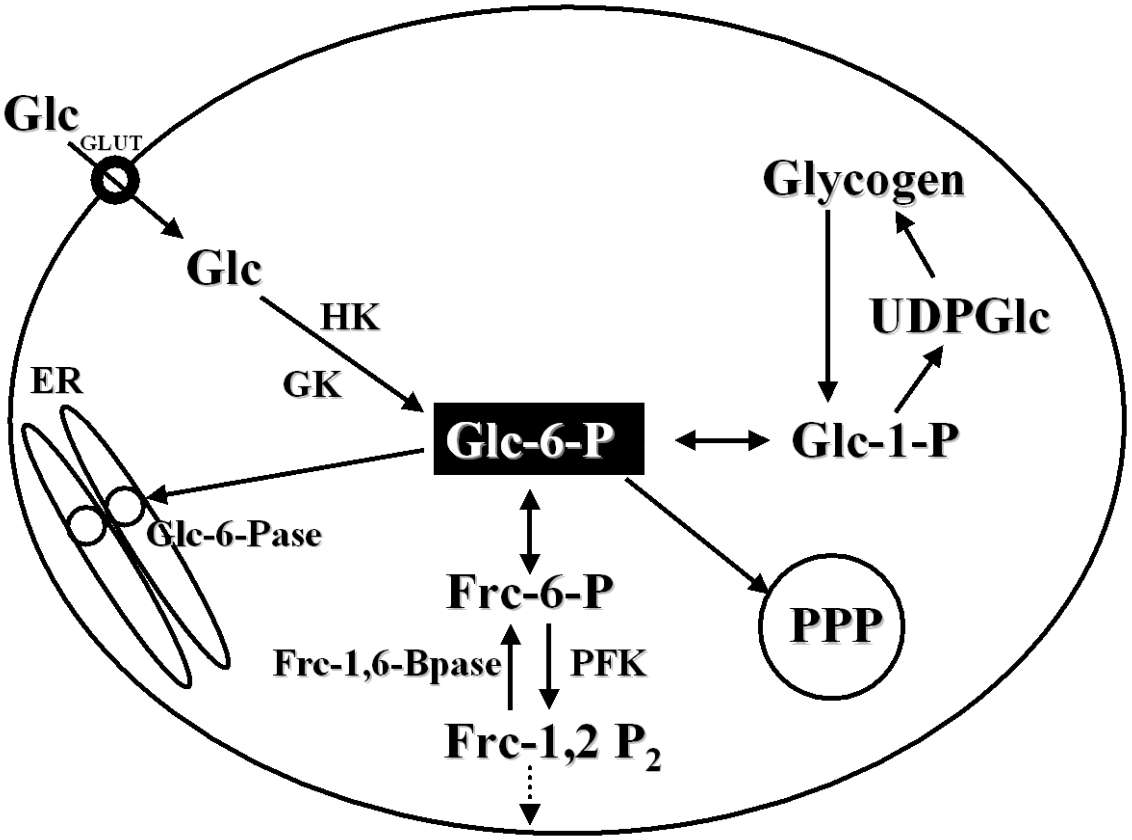
20) Glycoproteins: classification, synthesis, regulation

-Harper, Fig. 56-1, Fig. 56-2, Fig. 56-3, Fig. 56-4, Fig. 56-5, Fig. 56-6, Fig. 56-7, Fig. 56-8, Fig. 56-9.
-Glycoproteins can be divided in 4 major classes, based on the nature of the linkage between their polypeptide chains and their oligosaccharide chains:

i) those containing a serine (or threonine)- GalNAc linkage
ii) those containing a serine-xylose linkage
iii) collagens containing a hyl-gal linkage
iv) glycoproteins containing an Asn-GlcNAc linkage.

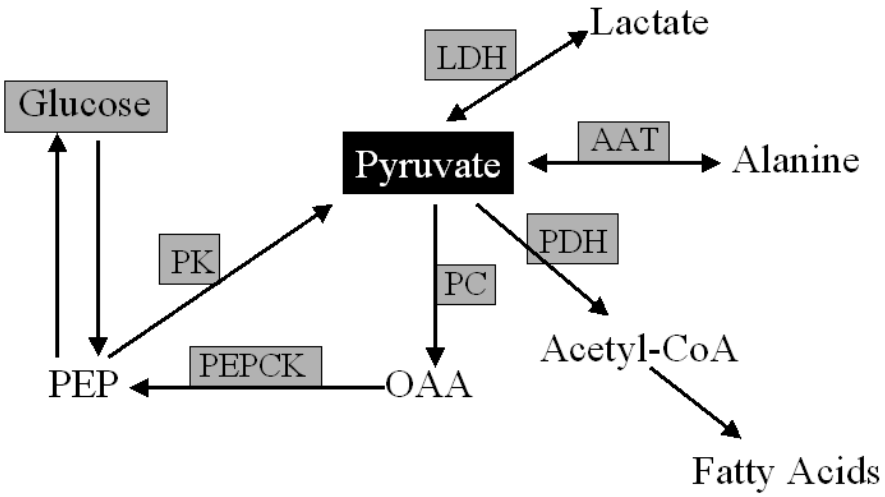
Classes i), ii) and iii) are joined to the corresponding O-glycosidic linkages (i.e., a linkage involving an –OH in the side chain of an aa and a sugar residue). The iv) class involves an N-glycosidic linkage (i.e., a linkage involving the N of the amide group of Asn and a sugar residue).
-Some glycoproteins contain both N- and O-glycosidic linkages.
-Mucins are rich in O-glycosidic linkages.
-The sugars of the oligosaccharide chains of the O-glycosidic type of glycoproteins are built up by the stepwise donation of sugars from nucleotide sugars, such as UDPGalNAc, UDPGal, and CMPNeuAc. The enzymes catalyzing this type of reaction are membrane bound glycoprotein glycosyltransferases. The synthesis of each such enzyme is controlled by one specific gene. Generally, synthesis of one specific type of linkage requires the activity of a corresponding specific transferase (the "one to one " hypothesis). The enzymes catalyzing the addition of the inner sugar residues are located in the ER, and addition of the first sugars occurs during translation (cotranslational modification). The enzymes adding the terminal sugars (such as NeuAc) are located in the Golgi apparatus.
-There are 3 major classes of N-linked glycoproteins: complex, hybrid, and high-mannose. Each type shares a common pentasaccharide ($[Man]_3[GlcNAc]_2$) but they differ in their outer branches. Glycoproteins of the complex type generally contain terminal NeuAc residues and underlying galactose and GlcNAc residues, the latter often constituting the disaccharide lactosamine. The presence of repeating lactosamine units characterizes a 4^{th} class of N-linked glycoproteins, the polylactosamine class. It is important, because the I/i blood group substances belong to this class.
-The oligosaccharide branches are often referred to as antennae. Other complex chains may terminate in galactose or fucose. High-mannose oligosaccharides typically have 2-6 additional mannose residues linked to the pentasaccharide core. Hybrid molecules contain features of both of the 2 other classes.
-The biosynthesis of N-linked glycoproteins involves oligosaccharide -P-P- Dolichol. The oligosaccharide chain is transferred en bloc to suitable Asn residues of acceptor glycoproteins.
–The phenomenon whereby the glycan chains of N-linked glycoproteins are first partially degraded and then in some cases rebuilt, is referred to as oligosaccharide processing.
-Regulation: Harper, Table 56-11.

21) Role of key junctions in the regulation of metabolism
Glucose-6-Phosphate



22) Role of key junctions in the regulation of metabolism

Pyruvate



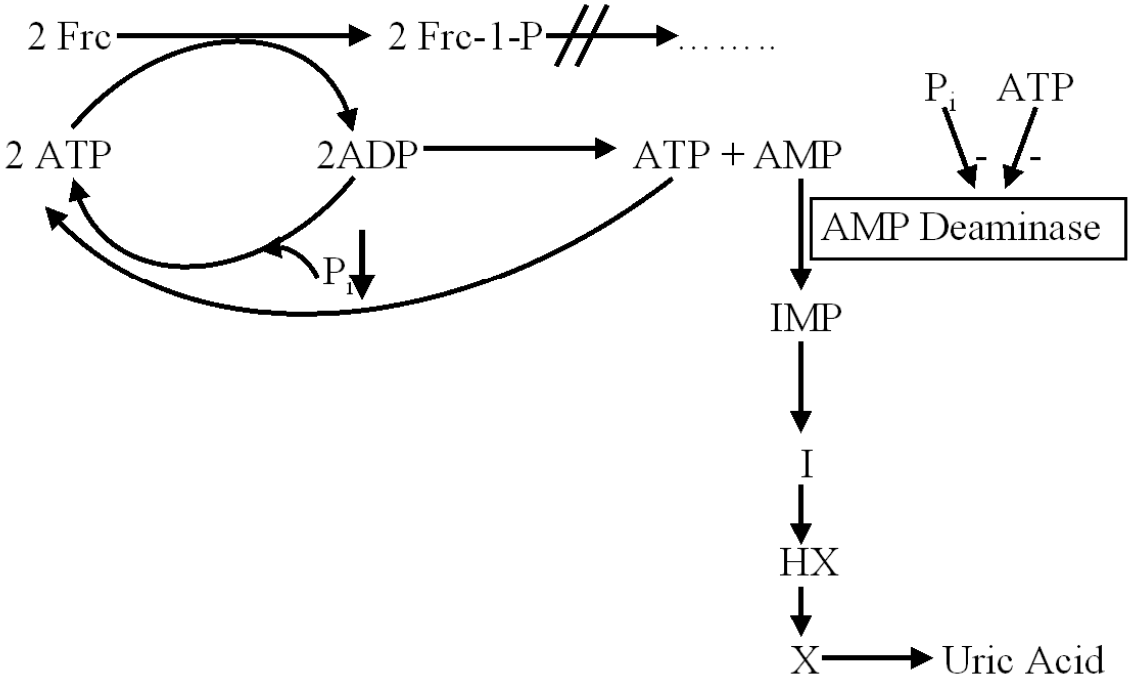
23) What is the pathomechanism of hypoglycemia and hyperuricemia in hereditary fructose intolerance?

-Hypoglycemia:

i) Inhibition of glycogenolysis at the phosphorylase step, due to lack of inorganic Phosphate, as it is trapped in the form of Fructose-1-Phosphate.

ii) Inhibition of gluconeogenesis by the absence of Aldolase B and inhibition of Frc-1,6-Bpase by the accumulation of Frc-1-P and Frc-1,6-BP_2.

-Hyperuricemia:



-As it is obvious from the above scheme, fructose will be continuously phosphorylated to Frc-1-P, by the Fructokinase, which is an unregulated enzyme. Therefore, Frc-1-P will built up, but the actual problem is the depletion of inorganic phosphate. Since Fructokinase is an unregulated step, it will indefinitely phosphorylate incoming fructose molecules, depleting ATP as well. There are two ways for a cell to replenish ATP stores: one by the oxidative phosphorylation, and the other by the adenylate kinase (myokinase) reaction. As the inorganic phosphate stores go low, ATP coming from oxidative phosphorylation slows down, and the adenylate kinase reaction gains metabolic importance. But the other side of the coin, is that adenylate kinase generates AMP which can only revert to ADP by reversal of this reaction (highly unfavorable at the time being as there is a need for ATP), or destined for catabolism by the AMP deaminase pathway. AMP deaminase normally, is inhibited by 95 %, by the allosteric actions of ATP and P_i. However, now that both of these molecules are in low concentrations,

the enzyme is significantly deinhibited, and therefore, AMP leaks down to the catabolic pathway, forming uric acid.

24) What is the effect of physical exercise on the lactate concentration of venous blood coming from working muscles in normal controls? And in patients suffering from McArdle`s disease?

-In normal controls it is high, because it is proven that glycogen breakdown in a vigorously contracting muscle results to lactate production.
-In McArdle patients it is very low or undetectable.
-“Second wind effect”: If you ask from a Mc Ardle patient to start exercising, he or she will experience pain. After some minutes, however, the patient will feel pain no more. This is the second wind effect. It is explained according to the following: By the time the patient starts to exercise, he will use his carbohydrate stores, namely the glycogen. But due to the absence of the muscle phosphorylase, the patient will get no energy out of it, and therefore, his muscle cells will start to rupture, due to the inefficient ATP amount for both muscle contraction and cell homeostasis. When the patient cease to feel pain, then alternative ATP sources come into play, that is the lipid stores, that they require no phosphorylase. So now the patient can have enough ATP for both muscle contraction and cell homeostasis. This mobilization of fat stores, is induced by the stress situation and the activation of the beta- mediated sympathetic system.

LIPIDS

1) Digestion and absorption of lipids

-Harper, Fig. 55-2.
-An adult man ingests about 60-150 gr. of lipid per day. Triacylglycerols (TG) constitute more than 90 % of the dietary fat. The rest is made up of phospholipids (PL), cholesterol, cholesterol esters (CE), and free fatty acids (FFA). In addition, 1-2 gr. of cholesterol and 7-22 gr. of phosphatidylcholine (lecithin) are secreted into the small intestine lumen as constituents of bile.
-The poor water solubility presents special problems for digestion because the substrates are not easily accessible to the digestive enzymes in the aqueous phase. In addition, even if ingested lipids are hydrolyzed into simple constituents, the products tend to aggregate to larger complexes that make poor contact with the cell surface and therefore are not easily absorbed. These problems are overcome by:
i) Increases in the interfacial area between the aqueous and lipid phase and
ii) "solubilization" of lipids with detergents.
-At least 5 phases can be distinguished in the process of lipid absorption:
i) hydrolysis of TG to FFA and monoacylglycerols (MG)
ii) solubilization by detergents (bile acids) and transport from the intestinal lumen toward the cell surface
iii) uptake of FFA and MG into the cell and resynthesis to TG
iv) packaging of newly synthesized TG into special lipid-rich globules, called chylomicrons
v) exocytosis of chylomicrons from cells and release into lymph.
-Digestion of lipids is initiated in the stomach by an acid-stable lipase, most of which is thought to originate from glands in the posterior part of the tongue. However, the major enzyme for TG hydrolysis is the pancreatic lipase. This enzyme is specific for esters in the α-position of glycerol, and therefore, the products are FFA and β-monoacylglycerols; (β-monoacylglycerols will nonenzymatically convert to α-monoacylglycerols, but this procedure takes a few more milliseconds, therefore, some β-monoacylglycerols remain).
-The purified form of this pancreatic lipase is strongly inhibited by the bile acids that normally are present in the small intestine during lipid digestion. The problem of inhibition is overcome by colipase, a small protein that binds to both the water-lipid interface and to lipase, thereby anchoring and activating the enzyme. It is secreted by the pancreas as procolipase and depends on tryptic removal of a NH_2- terminal decapeptide for full activity.
-PL are hydrolyzed by specific phospholipases. Pancreatic secretions are especially rich in the proenzyme for phospholipase A_2. As other pancreatic proenzymes, this one is also activated by trypsin. Phospholipase A_2 requires bile acids for activity.
-Bile acids are biological detergents that are synthesized by the liver and secreted as conjugates of glycine or taurine with the bile into the duodenum. At physiological pH values, (around pH=3 for the gastric secretions), they are present as anions, which have detergent properties. Therefore, the terms bile acids and bile salts are often used interchangeably. Bile acids at pH values above 3, reversibly form aggregates. These aggregates are called micelles, and the minimal concentration necessary for micelle formation is the critical micelle concentration. Micelles besides their property to transport lipids, they also transport cholesterol and the lipid-soluble vitamins A, D, E, K. Bile acids are absolutely essential for their absorption.

-Lipid globules within the intestinal cells, are excreted into the lacteals, instead of the venules of the intestinal villi, in the form of chylomicrons. This is because of their large diameter. The apolipoproteins of the chylomicrons are distinctly different from those of the liver, and they are designated as A-1 and B.
-While dietary medium-chain FA reach the liver directly with the portal blood, the long chain FA bypass the liver by being released in the form of chylomicrons into the lymphatics. The intestinal lymph vessels drain into the large body veins via the thoracic duct. Blood from the large veins first reaches the lungs and then the capillaries of the peripheral tissues, including adipose tissue and muscle, before it comes into contact with the liver. Fat and muscle cells in particular take up large amounts of dietary lipids for a storage or metabolism. The bypass of the liver may have evolved to protect this organ from a lipid overload after a meal.

2) ***General rules of lipid transport. Function of major lipoprotein classes***

-Harper, Fig. 27-1, Fig. 27-2, Fig. 27-3, Table 27-1, Table 27-2, Table 27-3.
-Since lipids (absorbed from the diet or synthesized in the liver or adipose tissue) are insoluble in water, the problem arises of how to transport them in an aqueous environment, the blood plasma. This is solved by associating nonpolar lipids (TG and CE) with amphipathic lipids (PL, cholesterol) and proteins to make them water miscible lipoproteins.
-Lipoproteins mediate the transport of lipids by transporting them from the intestines as chylomicrons (bypassing the portal vein) and from the liver as VLDL, to most tissues for oxidation and to adipose tissue for storage. Lipid is mobilized from adipose tissue as FFA attached to serum albumin.
-Pure fat is less dense than water; it follows that as the proportion of lipid to protein in lipoproteins increases, the density decreases.
-Chylomicrons are derived from the intestinal absorption of TG.
-VLDL is derived from the liver for the export of TG.
-LDL represents a final stage in the catabolism of VLDL.
-HDL is involved in VLDL and chylomicron metabolism, and also in cholesterol and reverse cholesterol transport.
-TG is the predominant lipid in chylomicrons and VLDL, whereas cholesterol and PL are the predominant ones in LDL and HDL, respectively.
-A typical lipoprotein-such as a chylomicron or a VLDL- consists of a lipid core of mainly nonpolar TG and CE surrounded by a single layer of amphipathic PL and cholesterol molecules. These are oriented so that their polar groups face outward to the aqueous medium, as in the cell membrane. The protein moiety of a lipoprotein is known as an apolipoprotein or apoprotein, constituting nearly 60 % of some HDL, and as little as 1 % of chylomicrons. Some apoproteins are integral and cannot be removed, whereas others are free to transfer to other lipoproteins (i.e. RBC`s).
-The main apoprotein of LDL is apo-B and is found also in VLDL and chylomicrons. However, apo-B of chylomicrons (B-48) is smaller than apo-B 100 of LDL and VLDL. B-48 is synthesized in the intestine and B-100 in the liver.
-Some lipoproteins contain also carbohydrate moieties.

-Apoproteins carry out several functions:
i) They are enzyme cofactors, i.e. C-II for LPL, A-I for LCAT.
ii) They can act as lipid transfer proteins.
iii) They act as ligands for interactions with lipoprotein receptors in tissues i.e. apo-B 100 and apo-E for the LDL receptor, apo-E for the remnant receptor, and apo-A-I for the HDL receptor.
-FFA (long chain) combine with albumin in the plasma, and to fatty acid binding protein (Z protein) in the cell, so they are never free. Long chain FFA are water insoluble. Short and medium chain FA are more water soluble and exist as the un-ionized acid or as a FA anion.
-The FFA turnover is related directly to [FFA]. Thus, the rate of FFA production in adipose tissue controls the [FFA] in plasma, which in turn determines the FFA uptake by other tissues. The nutritional conditions does not appear to have a great effect on the fractional uptake of FFA by tissues. It does, however, alter the proportion of the uptake which is oxidized compared to the fraction which is esterified, more being oxidized in the fasting than in the fed state.
-The liver plays a central role in lipid transport and metabolism, in terms of:
i) it facilitates the digestion and absorption of lipids by the production of bile, which contains cholesterol and bile salts synthesized within the liver de novo or from uptake of lipoprotein cholesterol.
ii) it possesses active enzyme systems for synthesizing and oxidizing FA, and for synthesizing TG and PL.
iii) it can convert FA to KB.
iv) it plays an integral part in the synthesis and metabolism of plasma lipoproteins.

3) Chylomicrons.Composition, formation, catabolism. Role of lipoprotein lipase

-Harper, Fig. 27-3, Fig. 27-4.
-The inability of particulate lipid of the size of chylomicrons to pass through endothelial cells of the capillaries without prior hydrolysis, is probably the reason that dietary fat enters the circulation via the lymphatics (thoracic duct) and not via the portal vein bypassing the liver.
-Although chylomicrons (and VLDL) isolated from blood contain apoproteins C and E, the newly secreted “nascent” lipoproteins contain little or none, and it would appear that the full complement of apo-C and apo-E polypeptides is taken up by transfer from HDL once the chylomicrons (and VLDL) have entered the circulation.
-TG of chylomicrons (and VLDL) are hydrolyzed by LPL. LPL is located on the walls of blood capillaries, anchored by proteoglycan chains of heparan sulfate. Normal blood does not contain appreciable quantities of the enzyme; however, following injection of heparin, LPL is released from its heparan sulfate binding into the circulation and is accompanied by the clearing of lipemia. A lipase is also released by the liver by large quantities of heparin (heparin-releasable-hepatic-lipase, HRHL), but this enzyme has properties different of those from LPL and does not readily react with chylomicrons.
-Both PL and apoprotein C-II are required as cofactors for LPL activity. Thus, chylomicrons (and VLDL) provide the enzyme for their metabolism with both its` substrate and cofactors. The TG is hydrolyzed progressively through a DG to a MG that is finally hydrolyzed to FFA

plus glycerol. Some of the released FFA return to the circulation, attached to albumin, but the bulk is transported into the tissue.
-Heart LPL has a low K_M for TG, whereas the K_M of the isoform in adipose tissue is 10 times greater. As the plasma [TG] decreases in the transition from the fed to the starved condition, the heart enzyme remains saturated with substrate but the saturation of the enzyme in adipose tissue diminishes, thus redirecting uptake from adipose tissue towards the heart.
-In adipose tissue, insulin enhances LPL synthesis in adipocytes and its translocation to the luminal surface of the capillary endothelium.
-The action of LPL forms remnant lipoproteins: chylomicron remnants (90 % loss of TG), and VLDL remnants (IDL). Subsequently, these remnants are taken up by the liver and the CE and TG are hydrolyzed and metabolized. Uptake appears to be mediated by a receptor on hepatic cells specific for apo-E.

4) Composition, formation and catabolism of VLDL

-Harper, Fig 27-3, Fig. 27-5, Fig. 27-7.
-See question ***4***.
-Most of the plasma VLDL are of hepatic origin (some come from the chyle of the intestines).
-VLDL is the vehicle of transport of TG from the liver to the extrahepatic tissues.

5) LDL and HDL. Composition, metabolic fate. Their role in the transport of cholesterol

-Harper, Fig. 27-5, Fig. 27-6.
-Most LDL appears to be formed from VLDL, but there is some evidence for some production directly by the liver. LDL is metabolized via the LDL receptor, which is a B-100, E receptor. It is so designated because it is specific for apo-B 100 but not apo-B 48, and under some circumstances it will take up lipoproteins rich in apo-E. Apo-B 48 lacks the carboxy-terminal domain of B 100 that contains the ligand for the LDL receptor. These receptors are defective for familial hypercholesterolemia. Approximately, 30 % of LDL is degraded in extrahepatic tissues and 70 % in the liver.
-HDL is synthesized and secreted from both liver and intestine. However, nascent HDL from intestine does not contain apo-C, nor apo-E, but only apo-A. Thus, apo-C and apo-E are synthesized in the liver and transferred to intestinal HDL when the latter enters the plasma. A major function of HDL is to act as a repository for apo-C and apo-E that are required in the metabolism of chylomicrons and VLDL.
-Nascent HDL consists of discoid PL bilayers containing apolipoproteins and free cholesterol. With the help of LCAT enzyme found in the plasma, PL and free cholesterol of the HDL particles are converted into CE and lysolecithin. The nonpolar CE move into the hydrophobic interior of the bilayer, whereas lysolecithin is transferred to plasma albumin. The reaction

continues, generating a nonpolar core that pushes the bilayer apart, until a spherical pseudomicellar HDL is formed, covered by a surface film of polar lipids and apoproteins. Thus, the LCAT system is involved in the removal of excess unesterified cholesterol from lipoproteins and from the tissues. The liver and possibly the intestines seem to be the final sites of degradation of HDL.
-An HDL cycle has been proposed to account for the transport of cholesterol from the tissues to the liver. This explains why [HDL_2] in the plasma varies reciprocally with the [chylomicron] and [VLDL] and directly with the activity of LPL.

6) Storage and mobilization of triacylglycerols

-See questions ***2, 3, 4, and 5.***
-Harper, Fig. 27-7, Fig. 27-8, Fig. 27-9.
-TG must be hydrolyzed by a suitable lipase to their constituent FA and glycerol before further catabolism can proceed.
-Factors that enhance both the synthesis of TG and the secretion of VLDL by the liver (these 2 processes are coupled) include:
i) the feeding of diets high in carbohydrate
ii) the fed rather than the fasting state
iii) high levels of circulating FFA
iv) ingestion of ethanol
v) the presence of high concentration of insulin and low concentrations of glucagon, which enhance FA synthesis and esterification and inhibit their oxidation.
-Imbalance in the rate of TG formation and export causes fatty liver. This imbalance can be brought about by 2 different mechanisms:
i) Raised FFA plasma levels
ii) metabolic block in the production of plasma lipoproteins, thus allowing TG to accumulate.
-Increased [FFA] results from mobilization of fat from adipose tissue (i.e. starvation) or from the hydrolysis of lipoprotein or chylomicron TG by LPL in extrahepatic tissues (i.e. high fat diets). Increasing amounts of FFA are taken up by the liver and esterified. The production of plasma lipoprotein does not keep pace with the influx of FFA, allowing TG to accumulate, causing a fatty liver.
-The lesion in the production of plasma lipoproteins may be due to:
a) a block in apolipoprotein synthesis
b) a block in the synthesis of the lipoprotein from lipid
c) a failure in provision of PL that are found in lipoproteins
d) a failure in the secretory mechanism itself.
-Ethanol causes fatty liver. A high influx of ethanol by the enzyme alcohol dehydrogenase will give a high NADH/NAD ratio, that will cause a shift to the left in the equilibrium
malate $\rightleftharpoons$ OAA reducing the activity of the TCA. Moreover, TG synthesis is favored, and FA oxidation is inhibited due to the presence of reducing equivalents in excess.
-Since the level of plasma FFA has most profound effects upon the metabolism of other tissues, particularly liver and muscle, the factors operating in adipose tissue that regulate the outflow of FFA exert an influence far beyond the tissue itself.

-In adipose tissue, since glycerol kinase is low in activity, this tissue is dependent on glycolysis and a supply of glucose for the provision of Glycerol-3-P.
-Increased glucose metabolism reduces the output of FFA. However, this reduced FFA output is not accompanied by a decrease in the release of glycerol, implicating that the effect of glucose is not mediated by reducing the rate of lipolysis. It is believed that this effect is due to the provision of Glycerol-3-P, which enhances the esterification of FA via Acyl-CoA.
-Adipose tissue is much more sensitive to insulin than are many other tissues, which points to adipose tissue as a major site of insulin action in vivo.
-For an optimal effect, most of the hormonally mediated lipolytic processes require the presence of glucocorticoids and thyroid hormones. On their own, these particular hormones do not increase lipolysis markedly, but act in a facilitatory or permissive capacity with respect to other lipolytic endocrine factors.

7) Transport of fatty acids through the inner mitochondrial membrane. Regulation of the transport

-Harper, Fig. 24-1, Fig. 24-10.
-Long chain fatty acids (only) penetrate the inner mitochondrial membrane as carnitine derivatives. Carnitine is widely distributed and is particularly abundant in muscle.
It is synthesized from lysine and methionine in liver and kidney.
-Together wit fructose and lactate, acetylcarnitine is an important fuel for sperm, supporting motility.

8) Beta-oxidation of saturated fatty acids. The energetics of oxidation. Regulation of beta-oxidation

-Harper, Fig. 24-3, Fig. 24-10.
-FA are both oxidized to acetyl-CoA and synthesized from acetyl-CoA; however, the two processes are entirely different (they also take place in different compartments: b-oxidation within mitochondria, while lipogenesis in the cytosol). The separation of FA oxidation from their biosynthesis allows each process to be individually controlled and integrated with tissue requirements.
-FA oxidation uses NAD^+ and FAD, while lipogenesis $NADP^+$. FA oxidation is a strictly aerobic process.
-Increased FA oxidation is characteristic of starvation and of Diabetes Mellitus (DM), leading to ketogenesis. Since we know that impairment of FA oxidation leads to hypoglycemia, a carnitine deficiency, or a carnitine palmitoyltransferase deficiency, would lead to hypoglycemia.
-FA are activated before being catabolized. There is only one step in the complete degradation of a FA that requires ATP: $\boxed{FA + ATP + CoA \longrightarrow Acyl\text{-}CoA + PP_i + AMP}$

-2 high-energy phosphates are expended during the activation of each FA molecule.
-In beta oxidation 2 carbons are cleaved at a time from Acyl-CoA molecules starting at the carboxyl end. The chain is broken between the $\alpha(_2)$ and $\beta(_3)$ carbon atoms, hence the name beta-oxidation.
-Energetics of beta-oxidation: 1 NADH plus 1 $FADH_2$, give 5 ATP, which is multiplied by 7 (7 Acetyl-CoA if we consider palmitate) give 7*5=35 ATP. A total of 8 mol of Acetyl-CoA are formed and each will give rise to a further 12 mol of ATP upon oxidation in the TCA, giving 8*12=96 ATP, minus 2 ATP for the initial activation of the FA, 96+35-2=129 ATP (68 % efficiency).
-Peroxisomes oxidize very long chain FA, leading to acetyl-CoA (through octanoyl-CoA) and H_2O_2. This system is not linked directly to phosphorylation and the generation of ATP, but it aids the oxidation of very long FA (20, 22 carbon atoms).
-Besides the beta-oxidation there is the alpha oxidation in the brain; the latter does not require CoA intermediates, nor does it generate high energy phosphates. There is also an ϖ-oxidation found in the liver, in the P450 system of the ER.

9) Oxidation of unsaturated fatty acids and fatty acids with odd number of carbon atoms

-Harper, Fig. 24-4, Fig. 21-2
-Oxidation of an unsaturated FA occurs by a modified beta oxidation pathway (Diels-Alder synthesis).
-Oxidation of a FA with an odd number of carbon atoms yields acetyl-CoA plus a molecule of proprionyl-CoA (3 carbon atoms). This compound is converted to succinyl-CoA, a constituent of the TCA. Hence, the propionyl residue from an odd-chain FA is the only part of a FA that is gluco(neo)genic.

10)Biosynthesis of saturated fatty acids

-Harper, Fig. 23-1, Fig 23-2, Fig, 23-3, Fig. 23-4, Fig. 23-6.
-Lipogenesis is present in the cytosol, and fatty acid elongation is present in the ER.
-Lipogenesis is present in the liver, kidney, brain, mammary gland, and adipose tissue.
-Cofactor requirements include NADPH, ATP, Mn^{2+}, biotin, and HCO_3^- (as a source of CO_2). Acetyl-CoA is the free substrate is the immediate substrate, and free palmitate is the end product.
-After the action of ATP citrate lyase, OAA formed in the cytosol must find its way back to the mitochondrion. This is accomplished in three different ways:

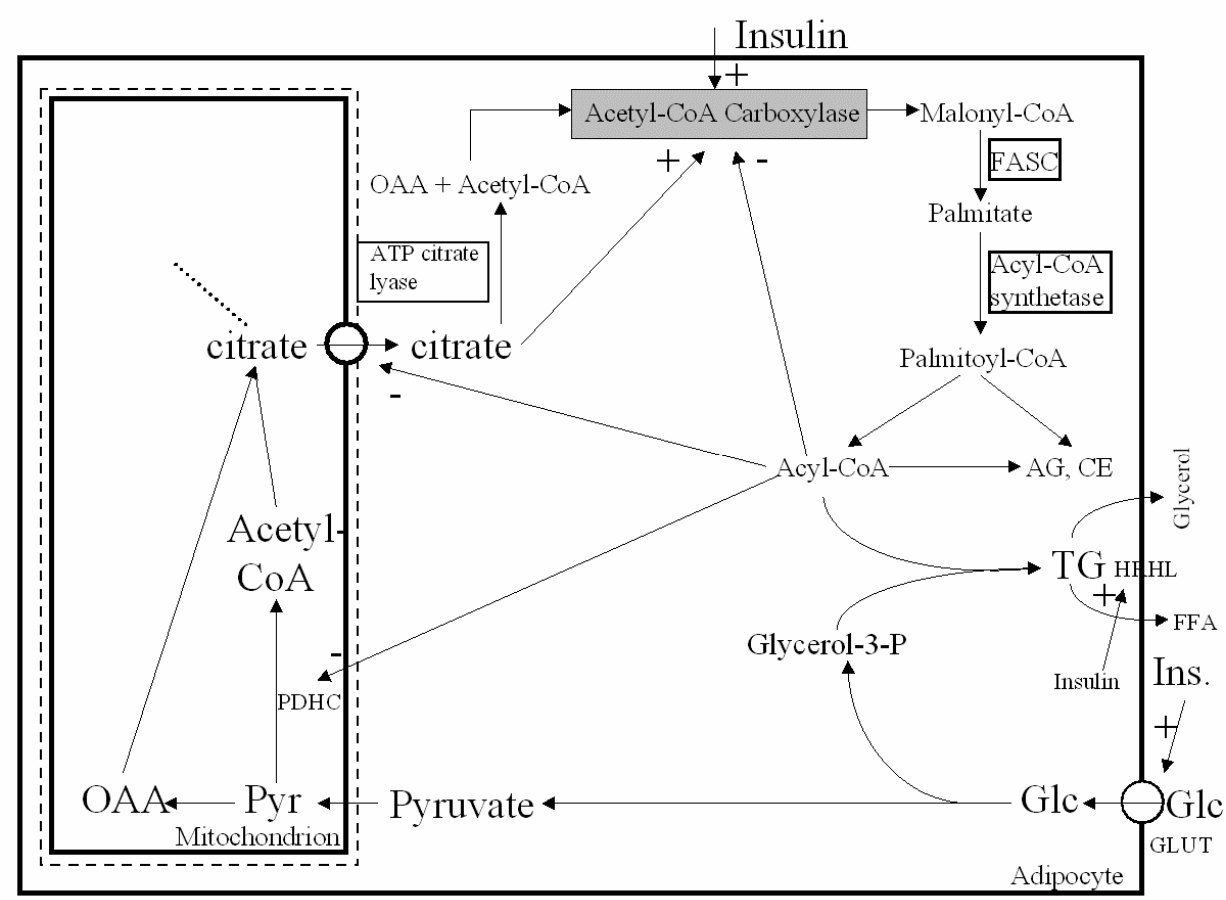
a) OAA becomes malate by cytosolic malate DHase and enters the mitochondrion through the malate/aspartate shuttle.

b) OAA becomes malate by cytosolic malate DHase and subsequently pyruvate by the malic DHase reaction and enters the mitochondrion by the pyruvate/proton cotransporter (or pyruvate/hydroxyl antiporter.
c) OAA becomes aspartate by the cytosolic aspartate transaminase and enters the mitochondrion by the malate/aspartate shuttle.

OAA FORMED BY THE ATP CITRATE LYASE WILL NEVER BECOME PEP BY THE PEPCK, AS THIS IS ENZYME IS ONLY FUNCTIONAL DURING GLUCONEOGENESIS. LIPOGENESIS AND GLUCONEOGENESIS ARE MUTUALLY EXCLUDED!

-Fatty acid synthase complex (FASC): The aggregation of all the enzymes of a particular pathway into one multienzyme functional unit offers great efficiency and freedom from interference by competing reactions, thus achieving the effect of compartmentalization of the process within the cell without the erection of permeability barriers. Another advantage of a single multienzyme polypeptide is, that synthesis of all enzymes in the complex is coordinated, since it is encoded by a single gene.
-The FASC is a dimer, in which each monomer is identical, lying in a "head" to "tail" configuration; only the dimer is active.
-Decarboxylation allows the reaction to go to completion, acting as a pulling force for the whole sequence of reactions.
-The free palmitate must be activated to Acyl-CoA before it can proceed via any other metabolic pathway. Its usual fate is esterification into acylglycerols.
-There are 2 centers of activity in one dimer complex that function independently and simultaneously to form 2 molecules of palmitate.
-Butyryl-CoA may act as a primer molecule in mammalian liver and mammary gland. If propionyl-CoA acts as a primer, long chain fatty acids having an odd number carbon atoms result.
-Translocation of citrate from the mitochondria to the cytosol is via the tricarboxylate transporter (requires malate for the antiport).
-After ATP citrate lyase, the resulting OAA can form aspartate, malate (via NADH-linked malate Dehydrogenase), followed by the generation of NADPH via the malic enzyme. This pathway is a means of transferring reducing equivalents from extramitochondrial NADH to $NADP^+$.
-Fasting largely abolishes chain elongation. Elongation of stearyl-CoA in brain increases rapidly during myelination in order to provide C_{22} and C_{24} FA that are present in sphingolipids.

11) Regulation of fatty acid synthesis



-The nutritional state of an organism is the main factor controlling the rate of lipogenesis. Thus, the rate is higher in the well-fed state, where diet contains a high proportion of carbohydrate (OAA availability for Acetyl-CoA provision). It is depressed under conditions of restricted caloric intake, on a high-fat diet, or when there is insulin deficiency (as in DM type I). All these conditions are associated with an increased plasma [FFA].

-There is an inverse relationship between hepatic lipogenesis and the serum [FFA]. Fat in the diet also causes depression of lipogenesis in the liver, and when there is more than 10 % of fat in the diet, there is little conversion of dietary carbohydrate to fat.

-Lipogenesis is higher when sucrose is fed instead of glucose, because fructose bypasses the PFK-1 control point in glycolysis, and floods the lipogenic pathway.

-<u>Citrate and Acyl-CoA regulate the Acetyl-CoA carboxylase</u>: The rate-limiting reaction in the lipogenic pathway is at the Acetyl-CoA carboxylase step. Acetyl-CoA carboxylase is activated by citrate, which increases in concentration in the well-fed state and is an indicator of a plentiful supply of Acetyl-CoA. However, it is inhibited by long chain Acyl-CoA molecules, an example of metabolic negative feedback inhibition by a product of a reaction sequence. Thus, if Acyl-CoA accumulates because it is not esterified quickly enough, it will automatically reduce the synthesis of new FA. Likewise, if Acyl-CoA accumulates as a result of increased lipolysis or an influx of FFA into the tissue, this will also inhibit synthesis of new FA. Acyl-CoA may also inhibit the mitochondrial tricarboxylate transporter, thus preventing egress of citrate from the mitochondria into the cytosol.

-PDH is also regulated by Acyl-CoA: Acyl-CoA causes an inhibition of PDH by inhibiting the ATP/ADP translocase, which leads to increased intramitochondrial [ATP]/[ADP] ratio with the result of activation of PDH. Also, oxidation of Acyl-CoA due to increased levels of FFA may increase the ratios of [Acyl-CoA]/[CoA] and [NADH]/[NAD^+] in mitochondria, thereby inhibiting the PDH.
-Hormones also regulate lipogenesis by several mechanisms. Insulin increases the transport of glucose into the cell (the statement stands for insulin-sensitive glucose transporters only), and thereby increases the availability of both pyruvate for FA synthesis and glycerol-3-Phosphate for esterification of the newly formed FA. Insulin converts the inactive form of PDH to the active form in adipose tissue, but not in the liver. Insulin also activates Acetyl-CoA carboxylase. (In its active form [dephosphorylated], Acetyl-CoA carboxylase polymerizes into long filaments; phosphorylation is accompanied by dissociation into monomeric subunits and loss of activity). Moreover, insulin by its ability to depress the level of cAMP (by inhibiting the cAMP phosphodiesterase), inhibits lipolysis in adipose tissue and therefore, reduces the plasma [FFA] and subsequently of the long chain FA, an inhibitor of lipogenesis. By the same mechanisms, insulin antagonizes the actions of glucagon and epinephrine, which inhibit Acetyl-CoA carboxylase, and therefore, lipogenesis, by increasing the cAMP.

12) Role of the adipose tissue in the carbohydrate and fatty acid metabolism

-See question ***6***.

13) Synthesis of mono-and polyunsaturated fatty acids

-Harper, Fig. 25-1, Fig. 25-2, Fig. 25-3.
-All double bonds present in naturally occurring unesterified FA (UFA) of mammals are of the *cis* configuration.
-Additional double bonds are introduced into existing monounsaturated FA between the existing double bond and the carboxyl group.
-The desaturation and chain elongation system is greatly diminished in the fasting state, upon glucagon and epinephrine action, and in the absence of insulin as in DM type I.
-*Trans*-unsaturated FA are found in ruminant fat, where they are arise from the action of microorganisms in the rumen, but the presence of large amounts of *trans*-unsaturated FA in partially hydrogenated vegetable oils (i.e. margarine), raises the question of their safety as food additives. They are metabolized more like saturated than like *cis*-unsaturated FA. This may be due to their similar straight conformation. In this respect, they tend to raise LDL levels and lower HDL levels. *Trans*-polyunsaturated FA do not possess Esterified FA (EFA) activity and may antagonize the metabolism of EFA and exacerbate EFA deficiency.

-A high ratio of polyunsaturated FA to saturated FA (P:S ratio) in the diet is a major factor in lowering plasma cholesterol level by dietary means, and is also considered to be beneficial in preventing coronary heart disease.

14) The essential fatty acids. Conversion of linoleate to arachidonate.

-Harper, Fig. 25-1, Fig. 25-4.
-The essential FA give rise to eicosanoids, which make up the PG, TX, LT, and LX.
-The content of UFA in fat determines its melting point and therefore, its fluidity.
-EFA: linoleic, linolenic, and arachidonic.
-EFA, except their role in PG, TX, LP, and LX formation, they are also used as structural lipids of the cell, and are concerned with the structural integrity of the mitochondrial membrane. Arachidonic acid is present in membranes and accounts for 5-15 % of the FA in phospholipids. Docosahexanoic acid (DHA), which is synthesized from α-linolenic acid or obtained directly from fish oils, is present in high concentrations in the retina, cerebral cortex, testis, and sperm. DHA is particularly for brain development, and it is supplied via the placenta and the maternal milk.
-ϖ-3 FA lower blood cholesterol level by decreasing the LDL/HDL ratio. They are found in fish coming from very cold water (e.g. cod).

15) Synthesis and metabolic fate of ketone bodies

-Harper, Fig. 24-5, Fig. 24-6, Fig. 24-7, Fig. 24-8, Fig. 24-9.
-Ketogenesis occurs when there is a high rate of FA oxidation in the liver. Acetoacetate and β-hydroxybutyrate are formed (acetone is exhaled in the lungs after spontaneous decarboxylation from acetoacetate), and they are interconverted depending on the intramitochondrial NAD^+/NADH ratio, i. e. the redox state. The ratio β-hydroxybutyrate/acetoacetate in the blood varies between 1:1 and 10:1.
-The net flow of KB from the liver to the extrahepatic tissues results from an active enzymatic mechanism in the liver for the production of KB coupled with very low activity of enzymes responsible for their utilization. The opposite stands true for the extrahepatic tissues.
-KB serve as a fuel for extrahepatic tissues. Acetoacetate once formed cannot be reactivated directly in the liver except in the cytosol, where it is a precursor in cholesterol synthesis, a much less active pathway.
-Moreover, the liver contains a limited amount of CoA*SH, and when most of it is tied up in Acetyl-CoA, β-oxidation of FA slows down for lack of the free coenzyme. The production and export of KB frees CoA*SH, allowing continued FA oxidation.
-Ketonemia ensues upon increased production of KB, or by their underutilization.
-Ketogenesis is regulated at 3 crucial steps:

i) Control is exercised initially in adipose tissue. Ketosis occurs only if the level of [FFA] in the blood arising from lipolysis of TG in adipose tissue increases. FA are the

precursors of KB. The liver, both in fed and in fasting conditions, retains the ability to extract about 30 % or more of the FA passing through it, so that at high [FFA] the flux passing through the liver is substantial. Therefore, the factors regulating the mobilization of FFA from adipose tissue are important in controlling ketogenesis.

ii) One of the 2 fates awaits the FFA upon uptake by the liver and after they are activated to Acyl-CoA: they are esterified mainly to TG and PL, or they are oxidized to CO_2 or KB. The capacity for esterification as an antiketogenic factor depends on the availability of precursors in the liver to supply sufficient glycerol-3-Phosphate (however, there will never be accumulation of FA nor any intermediate in their pathway of esterification to TG in the liver, as long as glycerol-3-Phosphate is decreased). The factor which would "decide" for which pathway the Acyl-CoA from the FFA will follow is the regulation of the transport of Acyl-CoA by carnitine palmitoyltransferase I. Its activity is low in the fed state, when FA oxidation is depressed, and high during fasting, when FA oxidation increases. Malonyl-CoA, the initial intermediate in FA synthesis, which increases in concentration in the fed state, inhibits this enzyme, withdrawing FFA (Acyl-CoA) from oxidation, and diverting Acyl-CoA to lipogenesis (TG, PL). FFA entering the liver cell in low concentrations are nearly all esterified to acylglycerols and transported out of the liver in the form of VLDL. However, as the [FFA] increases with the onset of starvation, Acetyl-CoA carboxylase is inhibited directly by Acyl-CoA, and [Malonyl-CoA] decreases, releasing the inhibition upon carnitine palmitoyltransferase I, and allowing more Acyl-CoA to be oxidized. These events are reinforced in starvation, by the [insulin]/[glucagon] ratio, which decreases, causing increased lipolysis in adipose tissue, the release of FFA, and inhibition of acetyl-CoA carboxylase in the liver.

iii) In turn, the Acetyl-CoA formed in beta-oxidation is oxidized in the TCA, or it enters the pathway of ketogenesis. As the levels of serum FFA raise, proportionally more FFA are converted to KB and less are oxidized via the TCA. The partition of Acetyl-CoA between the ketogenic pathway and the oxidation by the TCA is regulated in such a way that, the total free energy trapped in ATP, which results from the oxidation of FFA, remains constant. Complete oxidation of 1 mol of palmitate gives a net of 129 mol of ATP via beta-oxidation and CO_2 production via the TCA, whereas only 33 moles of ATP are produced when acetoacetate is the end product, while only 21 moles when β–hydroxybutyrate is the end product. Thus, ketogenesis may be regarded as a mechanism that allows the liver to oxidize increasing quantities of FA within a tightly coupled system of oxidative phosphorylation, without increasing its` total energy expenditure.

16) The physiological role of ketone bodies

-See question ***15***.

17) Synthesis of cholesterol. Synthesis and importance of cholesterol esters

-Harper, Fig. 28-1, Fig. 28-2, Fig. 28-3, Fig. 28-6.
-Cholesterol is present in tissues and in plasma lipoproteins either as free cholesterol, or combined with a long chain FA, as cholesteryl ester. It is synthesized in many tissues from Acetyl-CoA and it is ultimately eliminated from the body in the bile as cholesterol or bile salts.
-Cholesterol is the precursor of all steroids in the body, such as corticosteroids, sex hormones, bile acids, and vitamin D. It is typically a product of animal metabolism and therefore, occurs in foods of animal origin such as egg yolk, meat, liver, and brain.
-Cholesterol is an amphipathic lipid and as such is an essential structural component of membranes and of the outer layer of plasma lipoproteins. Lipoproteins transport free cholesterol in the circulation, where it readily equilibrates with cholesterol in other lipoproteins and in membranes. Cholesteryl ester is a storage form of cholesterol found in most tissues. It is transported as cargo in the hydrophobic core of lipoproteins. LDL is the mediator of cholesterol and cholesteryl ester uptake into many tissues. Free cholesterol is removed from tissues by HDL and transported to the liver for conversion to bile acids.
-Cholesterol is a major constituent of gallstones. However, its` chief role in pathologic processes is a factor in the genesis of atherosclerosis of vital arteries. Coronary atherosclerosis correlates with a high plasma LDL: HDL ratio.
-Approximately half the cholesterol of the body arises by synthesis (500 mg/day), and the remainder is provided by the average diet. The liver accounts for ~10 % of the total synthesis, the gut ~15 %, and the skin for a significant proportion of the remainder.
-Cholesterol synthesis takes place in the ER and the cytosol.
-Acetyl-CoA is the source of all carbon atoms in cholesterol.
-HMG-CoA is an intermediate of the cholesterol synthesis pathway; however, KB formation takes place within the mitochondria, while cholesterol synthesis in the cytosol. Initially, 2 molecules of Acetyl-CoA condense to form Acetoacetyl-CoA catalyzed by a cytosolic thiolase enzyme. Alternatively, in liver, acetoacetate made inside the mitochondria in the pathway of ketogenesis, diffuses into the cytosol and may be activated to Acetoacetyl-CoA by Acetoacetyl-CoA synthase, requiring ATP and CoA.
-Farnesyl pyrophosphate gives rise also to dolichol and ubiquinone.

18) Regulation of cholesterol synthesis

-Harper, Fig. 28-4, Fig. 28-5.
-At the tissue level, the following processes are considered to govern the cholesterol balance of cells:
A) Increase is due to:
i) Uptake of cholesterol-containing lipoproteins by receptors, i.e. the LDL receptor
ii) Uptake of cholesterol-containing lipoproteins by a non-receptor mediated pathway.
iii) Uptake of free cholesterol from cholesterol-rich lipoproteins to the cell membrane.
iv) Cholesterol synthesis
v) Hydrolysis of cholesteryl esters by the enzyme cholesteryl ester hydrolase.

B) Decrease is due to:
i) Efflux of cholesterol from the membranes to lipoproteins of low cholesterol potential, particularly to HDL_3, or nascent HDL, promoted by LCAT.
ii) Esterification of cholesterol by ACAT.
iii) Utilization of cholesterol for synthesis of other steroids, such as hormones or bile acids in liver.

19) LDL receptors and familial hypercholesterolemia /type II/

-Harper, Table 28-1.
-The LDL (apoB-100, E) receptors occur on the cell surface in pits that are coated on the cytosolic side of the cell membrane with a protein called clathrin. It reacts with the ligand on LDL apo B-100, and the LDL is taken up intact by endocytosis. It is broken down in the lysosomes, which involves hydrolysis of the apoprotein and cholesteryl ester followed by translocation of cholesterol into the cell. The receptors are not destroyed but return to the cell surface.
-The influx of cholesterol down-regulates the number of LDL receptors.

20) Metabolism of the bile acids. Synthesis, regulation of synthesis, enterohepatic circulation. Clinical aspects

-Harper, Fig.28-7.
-Approximately half of the cholesterol eliminated from the body is found in the feces after conversion of bile acids. The remainder is excreted as neutral steroids in the urine.
-Although fat digestion products, including cholesterol are absorbed in the first 100 cm of the small intestine, the primary and secondary bile acids are absorbed almost exclusively in the ileum, returning to the liver by way of the portal circulation about 98-99 % of the bile acids

secreted into the intestine. This is known as the **enterohepatic circulation**. However, lithocolic acid, because of its insolubility, is not reabsorbed to any significant extent. A small fraction of the bile acids escapes absorption and is therefore eliminated in the feces. Even though this is a very small amount, it nonetheless represents a major pathway for the elimination of cholesterol. Each day, an amount of bile acid equivalent to that lost in the feces is synthesized from cholesterol by the liver, so that a pool of bile acids of constant size is maintained. This is accomplished by a system of feedback inhibition.
-Since bile contains significant quantities of Na^+ and K^+ and the pH is alkaline, it is assumed that the bile acids and their conjugates are actually in a salt form, hence the term "bile salts".
-Primary bile acids are:
i) cholic acid
ii) chenodeoxycholic acid
iii) glycocholic acid
iv) taurocholic acid
-Secondary bile acids are:
i) deoxycholic acid (from cholic acid)
ii) lithocolic acid (from glycocholic and taurocholic acid)
-Bile acid synthesis is regulated at the 7α-hydroxylase step (by a negative feedback).

21) Biosynthesis of adrenal steroid hormones. Mineralocorticoid synthesis

-Harper, Fig. 48-1, Fig. 48-2, Fig. 48-3.
-There is an overlap of biologic activity since all natural glucocorticoids have mineralocorticoid activity and vice versa.
-Mineralocorticoids promote retention of Na^+ and excretion of K^+ and H^+ by the kidney.
-C_{21} hydroxylation is necessary for both mineralo- and glucocorticoid activity, but most steroids with a C_{17} hydroxyl group have more glucocorticoid and less mineralocorticoid action.

22) Biosynthesis of adrenal steroid hormones. Synthesis of glucocorticoid hormones

-Harper, Fig. 48-1, Fig. 48-2, Fig. 48-3.
-Glucocorticoids promote gluconeogenesis.
-Steroidogenesis involves the repeated shuttling of substrates into and out of the mitochondria of the fasciculata and reticularis cells.
-The conjugated steroids are water-soluble and do not bind to transport proteins; thus, they are readily excreted in the bile, feces, and urine.

23) Biosynthesis of androgen and estrogen hormones

-Harper, Fig. 50-1, Fig. 50-5, Fig. 50-6.

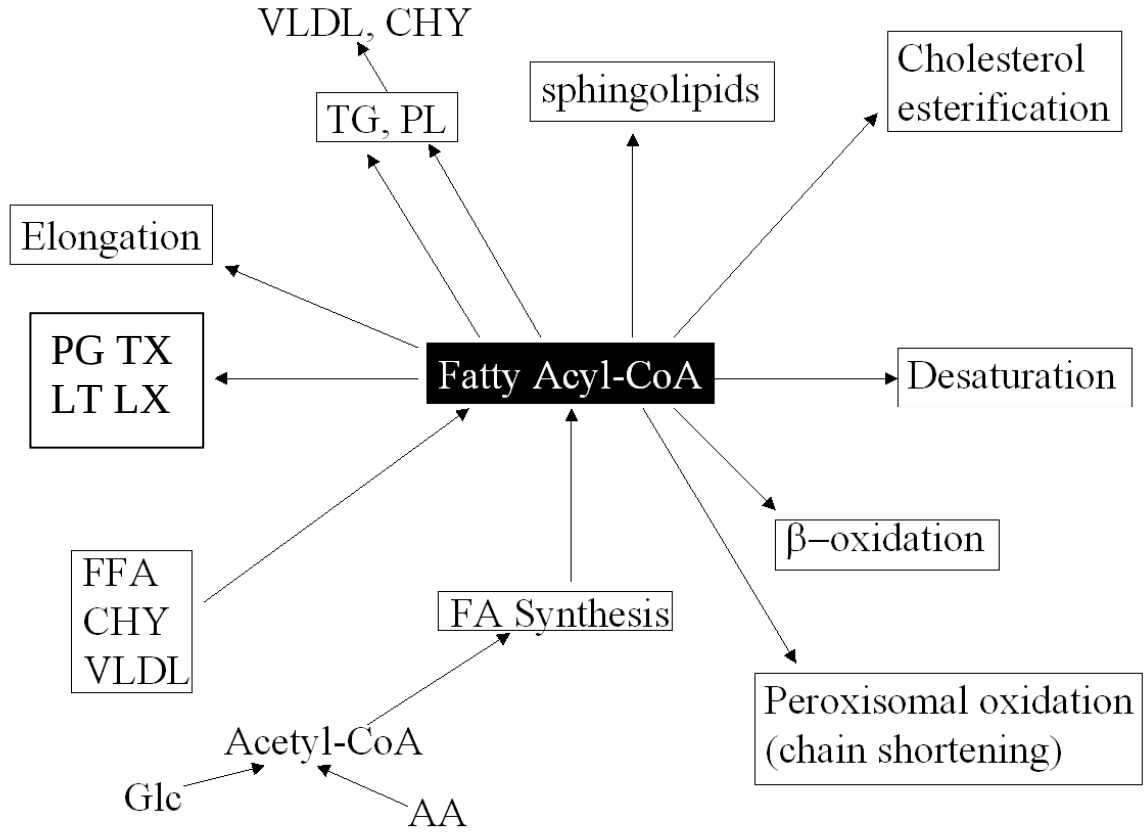
24) Biosynthesis, degradation, and turnover of phosphoglycerols

-Harper, Fig. 26-1, Fig. 26-2, Fig. 26-3, Fig. 26-5, Fig. 26-6.
-Phosphoglycerols, phosphosphingolipids, and glycosphingolipids are all amphipathic lipids and consequently ideally suited as the main lipid constituents of the plasma membrane.
-Some PL have specialized functions: lung surfactant, hormone second messengers, PAF, SAM, ABO blood groups.
-Although PL are actively degraded, each portion of the molecule turns over at a different rate; i.e. the turnover time of the phosphate group is different from that of the acyl group. This is due to the presence of enzymes that allow partial degradation followed by resynthesis.
-Lysolecithin may be formed by an alternative route that involves LCAT. This enzyme, found in the plasma and synthesized in the liver, catalyzes the transfer of a FA residue from the 2^{nd} position of lecithin to cholesterol to form cholesteryl ester and is considered to be responsible for much of the cholesteryl ester in plasma lipoproteins.
-Long chain FA are found predominantly in the 1^{st} position of PL, whereas the polyunsaturated FA (i. E. the precursors of PL) are incorporated more into the 2^{nd} position. The incorporation of FA into lecithin occurs by complete synthesis of the PL, by transacylation between cholesteryl ester and lysolecithin, and by direct acylation of lysolecithin by Acyl-CoA. Thus, a continuous exchange of the FA is possible, particularly in regard with introducing EFA into PL molecules.

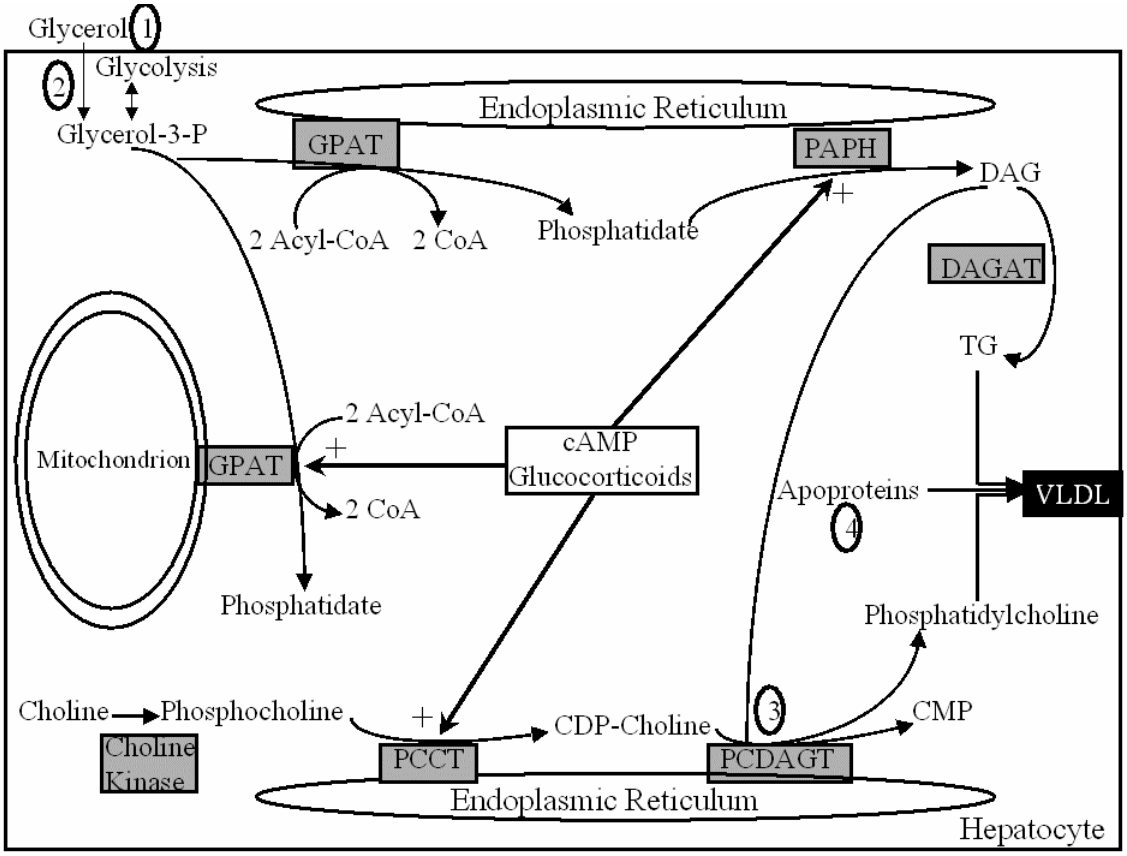
25) Metabolism of sphingolipids and glycolipids

-Harper, Fig. 26-7, Fig. 26-8, Fig. 26-9, Fig. 26-10.
-Galactocylceramide (GalCer) is the major lipid of myelin.

26) Role of key junctions in regulation of metabolism
Fatty Acyl-CoA



27) Coordinated regulation of triglyceride and phospholipid metabolism in the liver



-1: This glycerol comes from the adipose tissue as a product of the lipoprotein lipase.
-2: This reaction is brought about by glycerol kinase, which is only found in liver and kidney.
-3: PCDAGT has a very low K_M to assume phosphatidylcholine synthesis and not TG formation first, that would lead to unnecessary TG accumulation.
-4: Only apo B-100 is attached to the VLDL, as the rest apoproteins (apo C and apo E) will come from the HDL in the blood circulation.
Glucocorticoids activate only PAPH, but not GPAT or PCCT. cAMP is the mediator molecule for the actions of insulin (decrease of cAMP by activation of cAMP phosphodiesterase), and glucagon (increase of cAMP by activation of adenylate cyclase).

-Abbreviations:
GPAT: Glycerol Phosphate Acyl Transferase
PAPH: Phosphatidate Phosphohydrolase
DAGAT: Diacylglycerol Acyl Transferase
PCCT: Phosphocholine Cytidyl Transferase
PCDAGT: Phosphocholine Diacylglycerol Transferase
CMP: Cytidyl Monophosphate

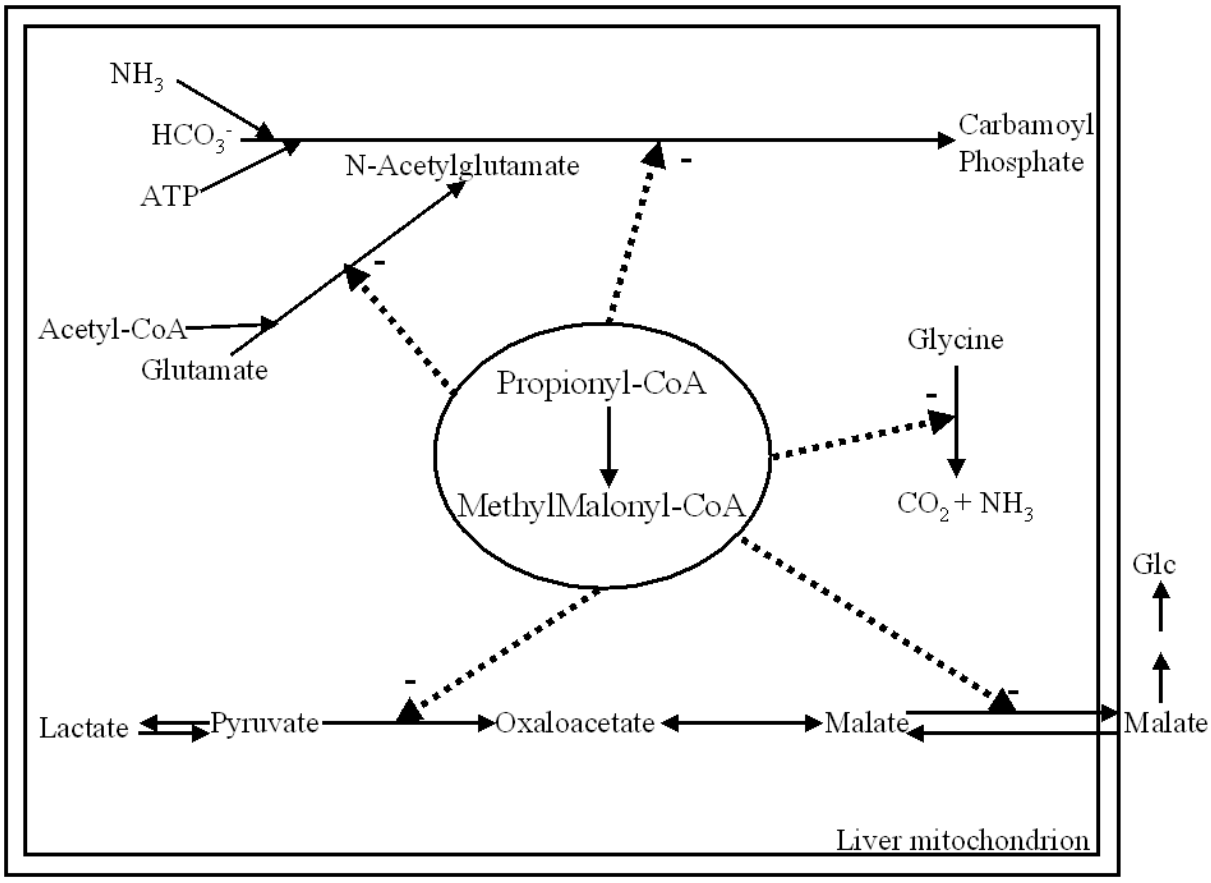
28) The role of carnitine. What forms and consequences of carnitine deficiency can be detected in humans?

-Role of carnitine (see also Harper, Fig. 24-1):

i) Transport of LCFA into the mitochondria (Medium- and short-chain fatty acids penetrate the inner mitochondrial membrane relatively easy).

ii) Transport of acetyl groups as acetylcarnitine between the matrix and the cytosol.

iii) Detoxification (see figure below).



-As it is shown above, accumulation of Propionyl-CoA and subsequently MethylMalonyl-CoA, (for reasons that will become apparent at the question ***31*** from the amino acids and nucleotides section) leads to:

a) Inhibition of Carbamoyl Phosphate synthesis leading to **hyperammonemia** (either by direct inhibition of Carbamoyl Phosphate Synthetase I [CPS I] or by inhibition of the formation of the activator of CPS I which is the N-Acetylglutamate).

b) Inhibition of the Glycine cleavage system leading to **hyperglycinemia**.

c) Inhibition of Pyruvate Carboxylase leading to accumulation of pyruvate and subsequently to accumulation of lactate leading to **hyperlactatemia**.

d) Inhibition of the malate-aspartate shuttle, leading to decreased gluconeogenesis, leading to **hypoglycemia**.

-Carnitine possesses detoxification role, as it is able to react with both Propionyl-CoA and MethylMalonyl-CoA and form **propionylcarnitine** and **methylmalonylcarnitine**, which can **both penetrate the inner mitochondrial membrane and excreted in the urine, relieving the inhibitions at the aforementioned steps**.

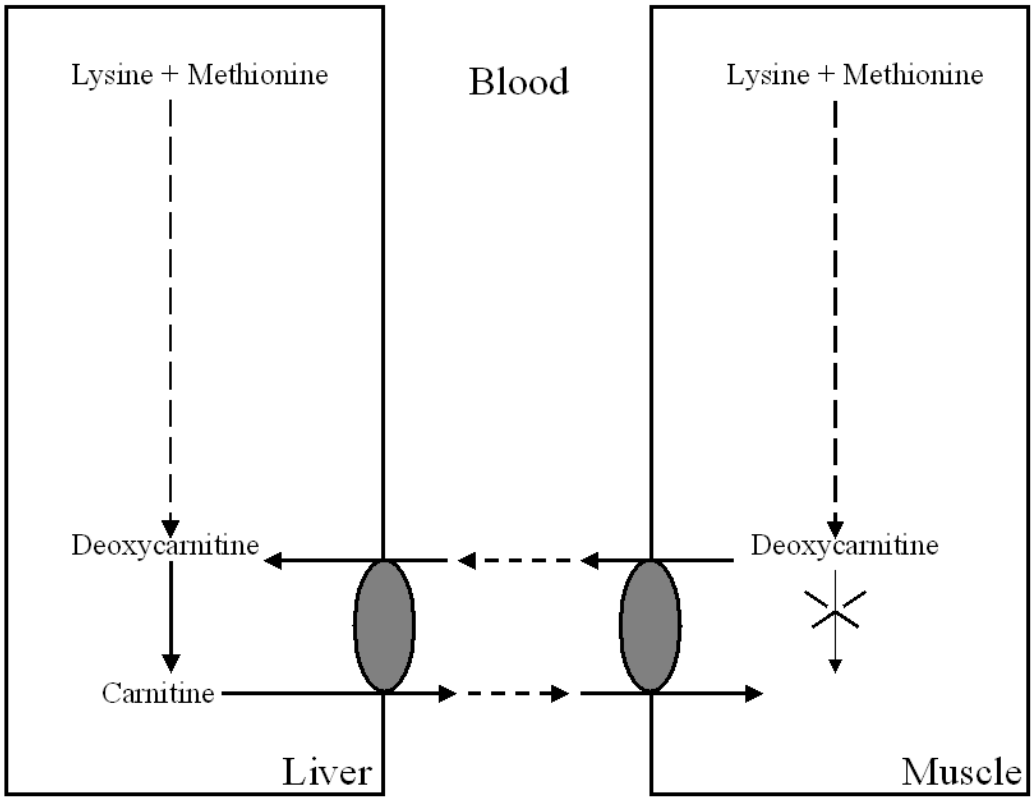
-CoA **intramitochondrially** is needed for:

a) α–Kg DHase.
b) PDHC.
c) Thiolase (terminal enzyme of beta-oxidation).
d) Beta-ketothiolase (tyrosine catabolism).
e) Lysine catabolism.
f) Leucine, Isoleucine, Valine catabolism.
g) Hippurate biosynthesis.

-CoA **extramitochondrially** is needed for:

a) ATP citrate lyase.
b) Acyl-CoA synthetase.
c) Acetyltransferase for Δ9- desaturase elongation system.

-Primary carnitine deficiency: muscle or liver type (see figure below).

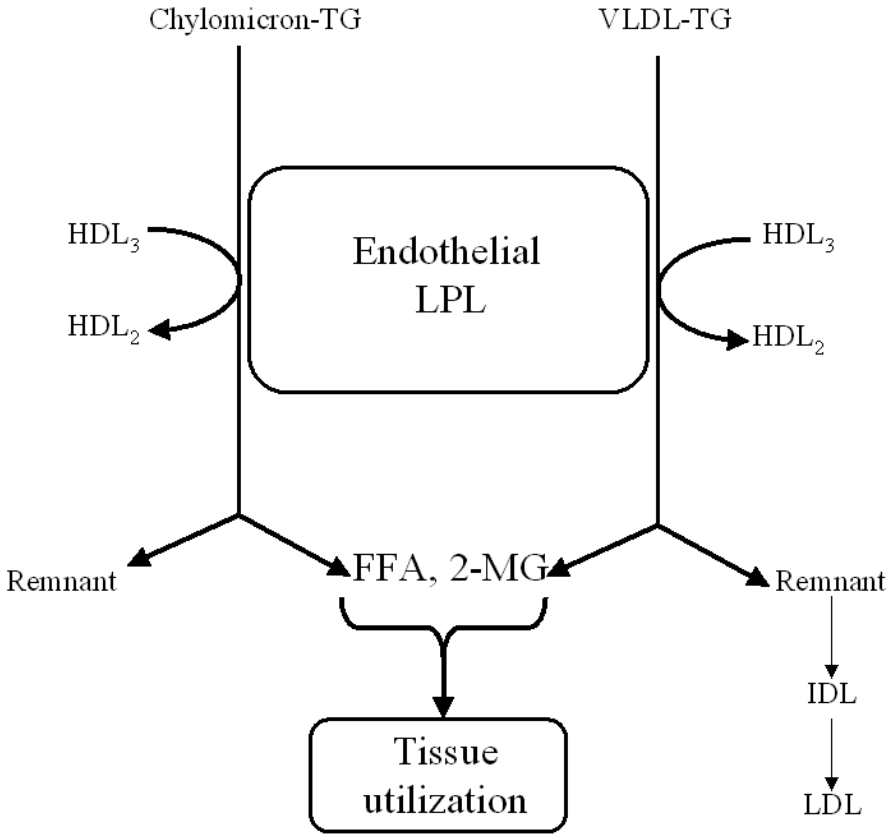


-As it is obvious from the above scheme, carnitine is synthesized only in the liver (and in the kidney), while in the muscle, normally, the terminal enzyme of the carnitine biosynthesis is missing. The muscle depends on its carnitine/deoxycarnitine antiporter in order to take

carnitine. In the liver there is the same antiporter but there it works on the opposite way. The reason for that is to maintain a gradient in the muscle for carnitine **influx**, while for the liver a carnitine **efflux**.

-Secondary carnitine deficiency: in case of excessive loss of carnitine by the kidney in the nephrotic syndrome.
-Muscle type of carnitine deficiency: Either deoxycarnitine is deficient or the transporter in the muscle is defective.
-Liver type of carnitine deficiency: Carnitine biosynthetic pathway is deficient or the liver transporter is defective.

29) The function of lipoprotein lipase. How are lipoprotein lipase isoenzymes regulated in various organs and in various physiological conditions?



-LPL function: see question ***2 to 11***.
-LPL requires Apo C II and phospholipids as cofactors.
-In LPL deficiency there is only hyperchylomicronemia, while the VLDL and LDL concentration in the blood remains normal. This is explained by the fact, that chylomicrons are too big to be processed by the hepatic lipase, and therefore their catabolism is absolutely dependent on a functional LPL. On the contrary, VLDL and LDL can be processed by the hepatic lipase, as they are much smaller than chylomicrons, and therefore they can easily reach in the vicinity where the hepatic lipase resides.
-LPL isoenzymes are found in brown adipose tissue (induced by cold, where the gene coding for thermogenin protein is upregulated), in the mammary gland (induced by prolactin), in the heart (induced by fasting), in the white adipose tissue (induced by insulin), and other organs but without a significant regulation.

AMINO ACID AND NUCLEOTIDE METABOLISM

1) Metabolism of proteins: nutritional requirements, deficiency states

-Harper, Fig. 31-1, Fig. 31-3, Table 54-1, Table 54-4.
-There are 9 essential aa in humans: histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine. Two other aa, cysteine and tyrosine, may be formed from the essential aa methionine and phenylalanine, respectively. Arginine and histidine are considered to be semi-essential aa, that means that they are needed in the diet when the organism is in a need for a positive nitrogen balance, i. e. pregnancy, lactation, tissue repair after injury, etc. Arginine is produced as a part of the urea cycle, while Histidine is made by the intestinal flora. As long as sufficient amounts of essential aa are present in the diet, the remaining aa required for protein synthesis and other purposes can be formed through transamination and other reactions.
-An adult animal in a state of metabolic equilibrium requires dietary protein to replace the essential aa and nitrogen lost during metabolic turnover. Nitrogen is lost in the urine, feces, saliva, desquamated skin, hair, and nails. For most situations, a diet in which 12 % of the energy is supplied as protein is adequate in humans.
-The quantity of protein required is affected by three major factors**: protein quality, energy intake, and physical activity.**

A. **Protein quality**: The quality of protein is measured by comparing the proportions of essential aa in a food with the proportions required for good nutrition. The closer the proportions are the higher the protein quality. Egg and milk proteins are high-quality proteins that are efficiently utilized by the body and are used as reference standards against which other proteins can be compared. Meat protein is of high protein quality, whereas several proteins from plants used as major food sources are relatively deficient in certain aa, i.e., tryptophan, lysine (especially from wheat), and methionine (especially from beans). In a mixed diet, a deficiency of an aa in one protein is made up by its abundance in another; such proteins are described as complementary; i.e. the protein of wheat and beans combined provides a satisfactory aa intake. Under such circumstances, a greater *total* amount of protein must be consumed to satisfy requirements. Aa that are not incorporated into new protein and are unnecessary for immediate requirements cannot be stored and are rapidly degraded, and the nitrogen is excreted as urea and other products.

B. **Energy intake**: The energy derived from carbohydrate and fat affects protein requirements because it spares the use of protein as an energy source. To use expensive (high-quality) dietary protein efficiently and to reduce requirements for it to a minimum, it is necessary to ensure adequate provision of energy from non-protein sources, some of which should be carbohydrate in order to spare protein from gluconeogenesis.

C. **Physical activity**: Physical activity increases nitrogen retention from dietary protein.

-Protein-energy malnutrition encompasses a range of disorders of starvation and malnutrition that involve deficiencies of other nutrients such as vitamins and minerals in addition to protein. In severe form, it occurs in growing children, usually under 5 years of age, in developing

countries. In **marasmus**, there is generalized wasting due to deficiency of both energy and protein. In **kwashiorkor**, which is characterized by edema, while energy intake may be adequate, there is a deficiency in both the quantity and the quality of protein. (Kwashiorkor is the word used by members of the Ga tribe in Ghana to describe "the sickness the older gets when the next child is born". It follows weaning from breast milk and exposure to a diet low in protein and high in carbohydrate).

2) Degradation of proteins in the gastrointestinal tract I. Endopeptidases /zymogens, activation, substrate specificity/

-As a result of contact with gastric HCl, proteins are denatured; i.e. the tertiary protein structure is lost as a result of the destruction of hydrogen bonds. This allows the polypeptide chain to unfold, making it more accessible to the actions of proteolytic enzymes (proteases).
-Protein digestion is initiated by pepsin. Pepsin is produced in the chief cells as the inactive zymogen, pepsinogen. This is activated to pepsin by H^+, which splits off a protective polypeptide to expose active pepsin; and by pepsin, which rapidly activates further molecules of pepsinogen (autocatalysis). Pepsin is an endopeptidase, since it hydrolyzes peptide bonds within the main polypeptide structure rather than adjacent to amino or carboxyl terminal residues, which is characteristic of exopeptidases. It is specific for peptide bonds formed by aromatic aa (i.e. tyrosine) or dicarboxylic aa (i.e. glutamate).
-Rennin (chymosin) is important in the digestive process of infants because it prevents the rapid passage of milk from the stomach. In the presence of calcium, rennin changes the casein of milk irreversibly to a paracasein, which is then acted on by pepsin. Rennin is reported to be absent from the stomach of adults. It is also used in cheese manufacture.
-Digestion continues in the intestine. The stomach contents, or **chyme**, are intermittently introduced during digestion into the duodenum through the pyloric valve. The alkaline content of pancreatic and biliary secretions neutralizes the acid of the chyme and changes the pH of this material to the alkaline side; this shift of pH is necessary for the activity of the enzymes contained in pancreatic and intestinal juice, but it inhibits the further action of pepsin.
-Trypsin, chymotrypsin, and elastase are endopeptidases. The proteolytic action of pancreatic secretion is due to those three endopeptidases.
-Trypsin is specific for peptide bonds of basic aa, and chymotrypsin is specific for peptide bonds containing uncharged aa residues, such as aromatic aa.
-Elastase, in spite of its name, has rather broad specificity in attacking bonds next to small aa residues, such as glycine, alanine, and serine.
-All three enzymes are secreted as zymogens. Activation of trypsinogen is due to another enzyme, enterokinase, secreted by the intestinal mucosa. Once trypsin is formed, it will attack not only additional molecules of trypsinogen, but also the other zymogens in the pancreatic secretion, chymotrypsinogen, proelastase, and procarboxypeptidase, liberating their active counterparts.

3) Degradation of proteins in the gastrointestinal tract II. Exopeptidases and absorption of amino acids

-Carboxypeptidase is an exopeptidase. The further attack on the polypeptides produced by the action of endopeptidases is carried on by this exopeptidase, which attacks the carboxyl terminal bond, liberating single aa.
-Aminopeptidase, which is secreted by the glands of Lieberkuhn, attacks bonds next to the amino terminals of aa of polypeptides and oligopeptides; these glands also secrete dipeptidases of various specificities, some of which may be within the intestinal epithelium. The latter complete digestion of dipeptides to free aa.
-Under normal circumstances, the dietary proteins are almost completely digested to their constituent aa, and these end products of protein digestion are then rapidly absorbed from the intestine into the portal blood. It is possible that some hydrolysis, i.e. of dipeptides, is completed in the intestinal wall.
The natural (L) isomer – but not the (D) isomer – of an aa is actively transported across the intestine from the mucosa to the serosa; vitamin B_6 (PLP) is involved in this transfer. This active transport of the L-aa is an energy dependent process. Aa are transported through the brush border by a multiplicity of carriers (transporters), many having Na^+- dependent mechanisms similar to the glucose transporter system. Of the Na^+- dependent carriers, there is a neutral aa carrier, a phenylalanine and methionine carrier, and a carrier specific for imino acids such as proline and hydroxyproline. Na^+- independent carriers specializing in the transport of neutral and lipophilic aa (i.e. phenylalanine and leucine) or of cationic aa (i.e. lysine) have been characterized.

4) L-glutamate dehydrogenase. The role of glutamic acid in the deamination of amino acids, transdeamination

-Harper, Fig. 30-1,Fig. 31-3, Fig. 31-4, Fig. 31-5, Fig. 31-6.
-Glutamine synthase and Glutamate Dehydrogenase (GluDHase) both fix N_i – one into amino and the other into the amide linkage. Both reactions are coupled to highly exergonic reactions, for GluDHase the oxidation of NAD(P)H and for Gln synthetase the hydrolysis of ATP. The GluDHase is a reversible reaction but glutamate (Glu) formation is favored.
-NH_3 is derived mainly from deamination of the α- amino N of aa, and is toxic to all animals. Human tissues therefore, initially detoxify NH_3 by converting it to certain aa (Gln, Ala, Ser) for transport to the liver, kidney, and gut. Deamination of them in the liver releases NH_3, which then is efficiently converted to the non-toxic, N – rich compound urea.
-Each transaminase is specific for one pair of substrates but non-specific for the other pair. Since Ala is also a substrate for Glu transaminase, all the amino from aa that can undergo transamination can be concentrated in Glu. This is important, because **L –Glu is the only aa in mammalian tissues that undergoes oxidative deamination at an appreciable rate**. Alanine serves as a key gluconeogenic aa (Glc-Ala cycle). In the liver, the rate of glucose synthesis from alanine is far higher than that observed from all other aa. The capacity of the liver for

gluconeogenesis from alanine, which is enormous, does not reach saturation 20-30 times its physiological level. The formation of NH_3 from α -amino groups thus, occurs mainly via conversion to the α -amino N of L –Glu.
L –Glu DHase occupies a central position in N metabolism. The α -amino groups of most aa ultimately are transferred to α -ketoglutarate (α -Kg) by transamination forming L –Glu. Release of this N as NH_3 is then catalyzed by L –GluDHase, an enzyme that uses either NAD^+ or $NADP^+$ as oxidant. Liver GluDHase activity is regulated by the allosteric inhibitors ATP, GTP, and NADH, and by the activator ADP. This freely reversible reaction functions both in aa catabolism and biosynthesis. Catabolically, it conveys N from Glu to urea. Anabolically, it catalyzes amination of α-Kg by free NH_3.
-Most aa undergo transamination, except Lys, Thr, Pro, and HyPro. Since transaminations are freely reversible, transaminases can function both in aa catabolism and biosynthesis. PLP (vitamin B_6) resides at the catalytic site of all transaminases (and also of many other enzymes with aa as substrates).
-As a synopsis, we may say that, L –Glu DHase plays a role in:
i) The elimination of the α-amino group of aa.
ii) The biosynthesis of nutritionally non-essential aa.
iii) Oxidative deamination.

5) The role of aspartate in the metabolism of N-containing compounds. Purine nucleotide cycle.

-Harper, Fig. 35-17.
-See also question ***8***.
-Aspartate donates one nitrogen in the urea cycle in order for urea to be formed (urea possesses 2 nitrogen atoms, while ammonia only one; one nitrogen of urea comes from ammonia and the other from aspartate).
-Aspartate also plays an important role in the purine nucleotide cycle especially in the muscle, as it is an intermediate of this cycle. Therefore, the muscle is able to deaminate (as the purine nucleotide cycle releases ammonia, and Glutamate DHase has a very low activity in the muscle), and also it couples this cycle with the Krebs cycle, as it forms fumarate.

6) Transamination. Mechanism, metabolic importance, the role of pyridoxal phosphate

-Harper, Fig. 31-4, Fig. 31-5, Fig. 52-3, Fig. 52-10, Fig. 17-5.
-See also question ***4***.

7) The transport of ammonia, the role of glutamine and alanine. Glutamine synthesis and the loss of amide-group

-Harper, Fig. 31-11, Fig. 31-12, Fig. 31-13, Fig. 31-8, Fig. 31-9.
-NH_3 generated by enteric bacteria is absorbed into the portal venous blood, which thus contains higher levels of NH_3 than does systemic blood. Since the liver normally removes this NH_3 promptly from the portal blood, peripheral blood is virtually NH_3–free. Should portal blood bypass the liver, NH_3 then may rise to toxic levels in the systemic blood (i.e. cirrhosis).
-Formation and secretion of NH_3 maintains acid-base balance. NH3 production increases in metabolic acidosis and decreases in alkalosis.
-Although brain tissue can form urea, this does not appear to play a significant role in NH_3 removal. In brain tissue, the major mechanism for detoxification of NH_3 is Gln formation, by the use of Glutamine Synthase reaction (Harper, Fig. 31-8). However, if blood NH_3 levels are elevated, the supply of blood Glu available to the brain is inadequate for formation of Gln. The brain therefore, also must synthesize Glu from α-Kg. This would rapidly deplete the TCA intermediates, unless they were replaced by CO_2 fixation, converting pyruvate to OAA. Fixation of CO_2 into aa indeed occurs in brain tissue. After infusion of NH_3, TCA cycle intermediates are diverted to the synthesis of α-Kg and subsequently of Gln.
-NH_3 is produced in the following ways:

i) degradation of aa (transdeamination, plus minor pathways: Thr, Cys catabolism)
ii) deamination of other compounds (N-containing side chains of nucleotides, neurotransmitters).
iii) NH_3 produced in the large intestine by bacteria.

-The use of alanine to transport NH_3 from hard-working skeletal muscles to the liver is another example of the intrinsic economy of living organisms. Vigorously contracting skeletal muscle operate anaerobically producing not only NH_3 from protein breakdown, but also large amounts of pyruvate from glycolysis. Both these products must find their way to the liver, for NH_3 to be incorporated into urea for excretion, and for pyruvate to be rebuilt into glucose and returned to the muscles. Animals thus solve these 2 problems with one cycle: they move the carbon atoms of pyruvate, as well as excess NH_3 from muscle to liver as alanine. In the liver, alanine yields pyruvate, the starting material for gluconeogenesis, and releases NH_4^+ for urea synthesis. The energetic burden of gluconeogenesis (PC, PEPCK, and PGK are ATP consuming reactions) is thus imposed on the liver rather than the muscle, so that the available ATP in the muscle can be devoted to muscle contraction.

8) Urea cycle. Reactions, regulation, metabolic disorders

-Harper, Fig. 31-14, ***pages 305-308***.
-Addendum: Biosynthesis of urea requires:
A) Permanent synthesis of Carbamoyl-P.
B) Permanent function of the urea cycle.

C) A slow filling up reaction of the urea cycle: *de novo* synthesis of ornithine from glutamate (in the presence of sufficient amount of dietary Arginine, this is not necessary).

9) Glucoplastic and ketoplastic amino acids: the metabolic intermediates produced from the carbon skeletons of amino acids

-Harper, Table 32-1, Fig. 32-1, Fig. 32-2.

10) Amino acids forming pyruvate /mechanisms/

-Harper, Fig. 32-7, Fig. 32-8, Fig. 32-9, Fig. 32-11, Fig. 32-12, Fig. 31-5.

11) Amino acids forming succinyl-CoA /mechanisms/

-Harper, Fig. 32-21, Fig. 32-22, Fig. 21-2, Fig. 32-25, Fig. 32-27, Fig. 32-28.

12) Amino acids forming alpha-ketoglutarate and oxaloacetate /mechanisms/

-Harper, Fig. 32-2, Fig. 32-3, Fig. 32-4.

13) The catabolism of phenylalanine and tyrosine /main intermediates and enzyme defects/

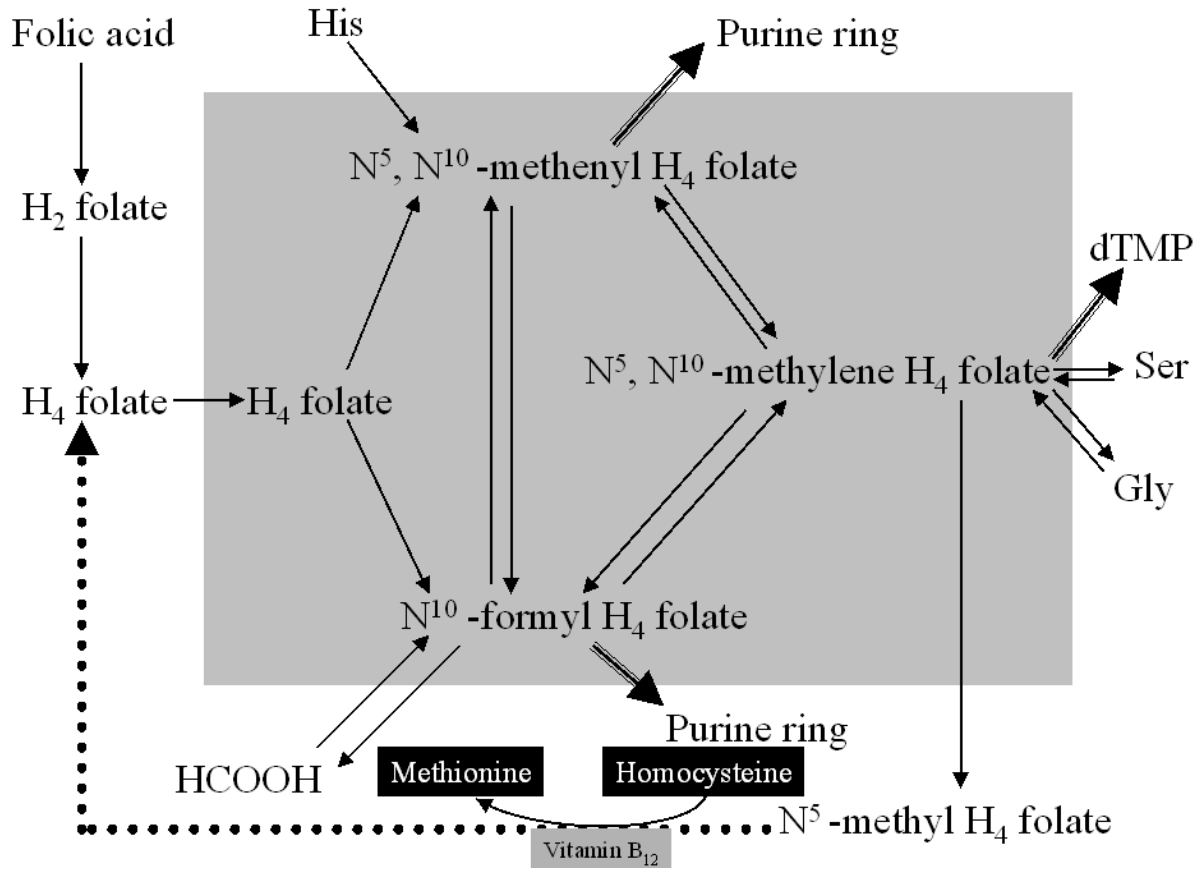
-Harper, Fig. 30-10, Fig. 32-13, Fig. 32-14, Fig. 32-15, Fig. 32-16, Table 32-3, **pages 316-321**.

14) The role of methionine as a methyl group donor in biosynthetic reactions, the metabolism of methionine

-Harper, Fig. 32-21, Fig. 32-22, Fig. 21-2.
-SAM is required in the following reactions:

i) Biosynthesis of phosphatidylcholine from phosphatidylethanolamine and SAM and SAH (S-adenosylhomocysteine).
ii) Formation of creatine.
iii) Cysteine formation.
iv) Spermine and spermidine formation.
v) B_{12} and folate metabolism.
vi) Epinephrine formation.
vii) Methylation of xenobiotics.
viii) DNA methylation.

15) The role of folate coenzymes in the metabolism

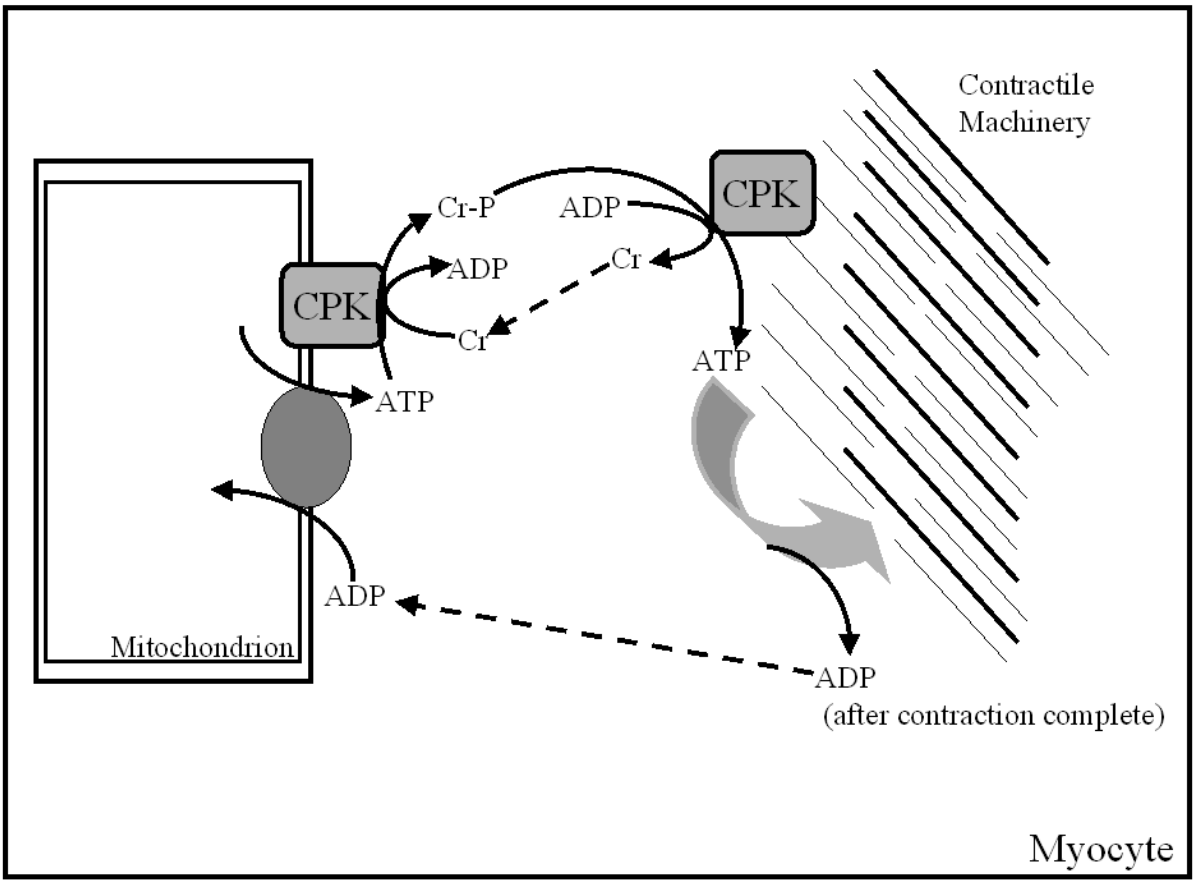


-The metabolic interconversions of folic acid and its derivatives are indicated with thin arrows. Pathways relying exclusively on folate are shown with triple thick arrows. The important B_{12}-dependent reaction regenerating H_4 folate is shown with a thick dotted arrow. The gray shaded box encloses the "pool" of C_1 derivatives of H_4 folate.

-Folate is taken up by the intestinal mucosal cells and reduced to tetrahydrofolate by the enzyme dihydrofolate reductase and circulates in the plasma primarily as the free N^5 – methyl derivative of tetrahydrofolate (THF). Inside cells, THF is found as polyglutamate derivatives, and these appear to be the biologically most potent forms.
-THF derivatives are used in many biosynthetic reactions: i.e. in the synthesis of choline, serine, glycine, purines, and dTMP. Since adequate amounts of choline and the aa can usually be obtained from the diet, the participation of folates in purine and dTMP synthesis appears to be metabolically the most significant of those reactions. In addition, THF and vitamin B_{12} are required, along with vitamin B_6, for the conversion of homocysteine to methionine. The most pronounced effect of folate deficiency, is inhibition of DNA synthesis due to decreased availability of purines and dTMP. This leads to arrest of cells in S phase and a characteristic "megaloblastic" change in size and shape of nuclei of rapidly dividing cells (i.e. bone marrow cells). The block in DNA synthesis slows down maturation of RBC`s, causing production of abnormally large "macrocytic" RBC`s with fragile membranes.

16) The synthesis of creatine and polyamines

-Harper, Fig. 33-5, Fig. 33-10.
Creatine Phosphate is a small molecule that can migrate very fast within the cell, faster than ATP. This physicochemical property is important for the contracting muscle (see figure below).



-Polyamines must be synthesized before mitosis occurs. Polyamines are very positively charged, helping the DNA to pack very densely and form chromosomes.

17) The synthesis of nonessential amino acids: serine, cysteine, aspartate, alanine, glutamate

-Harper, Fig. 32-11, Fig. 30-5, Fig. 30-9, Fig. 32-8, Fig. 32-22, Fig. 31-9, Fig. 30-3, Fig. 30-1.

18) The "de novo" synthesis of ornithine and proline

-Harper, Fig. 30-8, Fig. 31-14.

19) The metabolism of glycine

-Harper, Fig. 30-6, Fig. 32-6, Fig. 32-11.

20) The biosynthesis of heme, regulation of heme biosynthesis

-Harper, **CHAPTER 34 (there is no point on copying the book, since the WHOLE chapter is required)**.

21) The catabolism of heme, formation and metabolism of bile pigments

-Harper, **CHAPTER 34 (there is no point on copying the book, since the WHOLE chapter is required)**.

22) The catabolism of sulfur-containing amino acids. The formation and metabolic importance of "active sulfate"

-Harper, Fig. 32-8, Fig. 32-9, Fig. 32-21, Fig. 32-22, Fig. 21-2, Fig. 35-14, Fig. 26-9.
-Active sulfate, (PAPS) is the sulfate donor for the formation of sulfated proteoglycans, or urinary metabolites of drugs excreted as sulfate conjugates.

23) Cytoplasmic and mitochondrial synthesis of carbamoyl phosphate

-Harper, Fig. 31-14, Fig. 36-11.

24) The sources of the constituent atoms of the purine ring. Synthesis and regulation of purine nucleotides

-Harper, Fig. 36-1, Fig. 36-2, Fig. 36-3, Fig. 36-7, Fig. 36-8.

25) The structure and role of the active form of folic acid in the synthesis of nucleotides. The synthesis of thymidylate. The molecular basis of the actions of the antimetabolites of folic acid

-Harper, Fig. 52-15, Fig. 52-16, Fig. 52-17, Fig. 52-18, Fig. 52-19.
-See also question ***15***.

26) The biosynthesis of pyrimidine nucleotides and its regulation in mammalian tissues

-Harper, Fig. 36-11, Fig. 36-13.
-Since the biosynthesis of IMP from amphibolic intermediates consumes glycine, Gln, THF derivatives, Asp, and ATP, it is imperative that cells regulate purine biosynthesis. The major determinant of the overall rate of "de novo" purine nucleotide biosynthesis is the PRPP concentration, a parameter that reflects the relative rates of PRPP synthesis, utilization, and degradation. The rate of PRPP synthesis depends both on the availability of Ribose-5-Phosphate and on the activity of PRPP synthetase, an enzyme sensitive both to [Phosphate] and to the purine ribonucleotides that act as its allosteric regulators.
-Cross regulation between the pathways of IMP metabolism serves to decrease synthesis of one purine nucleotide when there is a deficiency of the other nucleotide.
-Purine and pyrimidine nucleotide biosynthesis are coordinately regulated processes (at the PRPP synthase step).

27) The biosynthesis and regulation of deoxyribonucleotides

-Harper, Fig. 36-9, Fig. 36-10.

28) Salvage reactions in the metabolism of purine nucleotides. The Lesch-Nyhan syndrome

-Harper, Fig. 36-5, Fig. 36-6.
-Conversion of purines, purine ribonucleotides, and purine deoxyribonucleotides to mononucleotides involves the so-called salvage reactions, that require far less energy than does the "de novo" synthesis. Quantitatively, the more important mechanism involves phosphoribosylation of a free purine by PRPP, forming a purine 5`mononucleotide. A second salvage mechanism involves direct phosphorylation of a purine ribonucleoside by ATP.
-Mammalian liver, the major site of purine nucleotide biosynthesis, provides purines and their nucleosides for salvage and utilization by tissues incapable of their biosynthesis. For example, brain has a low level of PRPP amidotransferase and hence, depends on –partly – on exogenous purines. RBC`s cannot synthesize 5`phosphoribosylamine and hence, utilize exogenous purines to form nucleotides.
-Lesch –Nyhan syndrome, an overproduction hyperuricemia with frequent uric acid lithiasis and a bizarre syndrome of self-mutilation, is due to a non-functional HGPRT, an enzyme of purine salvage. The accompanying rise in intracellular PRPP, which has been spared from purine salvage, results in purine overproduction.

29) The degradation of purine and pyrimidine nucleotides. Metabolic diseases in the catabolism of purines

-Harper, Fig. 36-14, Fig. 36-16, Table 36-1, Table 36-2.

30) The pathogenesis and consequences of hyperammonemias

Theoretical Background:
-How is NH_3 formed?
A) By degradation of aa (transdeamination, minor catabolic pathways, such as Thr, Cys catabolism).
B) By deamination of other than aa compounds (N-containing side chains of nucleotides, neurotransmitters), plus the purine nucleotide cycle.
C) By the intestinal flora; however, this ammonia will be processed extensively by the liver, unless this organ is bypassed, i.e. in cirrhosis by means of the portocaval anastomosis.
-How is NH_3 eliminated?
A) By the urea cycle in the liver.
B) By the glutamine synthase reaction (VIP [Very Important Pathway] in the brain)
-Why is there a need for NH_3 to be eliminated?
-Because it is toxic, especially for the brain. Accumulation of ammonia will lead to an excessive production of glutamine in the astrocytes (by the glutamine synthase reaction, Harper, Fig. 31-8). Glutamine is relatively impermeable, and therefore, by osmosis it will cause a lot of water to migrate towards the interior milieu of the astrocytes causing them to

swell. As the astrocytes wrap around the brain capillaries, they will be constricted, causing a major obstruction to the outflow of blood stream from the skull, leading to an intracranial edema. Moreover, the glutamine synthase reaction requires ATP, and since it is an unregulated one, it will continue indefinitely, as long as there is NH_3 and glutamate. Glutamate stores are limited within the neurons, and therefore, this will cause a shift of the formation of glutamate from metabolic intermediates, that is α–Kg, and that, in turn, will deplete the citric acid cycle intermediates. So, there will be a generalized depletion of energy due to increased consumption of ATP by the glutamine synthase reaction, and also a loss of important metabolic intermediates that would have led to provision of ATP and reducing equivalents ($NADH+H^+$ and $FADH_2$), at the level of the Krebs cycle.

-<u>Pathogenetic causes</u>:

a) Urea cycle enzyme(s) deficiency.
b) Aspartate deficiency (as an essential component of the urea cycle).
c) Glutamate deficiency.
d) B_{12} deficiency (see question ***31***).
e) MMCoA isomerase deficiency (see question ***31***).
f) PC deficiency. This is due to the fact, that OAA formed by the PC, is also converted to Aspartate by a transamination at an appreciable rate. Since Aspartate is an essential component of the urea cycle for the donation of a Nitrogen atom, it is obvious, that in Aspartate deficiency hyperammonemia would ensue, and this is the case of PC deficiency.
g) PDH deficiency. Since the product of PDH is Acetyl-CoA, and it is known that Acetyl-CoA is an absolutely necessary activator of PC, it follows that in tissues where Acetyl-CoA can come only from PDH (and not FA catabolism as it is the case in the brain), a PDH deficiency would lead to a secondary PC deficiency as well. See f).
h) Cirrhosis of the liver.

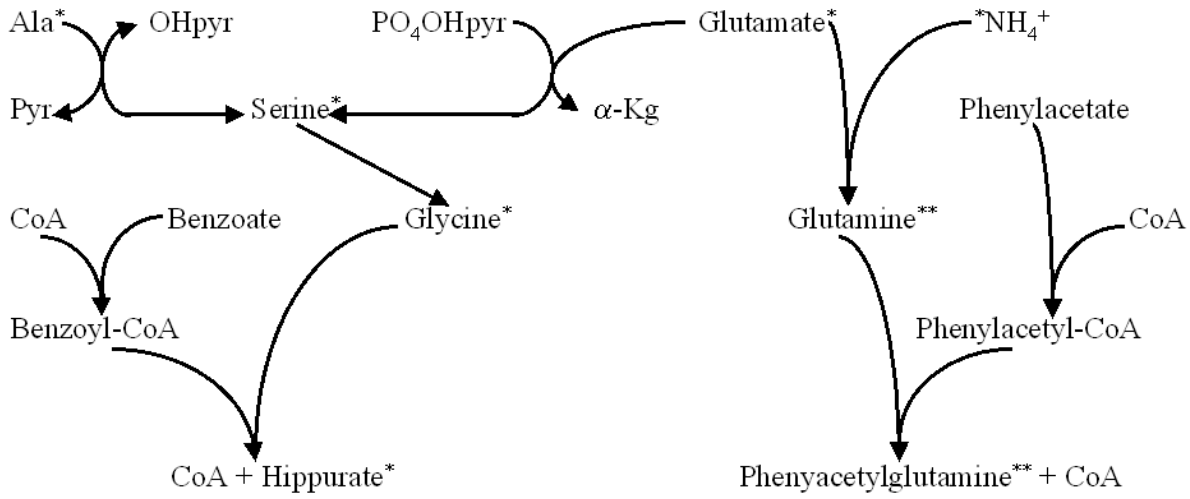
-<u>Consequences</u>:

a) Elevated urine pH. That is due to the fact that excessive NH_3 would combine with protons forming NH_4^+, and therefore lowering the proton concentration in the urine.
b) Elevated Glutamine level. See above.
c) Elevated uracil level. This is valid only in case of a urea cycle deficiency below the level of CPS I. In this case, carbamoyl-P accumulates and it is transported out of the mitochondria. In the cytosol, it will meet with the biosynthetic complex of the pyrimidine nucleotides, leading to an increased production of uracil, which is the terminal product. (Remember that in the cytosol there is also a CPS II which is the first enzyme of the pyrimidine nucleotides biosynthetic pathway, producing also carbamoyl-P, Harper, Fig. 36-11).
d) Encephalopathy (hepatic coma). This is mainly due to the cerebral edema, but also to the underlying factors, such as the ATP loss and the intermediates of the Krebs cycle loss.
e) Liver dysregulation and associated syndromes and /or diseases. See question ***13*** in **<u>ORGAN BIOCHEMISTRY</u>**.

-<u>What kind of therapeutic measures are taken in case of a hyperammonemia</u>?

A) Slight restriction of protein intake and replacement by Branched Chain Keto Acids (BCKAs). That is to decrease the ammonia load by dietary means.
B) Administration of NaBenzoate and/or NaPhenylacetate (see figure below). This approach is rational only if the liver is functional. The mode of action is that these two aforementioned compounds bind to glycine and glutamine, respectively, forming hippurate and phenylacetylglutamine, respectively, which are both excreted in the urine, “dragging” with

them Nitrogens. Because these reactions take place in the liver, it was mentioned before that only in a fully functional liver this method of approach would be applicable.



Asterisks denote Nitrogen atoms destined for waste Nitrogen excretion in hippurate and phenylacetylglutamine.

-However, nowadays the use of NaBenzoate and NaPhenylacetate has been abandoned.

C) Arginine, and /or Citruline supplementation. This is valid only in case of urea cycle deficiencies, above or at the level of ornithine transcarbamoylase (in case of citrulline supplementation), and above or at the level of Argininosuccinase (in case of Arginine supplementation) in order to "bypass" the defect.

D) Adequate caloric supplementation in form of carbohydrates, so insulin levels will remain high, and that would prevent deleterious proteolysis in the muscle.

Administration of Lactulose by mouth (per os). This is only valid in liver cirrhosis. The rationale for this approach is the following: Lactulose is a trisaccharide, which cannot be digested by the human body, as we lack the appropriate enzyme. However, bacteria are able to ferment it, with the result of organic acids, such as HCOOH, CH_3COOH etc. This creates a local acidic environment within the intestine. Remember that the intestinal flora can make up to 35 % of the total NH_3 produced in the body. In live cirrhosis, the liver is bypassed by the portocaval anastomosis, and therefore this huge amount of NH_3 reaches the systemic circulation bypassing the filtering function of the liver. In case of per os lactulose administration, the acidic products of its fermentation by the intestinal flora causes the reaction $NH_3 + H^+ \longrightarrow NH_4^+$ to shift towards the right. However, NH_4^+ cannot penetrate membranes, and therefore cannot reach the systemic circulation, while it is safely excreted in the feces.

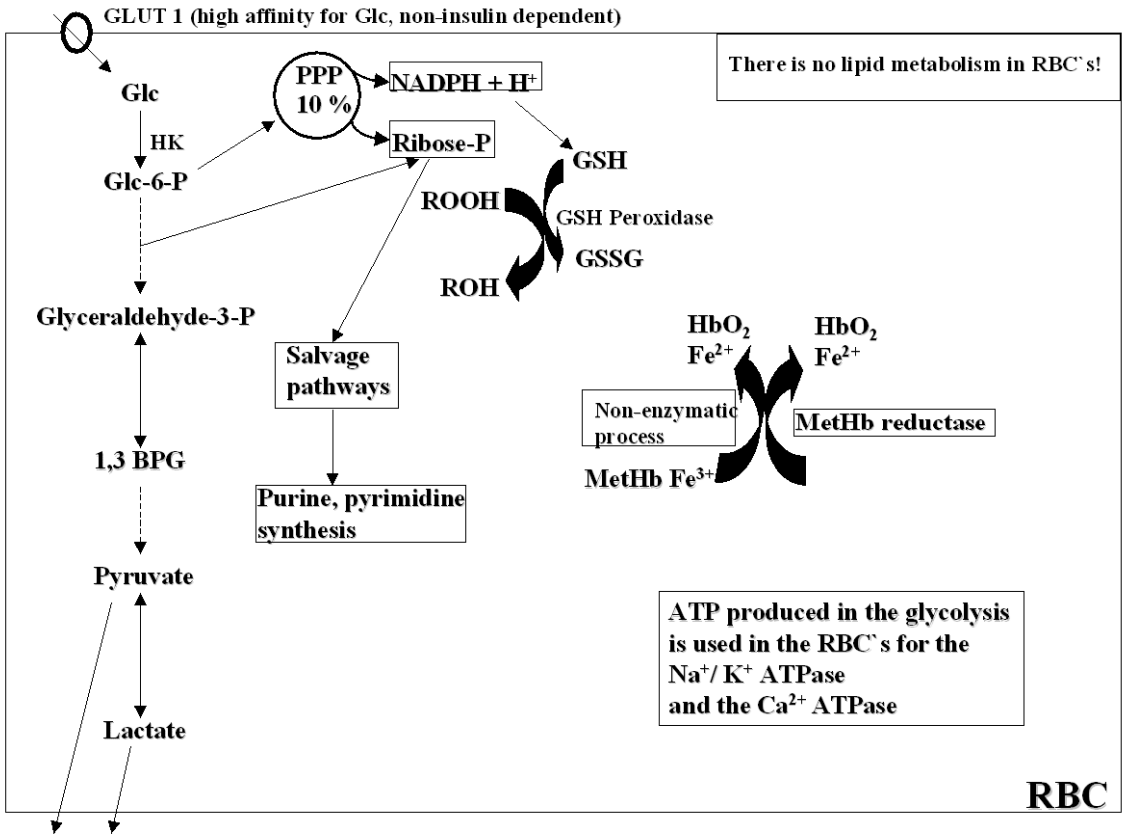
31) The effect of per os and parenteral administration of cyanocobalamin on the amino acid metabolism. The pathomechanism of vitamin B_{12} derivative deficiencies.

-See questions ***15 and 25*** for theoretical background. See also figure on detoxification function of carnitine.
-Vitamin B_{12} is a cofactor for two very important reactions:
A) The conversion of homocysteine to Methionine accompanied by the regeneration of THF (H_4 folate) [Harper, Fig. 52-15].
B) The conversion of L-MethylMalonylCoA to Succinyl-CoA [Harper, Fig 21-2].
-Since THF can only be regenerated in the presence of B_{12} (and of course fully functional enzymes), it follows that a deficiency of B_{12} would lead to a secondary deficiency of THF in the form of N^5-methyl H_4 folate, the so-called "folate trap" hypothesis. Therefore, in vitamin B_{12} deficiency, as in folate deficiency, a megaloblastic anemia would develop, due to a deficiency of dTMP (thymidylate), which is formed only if there are adequate stores of H_4 folate. In case of vitamin B_{12} deficiency, there would be also an accumulation of Propionyl-CoA and MethylMalonyl-CoA, that in turn would inhibit several reactions within mitochondria (see figure in question ***28***, concerning the detoxification function of carnitine).
-Vitamin B_{12} in order to be absorbed by the intestinal mucosa, it requires to bind on a protein, the so-called "Intrinsic Factor". The Intrinsic Factor is produced in the parietal cells found in the stomach. Absence of the Intrinsic Factor causes a severe malabsorption of Vitamin B_{12}, leading to **pernicious anemia**.
- Since Methionine is a precursor for SAM, it is obvious also that the importance of vitamin B_{12} resides also to SAM availability, for a variety of reactions (see question ***14***).

ORGAN BIOCHEMISTRY

1) *The metabolism of red blood cells, muscle tissue and kidney*

-Metabolism of RBC`s:



-Metabolism of muscle:

A) CHARACTERIZATION
B) CARBOHYDRATE, LIPID, AA, NUCLEOTIDE METABOLISM
C) ENERGETIC REQUIREMENTS
D) CREATINE-P
E) PREFERENTIAL OXIDATION

Concerning A)

	Fast twitch		Slow twitch	Heart muscle
	Red	White	Red	
O_2 consumption	↑	↓ (few mitochondria)	↑ (a lot of mitochondria)	
Rate of glycolysis	↑	↑ ↑	↓ (more sustained contractions)	
Myosin ATPase	↑	↑	↓	

Concerning B)

CARBOHYDRATES	Skeletal	Heart
Glycogenolysis	Contraction is associated with glycogen breakdown (δ subunit [calmodulin] is bound by calcium, therefore, activating phosphorylase kinase	Only in emergency (i.e. low oxygen saturation), or very heavy exercise, then epinephrine stimulates via cAMP the phosphorylase kinase)
Glycogenesis	Forget it.	Substrates: lactate, pyruvate
Glycolysis	If PFK-2 is phosphorylated it works as Frc-1.6-BPase If PFK-2 is dephosphorylated it works as PFK-2	Epinephrine stimulates the conversion of PFK-2 to Frc-1.6-Bpase, but not in the skeletal muscle

-PPP accounts for less than 2 % of the total in skeletal muscle.

LIPIDS	Skeletal	Heart
	FA synthesis is low	FA synthase has low activity
	High beta-oxidation	High beta-oxidation
	High consumption of KB (KB consumption inhibits proteolysis)	High consumption of KB (KB consumption inhibits proteolysis)

-AA metabolism

-Branched Chain AA (BCAA) are preferentially utilized by the muscle (and the brain). As we consume BCAA from the diet, they are passing through the liver without to be taken up, as they are spared for the muscle (and the brain).

-Nucleotide metabolism

-Skeletal muscle possesses a very high activity of purine nucleotide cycle, which has a deaminating capacity; this is a compensatory mechanism for the muscle, as it has a very low GluDHase activity.

Concerning C)

-Concerning the energetic requirements, we recognize three different conditions:

i) The well-fed state, in which the muscle consumes glucose exclusively, while the heart glucose, pyruvate, and alanine.
ii) The early fasting state, in which both tissues consume BCAA and FA.
iii) The long term fasting state, in which both tissues consume KB.

Concerning D)

-Cr-P concentration is by an order of magnitude higher than that of ATP.

-See also question ***16*** from **amino acid and nucleotide metabolism**.

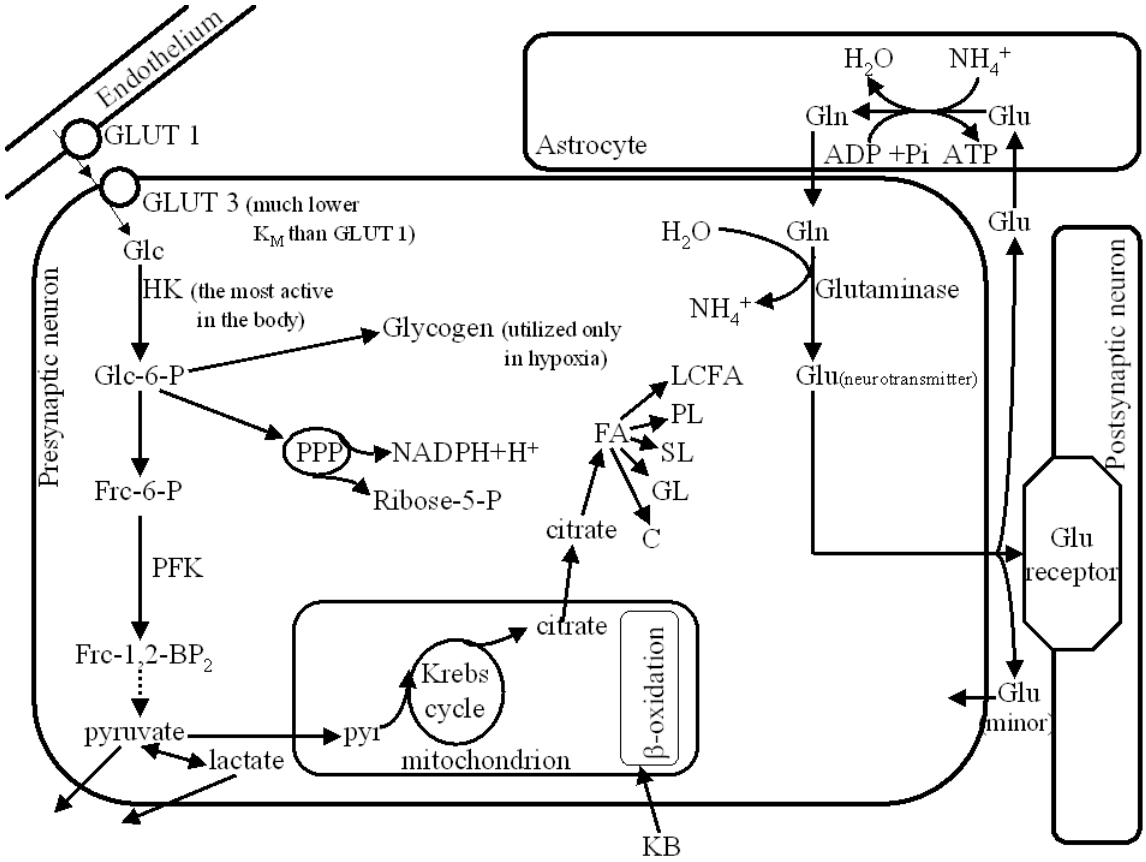
Concerning E)

-Glucose is preferentially oxidized by the muscle in case of well-fed state, rather than FA.

-Metabolism of kidney:
Kidney is divide in the cortex and the medulla. This is not only an anatomic categorization, but a biochemical too, as it will be apparent shortly the great differences among those 2 compartments in their metabolic profiles. The major difference, resides in the fact that kidney cortex is rich in mitochondria, while the medulla has only a few of them. Therefore, the medulla has a very high glycolytic capacity and a glycogen synthesis (but only from glucose), while the cortex possess gluconeogenesis, glycogenesis, glycolysis, beta-oxidation, and KB production.
-The kidney also makes arginine from citrulline and aspartate. This is very important, because arginine is taken up avidly by the liver as it enters from the diet, and nothing is left for the rest of the tissues. Therefore, kidney will compensate for the needs of the rest of the tissues for arginine, by manufacturing it from aspartate and citrulline, bypassing the liver.

2) ***The metabolism of brain and adipose tissue***

-Metabolism of brain:



-In neurons, pyruvate carboxylase has a very low activity (approximately one tenth of that in the astrocytes).
-Although neurons cannot utilize lipid coming from the blood as fuels (since they cannot penetrate the blood brain barrier as they are buried within the core of the bulky lipoproteins or they are bound on albumin, which is a huge molecule), they still possess a functional β–oxidation. This β–oxidation is for the turnover of the internally produced fatty acids and fatty acid products. 25 % of total body cholesterol is localized in the brain (indeed, brain is a very unhealthy organ to eat!).
-KB utilization takes place in the brain in case of prolonged starvation periods. Brain can account for over 70 % of KB utilization after long-term starvation.
-Concerning aa metabolism in the brain: neurons utilize very effectively the branched chain aa (valine, isoleucine, leucine) **and only** (glutamate as well, but mostly as a neurotransmitter).

-Concerning nucleotide metabolism in the brain: There is de novo synthesis of purines. Pyrimidine synthesis takes place only by the salvage pathways. Deficiency of those pathways leads to Lesch-Nyhan syndrome.
-Astrocytes contribute in pH homeostasis (in case of excessive lactate production by the neurons), as well as in ionic homeostasis in general. They are connected through gap junctions, so the signal spreading i.e. calcium signal, is vastly spread and homogeneous.
-Astrocytes possess an active glycogen metabolism.

Metabolism of adipose tissue:

-Harper, Fig. 27-8.

3) Biochemistry of the resorptive /well fed/ state
A/ Metabolic interrelationship of tissues

4) Biochemistry of the resorptive /well fed/ state
B/ Regulation of the main metabolic pathways in the liver cell

5) Biochemistry of fasting
A/Metabolic interrelationship of tissues

6) Biochemistry of fasting
B/ Regulation of the main metabolic pathways in the liver cell

7) The Cori cycle

-Harper, Fig. 21-4. Cori cycle is also known as the lactic acid cycle.

8) Regulation of the blood sugar level.

-The blood glucose level represents one of the best examples of homeostatic mechanisms; it must be kept within the narrow range of 4.5 to 5.5 mmol/L. Less than that marks hypoglycemia with deleterious consequences for the brain, RBC`s and retina, while more than that – if prolonged – is known as diabetes.
-There are mainly three organs that will contribute to normoglycemia, and these are the liver, the kidney (having exactly the same normoglycemic buffering functions as to the liver, though it is smaller in bulk), and the pancreas. The liver (and kidney) will play their role through a magnificent interplay of the biochemical pathways, namely glycolysis, glycogenesis, glycogenolysis, and gluconeogenesis, while the pancreas as an endocrine organ, by means of the insulin/glucagon ratio secreted by this organ to the blood circulation.
-Regulation of blood glucose level has five phases: Phase I is the well-fed state in which glucose is provided by dietary carbohydrate. Once this supply is exhausted, hepatic glycogenolysis maintains blood glucose levels during phase II. As this supply of glucose starts to dwindle, hepatic gluconeogenesis from lactate (coming from the brain, RBC`s, white skeletal muscle, kidney medulla, and skin), glycerol (coming from the hydrolysis of TG by LPL; this glycerol escapes in the blood circulation as it cannot be utilized by the adipose tissue, since adipose tissue doesn't have glycerol kinase), and alanine becomes increasingly important until, in phase III, it is the major source of blood glucose. These changes occur within not more than 20 hours of fasting, depending on how well fed the individual was prior to the fast, how much hepatic glycogen was present, and the sort of physical activity occurring during the fast. Several days of fasting move one into phase IV, where the dependence on gluconeogenesis actually decreases. At that time, KB have accumulated to high enough concentrations to enter the brain and meet some of its energy needs. Renal gluconeogenesis becomes also significant in this phase. Phase V occurs after very prolonged starvation of extremely obese individuals and it is characterized by even less dependence on gluconeogenesis. The energy needs of almost every tissue are met to a large extent by either fatty acid or ketone body oxidation in this phase.
-As long KB concentrations are high, proteolysis will be somewhat restricted, and conservation of muscle proteins and enzymes will occur. This continues until practically all of the fat is gone as a consequence of starvation. After all of the fat is gone, the body has to use muscle protein; and before it is gone – you are gone.

PHASE	ORIGIN OF BLOOD GLUCOSE	TISSUES USING GLUCOSE	MAJOR FUEL OF BRAIN
I	Exogenous	All	Glucose
II	Glycogen Hepatic gluconeogenesis	All except liver. Muscle and adipose tissue at diminished rates	Glucose
III	Hepatic gluconeogenesis Glycogen	All except liver. Muscle and adipose tissue at rates intermediate between II and IV	Glucose
IV	Gluconeogenesis, hepatic and renal	Brain, RBC`s, renal medulla. Small amount by muscle.	Glucose, KB
V	Gluconeogenesis, hepatic and renal	Brain at a diminished rate, RBC`s, renal medulla	KB, Glucose

-During gluconeogenesis, glycolysis should be simultaneously inhibited, and this is actually the case. This reciprocal mechanism is brought about by the release of inhibition of Frc-1,6-BPase, by Frc-2,6-BP_2, and the concomitant release of activation of PFK-1 again by Frc-2,6-BP_2. Exactly the same mechanism operates for glycolysis to be inhibited while glycogenolysis is activated, this regulation being mediated at the hormonal level by glucagon. The latter mechanism operates, so that glucose released from glycogen is not spent in the pathway of glycolysis, but it is diverted to the blood circulation for more important organs, i.e. the brain or the RBC`s.

-As it is obvious from the aforementioned 5 phases, a minimal supply of glucose is probably necessary in all extrahepatic tissues to maintain OAA concentration and the integrity of the citric acid cycle. In addition, Glucose appears to be the main source of Glycerol-3-P in tissues devoid of Glycerol kinase, such as adipose tissue and skeletal muscle. There is therefore, a minimal and obligatory rate of glucose oxidation under all conditions.

-KB and FFA spare the oxidation of glucose in muscle by impairing its entry into the cell, its phosphorylation by hexokinase and PFK-1, and its oxidative decarboxylation to pyruvate. Oxidation of FFA and KB cause an increase in the intracellular concentration of citrate that in turn inhibits PFK-1 allosterically. Oxidation of these substrates also causes [acetyl-CoA]/[CoA] and [ATP]/[ADP] ratios to increase, inhibiting PDH activity.

The combination of the effects of FFA in sparing glucose utilization in muscle and heart and the feedback effect of the spared glucose in inhibiting FFA mobilization in adipose tissue has been called the **glucose-fatty acid cycle**.
-In moderate endurance exercise, lipid is the main fuel, but in intense endurance exercise it becomes less adequate (due to inadequate O_2 availability which is absolutely necessary for β−oxidation) and carbohydrate assumes the role of major fuel until muscle glycogen is exhausted. Of importance are the TG stores in muscle cells themselves.
-TG provide more than half the energy requirements of the liver, heart, and resting skeletal muscle.
-**GLUCOSE PARADOX**: suppose that after long-term starvation, refeeding starts. ^{14}C is given by mouth incorporated in a candy. After sacrifice of the guinea pig, no radioactivity is found in liver glycogen. In another guinea pig, ^{14}C is given by mouth, as lactate. After sacrifice of the animal, radioactivity is detected in the liver glycogen. Explanation: During early refeeding, glucokinase is not yet active, so glucose escapes the liver. As this escaped glucose goes to the peripheral tissues, it is metabolized to lactate, then this lactate goes to the liver, and there as the PDH is still inactive, the PC and PEPCK – as still active gluconeogenetic enzymes - are still working, and therefore lactate will become ultimately glycogen.
-Several other hormones will participate in the regulation of blood glucose level, but only under special circumstances, i.e. glucocorticoids under stressful situations, thyroid hormones, growth hormone, E, NE. See Harper, pages 202-203: Other Hormones Affect the Blood Glucose.
-Ultimately, what should be appreciated is all of the aforementioned conditions are just the mediators to the contribution of the maintenance of blood glucose level, the most important factor regulating the blood glucose level is **THE BLOOD GLUCOSE LEVEL ITSELF.**

9) The effects of glucagon in the metabolism

-Harper, Table 51-6, **pages 595- 596**.

10) The metabolic effects of insulin

-Harper, Fig. 51-2, Fig. 51-3, Fig. 51-7, Fig. 51-8, Fig. 51-9, Fig. 51-10, Fig. 51-11, Fig. 51-12, Fig. 51-13, 51-14, Table 51-3, Table 51-4, **pages 582-594**.

11) <u>The biochemistry of diabetes</u>

-Diabetes mellitus (DM) is the most common endocrine disease, accounting for over 2% of the **<u>general population</u>**. Originally recognized from the age of Hippocrates, it adapted its name on the basis of its consequences: frequent drinking, frequent enuresis, and sweet urine (in Greek diabeno=διαβαινϖ, means "to go through" and in Latin mellitus means "sweet"). There are two types of diabetes: Insulin Dependent (IDDM) and Non-Insulin Dependent (NIDDM).

-In IDDM there is a complete absence of insulin production by the pancreas. Because of defective β–cell production of insulin, blood levels of insulin do not increase in response to elevated blood glucose levels. Even when dietary glucose is being diverted from the gut, the insulin/glucagon ratio cannot increase, and the liver remains gluconeogenic and ketogenic. Since it is impossible to switch to the process of glycolysis, glycogenesis, and lipogenesis, the liver cannot properly buffer blood glucose levels. Indeed, since hepatic gluconeogenesis is continuous, the live contributes to hyperglycemia in the well-fed state. Failure of some tissues, especially muscle, to take up glucose in the absence of insulin contributes further to the hyperglycemia. Accelerated gluconeogenesis, fueled by substrate made available by tissue protein degradation, maintains the hyperglycemia even in the starved state. The absence of insulin in patients with IDDM results in uncontrolled rates of lipolysis in adipose tissue. This increases blood levels of fatty acids and results in accelerated KB production by the liver. If KB are not used as rapidly as they are formed, diabetic ketoacidosis develops due to accumulation of KB and hydrogen ions. What also contributes to the development of ketoacidosis in IDDM patients, is the fact, that while brain would be able to utilize KB and lower their blood concentration it doesn't, because there is a lot of glucose. Therefore, the problem of the ketoacidosis resides also to the decreased KB utilization by the peripheral tissues in view of the fact that there are a lot of fuels in the blood. This is called "starvation in the midst of plenty". Not all the fatty acid taken up by the liver can be handled by the β–oxidation and ketogenesis. The excess is esterified and directed into VLDL synthesis. Hypertriglyceridemia results because VLDL is synthesized and released by the liver more rapidly than it can be cleared from the blood by LPL. The quantity of this enzyme is dependent on the blood insulin level. The defect in LPL also results in hyperchylomicronemia, since LPL is required for chylomicron catabolism in adipose tissue.

-In NIDDM, which accounts for 90 % of the diagnosed cases of diabetes, the problem resides to the fact that there is a resistance to the action of insulin, even though there are higher levels of insulin than in IDDM. Very recent data implicate increased levels of expression of TNF-α in adipocytes of obese individuals as a cause of the resistance. The greater the adipose tissue mass, the greater the production of TNF-α, which acts to impair insulin receptor function. It therefore follows that NIDDM patients are obese, in contrast to IDDM, which are very lean. However, the pancreases of NIDDM patients do not produce enough insulin to overcome the insulin resistance induced by their obesity. In contrast to IDDM, ketoacidosis does not develop because the adipocytes remain sensitive to the effect of insulin on lipolysis. Hypertriglyceridemia is characteristic of NIDDM but usually results from an increase in VLDL without hyperchylomicronemia. This is most likely explained by rapid rates of *de novo* hepatic synthesis of fatty acids and VLDL, rather than increased delivery of fatty acids from the adipose tissue.

-Diabetes is complicated by several disorders that may share a common pathogenesis. The lens, peripheral nerves, glomeruli, renal papillae, and possibly retinal capillaries contain 2 enzymes that constitute the polyol pathway (see Harper, Fig. 22.5). The first is aldose reductase, an NADPH-requiring enzyme. It reduces glucose to form sorbitol. Sorbitol is further metabolized by sorbitol DHase, an NAD^+-requiring enzyme that oxidizes sorbitol to fructose. Aldose reductase has a high K_M for glucose; therefore, this pathway is only quantitatively important during hyperglycemia. It is known that in diabetic animals the sorbitol content of lens, nerve, and glomeruli is elevated. Sorbitol accumulation may damage these tissues by causing them to swell.

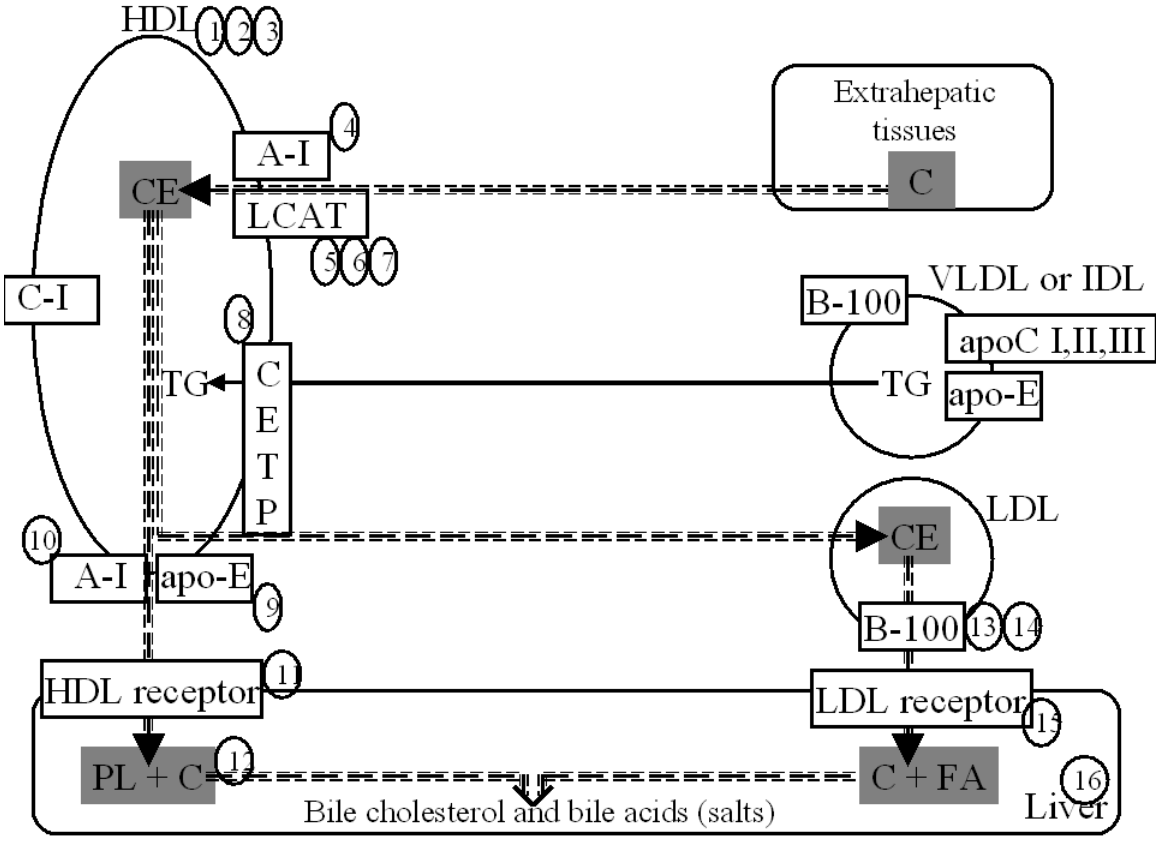
12) The biochemistry of acute liver failure

-See consultation material under the title " Acute and Chronic Liver Disease".

13) The most important biochemical functions of the liver. Biochemical deregulation in chronic liver disease

-See consultation material under the title " Acute and Chronic Liver Disease".

14) <u>Reverse cholesterol transport, biochemistry of atherosclerosis</u>



Abbreviations:
LCAT: Lecithin Cholesteryl Acyl Transferase
CETP: Cholesteryl Ester Transfer Protein
PL: Phospholipids
C: Cholesterol
TG: Triacylglycerols
CE: Cholesteryl Esters
FA: Fatty Acids

-Reverse cholesterol transport is shown in dashed arrows.
1) Glycation (non-enzymatic glycosylation) of HDL, i.e. in diabetes due to chronically elevation of blood glucose level, hinders its function, thereby preventing reverse cholesterol transport.
2) HDL is synthesized and secreted in both liver and intestine. However, nascent HDL from the intestine doesn't contain apo-C and apo-E, but only apo-A. Thus, apo-C and apo-E are synthesized in the liver and transferred to intestinal HDL when the latte enters the plasma. **A major function of HDL besides the reverse cholesterol transport is to**

act as a repository for apo-C and apo-E that are required in the metabolism of chylomicrons and VLDL.
3) Nascent HDL consists of discoid PL bilayers containing apolipoproteins and cholesterol.
4) A-I is found on HDL and chylomicrons; it is an activator of LCAT, and a ligand for the HDL receptor.
5) Cholesterol from extrahepatic tissues and other lipoproteins is taken up by HDL and esterified by LCAT. A great part of the CE formed, is subsequently transported to apo-B 100 containing lipoproteins, such as VLDL, IDL, LDL by CETP. These lipoproteins are catabolized in the liver via the LDL receptor. The rest of the CE is transported directly to the liver by HDL.
6) The overall reaction is: PL + C $\longrightarrow$ CE + LL (Lysolecithin). LL, which is quite toxic due to its ability to act as a detergent, is transported to LDL and plasma albumin.
7) As C in HDL becomes esterified, it creates a concentration gradient and draws in C from the tissues and other lipoproteins, becoming less dense (forming HDL_2), which delivers C to the liver.
8) TG is hydrolyzed by HRHL (Heparin Releasable Hepatic Lipase) in the liver.
9) Apo-E is important in C translocation to the plasma membrane, which is the initial step in reverse cholesterol transport.
10) Ethanol stimulates apo-A-I secretion by hepatocytes, with the implication of a mechanism for atherosclerosis production.
11) The HDL receptor accepts apo-A-I and apo-E.
12) In addition to TG, HRHL hydrolyses PL on the surface of HDL releasing C for uptake into the liver, allowing formation of smaller and denser HDL particles. HRHL activity is increased by androgens and decreased by estrogens, which may account for higher HDL concentration in plasma of women.
13) Apo-B-100 is synthesized in the liver, and it is a ligand for the LDL receptor; apo-B-48 is synthesized in the intestines and it doesn't bind to the LDL receptor.
15) Apo-B-100 is much bigger in size than the apo-B-48.
16) The final degradation site of HDL is mostly in the liver, but probably also in the intestines.

-Atherosclerosis is the leading cause of death in Western industrialized countries. The risk of developing it is directly related to the plasma concentration of LDL cholesterol and inversely related to that of HDL cholesterol. This explains why the former is frequently called "bad" cholesterol and the latter the "good" cholesterol, though chemically there is only one cholesterol. Atherosclerosis is a disorder of the arterial wall characterized by accumulation of cholesteryl esters in cells derived from the monocyte-macrophage line, smooth muscle cell proliferation, and fibrosis. The earliest abnormality is migration of blood monocytes to the subendothelium of the artery. Once there, they differentiate into macrophages. These cells accumulate cholesteryl esters derived from plasma LDL. Distortion of the subendothelium leads to platelet aggregation on the endothelial surface and release of plasma derived mitogens such as platelet derived growth factor (PDGF). This is thought to stimulate smooth muscle cell growth. Death of the foam cells results in the accumulation of a cellular lipid that can stimulate fibrosis.

The resulting atherosclerotic plaque narrows the blood vessel and serves as the site of thrombus formation, which precipitates distal infarction.

17) Biotransformation: reactions of the first and second phase, biological significance, clinical aspects

-Questions ***17***, ***18***, and ***19*** refer to chapter 61 in Harper. The WHOLE chapter is required.

18) Reactions and significance of cytochrome P_{450} isozymes

-Questions ***17***, ***18***, and ***19*** refer to chapter 61 in Harper. The WHOLE chapter is required.

19) Conjugation of xenobiotics and of endogenous substrates

-Questions ***17***, ***18***, and ***19*** refer to chapter 61 in Harper. The WHOLE chapter is required.

27) Synthesis of eicosanoids and their role in intercellular communication

-Harper, Fig. 25-5, Fig. 25-6.
-The essential FA give rise to eicosanoids, which make up the PGs, TXs, LTs, and LXs.
-Aspirin blocks PG synthesis.
-The major physiologic roles played by PGs are as modulators of adenylate cyclase activity, e.g. in controlling platelet aggregation, or inhibiting the effect of ADH in the kidney.
-LTs have muscle contractant and chemotactic properties, suggesting that they could be important in allergic reactions and inflammation. A mixture of LTs has been identified as the slow-reacting substance of anaphylaxis (SRS-A).

-By varying the proportions of the different PFA (Polyunsaturated Fatty Acids) in the diet, it is possible to influence the course of disease by dietary means.
-Each cell type produces only one type of prostanoid (PG plus TX). Prostanoid formation is inhibited irreversibly by aspirin. Corticosteroids are more effective because they block Phospholipase A_2 (PLA_2), so LT, LX, plus prostanoids are all blocked.
-"Switching off" of PG formation is partly achieved by a property of cyclooxygenase- that is self catalyzed destruction; it is a suicide enzyme. The inactivation of PGs, once formed, is rapid.
-PGs **increase** cAMP in platelets, thyroid, corpus luteum, adenohypophysis and lung, but **lower** cAMP in renal tubular cells and adipose tissue.

Note: These answers were written from the following reference books (using lecture notes as a guide): i) Biochemistry, Stryer 3d edition

ii) Harper's Biochemistry, Murray RK et al 23d edition
iii) Molecular Biology of the Cell, Alberts B.et al 2nd edition
iv) Molecular Cell Biology, Darnell J. 2nd edition
v) Biology, Vilee C. 2nd edition

During reading the text, you may have to see some figures that their number is presented here; unless there is no indication to which book the figure refers, this will be the Stryer's Biochemistry (e.g. for the answer 1/c. first remark, second line → fig. 27-35). The same is valid in case of referring to a page. If you see a reference such as MBC, fig. 10-40, this refers to the Molecular Biology of the Cell fig. 10-40, but if you see MCB fig. 11-47b for example, this refers to the Molecular Cell Biology fig. 11-47b. This book can be found only in the library of the Biochemistry Department (it can be found also in the central library of SOTE at Üllői út 26. but only the old edition which is not valid.) It is important to go through all the figures, that's why their number has been decreased to the minimum possible. On the left side of some pages, some additional notes may be seen. These points have been mentioned at the lectures and cannot be found anywhere.
Having these answers in your hands does not mean that attendance to the lectures of the corresponding topics discussed here, will be a waste of time for the simple reasons that by attending the lectures you will get a much better grasp on the material and also because the field of Molecular Biology is a very rapidly expanding one and new information can be drawn only by the lectures. Enjoy your studies.

Preface to the third edition.
In this third edition of this booklet, only the pictures are added - separately. The Chinese say that one picture is worth as 1000 words; we don't know about the validity of the ratio, but we do know that these 138 added pictures will contribute to your understanding concerning the field of molecular biology.

Preface to the second edition.
This second edition is just a small expansion of the answered material of the "Molecular Biology" section of the questions for the Biochemistry final exam, by simply adding the answer for question No..18 which is "the molecular genetics of cancer".
Although only 7 pages were added, a great effort was needed to collect the information included and organize then in a way which would answer the 18th question, because it deals with probably the hottest and most diachronic point of the history of medicine. Despite of the extensive research performed during the recent years, it seems that many more information are needed in order to understand and prevent or treat cancer. I hope that some of the readers of these lines will become contributors in the elucidation of the molecular genetics of cancer.

Christos Chinopoulos

1. The mechanism of DNA replication in prokaryotic cells.

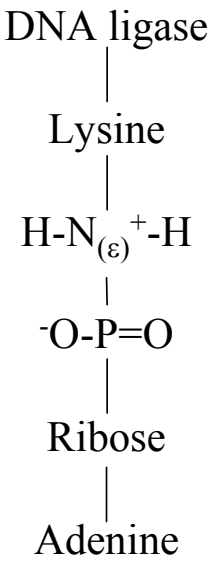
1/a. The properties of DNA polymerases (I, II, III) and the reactions catalyzed by these enzymes.

- DNA polymerase I catalyses the step-by-step addition of deoxyribonucleotide units to the 3' end of a DNA chain:
$(DNA)_n$ + dNTP $\Leftrightarrow$ $(DNA)_{n+1}$ + PPi (equation 1) n = {A,G,T,C}
PPi $\longrightarrow$ inorganic 2Pi (equation 2) reaction catalysed by pyrophosphatase.
DNA polymerase I requires all 4 deoxyribonucleotides - dATP, dGTP, dCTP, dTTP- and Mg^{2+} (for the hydrolysis and release of pyrophosphate) to synthesize DNA. The enzyme adds deoxyribonucleotides to the free 3'-OH of the chain undergoing elongation, which proceeds in the 5' $\rightarrow$ 3' direction.
-A primer chain with a free 3'-OH group is needed at the start. A DNA template containing a single- stranded region is also essential, because the template is the one that will determine which deoxyribonucleotide should be the appropriate building block for the newly synthesized DNA molecule, and not the polymerase. The polymerase catalyses the nucleophilic attack of the 3'-OH terminus of the primer on the innermost phosphorus atom of a dNTP. Pyrophosphate (PPi) is released and its subsequent hydrolysis to 2 Pi molecules (equation 2) pulls the reaction to the right.
- DNA polymerase 1 is a moderately processive enzyme (in comparison to the other two) - it catalyses multiple polymerisation steps (~20) before dissociating from the template DNA.
- DNA polymerase I can catalyse the hydrolysis of DNA chains as well as their polymerisation (sec double arrow in equation (1)). The enzyme catalyses the hydrolysis of nucleotides at the 3'-end of DNA chains. Thus, DNA polymerase I is a 3'$\rightarrow$5' exonuclease (editing function in polymerisation). To be removed, a nucleotide must fulfil two conditions: a) it must have a free 3'-OH terminus and b) it must not be part of a double helix in general; DNA polymerase I removes mismatched residues at the primer terminus before proceeding with polymerisation.
- The 3'$\rightarrow$5' exonuclease activity of the DNA polymerase I markedly enhances the accuracy of DNA replication by serving as a second test of the correctness of base pairing (but it also slows down the replication speed). In effect, DNA polymerase I examines the result of each polymerisation it catalyses, before proceeding to the next (proofreading).
- DNA polymerase I can also hydrolyse DNA starting from the 5' end of a chain. This "error-correcting" 5'$\rightarrow$3' exonuclease activity is different from the 3'$\rightarrow$5' exonuclease activity in four ways:
i) the cleaved bond must be in a double-helix.
ii) cleavage can occur at the terminal phosphodiester bond or at a bond several residues away from the 5' terminus (which can bear a free-OH group or be phosphorylated).
iii) 5'-3' exonuclease activity is enhanced by concomitant DNA synthesis.
iv) the active site for exonuclease action is clearly separate from the active site for polymerisation and the 3'$\rightarrow$ 5' hydrolysis (DNA polymerase I contains 3 different active sites on a single polypeptide chain).
- 5'$\rightarrow$3' exonuclease activity plays a key role in DNA replication by removing RNA primer (vide infra). Moreover, the 5'$\rightarrow$3' exonuclease complements the 3'$\rightarrow$5' exonuclease activity by correcting errors of a different type. For example, the 5'$\rightarrow$3' exonuclease participates in the excision of pyrimidine dimers formed by exposure of DNA to ultraviolet light.
- DNA polymerase I contains a deep cleft for binding double-helical DNA, and the moderate processivity of it may be due to the closure of this cleft, which would slow the dissociation of DNA from the enzyme during polymerisation.
- DNA polymerases II and III are like polymerase I in the following respects:
i) They catalyse a template- directed synthesis of DNA from deoxyribonucleotide precursors.
ii) A primer with a free 3'-OH group is required.
iii) Synthesis is in the 5'$\rightarrow$3' direction.
iv) They possess a 3'$\rightarrow$5' exonuclease activity.
- The role of these polymerases in vivo are the following: A multisubunit assembly containing polymerase III synthesizes most of the new DNA, whereas polymerase I erases the primer and fills gaps. The role of polymerase II is not yet known.
- The multisubunit assembly of DNA polymerase III holoenzyme is characterized by very high processivity, catalytic potency, and fidelity. The holoenzyme catalyses the formation of many thousands of phosphodiester bonds before releasing its template, compared with only 20 for DNA polymerase I. The much lower degree of processivity of DNA polymerase I is well suited to its role as a gap filler (vide infra) and repair enzyme. DNA polymerase III holoenzyme, on the other hand, is designed to grasp its template and not let go until it has been completely replicated. A second distinctive feature of the holoenzyme is its catalytic prowess, 1000 nucleotides added per second compared with only 10 by DNA polymerase I. This accelerated catalysis is not accomplished at the cost of accuracy. The greater catalytic prowess of polymerase III is largely due to its processivity; no time is lost in repeatedly coming on and off the template. These striking features of DNA polymerase III do not come cheaply. Indeed, many of the steps in DNA replication carried out by this enzyme, require the expenditure of ~ P to sequentially bind and release reaction partners and to move the apparatus of replication along the template strand. (For further description of DNA polymerase III see 1/c-1/d).

1/b. The reaction catalysed by DNA ligase, please mention several processes where this enzyme has an important role.

- DNA ligase joins ends of DNA in duplex regions by catalysing the formation of a phosphodiester bond between the 3'-OH group of the end of one DNA chain and the 5'-phosphate group at the end of the other. An energy source is required to drive this endergonic reaction. In E. coli and other bacteria, NAD^+ serves this role; in animal cells and bacteriophage, ATP is the energy source. The joining process carried out by DNA ligase is essential for the normal synthesis of DNA (see 1/c-1/d), the repair of

damaged DNA (see question 3), and the splicing of DNA chains in genetic recombination (see 6/c-6/d).
- DNA ligase cannot link 2 molecules of single-stranded DNA. Rather, the DNA chains joined by DNA ligase must belong to a double-helical molecule (or molecules).
- Mechanism of joining: ATP (or NAD^+) reacts with DNA ligase to form a covalent enzyme-AMP complex (enzyme-adenylate complex) in which AMP is linked to the ε-amino group of a lysine residue of the enzyme through a phosphoamide bond :



- Pyrophosphate (or nicotinamide mononucleotide NMN) is concomitantly released. The activated AMP moiety is then transferred from the lysine residue to the phosphate group at the 5' terminus of a DNA chain, forming a DNA-adenylate complex. The final step is a nucleophilic attack by the 3'-OH group on this activated 5'-phosphorus atom. This sequence of reactions is driven by the hydrolysis of pyrophosphate that was released in the formation of the enzyme-adenylate complex. Thus, 2 ~ P are spent in constructing a phosphodiester bridge in the DNA backbone if ATP is the energy source and PPi is hydrolysed:

E + ATP (or NAD^+) ⇔ E-AMP + PPi→ 2Pi (or NMN)

E-AMP + P-5'-DNA ⇔ E + AMP-P-5'-DNA

DNA-3'-OH + AMP-P-5'-DNA ⇔ DNA-3'-O-P-5';DNA + AMP.

DNA-3'-OH+P-5'-DNA+ATP (or NAD^+) ⇔ DNA-3'-O-P-5'-DNA+AMP + PPi (or NMN)

1/c. The initiation of DNA replication in procaryotes, the replication fork, please characterize the leading and lagging strand.

- The form of a replicating DNA molecule of an E. coli is a closed circle with an internal loop (fig. 27-35). It resembles a θ, hence the name theta structure. This structure shows that the DNA molecule maintains its circular form while it is being replicated. In a theta structure it is sure that long stretches of single-stranded DNA are absent. Thus, these pictures rule out a replication mechanism in which the parental DNA strands unwind completely before serving as templates for the synthesis of new DNA. Rather, the synthesis of new DNA is closely coupled to the unwinding of parental DNA. A site of simultaneous unwinding and synthesis is called a replication fork.
- Replication starts at a unique site. (In case of E. coli this is the ori (locus)
- Replication is bi-directional, meaning that it proceeds simultaneously in opposite directions, at about the same velocity. In other words, there are two replication forks: one moves clockwise, the other counter clockwise. The two replication forks meet at tre, the termination region, which is opposite the origin of replication.
- Replication begins with the unwinding of the ori C (origin) site; the binding of the dnA protein to 4 sites on oriC initiates an intricate sequence steps leading to the unwinding of the template DNA and the synthesis of a primer. The DNA must be negatively supercoiled to enable dnaA to bind. A complex of the dnaB and dnaC proteins joins dnA to bend and open the double helix. dnaB protein is a helicase: it catalyses the ATP-driven unwinding of double- helical DNA. The unwound portion of DNA is then stabilized by single-strand-binding protein (SSB), which binds to single-stranded DNA (in a cooperative manner since the protein is a tetramer). The unwinding of DNA at the origin would lead to the positive supercoiling of the circular DNA if the linking number remained constant, and unwinding would soon stop. This is prevented by the compensatory action of DNA gyrase, which introduces negative supercoils as it hydrolyses ATP.
- The DNA template is now exposed, but new DNA cannot be synthesized until a primer is constructed; a specific RNA polymerase called primase joins the prepriming complex in a multisubunit assembly called the primosome. Primase synthesizes a short stretch of RNA that is complementary to one of the template DNA strands (fig. 27-41). This primer RNA is removed at the end of replication by the 5'→3' exonuclease activity of DNA polymerase I.
- One strand of DNA is made in fragments and the other strand is synthesized continuously. In the former strand, as replication proceeds, these fragments (Okazaki fragments) become covalently joined by DNA ligase to form one of the daughter strands (fig. 27-39). The strand formed from Okazaki fragments is termed the lagging strand, whereas the one synthesized without interruption is the leading strand. Both the Okazaki fragments and the leading strand are synthesized in the 5'→ 3' direction. The discontinuous assembly of the lagging strand enables 5'→ 3' polymerisation at the nucleotide level to give rise to overall growth in the 3'→5' direction.

1/d. Please describe the nascent Okazaki fragment, what kind of reactions (and enzymes) are necessary to form the final DNA strand from Okazaki fragments.

- The mode of the synthesis of the lagging strand is complex. This strand is synthesized in fragments (~ 1000 nucleotides long), so that 5'→3' polymerisation leads to overall growth in the 3'→ 5' direction. This may be accomplished by a looping of the template for the lagging strand (fig. 27-43). The lagging strand template would then pass through the polymerase site in one subunit of a dimeric polymerase III in the same direction as the leading strand template in the other subunit. DNA polymerase III would have to let go of the lagging strand template after about 1000 nucleotides have been added to the lagging

strand. A new loop would then be formed, and primase would again synthesize a short stretch of RNA primer to initiate the formation of another Okazaki fragment. The gaps between fragments of the nascent lagging strand are filled by DNA polymerase I, which also uses its 5'→ 3' exonuclease activity to remove the ribonucleotide primer lying ahead of the polymerase site, (fig. 27-44). Finally, DNA ligase joins the fragments.

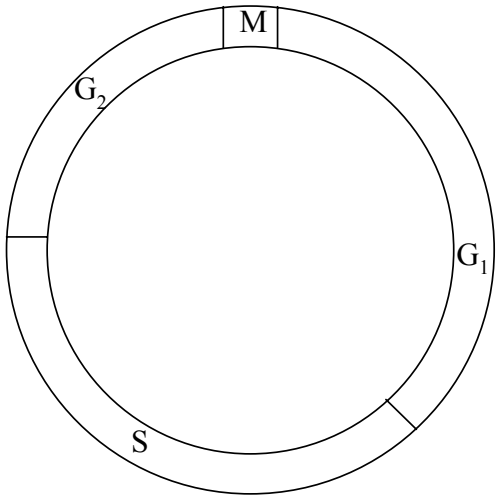
1/e. The proteins participating in the replication of the prokaryotic genome, and their roles.

- dnaB protein: Begins unwinding the double helix.
- Primase: Synthesizes RNA primers.
- rep protein: Unwinds the double helix.
- SSB: stabilizes single-stranded regions.
- DNA gyrase: Introduces negative supercoils.
- DNA polymerase III holoenzyme: Synthesizes DNA.
- DNA polymerase I: Erases primer and fills gaps.
- DNA ligase: Joins the ends of DNA.

2. The mechanism of the replication of DNA in eucaryotic cells.

2/a. The properties of the DNA polymerases present in eucaryotes. The time of DNA replication in the cell cycle.

- Eucaryotic cells contain several types of DNA polymerases. The α polymerase (found in the nucleus) plays a major role in chromosome replication whereas the b polymerase (found also in the nucleus) participates in the repair of DNA. The amount of the α polymerase increases more than 10- fold when quiescent cells begin to divide rapidly. The α polymerase is a multisubunit complex, whereas the b enzyme is a single polypeptide chain. The γ polymerase (found in mitochondria) is responsible for the replication of mitochondrial DNA. These enzymes, like procaryotic DNA polymerases, use dNTPs as activated intermediates and carry out template-directed elongation of a primer in the 5'→ 3'direction. The α, b and γ polymerases lack nuclease activities. It seems likely that proofreading is performed by nucleases that are associated with these polymerases. A recently discovered eucaryotic δ polymerase possesses a 3' exonuclease activity with editing capabilities akin to those of procaryotic polymerases. α polymerases can be specifically inhibited by aphidocolin.
- The time of DNA replication in a cell cycle varies, depending on the cell type and conditions of growth. In general, it is on the order of minutes or hours. The much greater length replication time of eucaryotes with that compared of the procaryotes, is not due to the greater amount of DNA to be replicated (eucaryotic genome presents multiple replication origins, vide infra), but it is due to the enormously advanced proofreading and repair mechanisms of the eucaryotes, which are time-consuming. Mitosis is usually the shortest phase.



2/b. The replication wobbles (the movement of the replication forks: leading strand and lagging strand). The distribution of old and new histones.

- Eucaryotic DNA, like all other DNA molecules, is replicated semiconservatively; moreover it is replicated bidirectionally from many origins. The use of many initiation points is necessary for rapid replication because of the great length of eucaryotic DNAs. A DNA molecule from the cleavage nuclei of Drosophila exhibits a serial array of replicated regions, or "eye form" (fig. 33-14). The activation of each initiation point generates 2 diverging replication forms. The eye forms expanding in both directions merge to form the 2 daughter DNA molecules, (fig. 33-15). An eye form within an eye form has not been observed, indicating that an origin cannot be reactivated until after the entire DNA molecule is replicated.
- A yeast chromosome contains ~ 400 initiation sites for DNA replication. These origins of replication share a consensus sequence called ars for autonomous replication sequence. Insertion of an ars sequence into a bacterial plasmid enables it to replicate autonomously in yeast cells.
- The ends of chromosomes (telomeres) are replicated by adding blocks of preformed oligonucleotides: The linearity of eucaryotic chromosomes poses a problem not encountered with circular DNA molecules such as those of procaryotes. Eucaryotic DNA polymerases, like procaryotic ones, are unable to synthesize in the 3'→5' direction or start chains de novo, and so erasure of the RNA primers leaves the 5' ends of nascent daughter DNA strands incomplete, fig. 33-17. How are the 5' ends finished? First, preformed oligonucleotide blocks are added to the overhanging 3' ends of newly synthesized daughter strands. In Tetrahymena, a few copies of a specific base sequence are added to each 3'-OH end. These extensions then loop around so that their 3'-OH serves as a primer for the synthesis of complementary sequences. DNA ligase joins each of these lengthened 3' extensions to the 5' ends, and a nuclease cleaves the unpaired loops to fashion a linear duplex with flush ends. Thus, Tetrahymena completes the replication of its linear

chromosome by transiently joining the 2 strands at each end of the double helix.
- Parental histones are associated with one of the daughter DNA duplexes, the other being bare because of the absence of new histones. This interpretation is directly supported by EM photographs showing that one of the daughter duplexes at a replication fork is beaded, whereas the other is naked, fig. 33-18. In other words, parental histones segregate conservatively during replication. This arrangement indicates that histones do not dissociate from DNA during replication. In fact, old histones stay with the DNA duplex containing the leading strand, whereas new histones assemble on the DNA duplex containing the lagging strand. A likely reason for this difference between the daughter DNA molecules is that histones bind much more strongly to double-stranded than to single-stranded DNA. Old histones probably do not follow the lagging duplex because it contains single-stranded regions prior to the joining of its Okazaki fragments.

2/c. The organization of the eucaryotic genome (nucleosome structure, packaged nucleosomes, extended chromatin, condensed chromatin, chromatids in the metaphase chromosome). The number of DNA molecules in G2 phase of the human cell.

- Nucleosomes are the repeating units of chromatin. Most of the DNA is wound around the outside of a core of histones. The remainder of the DNA, called the linker, joins adjacent nucleosomes and contributes to the flexibility of the chromatin fiber, fig. 33-9. Thus, a chromatin fiber is a flexibly jointed chains of nucleosomes, rather like beads on a string. 140 base pairs of DNA are wound on the outside of the core to form $1^{3/4}$ turns of a left-handed supercoil. A dimer of histones H3 and H4 occupy the center of the nucleosome, and that one each of H2A and H2B are at either end, fig. 33-10. Note that H1 is not part of the core.
- DNA histone contacts are made on nearly every turn of the DNA double helix and are confined to the inner surface of the superhelix. The histones do not embrace the DNA nor do they protrude between the turns of the superhelix. The most substantial contacts are made between the H3-H4 tetramer and the central part of the DNA double helix. α-helical rods projecting from the H3 dimer fit snugly into the minor grooves of DNA, on either site of the 2-fold symmetry axis of the nucleosome, fig. 33-11. The histones can interact with most DNA sequence, in keeping with their role as a device for packaging DNA.
- The H2A- H2B dimers are attached to each exposed end of the H3-H4 tetramer. Their binding to the last half turns of the superhelix further stabilizes the nucleosome core. The disassembly of a nucleosome, as in DNA replication, is probably initiated by the dissociation or pulling away of an H2A-H2B dimer.
- Histone H1 plays a key role at the next of chromosome structure by serving as a bridge between different nucleosomes. H1 is located on the outside of the nucleosome, near the linker DNA, where it interacts with H2A subunits of the core. It is released from nucleosomes when the DNA per particle is trimmed from 160 to 140 base pairs. H1 differs also in its stoichiometry of one per nucleosome compared with 2 for the other histones. It is noteworthy that several types of H1 have been found, in contrast with the constancy of the other histones. Moreover, H1 is phosphorylated just before mitosis, and is dephosphorylated following mitosis, suggesting that this covalent modification regulates its capacity to make DNA compact.
- The packing ratio (degree of condensation) of the nucleosome is about 7. However, the packing ratio of DNA in a metaphase chromosome is about 10^4 . DNA in interphase nuclei, where the chromatin is more dispersed, has a packing ratio of about 10^2 to 10^3. Clearly, the nucleosome is just the first step in the compaction of DNA. In the next level of organization of DNA, the nucleosomes themselves form a helical array, fig. 33-12. Metaphase chromosomes stripped of histones display a central protein scaffold that is surrounded by many very long loops of DNA, fig. 33-13. These DNA loops are probably bound to the scaffold. The presence of topoisomerase II in the scaffold makes it likely that changes in supercoiling are important in altering the architecture and accessibility of large segments of DNA in mitosis and meiosis. The most compacted DNA is found in sperm heads, in which histones are replaced by protamines, a series of Arginine-rich proteins that become highly α-helical on binding to DNA. The α helices of protamine probably lie in the major grooves of DNA, where they neutralize the negatively charged phosphate backbone and so enable DNA duplexes to pack tightly together.
- Extended chromatin (or euchromatin) is the term applied for the functional form of DNA, i.e. for the DNA strand capable of transmitting the encoded information. Condensed chromatin (or heterochromatin) has 2 forms, constitutive and facultative. The constitutive heterochromatin is inhibited for the entire life span of the cell, whereas the facultative heterochromatin is inhibited only for temporary periods and can therefore participate in the common life functions of the cell. Morphologically, the constitutive and facultative forms of heterochromatin differ in the degree of condensation. The quantities of hetero-and euchromatin vary with the stages of the cell cycle. Euchromatin predominates in the active stages and heterochromatin in the resting stages of cell life. The quantitative relations of hetero-and euchromatin also change during cell development; certain cells ultimately become entirely heterochromatic. In general, euchromatin is undermethylated and hypersensitive to DNase I in contrast with heterochromatin.
- The number of DNA molecules in G2 phase of the human cell is doubled.

3. The repair of DNA lesions.

3/a. The most frequent lesions of DNA and the agents resulting these lesions. The significance of the presence of thymine in the DNA instead of uracil.

- Bases in the DNA can be altered or lost, phosphodiester bonds in the backbone can be broken, and strands can become covalently cross-linked. These lesions are produced by ionizing radiation, UV light, and a variety of chemicals (e.g. base analogues: 5-bromouracil and 2-aminopurine, nitrous acid, hydroxylamine NH_2OH, acridines).

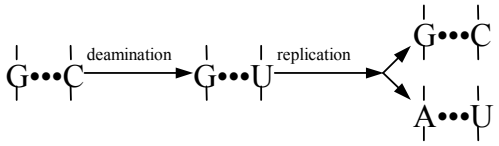
- The most frequent DNA lesions are the following:
I: Single-base alteration.
A) Depurination.
B) Deamination of C to V.
C) Deamination of A to hypoxanthine.
D) Alkylation of a base.
E) Insertion or deletion of nucleotide.
F) Base-analogue incorporation.

II: 2-base alteration.
A) UV-light induced T-T dimer.
B) Bifunctional alkylating agent cross-linkage.

III: Chain breaks.
A) Ionising radiation.
B) Radioactive disintegration of backbone element.

IV: Cross linkage.
A) Between bases. In the same or opposite strands.
B) Between DNA and protein molecules (e.g. histones).

- DNA contains T instead of U to permit repair of deaminated C: The deamination of C is potentially mutagenic because U pairs with A, and so one of the daughter strands will contain an AU base pair rather than the original GC base pair:



This mutation is prevented by a repair system that recognizes U to be foreign to DNA, fig. 27-52. First, a Uracil-DNA glycosidase hydrolyses the glycosidic bond between the U and deoxyribose moieties. At this stage, the DNA backbone is intact, but a base is missing. This hole is called AP site because it is apurinic (devoid of A or G) or apyrimidinic (devoid of C or T). An AP endonuclease then recognizes this defect and nicks the backbone adjacent to the missing base. DNA polymerase I excises the residual deoxyribose phosphate unit and inserts cytosine, as dictated by the presence of G on the undamaged complementary strand. Finally, the repaired strand is sealed by DNA ligase.
- The only difference between T and U is a $-CH_3$ in T in place of the C-5 hydrogen in U. Why is a methylated base employed in DNA and not in RNA? Recall that the methylation of deoxyuridylate to form deoxythymidilate is energetically expensive. The discovery of an active repair system to correct the deamination of C provides a convincing solution to this puzzle. Uracil-DNA glycosidase does not remove T from DNA. Thus, the methyl group on T is a tag that distinguishes it from deaminated C. If this tag were absent, U correctly in place would be indistinguishable from U formed by deamination. The defect would persist unnoticed, and so a GC base pair would necessarily be mutated to AU in one of the daughter DNA molecules. This mutation is prevented by a repair system that searches U in DNA and leaves T alone. It seems likely that T is used instead of U to enhance the fidelity of the genetic message in contrast, RNA is not repaired, and so U is used in RNA because it is a less expensive building block.

3/b. The enzymes and reactions in the repair of depurination.

- Depurination: Loss of A or G from their deoxyriboses by thermal disruption of their N-glycosyl linkages.
In case of depurination, on DNA there will a deoxuribose sugar with a missing base,(fig. 5-31 MBC). This sugar is rapidly recognized by the enzyme AP endonuclease, which cuts the DNA phosphodiester backbone at the altered site. By the action of this enzyme (which belongs to the family of DNA repair nucleases), a gap in the DNA helix will be formed. Another enzyme, DNA polymerase, binds to the 3'-OH and of the cut DNA strand and fills in the gap one nucleotide at a time by copying the information stored in the "good" (template) strand. The break (nick) in the damaged strand left when the DNA polymerase has filled in the gap is sealed by DNA ligase, which completes the restoration process, (fig. 5-33 MBC).

3/c. The enzymes and reactions in the repair of deamination.

- Deamination of C gives U, deamination of A gives hypoxanthine, deamination of G gives xanthine; T cannot be deaminated. In case of C deamination: first, the enzyme Uracil-DNA glycosylase removes the altered base (uracil see 3/a.), producing a deoxyribose sugar with a missing base. Because this sugar is the same substrate recognized by the AP endonuclease, the subsequent steps in the repair process proceed in the same way as for depurinated sites. The net result is that the U that was created by accidental deamination is restored to a C.

3/d. The formation of thymine dimers and the enzymes and reactions in the repair of this lesion.
- T-T dimers are formed on exposure of DNA to UV light. Adjacent T residues on a DNA strand can become covalently linked under these conditions:
Three enzymes are needed for this repair in E. coli, fig. 27-51: First, an enzyme complex uvrABC detects the distortion caused by the T-T dimer and cuts it. The removed oligonucleotide diffuses away by the action of this highly specific exonuclease. DNA polymerase I enters the gap to carry out repair synthesis. The 3' end of the nicked strand is the primer, and the intact complementary strand is the template. Finally, the 3' and of the newly synthesized stretch of DNA and the original portion of the DNA chain are joined by DNA ligase.
- Alternatively, the T-T dimer can be photochemically split. Nearly all cells contain a photoreactivating enzyme called DNA photolyase. The absorption of photon creates an excited state that cleaves the dimer into its original bases.
- See 3/e for one more ways of repairing T-T dimers.

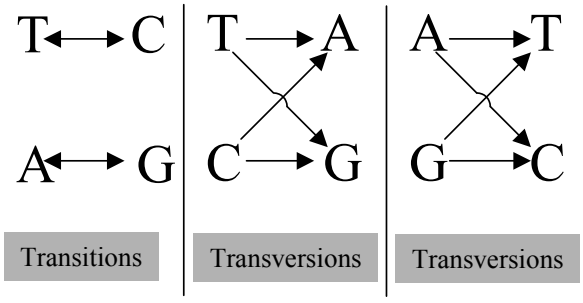
3/e. The possibility of the correction of DNA lesions of both strands.

- Cells have a separate "bulky lesion" repair pathway capable of removing almost any type of damage that creates a large change in the DNA double helix. Such bulky lesions include those created by the covalent reaction of DNA bases with large hydrocarbons (such as the carcinogen benzpyrene) as well as various pyrimidine dimers (T-T, T-C, C-C) caused by UV light, fig. 5-32 MBC. In these cases a large multienzyme complex scans the DNA for a distortion in the double helix rather than for a specific base change. Once a bulky lesion is found, the phosphodiester backbone of the abnormal strand is cleaved on both sides of the distortion and the entire lesion is excised. The small gap produced in the DNA helix is then repaired as described.
- The aforementioned method might correct a DNA lesion in both strands only when the following condition is fulfilled; in the case where both copies are damaged simultaneously, a second DNA helix of the same sequence must be present in the cell, and the repair mechanism involves genetic recombination events.
- Note that the bulky lesion repair pathway is only a subtype of DNA repair mechanisms involving lesions in both DNA strands; only the second remark of 3/e. can be regarded as the answer.

4. Point mutation of DNA.

4/a. Please characterize the point mutations (substitutions and frame-shift mutations). What is the "nonsense-mutation"?

- Substitutions (point) mutations can be either a transition or a transversion. In the former, a given pyrimidine is changed to the other pyrimidine or a given purine is changed to the other purine. Transversion are changes from a purine to either of the 2 pyrimidines or the change of a pyrimidine into either of the 2 purines:



- Substitutions can have no effect (wobble), or be missence (acceptable, partially acceptable e.g. hemoglobin S, or unnacceptable eg. $HbFe^{+3}$ → lethal), or be nonsense (premature termination).
- Frame shift mutations result from deletion or insertion of nucleotides in DNA that generates altered mRNAs; not only is the sequence of amino acids distal to the insertion or deletion garbled, but reading of the message can also result in the appearance of a nonsense codon and thus the production of a polypeptide both garbled and prematurely terminated near its carboxyl terminus.
- A nonsense mutation can arise from a substitution or frame-shift mutation, where a nonsense codon (stop) may appear that would then result in the premature termination of amino acid incorporation into a peptide chain and the production of only a fragment of the intended protein molecule. The probability is high that a prematurely terminated protein molecule or peptide fragment would not function in this assigned role.

4/b. Please describe the mechanism of spontaneous mutation (for example the transition from T-A to C-G).

Some of the H atoms on each of the 4 bases can change their location to produce a tautomer. An amino group (-NH_2) can tautomerize to an imino form (=NH). Likewise, a keto group (-C=O) can tautomerize to an enol form (=C-OH); (possibility ~10^{-4}). These transient tautomers can form non- standard base pairs that fit into a double helix. For example, the imino tautomer of adenine can pair with cytosine:
This A*-C pairing (the asterisk denotes the imino tautomer) would allow C to become incorporated into a growing DNA strand where T was expected, so that it would lead to a mutation if left uncorrected. In the next round replication, this adenine will probably reautomerize to the standard form, which pairs as usual with T but the C will pair with G. Hence, one of the daughter DNA molecules will contain a GC base pair in place of the normal AT base pair.

4/c. What is a suppressor mutation? What is the principle of Ames-test?

- A mutation in a different gene might overcome the deleterious effect of the first mutation. Such a mutation is called an intergenic suppressor. Most of these suppressor act by altering the reading of mRNA; i.e. consider a mutation that leads to the premature appearance of the codon UAG, a termination signal, producing an incomplete polypeptide chain. Such mutations are called nonsense mutations because incomplete polypeptides are usually inactive. The UAG nonsense mutation can be suppressed by mutations in several different genes. One of these suppressors causes UAG to be read as a codon for Tyrosine, which may read to the synthesis of a functional protein rather than to a incomplete polypeptide chain, fig. 30-15. Why did this mutant tRNA insert Tyr in response to the UAG codon? Tyr tRNA normally recognizes the codons UAC and UAU. This mutant tRNA proved to be identical with the normal Tyr tRNA expect for a single base change in this anticodon: GUA→ CUA. The mutation G→ C in the first base of its anticodon changed its recognition properties. The altered anticodon recognizes only UAG, as predicted by the wobble hypothesis. This kind of suppressor mutation is more likely to be selected if the tRNA undergoing mutation is not essential in other words, there must be another tRNA species that recognizes the same codons as does the one undergoing mutation. In fact, E. coli contains 2 different tRNAs that normally recognize both UAC and UAU. The minor species is the one that is altered in a suppressor mutation. The occurrence of this suppressor mutation raises a second question. If the mutation to UAG is read as Tyr rather than as a stop signal, what happens to normal chain termination? Surprisingly, most polypeptide chains terminate normally in the suppressor mutant, possibly because the termination signal

may be more than just the trinucleotide UAG. Indeed, some messages are known to end with 2 different stop codons. Other kinds of mutant tRNAs have been identified. Missense suppressors alter the reading of mRNA so that one amin acid is inserted instead of another at certain codons. Frameshift suppressors are especially interesting. One of them has an extra base in its anticodon loop. Consequently, it reads 4 nucleotides as a codon rather than 3. In particular, UUUC rather than UUU is read as the codon for Phenylalanine. A mutation caused by the insertion of an extra base can be suppressed by this altered tRNA.

- Bruce Ames has devised a simple and sensitive test for detecting chemical mutagens. A thin layer of agar containing about 10^9 bacteria of a specially constructed tester strain of Salmonella is placed on a Petri dish. These bacteria are unable to grow in the absence of Histidine because of a mutation in one of their genes for the biosynthesis of this amino acid. The addition of a mutagen to the center of a plate results in many new mutations. A small proportion of them reverse the original mutation so that His is synthesized. These revertants multiply in the absence of exogenous His and appear as discrete colonies after the plate is incubated at 37^oC for 2 days, fig. 27-53; i.e. 0,5μg of 2-N-anthrakene gives 11000 revertant colonies compared with only 30 spontaneous revertants in its absence. A series of concentrations of a chemical can readily be tested to generate a dose-response curve. These curves are usually linear which suggests that there is no threshold concentration for mutagenesis. Some of the tester strains are responsive to base-pair substitutions, whereas others detect deletions or additions of base pairs (frameshifts). The sensitivity of these specially designed strains has been enhanced by genetically deleting their excision - repair systems. Also, the lipopolysaccharide barrier that normally coats the surface of Salmonella is incomplete in the tester strains, which facilitates the entry of potential mutagens. Another important feature of this detection system is a mammalian liver homogenate. Recall that some potential carcinogens are converted into their active forms by enzyme systems in the liver (or other tissues). Bacteria lack these enzymes, and so the test plate requires a few milligrams of a liver homogenate to activate this group of mutagens; i.e. an activated compound with both a reactive side chain (enabling to form a covalent bond with DNA) and an aromatic region (allowing it to intercalate) is far more mutagenic than one with just the aromatic region.

5. Transcription of DNA in procaryotic cells.

5/a. Fundamental differences between the transcription of procaryotic cells and the transcription of eucaryotes.

- First, eucaryotic DNA templates present enhancer sequences that can stimulate initiation more than a 1000 base pairs away. Second, primary transcripts in eucaryotes are extensively modified, as exemplified by the capping of the 5' end of an mRNA precursor and the addition of a long poly A tail to its 3' end. Third, eucaryotic mRNA precursors are spliced by the catalytic action of spliceosomes consisting of ribonucleoproteins (snRNPs); other RNA molecules splice themselves (RNA catalytic activity). Fourth, eucaryotic transcription is separated in space and time from translation, due to the existence of the nucleus.

5/b. The properties of the DNA directed RNA polymerase of E. coli.

- RNA polymerase from E. coli consists of 4 kinds of subunits. The subunit composition of the entire enzyme, called the holoenzyme, is $a_2bb'\sigma$. The σ subunit finds a promoter site where transcription begins; it initiates RNA synthesis and then dissociates from the rest of the enzyme (a_2bb'); it contains the catalytic site. The b' subunit binds the DNA template, and the b subunit binds ribonucleotide substrates. RNA polymerase performs multiple functions in all three processes: initiation, elongation and termination:

i) it searches DNA for initiation sites,

ii) it unwinds a short stretch of double- helical DNA to produce a single- stranded DNA template from which it takes instructions,

iii) it selects the correct ribonucleotide and catalyzes the formation of a phosphodiester bond. This process is repeated many times as the enzyme moves undirectionally along the DNA template. RNA polymerase is totally processive - a transcript is synthesized from start to end by a single RNA polymerase molecule,

iv) it detects termination signals that specify where a transcript ends,

v) it interacts with activator and repressor proteins that modulate the rate of transcription over a wide dynamic range.

5/c. The promoter. The inhibition of transcription in procaryotes. Strong and weak promoters.

- Transcription is initiated at promoter sites on the DNA template. 2 common motives are present on the 5' side (upstream) of the start site. They are known as the -10 sequence (Pribnow box) and the -35 sequence because they are centered at about 10 and 35 nucleotides upstream of the start site. Their consensus sequences are:

```
          -35          -10          +1
5' ~~~ TTGACA ~~~~~~TATAAT ~~~~ Start site
```

- Promoters differ markedly in their efficacy. Strong promoters cause frequent initiation of transcription, as often as every 2 sec in E. coli. In contrast, genes with very weak promoters are transcribed about once in 10 min. The -10 and -35 regions of most strong promoters have sequences that correspond closely to the consensus ones, where as weak promoters tend to have multiple substitutions at these sites. The distance between these conserved sequences is also important; a separation of 17 nucleotides is optimal. The frequency of transcription of many genes is also markedly influenced by regulatory proteins that bind to specific sequences near promoter sites and interact with RNA polymerase.

- Sigma (σ) enables RNA polymerase to recognize promoters. Promoters for heat-shock genes are recognized by a special kind of sigma subunit. Negative supercoiling of circular DNA templates favors the transcription of many genes because it facilitates unwinding. Thus, the pumping

the negative supercoils into DNA by DNA gyrase can increase the efficiency of promoters located at distant sites. However, not all promoters are stimulated by negative supercoiling. The promoter for DNA gyrase is a noteworthy exception: the rate of transcription of this gene is decreased by negative supercoiling, a nice feedback control assuring that DNA does not become excessively supercoiled. The transition from the closed promoter complex (in which DNA is double helical) to the open promoter complex (in which a DNA segment is unwound) is a key event in transcription, setting the stage for the formation of the first phosphodiester bond of the new RNA chain.

- Rifamycin and rifampicin specifically inhibit the initiation of RNA synthesis. They act on the b subunit of RNA polymerase interfering with the formation of the first phosphodiester bond in the RNA chain. Actinomycin D binds tightly and specifically to double-helical DNA and thereby prevents it from being an effective template for RNA synthesis. Its action is enhanced by the presence of G residues. Actinomycin D binds intercalatively to both procaryotic and eucaryotic DNA. The symmetry of actinomycin D matches the symmetry of a specific base sequence in DNA.

5/d. The elongation and termination of transcription in procaryotes.

- Elongation takes place at transcription bubbles that move along the DNA template in the 5'→ 3' direction. In a transcription bubble there are DNA, nascent RNA, and the core RNA polymerase (the σ subunit was dissociated from the holoenzyme after the formation of the first phosphodister bond, which is a necessary condition for the RNA polymerase to bind strongly to the DNA template). RNA polymerase stays bound to its template until a termination signal is reached.
- The length of the RNA-DNA hybrid and of the unwound region of DNA stay constant as RNA polymerase moves along the DNA template. This finding indicates that DNA is rewound at the same rate at the near of RNA polymerase, as it is unwound at the front of the enzyme, fig. 29-11. The RNA- DNA hybrid must also rotate each time a nucleotide is added so that the 3'-OH and of RNA stays at the catalytic site.
- Note that RNA polymerase lacks nuclease activity. In contrast with DNA polymerase, it does not edit the nascent polynucleotide chain. Consequently, the fidelity of transcription is much lower than that of replication.
- In the termination phase of transcription, the formation of phosphodiester bonds ceases, the RNA- DNA hybrid dissociates, the melted region of DNA rewinds, and RNA polymerase releases the DNA. The termination of transcription is as precisely controlled as its initiation. The transcribed regions of DNA templates contain stop signals. The simplest one is a palindromic GC- rich region followed by an A-T rich region, fig. 29-12. The RNA transcript of this DNA palindrome is self-complementary. Hence, its bases can pair to form a hairpin structure with a stem and loop, a structure favoured by its high content of GC residues, fig. 29-13. Recall that GC base pairs are move stable than A-T pairs (Stryer, pg. 82). This stable hairpin is followed by a sequence of 4 or more U residues. The RNA transcript ends within or just after them. How does its hairpin oligo-U structure terminate transcription? First, RNA polymerase pauses when it encounters such a hairpin. Furthermore, the RNA-DNA hybrid helix produced after the hairpin is unstable because of its content of rU-dA base pairs, which are the weakest of the 4 kinds. Hence, nascent RNA dissociates from the DNA template and then from the enzyme. The solitary DNA template now rejoins its partner to reform the DNA duplex in the bubble region. The core enzyme (devoid of σ) has much less affinity for duplex DNA than for single-stranded DNA, and so the DNA is released. Sigma rejoins the core enzyme to form holoenzyme that can again search for a promoter site to initiate a new transcript.
- At some sites, termination requires the participation of a ρ protein. ρ protein hydrolyses ATP in the presence of single-stranded RNA but not of DNA or duplex RNA. It breaks RNA-DNA hybrid helix by pulling RNA away, fig. 29-15. A common feature of ρ -independent and ρ -dependent termination is that the active signals lie in newly synthesized RNA rather than in the DNA template.

5/e. The structure of polycistronic mRNA.

- Procaryotic mRNA molecules are polycistronic (polygenic)- that is they code for 2 or more polypeptide chains. For example, a single mRNA molecule about 7000 nucleotides long specifies 5 enzymes in the biosynthetic pathway for Tryptophan in E. coli. Each of these 5 proteins has its own start and stop signals on the mRNA. In fact, all known bacterial mRNA molecules contain signals that define the beginning and end of each encoded polypeptide chain.

6. Transcription of DNA in eucaryotic cells.

6/a. The structure of the gene in eucaryotes, the eucaryotic promoter and the role of enhancers.

- Eucaruotic gene characteristics:
◦ presence of enhancer sequences
◦ introns/exons (introns are many more than exons)
◦ presence of regulatory regions (for specific eucaryotic genes) usually located in the DNA that flanks the transcription initiation site at its 5' end (5' flanking-sequence DNA). Occasionally, such sequences are found within the gene itself or in the region that flanks the 3' end of the gene. In mammalian cells, each gene has its own regulatory region
◦ presence of silencer sequences (some genes) diminishing transcription.
- Promoters for RNA polymerase II (for eucaryotic RNA polymerases classification see Stryer pg. 716-717) are located on the 5' side of the start site for transcription. The one closest to the start site, centered at about -25, is called the TATA box (Hogness box). The consensus sequence is a heptanucleotide of A and T residues, fig. 29-22; (in yeast: TATAMINO ACIDA). The TATA box is present in nearly all eucaryotic genes giving rise to mRNA. Note that the TATA box of eucaryotes closely resembles the -10 sequence of procaryotes (TATAMINO ACIDT) but is further from the start site. Mutation of a single base in the TATA box markedly impairs promoter activity. Likewise,

interchanges of A and T bases lead to loss of activity, showing that the precise sequence, not just a high content of AT pairs, is essential. Most TATA boxes are flanked by GC-rich sequences.
- The TATA box is necessary but usually not sufficient for promoter activity. Additional elements are located between -40 and -110. Many promoters contain a CAMINO ACIDT box and some contain a GC box, fig. 29-23. Constitutive genes (genes that are continuously expressed rather than developmentally regulated) tend to have GC boxes in their promoters. The positions of these upstream sequences vary from one promoter to another, in contrast with the quite constant location of the -35 region in procaryotes. Another difference is that the CAAT box and the GC box can be effective when present on the template strand, unlike the -35 region.
- RNA polymerase II by itself cannot recognize promoter site and begin transcription; it needs the presence of specific proteins called transcription factors.
- The activities of many promoters in higher eucaryotes are greatly increased by sequences called enhancers that have no promoter activity of their own. Remarkably, enhancers can work even when located long distances from the promoter, or when they are upstream, downstream or in the midst of a transcribed gene, or when they are oriented in either direction; they also work through heterologous promoters, and even they bind one or more proteins. They are also effective if they are located on either the coding or the template strand. Enhancer elements do appear to convey nuclease hypersensitivity to those regions where they reside.
- A particular enhancer is effective only in certain kinds of cells. The actions of enhancers requires proteins that are expressed only in some cells (eg. glycocorticoids, fig. 29-26).
- As might be expected, the DNA of viruses infecting eucaryotic cells usually contains enhancers that are activated by host-cell proteins. The restricted host range of viruses is partly a consequence of their having tissue - specific and species- specific enhancers.
- How do enhancers act over distances of several kilobases? They serve as docking sites for the assembly of initiation complexes containing RNA polymerase II. DNA can be looped by the binding of regulatory proteins so that sites distant in the linear sequence are brought into proximity.

6/b. Modifications of the primary transcript at the 5' end and at the 3' end.

- The 5' triphosphate end of the nascent eucaryotic RNA chain is almost immediately modified. A ~P is released by hydrolysis. The diphosphate 5' end then attacks the α phosphorus atom of GTP to form an unusual 5'-5' triphosphate linkage. This highly distinctive terminus is called a cap. The N-7 nitrogen of the terminal G is then methylated by S- adenosylmethionine to form cap0. The adjacent riboses may be methylated to form cap1 and cap2, fig. 29-27. Caps are important for subsequent splicing reactions. They also contribute to the stability of mRNAs by protecting their 5' ends from phosphatases and nucleases. In addition, caps enhance the translation of mRNA by eucaryotic protein-synthesizing systems. Transfer RNA and ribosomal RNA molecules do not have caps.
- Some eucaryotic termination signals contain a hairpin followed by a series of U residues, as in procaryotic transcripts. Most eucaryotic mRNA molecules contain a polyadenylate (poly A) tail at their 3' end. This poly A tail is not encoded by DNA. Indeed, the nucleotide preceding poly A is not the last nucleotide to be transcribed. Some primary transcripts contain hundreds of nucleotides beyond the 3' end of the mature mRNA.
- Eucaryotic primary transcripts are cleaved by a specific endonuclease that recognizes the sequence AMINO ACIDUAMINO ACIDA, fig. 29-28. Cleavage does not occur if this sequence is impaired. Some mature mRNAs contain internal AMINO ACIDUAMINO ACIDA sequences. Thus, AMINO ACIDUAMINO ACIDA is only part of the cleavage signal, its context is also important. After cleavage by the endonuclease, a poly A polymerase adds about 250 A residues to the 3' end of the transcript; ATP is the donor in this reaction. This poly A tail then wraps around several copies of a binding protein.
- The role of the poly A tail is an enigma; mRNA devoid of a poly A tail can be transported out of the nucleus, and it serves as an effective template for protein synthesis. An important class of mRNAs, those encoding histone proteins, do not have a poly A terminus.

6/c. The mechanims of splicing. The components participating in this mechanism.

- Introns are precisely spliced out of mRNA precursors (hnRNAs), with a very high fidelity. This fidelity is maintained by the following structural motifs: All eucaryotes have a common structural motif: the base sequence of an intron begins with GU and ends with AG, table 29-3. The consensus sequence at the 5' end of vertebrate introns is AGGUAMINO ACIDGU, fig. 29-29. At the 3' end of introns, the consensus sequence is a stretch of 10 pyrimidines (U or C), followed by any base and then by C, and ending with the invariant AG. Introns also have an important internal site, called the branch site.
- Splicing begins with the cleavage of the phosphodiester bond between the upstream exon (exon1) and the 5' end of the intron, fig. 29-31. The attacking group in this reaction is the 2'-OH of an A residue in the branch site. A 2',5'-phosphodiester bond is formed between this A residue and the 5'-terminal phosphate of the intron. Note that this A residue is also joined to 2 other nucleotides by normal 3',5'-phosphodiester bonds, fig. 29-32. Hence, a branch is generated at this site and a lariat intermediate is formed. The 3'-OH terminus of exon1 then attacks the phosphodiester bond between the intron and exon2. Exon1 and 2 become joined and the intron is released in lariat from. Note that splicing is accomplished by two transesterification reactions rather than by hydrolysis followed by ligation. The first reaction generates a free 3'-OH at the end of exon1, and the second reaction links this group to the 5'-phosphate of exon2. The number of phosphodiester bonds stays the same during these steps. Until the 2 exons are joined, the products of the first reaction are held together by a spliceosome, an assembly of ribonucleoproteins that recognizes the 5' splice site, 3' splice site and branch site of mRNA precursors.

- Small nuclear ribonucleoprotein particles (snRNPs) bind mRNA precursors (hnRNAs) to form spliceosomes).

6/d. The mechanism of alternative splicing. (see also 16/a.)

- Alternative patterns of RNA splicing are subject to developmental control. For example, alternative patterns are used to generate 2 different immunoglobulin heavy-chain mRNAs- one that codes for a membrane-bound heavy-chain protein and another that codes for a secreted heavy chain protein.
- In general, we have 7 types of alternative patterns of RNA splicing: i) casette, ii) mutually exclusive, iii) internal acceptor site, iv) internal donor site, v) retained intron, vi) alternative promoters, and vii) alternative polyadenylation sites.

7. Replication of bacteriophages.

7/a. The lytic cycle of the replication of bacteriophages (T_4 phage).

- The T_4 virion consists of a head, a tail and 6 tail fibers, fig. 34-10. The DNA molecule is tightly packed inside an icosahedral protein coat to form the head of the virus. The tail is made up of 2 coaxial hollow tubes that are connected to the head by a short neck. In the tail, a contractile sheath surrounds a central core through which the DNA is injected into the bacterial host. The tail terminates in a baseplate that has 6 short spikes and gives off 6 long, slender fibers. The tips of the tail fibers bind to a specific site on the outer membrane of E. coli. An ATP-driven contraction of the tail sheath pulls the phage head toward the baseplate and tail fibers, which causes the central core to penetrate the cell wall but not the cell membrane. The naked T_4 DNA then penetrates the cell membrane. A few minutes later, all synthesis of cellular DNA, RNA, and protein stops, and the synthesis of viral macromolecules commence. The infecting virus commands the biosynthetic machinery of the cell and substitutes its genes for bacterial ones. T_4 DNA contains 3 sets of genes that are transcribed at different times after infection: immediate-early, delayed early and late. The early genes are transcribed and translated before T_4 DNA is replicated. Some proteins coded by these genes are responsible for switching off the synthesis of cellular macromolecules. Soon after infection the host-cell DNA is degraded by a deoxyribonuclease that is specified by a T_4 early gene. This enzyme does not hydrolyze T_4 DNA because it lacks clusters of C. T_4 DNA contains 5-hydroxymethylcytosine (HMC) instead of C: Furthermore, some HMC residues in T_4 DNA are glycosylated. The DNA of T_4 bacteriophage contains these derivatives of C because of the action of several phage-specified enzymes that are synthesized in the early phase of infection. Synthesis of the late proteins is coupled to the replication of T_4 DNA. The capsid proteins and a lysozyme are formed at this stage. The lysozyme digests the bacterial cell wall and causes its rupture when the assembly of progeny virions has been completed. Some 20 minutes after infection ~ 200 new virions emerge.
- 3 major pathways in the construction of the virus lead independently to the head, tail, and tail fibers, fig. 34-11.
- A strict sequential order is followed in each pathway.
- The head and tail must be completed before they can combine.
- The construction of T_4 virions does not occur by self-assembly alone. Rather, scaffold proteins and proteases play essential roles at certain stages of the assembly process. Thus, T_4 is constructed by a combination of self-assembly, scaffold-assisted assembly, and enzyme-directed assembly.
- T_4 DNA is inserted into a preformed head.

7/b. The lysogenic pathway in the replication of bacteriophage lambda.

- Some bacteriophages have a choice of life styles: they can multiply and then lyse an infected cell (lytic pathway) or their DNA can join the infected cell, retaining the capacity for multiplication and lysis (lysogenic pathway). Viruses that do not always kill their hosts are called temparate or moderate, e.g. the lambda (λ) virus. The 5' ends of a λ DNA molecule are single-stranded sequences of 12 nucleotides. These sequences are called cohesive ends because they are mutually complementary and can base pair to each other. In fact, the cohesive ends come together after infection of a bacterium. The 5' phosphate terminus of each strand is then adjacent to the 3'-OH end of the same strand. A host DNA ligase seals these 2 gaps to yield a circular λ DNA molecule, fig. 34-16. The circular λ DNA molecule can be replicated by the λ proteins acting with the DNA replication machinery of the host. Alternatively, the λ DNA circle can be inserted by enzymes into the bacterial chromosome by recombination between specific loci on the λ and the E. coli DNA. This event is called site-specific recombination in contrast with general recombination; Interestingly, the DNA must be negatively supercoiled to undergo recombination.
- When the λ DNA will be part of the E. coli genome, it will be called a prophage, and the bacterium containing the prophage is called a lysogenic bacterium. The prophage is stable in the absence of excisionase, an enzyme encoded by bacterium, (xis gene). Transcription of xis is blocked by the λ repressor. When repression is released, excisionase and integrase (an enzyme encoded by the bacterium -int gene) together catalyse the removal of the λ DNA from the E. coli DNA.
- A key feature of this recombination system is that insertion of the λ DNA to the E. coli genome occurs when integrase alone is present, whereas excision occurs when both integrase and excisionase are present. Hence, integration and excision are ultimately controlled by the level of the λ repressor (and cro protein, 810 Stryer)

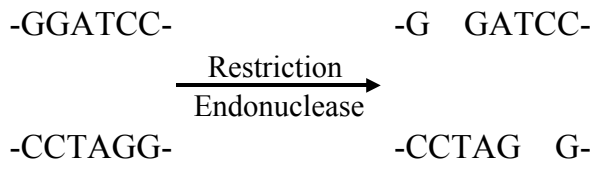
7/c. The role of the restriction endonuclease-methylase pairs in the bacteria.

- Bacteria contain enzymes called restriction endonucleases that cleave foreign DNA molecules such as viral DNAs. Methylation of a base in specific palindromic target sequences protects the bacterium's own DNA from cleavage by its restriction endonuclease. The essence of

this surveillance system is the possession by the bacterium of a methylase and a restriction endonuclease that recognize the same base sequence, fig. 34-13. Methylation of these specific sites is the password denoting that the DNA belongs to the bacterium and is not an intruder. Newly synthesized DNA methylated on only one strand is also protected from digestion. Methylases then convert hemimethylated DNA into fully methylated DNA before the next round of replication. Nearly all foreign DNA molecules containing unmethylated target sites are rapidly degraded by the endonuclease. However, this highly effective mechanism for distinguishing between self and non-self is not perfect. A small proportion of invading DNA molecule acquire a full complement of methyl groups before they can be cut by the restriction enzyme; the bacterial cell can no longer identify them as being foreign.
- 3 types of restriction- modification systems are found in bacteria.

7/d. Production of sticky DNA ends with the aid of a restriction endonuclease.

- Restriction endonucleases recognize and cleave specific double-stranded DNA sequences. These DNA cuts result in blunt ends or overlapping (sticky or cohesive or staggered) ends, depending on the mechanism used by the enzyme:



Sticky ends are particularly useful in constructing hybrid or chimeric DNA molecules.

8. Replication of animal viruses.

8/a. Replication of some DNA viruses in permissive and in non- premissive cells.

- In some cells (called permissive hosts) SV40 and polyoma virus go through a lytic cycle, which results in the production of many new virions, fig. 34-37. Productively infected cells are killed by these viruses. In other types of cells (called non-permissive hosts), some of the steps in viral expression are blocked. No progeny virus are formed but a small proportion of the cells, of the order of 1 in 10^5, are transformed following integration of the viral DNA into the host genome.

8/b. The strategy of replication of the "pure" RNA virus classes. The enzyme responsible for the replication of the RNA genome.

- DNA viruses recruit many host proteins in the replication and expression of their genomes. In the case of RNA viruses, a special problem arises because uninfected host cells lack enzymes for synthesizing RNA according to the instructions of an RNA template. Consequently, RNA viruses must contain genetic information for the synthesis of an RNA-directed RNA polymerase (also called an RNA replicase or an RNA synthetase) or for an RNA-directed DNA polymerase (also called a reverse transcriptase). It is informative to classify RNA viruses according to the relation between their virion RNA and mRNA. By convention, mRNA is defined as (+) RNA and its complement as (-) RNA. 4 pathways of replication and transcription of RNA viruses are known , fig. 34-22. Class 1 viruses are positive-strand RNA viruses. They synthesize(-) RNA, which then serves as the template for the formation of (+) mRNA. Class 2 viruses are negative-strand RNA viruses in which virion (-) RNA is the template for the synthesis of (+) mRNA. Class 3 viruses are double-strand RNA viruses in which the virion (±) RNA directs the asymmetric synthesis of (+) mRNA. Class 4 viruses, the most unusual, are retroviruses. They express the genetic information in their virion (+) RNA through a DNA intermediate that serves as the template for the synthesis of (+) RNA. Thus, the flow of information in retroviruses is from RNA to DNA and then back to RNA.

8/c. The replication of retroviruses.

- Fig. 34-41.
- The life cycle of a typical retrovirus starts when infecting virions bind to specific receptors on the surface of the host and enter the cell. The viral (+) RNA is uncoated in the cytosol. Reverse transcriptase brought in by the virus particle then synthesizes both the (-) and (+) strands of DNA and digests the viral (+) RNA. Thus, reverse transcriptase carries out 3 kinds of reactions: RNA-directed DNA synthesis, hydrolysis of RNA, and DNA-directed DNA synthesis.
- Reverse transcriptase, like other DNA polymerases, synthesizes DNA in the 5'→ 3' direction and is unable to initiate chains de novo. How then is viral DNA synthesis primed? Initiation is accomplished very economically: the (+) RNA viral genome contains a noncovalently bound transfer RNA that was acquired from the host during the preceding round of infection, fig. 34-40. The 3'-OH of this base paired tRNA acts as the primer for DNA synthesis. How is the entire (+) RNA strand replicated? Recall that a special problem is encountered in the replication of any linear DNA, in that the 5' ends of daughter strands are initially incomplete (Stryer pg. 832). Retroviruses have met this challenge in an ingenious way: their genomic (+) RNA contains the same sequence (called R) at the 5' and 3' ends. This terminal redundancy plays a critical role in the synthesis of duplex DNA, as shown in fig. 34-41. Nascent DNA strands undergo 2 shifts in base pairing before a complete duplex is formed. The resulting double-helical intermediate contains identical ends called LTR (long terminal repeats). LTRs are rich in signals for integration and transcription.
- The newly formed viral DNA duplex becomes circular and enters the nucleus. Transcription of retroviral DNA occurs only after it has been integrated into the host -cell DNA. Thus, integration is an obligatory step in the life cycle of retroviruses. In contrast, integration and productive infection are alternative pathways for oncogenic DNA viruses. Another difference is that the frequency of

integration of retroviral DNA is very high, as might be expected in view of its central role in productive infection.
- Productive infection by retroviruses, in contrast with oncogenic DNA viruses, is not lytic. Most retroviruses do not usually kill their hosts. Their DNA stays in the genome of the infected cell and continues to be expressed. Furthermore, the integrated viral DNA is replicated along with the host DNA, and so the viral genome is inherited by the daughter cells.

8/d. The reactions catalyzed by the reverse transcriptase.

See 8/c.

8/e. Viral oncogenes and cellular protooncogenes.

-Cancer-producing genes are called oncogenes. Analyses of more than 20 retroviral oncogenes have shown that they are closely related to normal cellular genes. They can be grouped in 5 classes:
i) Tyrosine kinases
ii) Growth factors
iii) Growth factor receptors
iv) Guanyl-nucleotide-binding proteins
v) Nuclear proteins
- Retroviruses do not appear to be a major cause of human cancers. However, their oncogenic actions in animals have been sources of insight into the molecular basis of cancer. Retroviruses transform susceptible cells by producing excessive quantities of certain key proteins in growth control, or by forming altered proteins that can no longer be controlled. Analysis of oncogenes has revealed that many of them encode proteins with key roles in controlling normal growth and development. It also shows that mutation, duplication or translocation of normal cellular genes involved in growth control can lead to cancer.
- A key issue raised by the discovery of viral oncogenes relates to their origin. Use of nucleic acid hybridization revealed that normal cells contained sequences similarly-if not identical- to those of the viral oncogenes. Thus, the viruses apparently incorporated cellular genes into their genomes during their passages through cells. The retention of such genes in their genomes indicated that they must confer a selective advantage on the affected viruses, presumably related to the altered growth properties of transformed cells. The cellular sequences were found to be conserved in a wide range of eucaryotic cells, suggesting that they were important components of normal cells. In addition, mRNA species and proteins derived from these normal sequences could be detected at various stages of the development or life cycles. The genes present in normal cells thus have been designated proto-oncogenes, and their products are believed to play important roles in normal differentiation and other cellular processes.

8/f. The replication cycle of human immunodeficiency viruses (HIV).

- The host cell for HIV is the T_4 lymphocyte (a helper and inducer cell).
- The HIV virion is enveloped by a lipid bilayer membrane containing 2 glycoproteins: gp41 spans the membrane and is disulphide-bonded to gp120, which is located on the external face, fig. 34-44. The core of the virus contains 2 copies of the RNA genome and associated tRNAs, and several molecules of reverse transcriptase. They are surrounded by many copies of 2 proteins called p18 and p24. HIV enters T_4 lymphocytes by the interaction of gp120 in the viral envelope with a plasma membrane receptor called CD4 (CD stands for Cluster of Differentiation). The 2 membranes fuse and the viral core is released directly into the cytosol, fig. 34-45. The replication cycle of HIV virus is similar to any other retrovirus' replication cycle (see 8/c.).
- The amino acid sequence of gp120 in virus particles obtained from different patients or even the same patient at different times exhibits a high degree of variability. Preparing an effective vaccine against HIV is likely to be much more difficult than for viruses that are genetically stable.

8/g. The antiviral effects of interferon.

- The resistance of animal cells to many viruses is markedly enhanced by interferons, a family of small proteins. They are synthesized and secreted by vertebrate cells following a virus infection. Double-stranded RNA molecules are particularly effective in stimulating the formation of interferon. Secreted interferons bind to the plasma membrane of other cells in the organism and induce an antiviral state in them. These cells acquire resistance to a broad spectrum of viruses. In contrast, immunity conferred by an antibody is highly specific. Interferons are very potent- as little as 10^{-11} M can have a significant antiviral effect.
- 3 classes of interferons have been isolated and purified: a, b, and γ.
- Interferons lead to an antiviral state by stimulating the production of 2 enzymes: a protein kinase and an oligoadenylate synthetase. These enzymes become active in the presence of double-stranded RNA, fig. 34-47. The target of the kinase is the a subunit of eIf2, the initiation factor that brings initiator tRNA to the small ribosomal subunit. Recall that phosphorylation of the a subunit blocks the recycling of eIf2 and hence inhibits translation (Stryer, pg. 762). In particular, the phosphorylation of only 30% of eIf2 leads to the complete cessation of protein synthesis. The second enzyme activated by interferon and double-stranded RNA is 2',5'- oligoadenylate synthetase. This enzyme catalyses the formation of oligoadenylates joined by 2',5' rather than by the usual 3',5' phosphodiester bonds. These 2 actions of interferon, in concert with double-stranded RNA, block cell growth and proliferation. The activated kinase interferes with the formation of new polypeptide chains, and the activated oligoadenylate synthetase leads to the destruction of mRNA templates and rRNA components of the protein-synthesizing machinery.

9. Protein synthesis I. Formation of aminoacyl -tRNA.

9/a. The structure and function of tRNA.

- Fig. 30-9.

- The 5' terminus is phosphorylated (pG), whereas the 3' terminus has a free hydroxyl group. A striking feature of this RNA is its high content of bases other than A, U, G and C. Many unusual nucleosides are present: inosine, pseudouridine, dihydrouridine, ribothymine, and methylated derivatives of guanosine and inosine. The amino acid attachment site is the 3'-OH group of the A residue at the 3' terminus of the molecule. All of the known tRNA base sequences can be written in a cloverleaf pattern in which about half the residues are base paired. Hence, tRNA molecules have many common structural features. This finding is not unexpected, because all tRNA molecules must be able to interact in nearly the same way with ribosomes and mRNAs. Specifically they must fit into the A and P sites on the ribosome and interact with the enzymatic site that catalyzes peptide-bond formation. All known tRNA molecules share the following features, fig. 30-10:
i) They are single chains containing between 73 and 93 ribonucleotides each.
ii) They contain many unusual bases; many of them are methylated or demethylated derivatives of A,U,C, and G that are formed by enzymatic modification of a precursor tRNA. Methylation prevents the formation of certain base pairs, thereby rendering some of the bases accessible for other interactions. Also, methylation impairs a hydrophobic character to some regions of tRNAs, which may be important for their interaction with synthetases and ribosomal proteins. Other modifications alter codon recognition (vide infra).
iii) The 5' end of tRNA is phosphorylated. The 5' terminal residue is usually pG.
iv) The base sequence at the 3' end of tRNAs is CCA. The activated amino acid is attached to the 3'-OH group of the terminal adenosine.
v) About half of the nucleotides in tRNAs are base paired to form double helices. 5 groups of bases are not base paired: the 3' CCA region; the TψC loop, which acquired its name from the sequence ribothymine-pseudouracil-cytosine; the "extra arm", which contains a variable number of residues; the DHU loop, which contains several dihydrouracil residues; and the anticodon loop.
vi) The anticodon loop consists of 7 bases, with the following sequence:
- Pyrimidine - Pyrimidine - X - Y - Z (anticodon) – Modified Purine – Variable Base -
- According to the 3-dimensional structure of tRNA, it has the following features:
i) The molecule is L-shaped, fig. 30-11.
ii) There are 2 segments of double helix. Each of these helices contains about 10 base pairs, which correspond to one turn of helix. The helical segments are perpendicular to each other, which gives the molecule its L-shape, fig. 30-12.
iii) Most of the bases in the nonhelical regions participate in unusual Hydrogen-bonding interactions. These tertiary interactions are between bases that are not usually complementary (e.g. GG, AA, and AC). Moreover, the ribose- phosphate backbone interacts with some bases and even with another region of the backbone itself. The 2' -OH groups of the ribose units act as Hydrogen-bond donors or acceptors in many of these interactions. In addition, most bases are stacked. These hydrophobic interactions between adjacent aromatic rings play a major role in stabilizing the architecture of the molecule.
iv) The CCA terminus containing the amino acid attachment site is at one end of the L. The other end of the L is the anticodon loop. Thus, the amino acid in aminoacyl-tRNA is far from the anticodon. The DHU and TψC loops form the comer of the L.
v) The CCA terminus and the adjacent helical region do not interact strongly with the rest of the molecule. This part of the molecule may change its conformation during amino acid activation and also during protein synthesis on the ribosome.
- tRNA has 3 functions:
i) to carry the amino acid to the ribosome;
ii) to activate the amino acid - by activating its -COOH group by attaching it to its 3'-OH terminus;
iii) to transfer the correct amino acid to the ribosome, using its anticodon loop.

9/b. Why is enough 31 tRNA for 61 codons?

- Some pure tRNA molecules can recognize more than one codon.
- In mitochondria: 22 tRNAs for 61 codons. The first two bases of a codon pair in the standard way. Recognition is precise. Hence, codons that differ in either of their first 2 bases must be recognized by different tRNAs, e.g. both UUA and CUA code for Leucine but are read by different tRNAs.
- The first base of an anticodon determines whether a particular tRNA molecule reads one, two three kinds of codons: C or A (1 codon), U or G (2 codons) or I (3 codons). Thus, part of the genetic code arises from imprecision (wobble) in the pairing of 3 base of the codon. We see here a strong reason for the frequent appearance of inosine (I), one of the unusual nucleosides, in anticodons. I maximizes the number of codons that can be read by a particular tRNA molecule, (fig. 30-14).

9/c. What determines the meaning of activated tRNA? Anticodon or amino acid? Experiment.

- This question was answered in the following way: First Cysteine was attached to its cognate tRNA (denoted by $tRNA^{Cys}$). The attached Cys unit was then converted into Ala by reacting Cys- $tRNA^{Cys}$ with Raney nickel, which removed the S atom from the activated Cys residue without affecting its linkage to tRNA. Thus, a hybrid aminoacyl-tRNA was produced in which Ala was covalently attached to a tRNA specific for Cys:

Does this hybrid tRNA recognize the codon for Ala or for Cys? The answer came on adding the tRNA to a cell-free protein-synthesizing system. The template was a random copolymer of U and G in the ratio of 5:1, which normally leads to the incorporation of Cys (UGU) but not of Ala (GCX). However, Ala was incorporated into a polypeptide when Ala-$tRNA^{Cys}$ was added to the incubation mixture, because it was attached to tRNA specific for Cys. The same result was obtained when mRNA for hemoglobin served as

the template in ^{14}C alanyl- $tRNA^{Cys}$ was used as the hybrid aminoacyl- tRNA. The only radioactive tryptic peptide produced was one that normally contained Cys but not Ala. On the other hand, peptides normally containing Ala but not Cys were devoid of radioactivity. Thus, the amino acid in aminoacyl-tRNA does not play a role in selecting a codon.

9/d. The formation of aminoacyl- tRNA.

- The formation of a peptide bond between the amino group of one amino acid and the carboxyl group of another is thermodynamically unfavourable. This barrier is overcome by activating the carboxyl group of the precursor amino acid. The activated intermediates in protein synthesis are amino acid esters, in which the -COOH group of an amino acid is linked to either the 2'-or the 3'-OH group of the ribose unit at the 3' and of tRNA. This activated intermediate is called an aminoacyl-tRNA, fig. 30-2.
- The activation of amino acid and their subsequent linkage to tRNAs are catalyzed by specific amino-acyl-tRNA synthetases, which are also called activating enzymes. The first step is the formation of an aminoacyl-adenylate from an amino acid and ATP. This activated species is a mixed anhydride in which the -COOH group of the amino acid is linked to the phosphoryl group of AMP; hence it is also known as aminoacyl-AMP:

$^{+}H_3N$-C-H-(R) +ATP⇔ $^{+}H_3N$-C-H-(R)-COO-PO_3^{-}Ribose-Adenine PP_i.

The next step is the transfer of the aminoacyl group of aminoacyl-AMP to a tRNA molecule to form aminoacyl-tRNA, the activated intermediate in protein synthesis. For some tRNAs, the amino acid is transferred to the 2'-OH of the ribose unit and, for others, to the 3'-OH. The activated amino acid can migrate very rapidly between the 2'-and 3'-OH groups:

Aminoacyl-AMP + tRNA ⇔ aminoacyl-tRNA + AMP

The sum of these activation and transfer steps is:

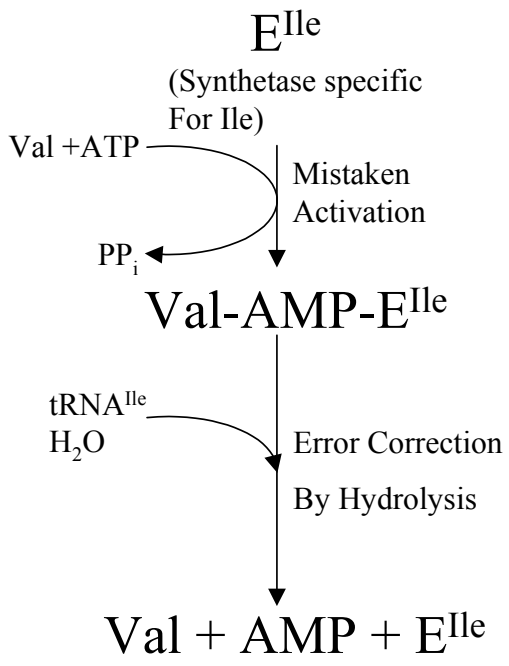
AMINO ACID + ATP + tRNA ⇔ amino-acyl-tRNA + AMP + PPi.

The reaction is driven by the hydrolysis of pyrophosphate.

- 2 high-energy phosphate bonds are consumed in the synthesis of an aminoacyl tRNA. One of them is consumed in forming the ester linkage of aminoacyl-tRNA, whereas the other is consumed in driving the reaction forward.

9/e. The specificity of aminoacyl- tRNA synthetases.

- Aminoacyl-tRNA synthetases are highly selective in their recognition of both the amino acid to be activated and the prospective tRNA acceptor. tRNA molecules that accept different amino acid have different base sequences, and so they can be readily distinguished by their synthetases. A much more demanding task for these enzymes is to discriminate between similar amino acid, e.g. ILe/Val. Actually, the synthetase can correct its own errors:



How does the synthetase avoid hydrolyzing ILe-AMP, the correct intermediate? Most likely, the hydrolytic site is just large enough to accommodate Val-AMP but too small to allow the entry of ILe-AMP.

- Most aminoacyl-tRNA synthetases contain hydrolytic sites in addition to synthetic sites. Complementary pairs of sites function as a double sieve to assure very high fidelity. The synthetic site rejects amino acid that are larger than the correct one because there is insufficient room for them, whereas the hydrolytic site destroys activated intermediates that are smaller than the correct species, fig. 30-5. Hydrolytic proofreading is central to the fidelity of many amino-acyl tRNA synthetases, as it is to DNA polymerases. However, a few synthetases achieve high accuracy without editing their covalently attached intermediates, e.g. tyrosyl-tRNA synthetase has no difficulty discriminating between Tyr and Phe; the -OH group on the Tyr ring enables it to be bound to the enzyme 10^4 times as strongly as Phe. Proofreading is costly in energy and time and hence is selected in the course of evolution only when fidelity must be enhanced beyond what can be obtained through an initial binding interaction.

10. Protein synthesis II. Initiation of protein synthesis.

10/a. The structure, function and assembly of ribosomes.

- Ribosomes are ribonucleoprotein particles (70S: sedimentation coefficient) consisting of a small (30S) and a large (50S) subunit.
- Ribosomes have 3 binding sites for RNAs: i) mRNA, ii) -P (peptidyl), iii) -A (aminoacyl)
- Ribosomal RNAs (5S, 16S, and 23S rRNA) contain many base-paired helical regions. A striking feature of these

rRNAs is their folding into defined structures with many short duplex regions.
- The 3' end of 16S rRNA-a constituent of the 30S ribosomal subunit, plays a key role in selecting the start site on the mRNA template.
- Fig. 30-18, 30-19.
- Ribosomes can be formed in vitro by self-assembly of their constituent proteins and RNAs: The 30S ribosomal subunit can be reconstituted from a mixture of its 21 constituent proteins and 16S RNA. The significance of such a reconstitution experiment is 2-fold:
i) it is demonstrated that all of the information needed for the correct assembly of this organelle is contained in the structure of its components; nonribosomal factors are not needed.
ii) Reconstitution can be used to ascertain whether a particular component is essential for the assembly of the ribosome or for a specific function; for example, the component responsible for sensitivity to streptomycin was identified in this way (Stryer pg. 759).
- Studies of the reconstitution of the 30S subunit have shown that 16S RNA is essential for its assembly and function. The requirement is quite specific because 16S RNA from yeast, which folds similarly, cannot substitute for 16S RNA from E. coli. Most of the 21 proteins are also needed for assembly, indicating that the 30S subunit is a cooperative functional entity. The assembly process is ordered and proceeds in stages, e.g. an intermediate containing 15 proteins and 16S RNA can be trapped by reconstituting the particle at 0°C. The other 6 proteins join this intermediate to form a functional 30S subunit when the temperature is raised to 40°C.
- The function of a ribosome is to bring the mRNA and tRNA together contributing to protein synthesis.

10/b. The direction of protein synthesis and the mRNA reading. Prove one of them experimentally!

- Proteins are synthesized in the amino-to carboxyl direction: Reticulocytes that were actively synthesizing hemoglobin were exposed to 3H-Leu. Completed hemoglobin was sampled frequently during a period shorter than required to synthesize a complete chain. Each sample was separated into a and b chains and then treated with trypsin. In the earliest samples, only peptides from the carboxyl ends were labelled. Later samples yielded labelled peptides closer and closer to the amino ends. Over all the samples, a gradient of radioactivity increasing from the amino to the carboxyl end of each chain was found, fig. 30-24. This would be expected if the amino part of the sampled chains was already synthesized prior to the addition of radioactive Leu. If the carboxyl end was synthesized last, radioactive label would appear there first, in chains that were almost complete when label was added to the medium; so, the direction of chain growth is from the amino to the carboxyl end.
- The direction of reading of mRNA was determined by using the synthetic polynucleotide
$^{5'}$ A- A-A (-A-A-A)$_n$ -A-A-C $^{3'}$
as the template in a cell-free protein-synthesizing system. AAA codes for Lys, whereas AAC codes for asparagine. The polypeptide product was
$^{+}$ H_3N -Lys -(Lys)$_n$ -Asn- COO^-
Because Asn was the carboxyl-terminal residue, the codon AAC was the last to be read. Hence, the direction of translation is 5' $\rightarrow$ 3'.
- If the direction of translation were opposite to that of transcription, only fully synthesized RNA could be translated. In contrast, if the directions are the same, mRNA can be translated while it is being synthesized. In fact, mRNA is synthesized also in the 5' $\rightarrow$ 3' direction (Stryer, pg. 709). In E. coli, almost no time is lost between transcription and translation. The 5' end of mRNA interacts with ribosomes (polyribosome formation) very soon after it is made, fig. 30-25. An important feature of procaryotic gene expression is that translation and transcription are closely coupled in space and time.

10/c. The recognition of initiation codon(s) is procaryotes. The role of IF1, IF2, IF3.

- Protein synthesis in bacteria starts with formylmethionine (fMet). A special tRNA brings fMet to the ribosome to initiate protein synthesis. This initiator tRNA (abbreviated as tRNAf) is different from the one that inserts Met in internal positions (abbreviated as tRNAm). The subscript f indicates that Met attached to the initiator tRNA can be formylated, whereas it cannot be formylated when attached to tRNAm. (Met is linked to these 2 kinds of tRNAs by the same amino-acyl-tRNA synthetase, fig. 30-27). A specific enzyme (transformylase) then formylates the amino group of Met that is attached to tRNAf. The activated formyl donor in this reaction is N^{10} -formyl H_4 folate. It is significant that free Met and methionyl -tRNAm are not substrates for this formylase. Met-tRNAf enters directly into the partial P site; only the regular tRNAs can enter the complete A site.
- 2 kinds of interactions determine where protein synthesis starts: the pairing of mRNA bases (Shine- Dalgarno sequence) with the 3' end of 16S rRNA, and the pairing of the initiator codon on mRNA with the anticodon of fMet initiator tRNA. The initiator codon is AUG (or GUG), fig. 30-29.
- How are mRNA and fMet-tRNAf brought to the ribosome to initiate protein synthesis? 3 protein initiation factors (IF1, IF2 and IF3) are essential. The 30S ribosomal subunit first forms a complex with these 3 factors, fig. 30-30. The binding of GTP to IF2 enables mRNA and the initiator tRNA to join the complex as IF3 is released. fMet-tRNAf is specifically recognized by IF2, and the release of IF3 allows the 50S subunit to join the complex hydrolysis of GTP bound to IF2 on entry of the 50S subunit leads to the release of IF1 and IF2. The result is a 70S initiation complex.
The role of IF-1 is probably recycling. The role of IF-2 is the binding of fMet-tRNAf and the hydrolysis of GTP. The role of IF-3 is subunit dissociation and mRNA binding.

10/d. The recognition of translation start in eucaryotes. The role of eIF2, eIF3, eIF4.

- Initiation can be divided into 4 steps:
i) dissociation of the ribosome into its 40S and 60S subunits; 2 initiation factors, eIF-3 and eIF-1A, bind to the 40S subunit. This favours dissociation of the 80S ribosome into its 40S and 60S subunits and prevents reassociation.

The binding of eIF-3A to the 60S subunit may also prevent reassociation.
ii) Formation of the 40S preinitiation complex; the first step in this process involves the binding of GTP by eIF-2 (3 subunits comprise eIF2: α= binds to GTP [controlled by phosphorylation], β= recycling factor, γ= binds to Met-$tRNA_i$). This binary complex then binds to met-tRNA; (initiator)), a tRNA specifically involved in binding to the initiation codon AUG. This ternary complex binds to the 40S ribosomal subunit to form the 40S preinitiation complex, which is stabilized by association with eIF3, and eIF-1A.
iii) Formation of the 40S initiation complex; the 5' termini of most mRNA molecules in eucaryotic cells are "capped". This methyl-guanosyl triphosphate cap facilitates the binding of mRNA to the 40S preinitiation complex. A cap binding protein , eIF-4F, binds to the cap through one of its subunits. Then eIF-4A and eIF-4B bind and probably reduce the complex secondary structure of the 5' end of the mRNA through their respective ATPase and helicase activities. Certainly the association of mRNA with the 40S preinitiation complex to form the 40S initiation complex requires ATP hydrolysis. eIF-3 is a key protein because it binds with high affinity to both mRNA and the 40S ribosomal subunit. Following association of the 40S preinitiation complex with the mRNA cap and reduction ("melting") of the secondary structure near the 5' end of mRNA, the complex scans the mRNA for a suitable initiation codon. Generally this is the 5'-most AUG, but the precise initiation codon is determined by so-called Kozak consensus sequences that surround the AUG. Most preferred is the presence of a purine at positions -3 and +4 relative to the AUG.
iv) Formation of the 80S initiation complex; the binding of the 60S ribosomal subunit to the 40S initiation complex involves the hydrolysis of the GTP bound to eIF-2 by eIF-5. This reaction results in the release of the initiation factors bound to the 40S initiation complex (these factors then are recycled) and the rapid association of the 40S and 60S subunits to form the 80S ribosome. At this point, the met-tRNA is on the P site of the ribosome, ready for the elongation cycle to commence.
- eIF-2: The α binds to GTP (controlled by phosphorylation). The β is a recycling factor. The γ binds to Met-$tRNA_i$.

10/e. Comparison of initiation in eucaryotes and in procaryotes.

- Initiator tRNA. In eucaryotes, the initiating amino acid is Methionine rather than N-formylmethionine. However, as in procaryotes, a special tRNA participates in initiation. This aminoacyl-tRNA is called Met-tRNAf or Met-tRNAi (the subscript f indicates that it can be formylated in vitro, and i stands for initiation).
- Start signal. The initiating codon in eucaryotes is always AUG. Eucaryotes, in contrast with procaryotes (Stryer pg. 753), do not use a purine-rich sequence on the 5' side to distinguish initiator AUGs from internal ones. Instead, the AUG nearest the 5' end of mRNA is usually selected as the start site (Kozak sequences). 40S ribosomes attach to the cap at the 5' end of eucaryotic mRNAs and search for an AUG codon by moving stepwise in the 3' direction. This scanning process is driven by the hydrolysis of ATP. The anticodon of Met-tRNA; bound to the 40S subunit pairs with the AUG codon, signifying that the target has been found. A eucaryotic mRNA has only one start site and hence is the template for a single protein (monocistronic). In contrast, a procaryotic mRNA can have multiple start sites and be the template for the synthesis of several proteins (polycistronic).
- Initiation complexes. Eucaryotes contain many more initiation factors than do procaryotes, and their interplay is much more intricate. 10 are known, and several consist of multiple subunits. The GTP form of eIF2 brings the initiator tRNA to the 40S subunit. Cap-binding proteins (CBPs) bind the cap of mRNA. They are joined by eIF-3, which finds the initiator AUG. eIF-4 is the ATP-driven engine in this search. eIF-5 induces the release of eIF2 and eIF3 following the pairing of Met-tRNAi with the initiating AUG; eIF-5 exerts these actions by triggering the hydrolysis of GTP bound to eIF-2. Finally, the 60S subunit joins the complex of initiator tRNA, mRNA, and 40S subunit to form the 80S initiation complex.

11. Protein synthesis III.
Elongation and termination of protein synthesis.

11/a. Steps of elongation.

- Elongation, a cyclic process, involves several steps catalyzed by proteins called elongation factors (eEF). These steps are:
i) Binding of aminoacyl-tRNA to the A site of the ribosome: In the complete 80S ribosome formed during the process of initiation, the A site is free. The binding of the proper aminoacyl-tRNA in the A site requires proper codon recognition. Elongation factor eEF-1α forms a complex with GTP and the entering aminoacyl-tRNA (Harper, fig. 40-8). This complex then allows the aminoacyl-tRNA to enter the A site with the release of eEF-1α • GDP and phosphate. eEF-1α • GDP then recycles to eEF-1α • GDP with the aid of other soluble protein factors and GTP. Similar GTP/GDP cycling also occurs during hormonal action (vide infra) and during visual excitation.
ii) Peptide bond formation: The α-amino group of the new aminoacyl-tRNA in the A site carries out a nucleophilic attack on the esterified carboxyl group of the peptidyl-tRNA occupying the P site. This reaction is catalyzed by a protein component, peptidyl-transferase of the 60S subunit. Because the amino acid on the aminoacyl-tRNA is already "activated", no further energy source is required for this reaction. The reaction results in attachment of the growing peptide chain to the tRNA in the A site.
iii) Translation: Upon removal of the peptidyl moiety from the tRNA in the P site, the discharged tRNA quickly dissociates from the P site. Elongation factor 2 (eEF-2 or EF-G or translocase) and GTP are responsible for the translocation of the newly formed peptidyl-tRNA at the A site into the empty P site. The GTP required for eEF-2 is hydrolyzed to GDP and phosphate during the translocation process. The translocation of the newly formed peptidyl-tRNA and its corresponding codon into the P site then frees

the A site for another cycle of aminoacyl-tRNA codon recognition of elongation.

11/b. The function of elongation factors in procaryotes and eucaryotes.

Procaryotes	Eucaryotes	Function
EF-Tu	EF-1α	GTP binding
EF-Ts	EF 1β,γ	GTP-GDP exchange
EF-G	EF-2	translocation

11/c. The role of EF-Tu in setting the pace and fidelity of protein synthesis.

- The fidelity of protein synthesis depends on having the correct aminoacyl-tRNA in the A site when the peptide bond is formed. This is the point of no return, for cells lack a means of excising an incorrect amino acid residue once the peptide bond is made. As might be expected, the incoming aminoacyl-tRNA is carefully scrutinized to make sure that its anticodon is complementary to the mRNA codon occupying the A site. An aminoacyl-tRNA that can pair with 2 of the 3 bases of a codon may bind transiently but it will usually leave the A site before a peptide bond is formed. It takes time for the ribosome to decide whether the bound aminoacyl-tRNA is the right one. Protein synthesis would be inaccurate if a peptide bond formed immediately on arrival of an aminoacyl-tRNA. The timer that determines the period of scrutinity is the GTPase site of EF-Tu, fig. 30-33. A peptide bond cannot be formed until EF-Tu is released from the aminoacyl-tRNA; release requires that bound GTP be hydrolyzed to GDP. (EF-Tu does not react with fMet-tRNAf, but it binds to Met-tRNAm). On average, several msec pass before GTP is hydrolyzed, and another several msec pass before EF-Tu-GDP leaves the ribosome. An incorrect aminoacyl-tRNA usually leaves the ribosome during one of these intervals, whereas the correct one stays bound. The conformation of EF-Tu changes when it hydrolyzes GTP, and so the context of the codon-anticodon interaction is altered. The correct aminoacyl-tRNA interacts strongly with mRNA in both states, but an incorrect one does not. In effect, the codon-anticodon interaction is scrutinized twice, in different ways, to achieve higher accuracy, just as proofreading of a manuscript by 2 readers for an hour is much more effective than 2 hours of proofreading by one.

11/d. The termination of protein synthesis in pro-and eucaryotes.

- In eucaryotes, RF1,2 are substituted by eRF (it requires GTP to bind to ribosomes).
- Aminoacyl-tRNA does not normally bind to the A site of a ribosome if the codon is UAA, UGA, or UAG. Normal cells do not contain tRNAs with anticodons complementary to these stop signals. Instead, the signals are recognized by release factors, which are proteins. One of these release factors, RF1, recognizes UAA or UAG. A second factor, RF2, recognizes UAA or UGA. Thus, proteins can recognize trinucleotide sequences with high specificity. The binding of a release factor to a termination codon in the A site activates peptidyl transferase so that it hydrolyzes the bond between the polypeptide and the t-RNA in the P site. The specificity of peptidyl transferase is altered by the release factor so that H_2O rather than an amino group is the acceptor of the activated peptidyl moiety. The polypeptide chain leaves the ribosome, followed by tRNA and mRNA. Finally, the ribosome dissociates into 30S and 50S (or 40S and 60S in eucaryotes) subunits as the prelude to the synthesis of another protein molecule. The binding of IF3 to the 30S subunit prevents it from joining the 50S subunit to form a dead-end 70S complex devoid of mRNA.

12. Targeting of proteins to the cytosolic compartment and endoplasmic reticulum.

12/a. Covalent modification of proteins in the cytosol.

- Glycosylation: N-actylglycosamine is attached by an O-glycosidic linkage to a Ser of the protein.
- Fatty acylation: i) Myristic acid (necessary for oncogenic action of src $p6O^{src}$); it is attached to a Gly.
ii) Palmitic acid (ras oncogene); always membrane associated; it is attached to a Cys.
- Phosphorylation / Dephosphorylation.
- Methylation.
- Acetylation (eg. histone acetylation).

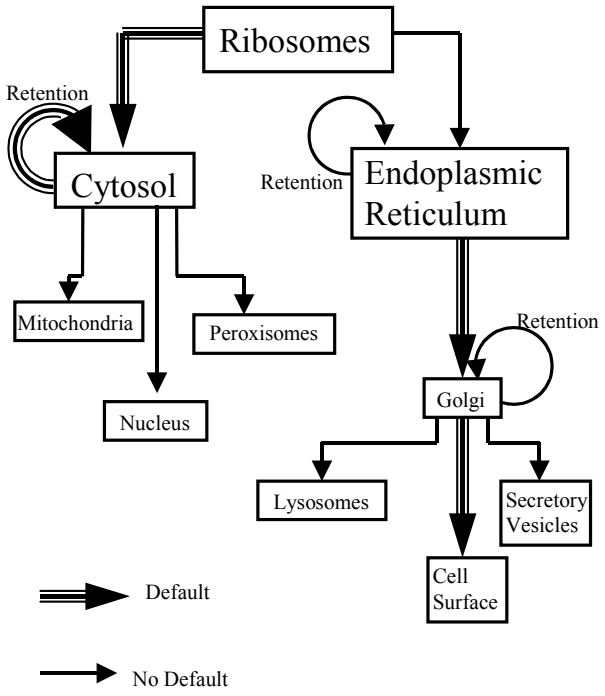
12/b. Ubiquitin-dependent selective protein turnover.

- Ubiquitin, a small protein present in all eucaryotic cells, plays an important role in tagging proteins for destruction. This protein is highly conserved in evolution. The carboxyl-terminal Gly of ubiquitin becomes covalently attached to the E-amino group of Lys residues of proteins destined to be degraded. The energy for the formation of these isopeptide bonds comes from ATP. 3 enzymes (E1, E2, and E3) participate in the conjugation of ubiquitin to proteins. First, the terminal carboxylate group of ubiquitin becomes linked to a sulfhydryl group of E1 by a thioester bond, fig. 31-39. Activated ubiquitin is then shuttled to a sulfhydryl of E2. Finally, E3 catalyzes the transfer of ubiquitin from E2 to the target protein. A protein tagged for destruction usually acquires several molecules of ubiquitin, fig. 8-18 MBC).
- The cytosolic proteins that are programmed for rapid destruction carry signals that are recognized by this proteolytic machinery responsible for their degradation. One of these signals is remarkably simple, consisting only of the first amino acid in the polypeptide chain. When present as the amino-terminal residue, the amino acids Met, Ser, Thr, Ala, Val, Cys, Gly, and Pro are stabilizing, while the remaining 12 amino acids attract a proteolytic attack. These destabilizing amino acids are virtually never found at the amino-terminus of stable cytosolic proteins. They are often present, however, at the amino terminus of proteins that are transported into other compartments, such as the ER; since the cytosolic degradation machinery is not present in the lumen of the ER or the Golgi apparatus, such proteins are normally long lived in their respective compartments. A destabilizing amino-terminal amino acid on these noncytosolic proteins may provide the cell with a simple way of disposing of copies that have been

misrouted: those molecules that fail to be rapidly translocated out of the cytosol are rapidly degraded. A similar single-residue code is apparently used in bacteria to signal the rapid destruction of specific proteins.

12/c. The role and characteristics of signal sequences in direction of proteins.

- There are thought to be 2 types of sorting signals on proteins that direct them step by step through the branching-pathways outlined in the following figure:



For some steps the sorting signal resides in a continuous stretch of amino acid sequence. This signal peptide is often (but not always) removed from the finished protein once the sorting decision has been executed. The signal for other sorting steps is thought to consist of a particular 3-dimensional arrangement of atoms on the protein's surface that forms when the protein folds up. The amino acid residues that comprise this signal patch may be quite distant from one another in the linear amino acid sequence, and they generally remain in the finished protein, (fig. 8-13, MBC). Signal peptides are used to direct proteins from the cytosol into the ER, mitochondria, chloroplasts, and nucleus; they are also used to retain certain proteins in the ER. Signal patches are thought to be used for some other sorting steps, including the recognition of certain lysosomal proteins by a special sorting enzyme in the Golgi apparatus (MBC, pg.463).
- Signal sequences exhibit several common features:
i) They range in length from 13 to 36 residues.
ii) The amino-terminal part of the signal contains, at least one (+) charged residue.
iii) A highly hydrophobic stretch, typically 10 to 15 residues long, forms the center of the signal sequence.
iv) The cleavage site at the carboxyl-terminal end is preceded by a sequence of about 5 residues that is more polar than the hydrophobic core.
- Different types of signal peptides are used to specify different destinations in the cell, (MBC table 8-3 pg 415). Proteins destined for initial transfer to the ER usually have amino-terminal signal peptides with a central part of the sequence composed of 10-15 hydrophobic amino acid residues. Most of these proteins will pass from the ER to the Golgi apparatus; those with a specific sequence of 4 amino acids at their carboxyl terminus, however, are retained as permanent ER residues. Proteins destined for mitochondria have signal peptides in which (+) amino acid residues alternate with hydrophobic ones. Many proteins destined for the nucleus carry signal peptides formed from a cluster of (+) charged amino acid residues. Finally, some cytosolic proteins have signal peptides that cause a fatty acid to be covalently attached to them, which directs the proteins to membranes without insertion into the ER, (MBC pg. 417).

12/d. Transport of proteins across the endoplasmic reticulum membrane. The structure and role of the signal recognition particle.

- The attachment of an actively synthesizing ribosome to the ER membrane is a key event in translocating a protein across this membrane. The signal for attachment is a sequence of amino acid residues near the amino-terminus of the nascent polypeptide chain. These signal sequences are absent from normally secreted proteins because they are cleaved by a signal peptidase on the luminal side of the ER membrane. Not all secretory and plasma membrane proteins contain an amino-terminal signal sequence that is cleaved following translocation across the ER membrane. Some proteins contain an internal signal sequence that serves the same role.
- A cytosolic protein can be redirected to the ER by joining a signal sequence to its amino-terminus.
- How are signal sequences recognized and the proteins containing them translocated across the ER membrane? The adaptor coupling the protein-synthesizing machinery in the cytosol to the protein-translocating machinery in the ER membrane is a ribonucleoprotein called signal recognition particle (SRP). It consists of a 300-nucleotide RNA molecule (called 7SL RNA) and 6 different polypeptide chains. SRP binds tightly to ribosomes containing a nascent chain with a signal sequence but not to other ribosomes, fig. 31-5. This binding occurs soon after the emergence of the amino-terminal signal sequence from the ribosome. Elongation of the polypeptide chain is arrested or slowed while SRP is bound. The SRP-ribosome complex then diffuses to the ER membrane, where SRP binds to the SRP receptor (docking protein). The ribosome containing the nascent polypeptide chain is delivered to the translocation machinery, which includes 2 integral membrane proteins called ribophorin I and II, and SRP is released into the cytosol. The nascent chain resumes signal sequence to the ER membrane.
- The structural backbone of SRP is its RNA component, which contains 2 Alu sequences (they recur frequently in mammalian DNA), (Stryer, pg.835). The RNA binds 6 proteins to form a bipartite SRP. One domain of this highly

elongated particle recognizes and binds the emerging signal sequence; the other domain of SRP sterically interferes with the entry of aminoacyl-tRNA and the peptidyl transferase step, and so elongation of the polypeptide chain is arrested. This arrest is relieved when SRP binds its receptor on the ER membrane. The cytosolic part of the α subunit of SRP receptor contains several highly positively charged loops that probably interact with the negatively charged RNA of SRP.

12/e. Connection between translocation across the ER membrane and the elongation of protein synthesis.

- Although translocation and elongation are mechanistically separate process, they occur simultaneously. This is because most fully synthesized proteins possessing signal sequences cannot be efficiently translocated, probably because they have folded up. A pair of interacting α helices, i.e., cannot fit in the ribosomal tunnel. Unfolded polypeptide chains are optimal substrates for translocation across the ER membrane. The elongation arrest or slowing induced by the binding of SRP to ribosomes prevents premature folding on the cytosolic side of the membrane. The ribosome, too, participates in maintaining the translocatability of the nascent chain by keeping its most recently synthesized part fully stretched out in the narrow tunnel of the large subunit.

13. Targeting proteins to the nucleus and mitochondria.

13/a. Transport of proteins into the nucleus. Nuclear import signal.

When proteins are extracted from the nucleus and microinjected back into the cytoplasm, even the very large ones efficiently reaccumulate in the nucleus; nuclear pores appear to function like a close-fitting diaphragm that opens to just the right extent when activated by a signal on an appropriate large protein. Proteins are actively transported through the pores, probably while still in their folded forms, and the energy required is derived from ATP hydrolysis.
- The selectivity of nuclear transport resides in nuclear import signals, which are present only in nuclear proteins. The signal, which can be located anywhere in the protein (except at the amino- or carboxyl- terminus so it can not be cleaved off), and it can occur more than once and in different types, consists of a short peptide, 4-8 amino acid residues long, that is rich in the (+) charged amino acid Lys and Arg and usually contains Pro. It was first identified in the SV40 virus-encoded protein called T-antigen, a large protein that is needed for viral DNA replication in the nucleus. The T-antigen normally accumulates in the nucleus shortly after being synthesized in the cytosol. However, a mutation in a single amino acid prevents nuclear transport and causes the mutant protein to remain in the cytoplasm, (MBC, fig. 8-25)).
- The mechanism of nuclear protein is fundamentally different from the import mechanisms for the other membrane-bounded organelles in that it occurs through a controlled aqueous pore rather than through one or more membranes. Moreover, when a nucleus disassembles in mitosis, its contents mix with those of the cytosol and the nuclear proteins escape. When reassembly occurs, groups of chromosomes are first enveloped within their own double membrane, which is so closely applied that it excludes soluble proteins, among them many former residues of the nucleus (MBC, pg. 777). The enveloped chromosomes then fuse to form a single nucleus that must reimport the proteins it requires from the cytosol. Possibly, because a nuclear protein molecule needs to be repeatedly imported, its nuclear import signal peptide is not cleaved off after transport into the nucleus. In contrast, once a protein molecule has been imported by any of the other membrane-bounded organelles, it is passed on from generation to generation within that compartment and need never be translocated again; the signal peptide on these molecules is often removed following protein translocation.

13/b. Sorting of proteins into the mitochondria.

- Proteins are imported into the mitochondria after they are synthesized in the cytosol. Furthermore, most mitochondrial proteins destined for the matrix, the intermembrane space, or the inner membrane are found to contain an additional 20-60 amino acids, called the uptake-targeting sequence, at the N-terminus that are not present in the mature protein (MCB, fig. 18-11).
- These precursor proteins, including hydrophobic integral membrane proteins, are soluble in the cytosol; there, they bind to one or more unfolding proteins (heat-shock proteins?), which use the energy released by ATP hydrolysis to keep the precursor proteins in such a state that can be taken up by mitochondria (MCB, fig. 18-12). During mitochondrial uptake, the precursors of integral membrane proteins undergo a major conformational change to bind to the hydrophobic core of the inner or outer membrane. This change may be due in part to cleavage of all or a portion of the N-terminal uptake- targeting sequence in the matrix.
- Many imported proteins do not fold spontaneously in the matrix but require a protein catalyst. One such catalyst is hsp60, a heat shock protein localized to the matrix.
- Targeting proteins to the matrix turns out to be the most direct procedure. Precursors of matrix proteins, such as alcohol dehydrogenase, have a matrix targeting sequence at their N-terminus (MCB, fig. 18-13) that contains all the information required to target a protein from the cytosol to the mitochondrial matrix. A matrix enzyme removes these N-terminal sequences as they arrive in the matrix. This enzyme, a 2 subunit metal-containing protease, specifically cleaves the N-terminal matrix targeting sequence from several different precursor proteins.
- The matrix-targeting sequences exhibit some common characteristics. They are rich in (+) Arg and Lys and hydroxylated Ser and Thr; they are also devoid of Asp and Glu, acidic residues. Apparently, the receptor (s) for matrix-targeting sequences on the outer mitochondrial membrane (MCB, fig. 18-12) can recognize a large number of related amino acid sequences.
- 2 different N-terminal uptake-targeting sequences on precursors to such proteins as cytochromes b_2 and c_1 target them to intermembrane space. A matrix-targeting sequence at the N-terminus of the precursor directs the protein to the matrix first (MCB, fig.18-13/18-14), where the sequence is

removed by the matrix protease. The second uptake-targeting sequence directs the protein from the matrix across the inner membrane to the intermembrane space.
- All mitochondrial proteins synthesized in the cytosol contain one or more uptake-targeting sequences, but these sequences are not removed in some cases. An important example is porin, the abundant outer mitochondrial membrane protein that forms channels through the phospholipid bilayer and accounts for the unusual permeability of this membrane to small proteins. The N-terminus of porin contains a short matrix-targeting sequence followed by a long stop-transfer stretch of hydrophobic amino acid (MCB, fig. 18-13); the latter causes the protein to divert from the matrix pathway and accumulate in the outer mitochondrial membrane. Neither sequence is normally removed. Proteins such as the ATP-ADP antiport, which are localized to the inner mitochondrial membrane, also utilize one or more uptake-targeting sequences that are not removed.
- N-terminal uptake-targeting sequences of precursors of mitochondrial proteins interact with one or more receptors and direct these proteins to their final destinations in the mitochondrion. 2 separate inputs of energy are required for import: ATP hydrolysis in the cytosol, and a proton-motive force across the inner membrane. ATP hydrolysis in the cytosol is required to unfold precursor proteins and keep them in an unfolded or partially folded state, so that they can interact with the mitochondrial uptake machinery (MCB, fig.18-12). Several different unfolding proteins have been identified. They bind to precursor mitochondrial proteins and also to precursor ER proteins (MCB, 17-16), and couple the energy released by ATP hydrolysis in some way to prevent the bound proteins from folding. The only role of cytosolic ATP is to keep the precursor proteins in an unfolded state. However, in all cases in which a protein is taken into the mitochondrion (except for cytochrome c), a proton-motive force across the inner membrane is required.

13/c. The role of heat shock proteins in targeting proteins into different intracellular compartments.

- Precursor proteins presumably unfold before crossing the 2 mitochondrial membranes at a contact site. It is difficult to envisage how a folded, H_2O -soluble protein could straddle two (or even one) lipid bilayers while retaining its native 3-dimensional conformation. It is also hard to imagine a protein pore opening to accomodate globular proteins that vary greatly in size and shape without letting protons pass through it and thereby collapsing the electrochemical gradient across the inner membrane. On the other hand, all proteins have a similar conformation when unfolded, so they could be translocated by a common mechanism. But since the folded state of a protein is of lower free energy than the unfolded state (which is why polypeptides spontaneously fold up), unfolding a protein requires energy. ATP hydrolysis is believed to provide this energy certain members of the hsp70 family of stress-response proteins (MBC, pg. 420) are the "couplers" for an ATP-dependent protein-unfolding reaction on the cytosolic site of mitochondria.
- As we have seen, translocation of proteins into mitochondria (and peroxisomes) occurs post-translationally-after the protein is completed and released into the cytosol- whereas translocation across the ER membrane usually occurs during translation (co-translationally). This explains why ribosomes are bound to the ER membrane but not to the cytoplasmic surface of the other organelles. Recent studies of ER translocation, however, have shown that selected protein precursors can be imported into the ER after their synthesis has been completed. The import requires ATP hydrolysis but not ongoing protein synthesis (MBC, fig. 8-44). As for mitochondrial protein import, the ATP hydrolysis is thought to be required to unfold the protein as it passes through the membrane and a subclass of hsp70 stress-response proteins is required (MBC, pg. 420).
- Most ER precursor proteins in mammalian cells, however, cannot be imported into the ER once their synthesis has progressed beyond a certain point. It seems that these proteins fold up in a way that either masks the signal peptide or makes it impossible for the ER translocation machinery to unfold the protein. Co-translational import may have allowed these proteins to evolve without the folding constraints that presumably exist for proteins imported into other organelles.

14. The lac operon.

14/a. Induction of beta galactosidase activity.

- E. coli can use lactose as its sole source of C. An essential enzyme in the metabolism of this disaccharide is b-galactosidase, which hydrolyzes Lac to Gal and Glc (fig. 32-2). An E. coli cell growing on Lac contains several thousand molecules of b-galactosidase. In contrast, the number of b-galactosidase molecules per cell is fewer than 10 if E. coli is grown on other sources of C, such as Glc or glycerol. The presence of Lac in a culture medium induces a large increase in the amount of b-galactosidase in E. coli by eliciting the synthesis of new enzyme molecules rather than by activating a proenzyme (fig. 32-3). Hence, b-galactosidase is an inducible enzyme. 2 other proteins are synthesized in concert with b-galactosidase- namely galactoside permease and thiogalactoside transacetylase. The permease is required for the transport of Lac across the bacterial cell membrane. The transacetylase is not essential for Lac metabolism (uncertain role). Within an E. coli cell, the physiologic inducer is allolactose, which is formed from lactose by transglycosylation. The synthesis of allolactose is catalyzed by the few b-galactosidase molecules that are present prior to induction. Studies of synthetic inducers showed that some b-galactosides are inducers without being substrates of b-galactosidase, whereas other compounds are substrates without being inducers. For example, isopropylthiogalactoside (IPTG) is a nonmetabolizable inducer.

14/b. Discovery of the regulatory gene by using mutant bacterial strains.

- An important clue concerning the nature of the induction process was the finding that the amounts of the permease and the transacetylase increased in direct proportion to that of b-galactosidase for all inducers tested. Further insight came from studies of mutants, which showed that b-

galactosidase, the permease, and the transacetylase are encoded by 3 contiguous (adjacent) genes, called z, y, and a, respectively. An interesting class of mutants affecting all 3 proteins was isolated. These constitutive mutants synthesize large amounts of all 3 proteins whether or not inducer is present. It was deduced that the rate of synthesis of these 3 proteins is governed by a common element that is different from the genes specifying their structures. The gene for this common regulatory element was named i. Wild-type inducible bacteria have the genotype $i^+ z^+ y^+ a^+$, whereas the constitutive Lac mutants have the genotype $i^- z^+ y^+ a^+$. How does the i^+ gene affect the rate of synthesis of the proteins encoded by the z, y, and a genes? The simplest hypothesis was that the i^+ gene determines the synthesis of a cytoplasmic substance called a repressor, which is missing or inactive i^-. This idea was tested in a series of experiments involving partly diploid bacteria that contained 2 sets of genes for the Lac region. One set was on the bacterial chromosome whereas the other was on an F' sex factor, an episome, introduced by conjugation. For example, an i^+, z^- /F i^-, z^+ diploid was isolated. In this diploid, i^+, z^- is on the chromosome, whereas i^-, z^+ is on the episome. Is this diploid inducible or constitutive for b-glactosidase? In other words, does i^+ on the bacterial chromosome repress the expression of z^+ on the episome? The experimental result was clear-cut; the diploid was inducible rather constitutive. The same result was obtained for the diploid i^-, z^+ / F i^+, z^- . Hence, a diffusible repressor is specified by the i^+ gene. A diffusible repressor is an example of a trans-acting factor, one that can be encoded by a locus on a DNA molecule different from the one containing its target.

14/c. The function of lac operon.

- The genetic elements of the lac operon model are a regulator gene, an operator site, and a set of structural genes, (fig. 32-4). The regulator gene produces a repressor that can interact with the operator. Subsequent work revealed that the repressor is a protein. The operator is adjacent to the structural genes it controls. The binding of the repressor to the operator prevents the transcription of the structural genes. The operator and its associated structural genes are called an operon. For the lac operon, the i gene is the regulator gene, o is the operator, and the z, y, and a genes are the structural genes. The operon also contains a promoter site (denoted by p) for the binding of RNA polymerase. This site for the initiation of transcription is next to the operator. An inducer such as IPTG binds to the repressor, which prevents it from interacting with the operator. The z, y, and a genes can then be transcribed to give a single mRNA molecule that codes for all 3 proteins, (fig. 32-5). An mRNA molecule coding for more than one protein is known as a polygenic (or polycistronic) transcript.
- The lac repressor is a protein that binds to DNA carrying the lac operon but not to other DNA molecules. As predicted, IPTG prevents the binding of lac repressor to lac operator DNA.
- The rate constant for association of the repressor to the operator is strikingly high, indicating that the repressor finds the operator site by diffusing along a DNA molecule (a one-dimensional search) rather than by encountering it from the aqueous medium (a 3-dimensional search). Recall that RNA polymerase finds promoter sites in a similar way, (pg. 706). The lac repressor binds 4×10^6 times as strongly to its operator as to other sites on the chromosome. This high degree of selectivity is necessary because the genome contains a vast excess ($1{,}6 \times 10^5$) of competing sites.
- The base sequence of the lac operator region is very interesting: a total of 28 base pairs are related by a 2-fold axis of symmetry, (fig. 32-8), just as the subunits of lac repressor are. Thus, the symmetry of the repressor molecule matches that of its target site in DNA. Symmetry matching is a recurring motif in protein-DNA interactions.

14/d. The role of CAP-protein and cAMP in dual control of lac operon.

- It has long been known that E. coli grown on Glc, a preferred energy source, have very low levels of catabolic enzymes, such as b-galactosidase, galactokinase, arabinose isomerase, and tryptophanase. Clearly, it would be wasteful to synthesize these enzymes when Glc is abundant. The molecular basis of this inhibitory effect of Glc, called catabolite repression, has been elucidated. A key clue was the observation that Glc lowers the concentration of cAMP in E. coli. It was then found that exogenous cAMP can relieve the repression exerted by Glc. Subsequent studies revealed that cAMP stimulates the initiation of transcription of many inducible operons. It is interesting to note that cAMP serves as a hunger signal both in bacteria and in mammals. Recall that a low level of blood sugar stimulates the secretion of glycagon, which leads to elevated cAMP levels inside hormone -sensitive cells (pg. 458). In mammalian cells, cAMP acts by stimulating a protein kinase that phophorylates many target proteins, such as those controlling glycogen synthesis and breakdown (pg. 462). The action of cAMP in bacteria is very different. cAMP binds to CAP (Catabolite gene Activation Protein), a dimer of identical subunits. Proteolytic digestion experiments have shown that each subunit contains a DNA-binding domain and a cAMP binding domain. The complex of CAP and cAMP, but not CAP alone, stimulates transcription by binding to certain promoter sites. In the lac operon, CAP binds next to the site for RNA polymerase, as shown by nuclease digestion studies. Specifically, CAP protects nucleotides -87 to -49 from digestion, whereas RNA polymerase protects nucleotides -48 to +5, (fig. 32-9). In this numbering system, the first transcribed nucleotide is +1. CAP exhibits 2-fold symmetry, (fig. 32-10) that matches that of its DNA binding site CAP stimulates the initiation of lac mRNA synthesis by a factor of 50. How? The contiguous and nonoverlapping arrangement of the binding sites for CAP and RNA polymerase suggested that the binding of CAP to DNA creates an additional interaction site for RNA polymerase. Indeed, the binding of RNA polymerase to the promoter is enhanced by its energetically favorable contacts with bound CAP. In contrast, the lac repressor binds to nucleotides -3 to +21, which significantly overlaps the RNA polymerase site. Hence, the repressor sterically interferes with the binding of RNA polymerase. The cAMP-CAP complex acts similarly at other inducible catabolic operons. All contain the sequence TGTGA (upstream). Another common feature is that their -35 and -10 sequences differ from the

consensus sequence of strong promoters, (fig. 32-11). Evolution has weakened these promoters to make their operons dependent on a helper protein for efficient initiation of transcription. Thus, inducible catabolic operons are under dual control. A high level of expression requires the simultaneous presence of cAMP and a specific inducer, such as a galactoside for the lac operon. The specific inducer acts on a single operon, whereas the cAMP-CAP complex affects many.

15. The control of gene expression in eucaryotes I. Transcriptional control.

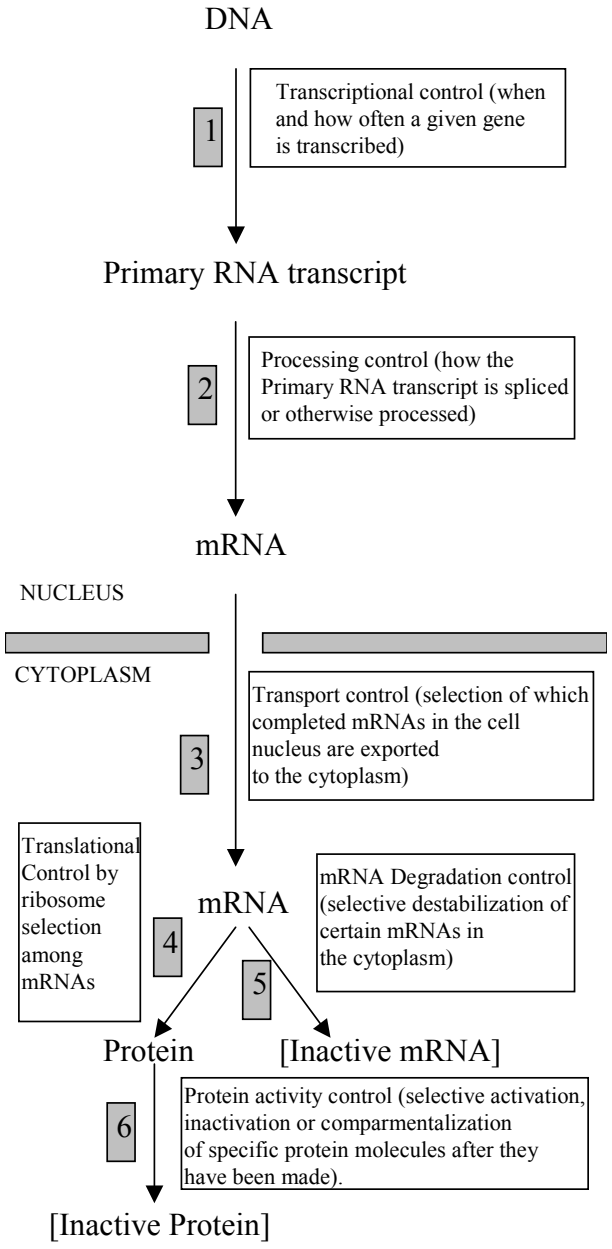
15/a. The components (signal, level, mechanism) and purpose of gene control in eucaryotic organisms.

- A logical analysis of gene control in eucaryotes requires answers to 3 questions, (MCB, fig. 11-2):
i) What are the molecular signals to which a specific gene responds?
ii) At which level (i.e. at which step or steps) in the chain of events leading from the transcription of DNA to the use of mRNA in protein synthesis is control exerted on the specific gene?
iii) What are the molecular mechanisms at each level of gene control and are they the same or different for different genes? That is, which individual molecules or cell structures are involved in the control of genes, and how do they exert their effects on individual genes?
Concerning i):

Class	Example	Target tissue / gene
Hormones:*	Growth hormone	Many cells.
Proteins	Prolactin	Secretory cells in breast tissue.
Steroids	Estrogens	Liver, brain, reproductive organs.
	Testosterone	Muscle, bone, skin, reproductive organs.
Circulating or secreted protein factors	Nerve growth factor	Axons (differentiating nerve cells)
	Epidermal growth factor	Many surface tissues- skin, eye, etc. as well as cultured cells of all types.
	Interleukins (lymphokines)	White blood cells.
	Erythroprotein	Red blood cell precursors.
	Interferon	Most epithelial cells, white blood cells.
	PDGFs (platalet-derived growth factors)	Many fibroblast cell types.
	Cell-cell or cell-matrix contacts**	Embryonic gut (endodermal cells interact with mesenchymal cells).
Environmental	Nutritional signals	Most genes for synthetic activity
	a) lower eucaryotes	(amino acid nucleic acid components) and for hydrolytic functions (phosphatases, saccharidases).
	b) animal cells	Genes for glyconeogenic enzymes in starvation; some genes for synthetic activity (there is some repression by excess products).
	Heat shock	Genes for specific proteins (induced); most other mRNAs (general translation suspended).
	Toxic substances:	
	Drugs, carcinogens (xenobiotics)	Genes for cytochrome P-450 and other detoxifying enzymes in liver.
	Heavy metals	Genes for metallothioneins in liver, kidney, and other tissues, and in single-celled eucaryotes.

Products of haemorrhage or inflammation	White blood cells, liver.

* Hormones have far-reaching effects on their target cells, both in adult organisms and during growth and development. Some but not all hormones operate by causing cells to change expression of individual genes or gene sets. Other hormones mainly affect the action of already existing enzymes or cell structures. It is clear that 2 general types of hormone molecules can serve as signals for gene control:
i) small fat-soluble molecules (e.g. steroid and thyroid hormones that dissolve in the plasma membrane and enter cells to cause their effect and
ii) polypeptides or proteins that bind to specific cells-surface receptors and in most cases exert their action without entering cells. (MCB, fig. 11-2) The receptors for these 2 types of gene signals are not alike. Steroid receptors, for example, are cytoplasmic molecules that bind steroids and move to the nucleus to act directly on transcription. In contrast, cell-surface receptors bind protein ligands outside the cells; this binding initiates a cascade of intracellular events that affect gene control inside the cell in a variety of ways.
** Cell-cell and cell-matrix contacts can act as signals to control genes. For example, in order for embryonic determination of most specialized cells to occur, 2 or more different cells types must make contact. Often the interacting cells come from different primordial cell layers (i.e. the mesoderm, endoderm, and ectoderm). One example is the primitive endodermal cells of the gut tube, which bud out at several regions and touch mesenchymal (mesodermal origin) cells. This pair wise interaction results in subsequent development of salivary glands, lungs, pancreas, and liver, depending on which section of the embryonic gut is involved. This type of gene-activating interaction sometimes depends on secretion of growth factors.
In many cases the interaction between cells probably depends on direct contact or takes place through intermediate contact. In the latter case, a cell of one type may form an extracellular matrix, and contact with this matrix may be the signal for a cell of another type. Although the molecules of the cell surface (e.g. integrins) and of the extracellular matrix (e.g. lamins, fibronectins, and cell-adhesion molecules) are being actively studied, it is not yet known which cell- surface contacts function as gene-control signals.
Concerning ii)and iii):



1: a) chromosome condensation
b) DNA methylation
c) gene regulatory proteins (upstream promoter elements, master proteins)
d) gene activated by external signals
2: a) Alternative RNA splicing
b) Change of poly-A addition
3: (Adenovirus RNA transport)
4: a) Translation repressor protein-ferritin
b) Translational enhancer -polyovirus
c) Translational frameshift-retrovirus
d) Phosphorylation of eIF2
5: a) Specific sequences at 3'-end
b) Untranslated regions (transferrin, histones)
c) Activation of nucleases (tubulin)

- Single-cell eucaryotes, such as yeast cells, seem to be designated only, or mainly for the purpose of growth; yeasts also possess many genes that are controlled in response to environmental variables (e.g., nutritional status, oxygen tension, and temperature). Even in the organs of higher animals -for example, in mammalian liver- some genes can respond reversibly to external stimuli such as noxious chemicals. In general, however, metazoan cells are protected from immediate outside influences; that is, most cells in metazoans experience a fairly constant environment. Perhaps for this reason, genes that are devoted to responses to environmental changes constitute a much smaller fraction of the total number of genes in multicellular organisms than in single-cell organisms.
- The most characteristic and exacting requirement of gene control in multicellular organisms is the execution of precise developmental decisions so that the right gene is activated in the right cell at the right time. In most cases, once a developmental step has been taken by a cell, it is not reversed. Thus, these decisions are fundamentally different from bacterial induction and repression. Many differentiated cells (e.g. skin cells, red blood cells, lens cells of the eye, and antibody-producing cells) march down a pathway to final cell death in carrying out their genetic programs, and they leave no progeny behind. These fixed patterns of gene control leading to differentiation serve the need of the whole organism and not the survival of an individual cell.

15/b. Gene regulatory sequences, (enhancer, upstream promoter element) and gene regulatory proteins.

- Initiation of transcription of eucaryotic genes by RNA polymerase II, even when the TATA box and TATA factor are present, is not a frequent event. The process is accelerated by activators, which are required for normal transcription of many, perhaps all, eucaryotic genes. In yeast and other simple organisms, a single required upstream activating sequence (UAS) can be sufficient; the UAS is usually located near a gene, and often a single protein factor binds to a 15-to 20-bp segment in the UAS, (MCB, fig. 11-10). In contrast, all mammalian genes contain multiple 15-to 20-bp protein-binding sites each of which is necessary for maximal transcription. Many of the binding sites in DNA have dyad symmetry, suggesting that binding proteins may act as dimers.
- As it is known, some gene-activating sites in eucaryotes are located at great distances (up to 50 kb) from RNA start sites, usually upstream but also sometimes downstream; these are called enhancer sites. Regulatory sites closer to the start are referred to as promoter sites. Now that the proteins which recognize many of these sites have been identified, it has become clear that there is no real distinction other than distance between these regions: the same proteins can bind close to or far from RNA start sites. It is widely assumed that even distantly bound proteins can affect transcription and that looping of the DNA can bring a distantly bound protein into proximity with the RNA start site.
- Eucaryotic cells contain a large set of sequence-specific DNA-binding proteins whose main function is to turn genes on or off. Each of these gene regulatory proteins is present in relatively few copies per cell (about 1 molecule per 3000 nucleosomes, or 10^4 copies per mammalian cell) and recognizes a particular DNA sequence that is usually about 8 to 15 nucleotides long. The binding of these proteins to the DNA can either facilitate (positive regulation) or inhibit (negative regulation) transcription of an adjacent gene, (MBC, fig. 10-6). The different cell types in a multicellular organism have different mixtures of gene regulatory proteins, which causes each cell type to transcribe different sets of genes.
- Combinations of a few gene regulatory proteins could specify a large number of cell types; the essence of combinatorial gene regulation is illustrated in fig. 10-7 (MBC), in which each numbered element represents a different gene regulatory protein. In this scheme, one initial cell type gives rise to 2 types of cells, A and B, which differ only in the presence of gene regulatory protein 1 in one but not in the other. The subsequent development of each of these cells leads to the additional production in some cells first of gene regulatory proteins 4 and 5. In the end, 8 cell types (G through cell N) have been created with 5 different gene regulatory proteins. With the addition of two more gene regulatory proteins to the scheme shown in fig. 10-7, (6 and 7) 16 cell types could be generated at the next step. By the time 10 more such steps occurred, slightly more than 10.000 cell types could have been specified, in principle, through the action of only 25 different gene regulatory proteins.
- Input from several different regulatory proteins usually determines the activity of a gene; the scheme for combinatorial gene regulation shown in fig. 10-7 (MBC), at first glance, seems to predict simple additive differences between the cells of successive generations. One might imagine, for example, that the addition of regulatory protein 2 to cell C and to cell E would add to each of these cells the same set of additional proteins- namely those encoded by genes activated by the regulatory protein 2. Such a view is incorrect for a simple reason: combinatorial gene regulation is more complicated than that because different gene regulatory proteins interact with one another. Even in bacteria the interaction of 2 different regulatory proteins is often needed to turn on a single gene. In higher eucaryotes whole clusters of gene activator proteins generally act in concert to determine whether a gene is to be transcribed. By interacting with gene regulatory protein 1, for example, protein 2 can turn on a different set of genes in cell E from those it turns on in cell C. This is presumably why a single steroid-hormone-receptor protein (an example of a gene regulatory protein) modulates the synthesis of different sets of proteins in different types of mammalian cells. The particular change in gene expression caused by the synthesis of a given gene regulatory protein in a developing cell will depend, in general, on the cell past history, since this history will determine which gene regulatory proteins are already present in the cell, (MBC, fig. 10-8).
- Not all gene regulatory proteins are equal, however. The regulatory network contains master gene regulatory proteins, each of which has a decisive coordinating effect in controlling many other genes (MBC, fig. 10-9). Early evidence for a role of master gene regulatory proteins in vertebrates came from the observation that the absence of a single gene regulatory protein (the receptor protein for the

steroid hormone testosterone) causes a human with a male (XY) genotype to develop as an almost perfect female.
- A single master gene regulatory protein can convert a fibroblast to a myoblast.
- Upstream promoter elements are DNA sequences near the RNA start site that are required for the initiation of gene transcription; they generally extend for about 100 nucleotide pairs and they include the TATA box. (see also answers for 5/c. and 6/a.).
- An enhancer is defined as a regulatory DNA sequence that:
i) Activates transcription from a promoter linked to it, with synthesis beginning at the normal RNA start site (it can also inhibit transcription, vide infra).
ii) Operates in both orientations (normal or flipped).
iii) Functions even when moved more than 1000 nucleotide pairs from the promoter and from either an upstream or a downstream position.
- We know that most genes in higher eucaryotes are regulated by the combination of a nearly upstream promoter element plus one or more enhancers farther away from the RNA start site, (MBC, fig. 10-22/a). Both enhancers and upstream promoter elements bind sequence-specific DNA- binding proteins that mediate their effects. As we shall see, some of these gene regulatory proteins seem to be present in all tissues, whereas others are confined to particular cell types.
- The activity of each enhancer and upstream promoter element can generally be pinpointed to a segment of DNA that is 100 to 200 nucleotide pairs long. An element of this size contains binding sites for multiple proteins. The function of both kinds of regulatory elements depends on the binding of specific gene regulatory proteins that are often restricted to particular cell types. Only some (such as the SV40 virus enhancer) will function in almost any cell type; most function best in the specific cell types that express the gene with which they are normally associated.
Studies of enhancers and upstream promoter elements have led to the following general conclusions:
i) Each of these regulatory elements has a modular design, consisting of a series of specific nucleotide sequences- each 8 to 15 nucleotides long -that bind a corresponding series of gene regulatory proteins. Some of these proteins are found only in a restricted set of cell types, whereas others are present in most or all cell types.
ii) Some gene regulatory proteins activate transcription when they bind; others inhibit it. The net effect of a regulatory element depends on the combination of proteins bound, and for a given element the effect can change as a cell develops. An enhancer element, for example, can either activate or inhibit transcription, (MBC, fig. 10-22b). In this sense, the original name given to these elements is a misnomer.
iii) Enhancer elements and upstream promoter elements seem to bind many of the same proteins, which may mean that both types of element affect transcription by a similar mechanism.
iv) When a newly defined regulatory element from a vertebrate gene is analyzed, many of the proteins that bind to it turn out to be previously described regulators of other genes. This suggests that a relatively small number of gene regulatory proteins may control transcription in higher eucaryotes, (MBC, table 10-1).
- The proteins bound to the upstream promoter element of a gene cooperate with those bound to enhancers in controlling transcription. In general, their net effect on gene activity will reflect the result of opposing activating and repressing influences, (MBC, fig. 10-22).

15/c. General procedure for detecting and isolating DNA regulatory sequences. Gel retardation, footprinting.

- General procedure for detecting and isolating DNA regulatory sequences: MCB, fig.11-9.
- Gel retardation: A DNA molecule is highly negatively charged and will therefore migrate more rapidly towards a positive electrode when it is subjected to an electric field. When analyzed polyacrylamide-gel electrophoresis, DNA molecules are separated according to their size because smaller molecules are able to penetrate the fine gel meshwork more easily than large ones (MBC, pg. 184). Protein molecules bound to a DNA molecule will cause it to move more slowly through the gel; in general, the larger the bound protein, the greater the retardation of the DNA molecule. This phenomenon provides the basis for a gel retardation assay (also called a gel-mobility shift assay), which allows even trace amounts of a sequence-specific DNA-binding protein to be readily detected. In this assay a short DNA fragment of specific length and sequence is radioactively labeled and mixed with a cell extract; the mixture is then loaded onto a polyacrylamide gel and subjected to electrophoresis. If the DNA fragment corresponds to a chromosomal region where multiple sequence-specific proteins bind, autoradiography will reveal a series of DNA bands, each retarded to a different extent and representing a distinct DNA-protein complex. The proteins responsible for each band on the gel can then be separated from one another by subsequent fractionations of the cell extract, (MBC, fig. 9-9).
- Footprinting: (Here, this procedure is presented but not according to fig.11-9 of MCB): Experiments with purified DNA fragments containing promoters have proved that RNA polymerase touches DNA at certain bases in the conserved regions of promoters. In these experiments, RNA polymerase is bound to DNA, which is then subjected to DNase attack or to chemical reagents that modify the chain at specific sites unprotected by the polymerase; further chemical treatment breaks the chain at the modified base. When the Dnase-digested or chemically broken DNA is analyzed by gel electrophoresis, the regions protected by RNA polymerase appear as gaps ("footprints") in the array of bands, (MCB, fig. 7-4, 7-5).

15/d. DNA binding proteins.

- Helix-turn-Helix proteins (and homeoboxes): Helix -turn helix proteins bind as dimers and all have within their structures 3 α-helical regions separated by short turns. A protein helix in each of the dimers occupies the major groove in 2 successive turns of the DNA helix.
- Example: These proteins are encoded by regulatory genes that control early segmentation and body part development in Drosophila embryos. The Drosophila genes include antennapedia (Antp), fushi tarazu (ftz), and ultrabithorax (Ubx). Mutations in the Antp or Ubx genes can cause substitution of one developmental pathway for another; for

this reason, they are called homeotic genes. When the nucleotide sequences of these and several other Drosophila homeotic genes were completed, it was clear that a region encoding about 60 amino acids was remarkably conserved in all these genes. This region called the homeobox, is also conserved well enough to be easily recognized in genes from frogs and mammals.

- Zinc-finger proteins: The first eucaryotic positive-acting regulatory protein to be well-characterized was transcription factor IIIA ($TF_{III}A$), which is required for RNA polymerase III transcription of the 55-rRNA genes. This protein has 9 repeated domains that contains Cys and His spaced at regular intervals, (MCB, fig. 11-12). The purified protein has Zn associated with it, and the Zn is required for activity. Furthermore, Cys and His are known to be the most common ligands in protein for Zn. In the proposed Zn-finger folded structure of $TF_{III}A$ shown in fig. 11-13 (MCB), the repeated domains form loops in such a way that a Zn ion is bound between a pair of Cys and a pair of His. A Dhe or Tyr residue and a Leu residue occur at nearly constant positions in the loops, which are known to be required for DNA binding. Many regulatory proteins have been found to contain Zn-fingers. These include a yeast protein that regulates a cytochrome c gene, and a whole class of proteins that bind steroid hormones in vertebrates. Thus, the Zn-finger motif appears to be a common one in eucaryotic regulatory proteins.

- Amphipathic helical proteins: Leucine-zipper and helix-loop-helix proteins: All of these proteins form dimers and have a region rich in basic amino acids near the dimerization domains. The first protein of this large group to be studied extensively was isolated from rat liver cell extracts. This protein was called C/EBP because of its ability to act as a viral Enhancer Binding Protein and because it was also thought initially to bind to CCAAT boxes. The carboxyl-terminal 100 amino acids of C/EBP, has in its sequence a series of residues that can form an α helix, that is necessary in DNA binding. If the amino acid residues between 315 and the C-terminus at 359 are positioned on a helical wheel, every 7th amino acid-a Leu-is brought into register on one side of the helix (MCB, fig. 11-14a). The C/EBP only binds to DNA as a dimer. The helical region is amphipathic: that is, it favours lipids on one side and H_2O on the other. One side of the helical region of C/EBP has charged amino acid; this charged side would favour facing the H_2O interface. The array of Leu on the other side presents a very hydrophobic surface. See fig. 11-14b MCB. The zipper region is required for dimer formation, whereas the basic region constitutes the DNA-binding domain. So far, only homodimers of C/EBP have been detected. However, in vitro studies have shown that heterodimers can form between C-Jun and C-Fos, 2 protooncogene proteins that contain Leu-zipper regions.

- A similar but clearly distinguishable design has been recognized in a group of proteins termed helix-loop-helix proteins (e.g. MyoD protein). The helical regions are required for dimerization and the basic amino acid regions are also required for DNA binding; these characteristics are reminiscent of the Leu-zipper proteins, but the amino acid sequences in the 2 groups of proteins are not alike. The helix-loop-helix proteins easily form heterodimers, (MCB, fig. 11-14c).

- Heterodimers of dissimilar proteins: One widely distributed group of such structurally distinct factors binds to the CCAAT consensus sequence found in some eucaryotic genes. These CCAAT-binding proteins are heterodimers of dissimilar proteins, one of which binds nearly to DNA on its own but strongly in the presence of the second protein. The overall structure of one of these protein pairs has apparently been conserved over the entire range of eucaryotic evolution.

- DNA-binding proteins can increase transcription in vitro and in vivo, (MCB, fig. 11-15, 11-16).

15/e. The role of chromatin condensation, decondensation and DNA methylation in gene expression in vertebrates.

- Some regions of chromosomes fail to decondense during interphase, appearing to maintain the highly condensed conformation of a metaphase chromosome. These regions were called heterochromatin to distinguish them from the rest of the chromatin in the interphase nucleus (which was called euchromatin). Some chromosomal regions are condensed into heterochromatin in all cells. In humans this constitutive heterochromatin is localized around the centromere of each mitotic chromosome, where it can be detected as darkly staining bands, (MBC, fig. 10-36), In some other mammals, constitutive heterochromatin is also located in specific bands on chromosome arms. During interphase the regions of constitutive heterochromatin can aggregate to form chromocenters, (MBC, fig. 9-43). In mammals, the number and arrangement of such chromocenters varies with all type and developmental stage. Most regions of constitutive heterochromatin contain relatively simple serially repeated DNA sequences, called satellite DNAs. These highly repeated DNA sequences are not transcribed and their function, and that of the condensed interphase chromatin structure that they form: unknown. Other regions of DNA are unusually condensed during interphase in some cell types of an organism but not in others. These regions also are thought not to be transcribed, but they do not consist of simple sequences. The total amount of this facultative heterochromatin is very different in different cell types: embryonic cells seem to have very little, whereas some highly specialized cells have a great deal. This suggests that, as cells develop, progressively more genes can be packaged in a condensed form in which they are no longer accessible to gene activator proteins.

- In mammals all female cells contain 2 X chromosomes, while male cells contain one X and one Y chromosome. Presumably, because a double dose of X-chromosome products would be lethal, the female cells have evolved mechanism for permanently inactivating one of the 2 X chromosomes in each cell forming a Barr body. This body replicates late in S phase, and most of its DNA is not transcribed. Because, the inactive X chromosome is faithfully inherited, every female is a mosaic composed of clonal groups of cells in which only the paternally inherited X chromosome (Xp) is active and a roughly equal number of groups of cells in which only the maternally inherited X chromosome (Xm) is inactive. In general, the cells expressing Xp and those expressing Xm are distributed in small clusters in the adult animal, reflecting the tendency of sister cells to remain close neighbours during the later

stages of embryonic development and growth after an early period when there is more mixing, (MBC, fig. 10-37). The condensation process that forms heterochromatin in an X chromosome has a tendency to spread continuously along a chromosome. This has been demonstrated by studies with mutant animals in which one of the X chromosomes has become joined to a portion of an autosome. In the mutant chromosomes, regions of the autosome adjacent to an inactivated X chromosome are often condensed into heterochromatin and the genes they contain thereby inactivated in a heritable way. This suggests that X-chromosome inactivation occurs by a cooperative process, which can be thought of as chromatin "crystallization" spreading from a nucleation site on the X chromosome. Once the condensed chromatin structure is established, a process resembling that in fig. 10-35 (MBC) could cause it to be faithfully inherited during all subsequent replications of the DNA. The condensed X chromosome is reactivated in the formation of germ cells in the female; thus, no permanent change can have occured in its DNA.

- A phenomenon analogous to X-chromosome inactivation occurs in Drosophila, where breakage and rejoining events that place the middle of a region of heterochromatin next to a region of euchromatin tend to inactivate the nearby euchromatic genes. Moreover, while the extent of the "spreading effect" is different in different cells, the inactivated zone established in an embryonic cell is stably inherited by all of the cell progeny, (MBC, fig. 10-38). Studies of this position effect variegation in Drosophila have revealed that the extent of spreading is reduced in flies bred to contain extra amounts of constitutive heterochromatin, which suggests that the supply of special proteins needed to produce heterochromatin can become limiting in flies with extra heterochromatin. Regardless of its molecular basis, the packing of selected regions of the genome into heterochromatin is a type of gene regulatory mechanism that is not available to bacteria. The crucial feature of this uniquely eucaryotic form of gene regulation is the storing of the stable memory of gene states in an inherited chromatin structure rather than in a stable feedback loop of self- regulating gene regulatory proteins that can diffuse from place to place in the nucleus. The evidence to be described next, however, suggests that the expression of individual genes is often regulated by nearby control sequences and is not absolutely dependent on the more global chromosomal environment.

- A specific chromosomal environment is often required for the optimal expression of a gene but not for it to be turned on in the correct cells; it has been discovered that the strong position effects observed for the human b-globin gene in the erythrocytes of transgenic mice are eliminated and full transcriptional activity recovered if a DNA fragment that is normally located 50,000 nucleotide pairs away from the promoter is attached to the gene. This fragment which contains a cluster of 6 nuclease-hypersensitive sites (see MBC, pg. 498), seems to affect an entire cluster of globin genes. We shall refer to as a domain control region. The general observation of position effects makes it likely that many vertebrate genes will require specific distant sequences for quantitatively correct expression. We next discuss what these sequences may do.

- An initial local chromatin decondensation step might be required before eucaryotic genes can be activated; people with a particular form of thalassemia, have a large deletion of DNA upstream from the b-globin gene. The deleted region (about 100,000 nucleotide pairs) contains several other b-like globin genes, as well as the domain control region identified in the transgenic mouse experiments, (MBC, fig. 10-39 A). Although the b-globin gene is intact, its level of transcription is greatly reduced. In nuclease digestion experiments this gene, unlike a normal b-globin gene, is digested at the same slow rate as bulk chromatin and is thus presumably packaged into active chromatin, (MBC, pg. 511). The normal homologue in the same erythrocyte lacks this deletion, and its entire cluster of b-like globin genes-including at least 90,000 nucleotide pairs of DNA-seems to be decondensed as active chromatin by the time the first of these genes (the E-globin gene) begins to be transcribed (see MBC, fig. 10-39 B). Results such as these have been used to support a 2-step model for the induction of transcription in higher eucaryotic genes. In stage 1 all chromatin in a region tens of thousands of nucleotide pairs long is converted into a relatively decondensed, "active" form, (MBC, fig. 10-40). This step might be triggered by a special type of gene regulatory protein that causes a structural change in neighbouring chromatin that is propagated from a domain control region throughout an entire looped domain (see MBC, pg. 503) of chromatin. In stage 2 gene regulatory proteins that act as enhancer and upstream promoter elements regulate the transcription of specific genes within regions of exposed active chromatin. Because of these more local controls, initially the human E-globin gene is expressed in the embryonic yolk sac, then the 2 γ-globin genes are expressed in the fetal liver, and finally, the b-globin gene is turned on near the time of birth, (see MBC, fig. 10-39 B).

- The model for gene activation outlined in fig. 10-40 (MBC) implies that some of the gene regulatory proteins in higher eucaryotes have functions that differ from those of their bacterial analogues. Instead of loading RNA polymerase (or its transcription factors) directly onto a nearby promoter sequence, (see MBC, fig. 10-27), some sequence- specific DNA-binding proteins may function solely to decondense the chromatin in a local chromosomal domain or to remove a nucleosome from an adjacent enhancer or promoter in order to provide access for gene regulatory proteins of a more familiar kind. We are not certain, however, that this is so; it remains possible that the observed differences in the chromatin structure of active genes are an automatic consequence of the assembly of transcription factors and / or RNA polymerase onto a promoter sequence rather than being a required prerequisite for any of these events.

- Vertebrate DNAs contain 5-methylcytosine (5-mc), which has the same relation to cytosine that T has to U and likewise has no effect on base-pairing, (MBC, fig. 10-44 A). The methylation is restricted to C bases in the sequence CG; since this sequence is base-paired to exactly the same sequence. (In opposite orientation) on the other strand of the DNA helix, a simple mechanism permits a preexisting pattern of DNA methylation to be inherited directly by a template process. An enzyme called a maintenance methylase acts only on those CG sequences that are base-paired with a GC sequence that is already methylated. As a result, the preexisting pattern of DNA methylation will be inherited directly following DNA replication. (MBC, fig.

10-45). The automatic inheritance of 5 mc residues raises a "chicken and egg" problem: where is the methyl group first added in a vertebrate organism? Experiments show that methyl groups will be added to nearly every CG site in a fully unmethylated DNA molecule that is infected into a fertilized mouse egg (although an important exception will be described below). Thus, the vast majority of a vertebrate genome starts out heavily methylated. Since the maintenance methylase normally cannot methylate fully unmethylated DNA, a novel establishment methylase activity must be present in the egg. Since the establishment methylase soon disappears, the DNA in the cells of developing tissues relies on the maintenance methylase for retention of its methylated nucleotides.

- What is the effect of GC methylation? Studies indicated that, in general the DNA of inactive genes is more heavily methylated than that of active genes. Moreover, an inactive gene that contains methylated DNA usually will lose many of its methyl groups after the gene has been activated. Evidence that the change in methylation affects gene expression comes from experiments in which a nucleoside containing the base analogue 5-azacytosine (5-azaC, MBC, fig. 10-44 B), is added for a brief period to cells in culture. The 5-azaC, which cannot be methylated, is incorporated into DNA, where it acts as an inhibitor of the maintenance methylase, thereby reducing the general level of DNA methylation. In cells treated in this way, selected genes that were previously inactive become active and at the same time, acquire unmethylated C residues. Once activated, the active state of these genes usually is maintained for many cell generations in the absence of 5-azaC, implying that the initial methylation of the genes helped to maintain their inactivity. Studies have shown that DNA methylation plays a quite subsidiary part in cell diversification. The important developmental decisions apparently are made by gene regulatory proteins that can turn genes on or off regardless of their methylation status. The female X chromosome, for example, is first condensed and inactivated and only later acquires an increased level of methylation on some of its genes. Conversely, several liver-specific genes are turned on during development while they are fully methylated; only later does their level of methylation decrease.

- DNA transfection experiments have helped to reconcile these contrasting observations on the role of DNA methylation in gene expression. A tissue-specific gene coding for muscle actin, for example, can be prepared in both its fully methylated and fully unmethylated form. When these 2 versions of the gene are introduced into cultured muscle cells, both are transcribed at the same high rate. When they are introduced into fibroblasts, which normally do not transcribed the gene, the unmethylated gene is transcribed at a low level but still one much higher than either the exogenously added methylated gene or the endogenous gene of the fibroblast (which is also methylated). These experiments suggest that DNA methylation is used in vertebrates to reinforce developmental decisions made in other ways.

- Because of the way DNA repair enzymes work, methylated C residues in the genome tend to be eliminated in the course of evolution. Accidental deamination of an unmethylated C gives rise to U, which is not normally present in DNA, and thus is recognized easily by the DNA repair enzyme uracil DNA glycosylase, excised, and then replaced with a C. But accidental deamination of a 5-mC cannot be repaired in this way, for the deamination product is a T and so indistinguishable from the other nonmutant T residues in the DNA. Thus, those C residues in the genome that are methylated tend to mutate to T over evolutionary time. Since the divergence of vertebrates from invertebrates about 400.000.000 years ago, more than 3 out of every 4 CGs have been lost in this way, leaving vertebrates with a remarkable deficiency of this dinucleotide. The CG sequences that remain are very unevenly distributed in the genome; they are present at 10 to 20 times their average density in selected regions that are 1000 to 2000 nucleotide pairs long, called CG islands. These islands surround the promoters of so-called "housekeeping genes" -those genes that encode the many proteins that are essential for cell viability and are therefore expressed in most cells, (MBC, fig. 10-46). These genes are to be contrasted tissue-specific genes, which encode proteins needed only selected types of cells.

- The distribution of CG islands can be explained readily as a secondary effect of the introduction of CG methylation as a way of reducing the expression of inactive gene in vertebrates, (MBC, fig. 10-47). In germ cells, all tissue-specific genes (except those specific to eggs and sperm) are inactive and methylated; over long periods of evolutionary time, their methylated CG sequences are lost through accidental deamination events that are not correctly repaired. The CG sequences in the regions surrounding the promoters of all active genes in germ cells, however, including the housekeeping genes, are kept demethylated, and so they can be readily repaired after spontaneous deamination events in the germ line. These genes are thought to be recognized by sequence-specific DNA-binding proteins present in the germ cells that remove any methylation near their promoters (this was the important exception said to be mentioned a few pages ago). Experiments with cloned genes show that only the CGs in CG islands remain unmethylated when fully unmethylated DNAs are injected into a fertilized mouse egg. The mammalian genome, (about $3x10^9$ nucleotide pairs), contains an estimated 30,000 CG islands, each about a thousand or so nucleotide pairs in length. Most of the islands mark the 5' ends of a transcription unit, and thus, presumably, a gene.

<u>15/f. Homeotic genes, master genes and master proteins in the development of multicellular organisms.</u>

- Homeotic genes have their site of gene activity in the embryo, and they control the identity of the segments, i.e. Drosophila. (In Drosophila these "segments" are called imaginal disks; imago is the name given to the adult form of an insect). Mutation in such a region will cause part of the fly to form structures normally formed in other imaginal disks (homeotic mutations). One such mutation in Drosophila, called antennapedia, produces a single master gene regulatory protein aberrantly in the group of cells that would normally make an antenna and thereby causes these cells to switch to making a leg instead, so that the adult fly has a leg growing out of its head. Early evidence for a similar role for master gene regulatory proteins in vertebrates came from the observation that the absence of a single gene regulatory protein (the receptor protein for the

testosterone) causes a human with a male (XY) genotype to develop as an almost perfect female, (see MBC, fig. 10-9 for mode of action of a master gene regulatory protein). And what is the connection between a homeotic gene and a master gene regulatory protein? First, we have to answer that a master gene regulatory protein (or simply master protein) is coded by a master gene and secondly, we have to add that a homeotic gene is a type of the master genes family. Other master genes are i) the maternal effect genes, which their site of activity is on the maternal tissues where they initiate pattern formation by activating regulatory genes in nuclei in certain locations of embryo; ii) the gap genes which act in the embryo by causing alternating segments (imaginal disks, since we are always speaking about Drosophila) to be missing; some of these gap genes may influence the activity of pair rule genes, segment polarity genes and homeotic genes; iii) pair rule genes which act in the embryo and they cause parts of segments to be missing; some of these pair rule genes may influence the activity of segment polarity genes, and homeotic genes; iv) the segment polarity genes which act in the embryo and they delete part of every segments they also replace with mirror image of remaining structure and they may influence the activity of the homeotic genes.

- (This parenthesis is opened for the shake of the homeotic genes; When the DNA sequences of a number of homeotic genes were analyzed, it was discovered that there is a short DNA sequence of approximately 180 base pairs that is characteristic of many homeotic genes. This sequence has been termed “the homeobox”, (Villee Fig. 16-12). Using the homeobox sequence of bases as a molecular probe it was made possible to clone new homeotic genes in Drosophila that had not been previously identified. Surprisingly, the homeobox probe also detected homologous DNA sequences in a wide range of other organisms, including humans. The hoemobox sequences of a large number of genes have been determined. Comparisons of those DNA sequences have shown that the sequence itself is highly conserved during evolution and shows remarkable similarities between organisms as diverse as sea urchins, yeast, and humans. The first clues to the function of the proteins that are coded for by genes containing homeoboxes came from computer-generated searches; those comparisons showed that the amino acid sequences coded for by the homeoboxes (known as homeodomains) are homologous to amino acid sequences in certain DNA-binding proteins. Most, if not all, of these "homeotic" proteins contain a helix-turn-helix sequence in the homeodomain that corresponds to a similar domain found in certain regulatory proteins in procaryotic cells and their viruses. Other evidence that the homeotic genes code for DNA-binding regulatory proteins comes from the finding that all of the Drosophila proteins that contain homeodomains accumulate in the nucleus. [It is worth mentioning here that most gene regulatory proteins contain domains with distinct functions, (MBC, fig. 10-24, fig. 10-25)]. Some of the segmentation genes that act earlier in development code for a "zinc-finger" type of DNA binding regulatory protein, while other segmentation gene proteins have other types of amino acid sequences that are common to several different protein).

- See also MBC, fig. 10-48 and fig. 10-49. (corresponding paragraph: complex patterns of gene regulation are required to produce a multicellular organism, pg. 586).

16. Gene expression in eucaryotes II. Processing control, transport control, mRNA degradation control.

16/a. Processing control-alternative RNA splicing-changes of poly A addition.

- A large proportion of higher eucaryotic genes produce multiple proteins by means of alternative RNA splicing. (see pg. 536 MBC- it is also the answer for question 6/d.). When different splicing possibilities exist at several positions of the transcript, a single gene can produce dozens of different proteins; usually, however, the splice alternatives are more limited, and only a few different proteins are synthesized from each transcription unit. In some cases alternative RNA splicing occurs because there is an "intron ambiguity": the standard spliceosome mechanism for removing introns is unable to distinguish clearly between 2 or more alternative pairings of 5' and 3' splice sites, so that different choices are made haphazardly on different occasions (constitutive form of alternative splicing). In many cases alternative RNA splicing is regulated rather than constitutive. Splice-site selections are determined by the cell. Thus, the same primary RNA transcript can produce a different protein (or set of proteins) in different cell types, according to the needs of the organism. Many proteins are produced in tissue - specific forms in this way, including components of the i) extracellular matrix (fibronectin), ii) cytoskeleton (tropomyosin), iii) plasma membrane (N-CAM Neuron-Cell Adhesion Molecule), and K^+ channels, iv) nucleus (MBC, pg. 818), and v) intracellular cell-signaling pathways, such as C-kinase and the Tyr kinase encoded by the src proto-oncogene, (MBC, fig. 10-51). In general, the exon changes caused by alternative RNA splicing do not produce radically different proteins. Instead, they produce a set of proteins of comparable function, called protein isoforms, which are modified to suit the particular tissue, (e.g. troponin T and α-tropomyosin in different muscle fibers). The changes may determine which other proteins the molecule interacts with, for example, leaving although the catalytic or structural domains unaltered. There are exception to this rule, however. For example, a major change of coding exons resulting from alternative RNA splicing leads to the production of 2 entirely different polypeptide hormones from the same transcription unit-calcitonin in the thyroid gland and calcitonin-gene-related peptide (CGRP) in neural tissue.

- Some genes are constantly transcribed in all cells, but because the constitutive RNA-splicing mechanism produces an mRNA that codes for a nonfunctional protein, the gene is expressed only in selected cells in which a specialized splicing reaction occurs. This type of gene regulation has been especially well characterized in Drosophila. The ability of the P element, (MBC, pg. 268), to transpose only in germ cells, for example, is due to its failure to produce a functional transposase in somatic cells; this failure has in turn been traced to the presence of an

intron in the transposase mRNA that seems to be removed only in germ cells.
- Regulated changes in the choice of RNA splice sites are presumed to be mediated by the binding of tissue-and gene-specific proteins or RNA molecules to the growing RNA transcripts. Since the splice -sites selected in both the constitutive and regulated pathways of RNA splicing all seem to share the standard consensus sequences described previously, (MBC, fig. 9-80), the binding of the specific component must change the conformation of the RNA transcript so as preferentially to mask or expose preexisting splice sites. An important new finding that may shed light on this subject concerns variation in the protein components of sn RNPs in the spliceosomes that process hnRNA in the nucleus. The snRNPs from many cells contain a protein called B; the snRNPs from brain cells contain another protein termed N (for neuron) which have a similar but not identical amino acid sequence (MCB, fig. 11-44). Tumor cell lines from the anterior pituitary that make only CGRP are known, as are thyroid tumor lines that make only calcitonin. The sn RNPs from the pituitary lines have the N protein, and those from the thyroid lines have the B protein. It is possible that this difference causes a difference in the RNA-processing ability of the different cell types. In addition to differences in the protein components of snRNPs, the RNA components may vary. A careful analysis of some of the several hundred genes encoding U1 snRNAs showed that about 1 in 10 of these genes contained a sequence variation. This variation is also found in the U1 snRNAs produced by cultured cells. These sequence differences have not yet been shown to correlate with any cell-specific differential processing, but it is an obvious possibility.
- A change in the site of RNA transcript cleavage and poly-A addition can change the carboxyl terminus of a protein; in eucaryotes, the 3' end of an mRNA molecule is not determined by the termination of RNA synthesis by the RNA polymerase; instead, it is determined by an RNA cleavage reaction that is catalyzed by additional factors while the transcript is elongating. The site of this cleavage can be controlled so as to change the carboxyl terminus of the resultant protein (which is encoded by the 3' end of the mRNA). In procaryotes, producing a longer RNA transcript can only add more amino acid onto a protein chain. In eucaryotes, however, RNA splicing can create mRNAs that cause the original carboxyl terminus of a protein to be removed entirely and to be replaced with a new one after a long transcript is made. A good example of this type of change mediates the switch from the synthesis of membrane-bound to secreted antibody molecules during the development of B lymphocytes.

16/b. Transport control of mRNA.

- RNA transport from the nucleus might be regulated; RNA export through the nuclear pores is an active process. If this export depends on the specific recognition of the transported RNA molecule (or of a protein or RNA molecule bound to it) by a receptor protein in the nuclear pore complex, RNAs that lack this recognition signal would be selectively retained in the nucleus. Alternatively, RNA export may not require recognition signals; all RNAs might be automatically transported unless they are specifically retained. A third possibility is that a combination of selective export and selective retention operates. Since an RNA molecule seems to be selectively retained in the nucleus until all of the spliceosome components have dissociated from it, selective retention could be caused by a mechanism that prevents the completion of RNA splicing on particular RNA molecules. You may ask now: and what kind of RNA is this that it is not let to be exported to the cytoplasm? Well, this is the portion of the RNA that is removed from the hnRNA after splicing and remains in the nucleus in order to be degraded.
- Because a virus parasitizes normal intracellular pathways, studies of viral development often help to decipher these pathways. Adenovirus, for example, has a double-stranded DNA genome, which is replicated and transcribed in the host-cell nucleus. Late in infection the transport of host-cell RNAs from the nucleus is blocked, so that most of the RNAs that reach the cytoplasm are encoded by the adenovirus. Genetic analysis has shown that 2 adenovirus proteins that are produced early in infection are required for this change in the selectivity of transport of RNAs from the nucleus providing a promising model system for analyzing how RNA transport is controlled.

16/c. Control of mRNA degradation.

- Measurement of mRNA half life: MCB, fig. 11-46.
- mRNA degradation rates seem to be related to the presence or absence of poly A tails and the presence of certain sequences in the 3' end. For example, if globin mRNA without poly A is injected into Hela cells or frog eggs, it is quickly degraded; studies on the synthesis of viral RNA by cells treated with drugs that block poly A formation also indicate that mRNA lacking poly A is lost quickly from cells. These findings correlate well with the naturally short half-life of Hela cell histone mRNAs which lack poly A at their 3' end.
- Numerous mRNAs that encode proteins thought to play a role in regulating cell growth and differentiation (protooncogene proteins and cell-growth factors) have short half-lives. A number of these mRNAs contain several A $(U)_n$ A sequences in their 3' untranslated regions, (MCB, fig. 11-47a). For example, the 3' untranslated end of human Granulocyte Monocyte- Colony Stimulating Factor (GM-CSF) contains seven AUUUA sequences, some of which are overlapping. When a segment of human GM-CSF DNA containing these sequences was inserted into a b-globin gene, the half-life of the globin mRNA decreased from more than 10h to less than 2h, (MCB, fig. 11-47b). A similar experiment with segments of the protooncogene C-fos showed that the poly A tail of the recombinant globin mRNA was shortened much faster when the A $(U)_n$ A-rich region was present than when it was absent. A reasonable hypothesis to explain generalized mRNA turnover is that all mRNAs are subject to nucleolytic attack near their 3' end. This attack, which is favoured by the presence of A $(U)_n$ A sequences, shortens poly A; if the nuclease removes all of the poly A (e.g. by cutting within the A $(U)_n$ A-rich region), an mRNA is converted to a completely nuclease-sensitive state. The rapid degradation typical of histone mRNAs and mRNAs containing A $(U)_n$ A sequences in their 3' ends does not occur when protein synthesis is blocked by drugs. As we will see, other mRNAs whose

half-lives are reduced under certain conditions also require normal protein synthesis to be subject to turnover. This association between mRNA degradation and protein synthesis suggests that normal mRNA turnover is caused by a ribosome-associated nuclease or nucleases. Access of this purported nuclease to various mRNAs would vary greatly depending on their structure and the presence or absence of molecules that protect or expose them. Thus, both normal variation in the half-lives of different mRNAs and differential regulation of mRNA stability may ultimately be explainable in terms of interactions that increase or decrease access of ribosome bound nucleases to mRNAs.

- Stability of specific mRNAs can be regulated by a variety of machanisms: i) Hormone-regulated stabilization of mRNA: Many of the known cases of differential stabilization of mRNAs in vertebrates are mediated by hormones. Because this hormone-dependent cytoplasmic control often is accompanied by hormone-dependent transcriptional control, the ability of many hormones to increase the synthesis of specific mRNAs result from 2 different effects. Casein, which is the most abundant protein in milk, is produced in the epithelial cells of breast tissue in response to hormones, including the polypeptide hormone prolactin. Small pieces of breast tissue that are cultured in the absence of prolactin contain only about 300 molecules of casein mRNA per cell, whereas cells cultured in the presence of prolactin contain about 30,000 casein mRNA molecules per cell. The rate of synthesis of casein mRNA molecules in cultured breast cells, however, is only about 3 times greater in the presence of prolactin than its absence. In contrast, prolactin treatment increases the half-life of the casein mRNA about 30 to 50-fold in cultured breast tissue, (MCB, fig. 11-48). These results strongly suggest that most of the prolactin-induced increase in casein mRNA at least in cultured cells is due to the increased stability of the mRNA rather to its increased synthesis.

ii) Autoregulation of tubulin: Microtubules are built of very long chains of heterodimers of 2 proteins α- and β-tubulin. When these polypeptide chains are finished they immediately associate and enter a pool of αβ tubulin heterodimers that are in equilibrium with the microtubules. When cells are treated with drugs that block tubulin polymerization (e.g. colchicine, which prevents mitotic spindle formation), the intracellular tubulin heterodimer concentration rises about 2-fold. When this happens both the α-and β-tubulin mRNAs are destroyed rapidly, reaching a new equilibrium at about 1/10 their usual concentrations. Conversely, there are drugs that remove free intracellular tubulin heterodimers; for example the anticancer drug vincristine causes all free heterodimers to precipitate in the cytoplasm as quasicrystalline structures. When cells are treated with this drug, the mRNA levels climb to 3 or 4 times their normal valves. None of these treatments that affect tubulin mRNA concentration have any affect on nuclear synthesis of tubulin pre-mRNA. Thus, they most likely affect the turnover of the tubulin mRNA itself. The sequence in the chicken b-tubulin mRNA that is required for b-tubulin destruction in colchicine-treated cells includes the 5' untranslated region and the first 4 codons of the mRNA. However, only tubulin mRNA that is associated with polyribosomes is degraded. Moreover, only if the mRNA has been translated for a distance of about 40 amino acids is the mRNA subject to destruction. Artificial mRNAs with a termination codon introduced before the 40th amino acid are <u>not</u> subject to autoregulation, (see MCB, fig. 11-49). It seems likely that the recognition sequence (the first 4 amino acids) is not present at the surface of the ribosome until a chain of about 40 amino acids has been made. The excess tubulin subunits then interact with the end of the nascent chain and lead to mRNA destruction.

iii) Regulation of mRNA stability by a nutrient: The mRNA encoding transferrin (an iron carrying serum protein) receptor is destroyed rapidly in cultured cells with abundant intracellular iron, whereas the half-life of this mRNA is increased substantially when the cultured medium contains low concentrations of free iron. The regulatory effect of the iron requires protein translation; that is, high intracellular levels of iron in the presence of translation-blocking drugs (e.g. cycloheximide) do not destabilize the transferrin-receptor mRNA. Transferrin-receptor mRNA have 5 stem-loops near its 3' end; each is about 12 base pairs long and has a CAGUG sequence in the loop, (see MBC, fig. 10-60). Cultured cells containing recombinant transferrin receptor genes that do not encode this stem-loop region produce mRNA whose half-life is not decreased by high iron concentration.

<u>17. The control of gene expression in eucaryotes III.</u>

<u>17/a. Translational repressor protein.</u>

- Not all mRNA molecules that reach the cytoplasm are translated into protein. The translation of some is blocked by specific translation repressor proteins that bind near their 5' end, where translation would otherwise begin, (MBC, fig. 10-55). In eucaryotic cells a particularly well-studied form of negative translational control allows the synthesis of the intracellular iron storage protein ferritin to be adjusted rapidly to the level of soluble iron atoms present. The ferritin mRNA in the cytoplasm can be shown to shift from an inactive ribonucleoprotein complex to a translationally active polyribosome complex after exposure of a cell to iron. Recombinant DNA experiments indicate that the iron regulation depends on a sequence of about 30 nucleotides in the 5' leader of the ferritin mRNA molecule. This iron-response element folds into a stem-loop structure (MBC, fig. 10-60 B) that binds a regulatory iron-binding protein when that protein is not bound to iron. When the protein is bound to the iron-response element, the translation of any RNA sequence downstream is repressed, (see MBC, fig. 10-55). The addition of iron dissociates the protein from the mRNA, increasing the rate of translation of the mRNA by as much as 100-fold.

- And how do we get from the previously mentioned stem-loop structure (MBC, fig. 10-60 B) to the initiation of protein synthesis in eucaryotes? As currently envisioned, the initiation of protein synthesis requires recognition by a complex of proteins of the cap structure at the 5' end of an mRNA, followed by unwinding of any stem-loop structure near the 5' end. This allows recognition of the AUG start codon by a 40S- ribosomal initiation complex. The

components that recognize and bind to the cap (MCB, fig. 11-51) include a cap-binding protein (CBP), which specifically recognizes the cap structure, and at least 3 other proteins, which together are referred to as eucaryotic initiation factor 4F, or eIF_{4F}. The ribosomal initiation complex consists of the 40S subunit eIF_3, and a "ternary complex" consisting of eIF_2, GTP, and initiator methionyl-$tRNA^{Met}$ (Met-$tRNA_1^{Met}$). Much evidence indicates that the greater the secondary structure at the 5' end of an mRNA, the more difficult is to start translation and the greater the requirement for eIF_{4F} activity. For example, oligonucleotides complementary to the first 15 bases of mRNAs greatly inhibit initial ribosome binding, and mRNAs with little or no secondary structure at their 5' end associate well with ribosomes without assistance of eIF_{4F} or at much lower than normal concentrations. The heat-shock mRNAs, the only mRNAs that are translated efficiently in heat-shocked cells, have a very A-rich 26 nucleotide stretch before their initiation codon; such A-rich stretches do not form extensive secondary structures. These results suggest that translation control might be affected by substances that increase or decrease the secondary structure at the 5' ends of mRNAs.

17/b. Translational enhancer.

- In principle, positive translational control could be mediated by a special "translation-enhancer" region in an mRNA molecule that preferentially attracts ribosomes. Certain RNA viruses -the picornaviruses-can be shown to contain such a region, whose presence can cause translation to begin at internal AUG sites that would not otherwise be used to start protein synthesis in a eucaryotic cell, (MBC, fig. 10-56). Positive translational control has also been demonstrated in yeast cells. Studies have identified specific proteins required to activate the translation of the mRNA produced from a yeast gene called GCN4; without them the mRNA remains untranslated. The GCN4 mRNA resembles a class of poorly translated mRNAs from higher eucaryotes, whose translation is suspected to be similarly controlled. These RNAs have an unusually long 5' leader sequence that contains a series of AUGs that interfere with the translation of the major coding sequence downstream by initiating the synthesis of short peptides; a stop codon occuring before the major coding sequence prevents readthrough. By analogy with the picornaviruses, the translation of the major coding sequence in these mRNAs might depend on the binding of translation-activator molecules to a translation-enhancer sequence next to the appropriate AUG, so that translation reinitiates, (MBC, fig. 10-57).

17/c. Trnaslational frameshifting.

- The translational controls thus far discussed, affect the rate at which new protein chains are initiated on an mRNA molecule. Usually, the completion of the synthesis of a protein is automatic once this synthesis has begun. In special cases, however, a process called translational frameshifting can alter the final protein that is made. Translational frameshifting is commonly used by retroviruses; it allows different amounts of 2 or more proteins to be synthesized from a single mRNA. These viruses commonly make a large polyprotein that is cleaved by a viral protease to produce a group of capsid proteins (gag proteins) and the viral reverse transcriptase and integrase (pol proteins). In many cases the gag and pol genes are in different reading frames, so that a translational frameshift is required to produce the much less abundant pol proteins. The frameshift occurs at a particular codon in the mRNA and requires specific sequences, some of which are upstream and some downstream from this site, (MBC, fig. 10-58).
- Another curious mechanism that extensively alters RNA transcripts that code for proteins has been found in the mitochondria of trypanosomes. In this RNA-editing process, one or more U nucleotides are either added or removed from selected regions of a transcript, causing modifications in the original reading frame that change the meaning of the message. RNA editing of a much more limited kind occurs in mammals, where it allows the apolipoprotein B gene to produce 2 types of transcripts ($apoB_{48}$ / $apoB_{100}$): in one of the transcripts a DNA-encoded C is changed to a U, creating a stop codon that causes a truncated version of this large protein to be made in a tissue-specific manner. Although this is the only example reported thus far, it seems very unlikely that RNA editing is confined to a single gene in mammals.

17/d. Control by phophorylation of eIF2.

- MCB, fig. 11-52. Another means of regulating the efficiency of translational initiation is modification of an initiation factor. One example of this mechanism is phosphorylation of initiation factors by protein kinases, which add phosphates to Ser, Thr, and Tyr residues. The biochemistry of translational initiation in mammalian cells, has been studied most extensively in reticulocytes. When eIF_2 is phosphorylated, it is inactive. In reticulocyte extracts, the trigger for phosphorylation by a protein kinase can be either the absence of heme, or the presence of a segment of double-stranded RNA (e.g. during RNA virus infection). Since hemoglobin contains both heme and the globin chains, a balance of the 2 is assured by this translational control; a decrease in heme leads to a decrease in globin mRNA translation. It is likely that all cells possess these protein kinases and that many factors other than heme and double-stranded RNA may affect their activity. The finding of an interferon-induced kinase that acts similarly in virus-infected cells (p. 881) suggests that this mechanism is widely used to control gene expression in eucaryotic cells.

18. The molecular genetics of cancer.

18/a. Mechanisms by which retroviruses can cause cancer.

- Retroviruses can transform cells and induce tumors by 2 quite distinct mechanisms:
i) one is utilized by transducing retroviruses (immediately transforming)
ii) the other, by slow-acting retroviruses
Concerning i) The key to our present understanding of the transforming and tumor inducing ability of some retroviruses came from the realization that such viruses contain genetic information other than gag, pol, and env.

The additional genetic material contained by these viruses are oncogenes appropriated from normal cellular DNA. Because phages that have acquired cellular genes are said to transduce the genes they acquire, retroviruses that have acquired cellular sequences are called transducing. Transducing retroviruses arise because of complex, apparently spontaneous rearrangements following the integration of a retrovirus near a cellular proto-oncogene (MCB, fig. 24-14). Most commonly, the acquired genetic information in a transducing retrovirus replaces part or all of gag, pol, and env, making the new retrovirus defective. However, the products of gag, pol, and env can be provided by a helper virus (a confecting wild-type retrovirus), so propagation of the defective virus is possible. The characteristics of different transducing retroviruses are determined in part by which oncogene(s) they contain. For instance, some transform only adherent cells, whereas others transform both adherent and nonadherent cells. In animals, the ones that transform nonadherent cells cause leukemia, the remainder cause sarcomas. [The oncogene in a retrovirus is denoted with the prefix v, e.g v-src, and the equivalent proto-oncogene is given the prefix c, e.g c-src].

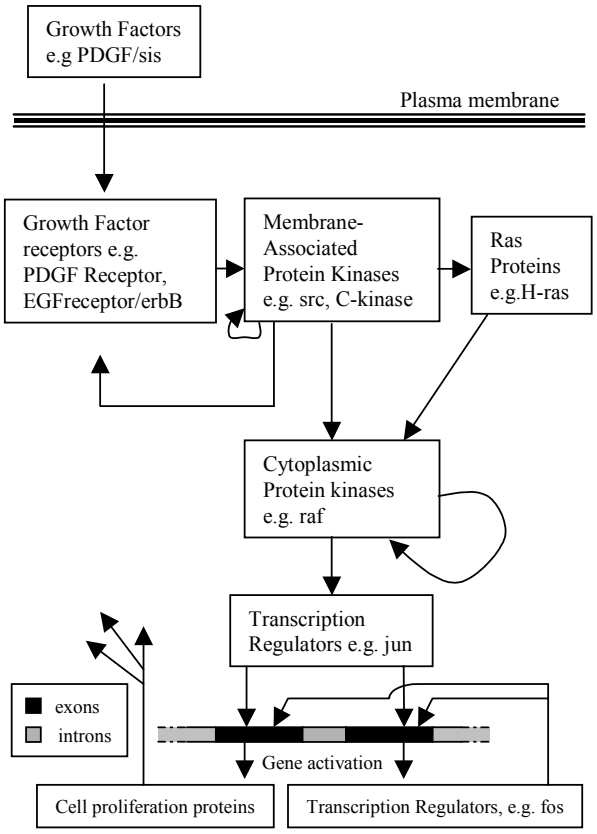
Although the transforming activity of a transducing retrovirus is provided by its acquired cellular genetic information, the genetic elements that allow the oncogenes to be expressed and tranferred from cell to cell are all inherited from the retroviral parent. The demonstration in transducing retroviruses that cellular genetic information can be agent of transformation, clearly implies that human genomes harbor genes that can, in the right circumstances, cause cancer. Furthermore, this notion also suggests that other agents, like chemicals and radiation could change normal genes into cancer-inducing genes by causing relatively simple mutations. 3 kinds of alterations, affecting either the coding region of the gene or its expression, may convert a proto-oncogene into an oncogene. First, transcription of a cellular gene acquired by a retrovirus is determined by the viral control elements, which can promote high transcription rates in many different cell types. This quantitative difference alone appears to explain why certain normal genes become transforming genes when they are captured by a virus. Indeed, proto-oncogenes cloned from cellular DNA can become transforming genes after being inserted into a retrovirus or otherwise placed in a high-transcription environment. A second kind of alteration is the loss of parts of the normal cellular gene, which apparently leaves the protein product with an unregulated activity that may allow it to transform cells. Third, a single point mutation or other small change may produce a crucial sequence difference between a transforming gene and its cellular counterpart. These alterations in a proto-oncogene may result in a quantitative change in the amount of the encoded protein or a qualitative change in the protein itself. Either one or both types of changes may play a role in converting a given proto-oncogene into an oncogene.

ii) The second type of oncogenic retroviruses- the slow-acting retroviruses- induces cancer over a period of months or years rather than the days or weeks required for cells to respond to a transducing virus. The genomes of slow-acting retroviruses differ from those of transducing viruses in one crucial respect: they lack an oncogene. The slow-acting retroviruses not only have no affect on growth of cells in culture, but they are also completely proficient to multiply themselves because they have suffered no debilitating deletions or insertions during acquisition of an oncogene. The mechanism by which avian leukemia viruses cause cancer appears to operate in all slow-acting retroviruses. Like all retroviruses, avian leukemia viruses generally integrate into cellular chromosomes more or less at random. However, the site of integration in the cells from tumors caused by these viruses was found to be near a gene called c-myc, an established proto-oncogene. This finding suggested that the slow-acting avian leukemia viruses cause disease by activating c-myc to become an oncogene. The viruses act slowly both because integration near c-myc is a random, rare event and because secondary events probably have to occur before a full-fledged tumor becomes evident.

- An integrated viral genome can act in a number of ways to cause a cellular gene to become an oncogene. In some tumors, the 5' end of the myc gene transcript includes a sequence from a retrovirus LTR. In such cases the right-hand LTR of the integrated retrovirus-which usually serves as a terminator-is beleived to act as a promoter, initiating synthesis of RNA transcripts from the c-myc gene (MCB, fig. 24-17a). Such c-myc transcripts apparently encode a perfectly normal c-myc product. The enhanced level of c-myc RNA resulting from the strong promoting activity of the retroviral LTR appears to be the explanation of oncogene activation in such cases. That is, the carcinogenic effect of the retrovirus results entirely from an increase in the amount of product from a normal gene. This mechanism is cabled promoter insertion.
- In a few tumors induced by avian leukemia viruses, the proviral DNA is located near the c-myc gene but at the 3' end of the gene; in others, it is at the 5' end of c-myc but oriented in the opposite transcriptional direction (MCB, fig. 24-17b). In such cases, the promoter-insertion model can not be applicable, and the proviral DNA is thought to exert an indirect enhancer activity that apparently activates c-myc by increasing its level of transcription, changing the cell type in which it is expressed and /or altering its time of expression in the cell cycle. This mechanism is called enhancer insertion. Either promoter insertion or enhancer insertion is the probable explanation for many tumors caused by retroviruses that do not carry oncogenes-meaning slow-acting viruses. These kind of viruses act as fingers pointing to the genomic location of potential oncogenes.
- The activation of a proto-oncogene following integration of a slow-acting retrovirus should be clearly distinguished from the acquisition of a proto-oncogene by a virus, converting it into an oncogene-containing transmissible virus. As it is evident from MCB fig. 24-17a, the first step in the formation of a transducing virus may well be promoter insertion, but several additional steps must occur before a defective, oncogene-containing retrovirus emerges. Those steps are all low-probability events, so very few promoter insertions progress to the formation of a transmissible, transducing virus. In fact, because defective transducing retroviruses depend on a helper virus for their cell -to-cell transmission, they are not maintained in natural animal populations. Most oncogene-containing retroviruses have arisen in laboratories or in domesticated animals and have been maintained for experimental purposes; the defective viruses are not readily spread from animal to

animal. In natural populations, insertional oncogene activation is probably the major cause of retrovirus-induced cancer.

18/b. Major classes of protooncogenes in the intracellular control network.



- This scheme is a tentative outline of the relationships of the major classes of proto-oncogenes in the intracellular control network through which external signals stimulate cell proliferation. A single representative proto-oncogene is indicated for each class. Note that each of the classes of regulatory molecules has many members, so that each arrow in the diagram potentially stands for many parallel arrows linking the individual members of one class to those of another. Moreover, the members of a given class in many instances can interact with one another-for example by mutual phosphorylation –as well as with the members of the classes upstream and downstream from them.

.18/c. Transformation of cells, detection of oncogenes (expt).

- Treatment of adherent cells with various agents (e.g. viruses, various chemicals, radiation) can dramatically change their growth properties in culture. Furthermore, such treated cells can form tumors after they are injected into susceptible animals. Such changes in the growth properties of cultured cells and their subsequent development of tumor-forming capacity are collectively referred to as malignant transformation, or just transformation. These changes are the following:

i) decreased growth factor requirements; some transformed cells
can produce growth factor analogues, so they may provide their own growth factors (called autocrine stimulation). (MCB, fig. 24-8, very important!)
ii) Loss of capacity for growth arrest; transformed cells may even kill themselves trying to continue growth in an impossible environment.
iii) Loss of dependence on anchorage for growth; this characteristic correlates extremely well with the ability of transformed cells to form tumors: cells that have lost anchorage dependence generally form tumors with high efficiency when they are injected into animals that cannot immunologically reject the cells.
iv) Changed cell morphology and growth habits.
v) Loss of contact inhibition of movement.
vi) Increased mobility of surface proteins.
vii) Easier agglutination by lectins (MCB, fig. 24-9).
viii) Increased glucose transport.
ix) Reduced or absent surface fibronectin.
x) Loss of actin microfilaments.
xi) Release of transforming growth factors.
xii) Protease secretion; transformed cells often secrete plasminogen activator.
xiii) Altered gene transcription.
xiv) Immortalization of cell strains.

- MCB fig. 24-29 (very important!)

18/d. The conversion of protooncogenes to oncogenes.

i) Promoter insertion (by a retrovirus).
ii) Enhancer insertion (by a retrovirus).
iii) Chromosomal translocations.
iv) Gene amplification (by chemicals).
v) Single point mutation.

18/e. Tumor suppressor genes.

- Genes other than oncogenes have been found to play a major role in the causation of at least some types of tumors. [These are tumor suppressor genes, sometimes called recessive oncogenes or antioncogenes. They operate quite differently from oncogenes (Harper, Table 62-6) in that their inactivation (as opposed to activation) removes certain mechanism of growth control. An important example is the retinoblastoma. Another example is the gene p53 encoding the protein p53. This protein acts as a transcriptional activator and as a G_1 checkpoint control for DNA damage. If excess damage to DNA has occurred, the amount of p53 could increase and thus inhibit cell division, allowing time for repair. On the other hand, if in tumor cells it is inactivated by mutation, then it would not fulfil this function, DNA damage could accumulate in the cell, and the cell would be genetically less stable. If this speculation is correct p53 may be a "guardian of the genome" or "molecular policeman".