Learning Objectives

- Identify the two types of nucleic acids and the function of each type.
- Describe how nucleotides are linked together to form nucleic acids.
- Describe the secondary structure of DNA and the importance of complementary base pairing.

Nucleic acids are large polymers formed by linking nucleotides together and are found in every cell. Deoxyribonucleic acid (DNA) is the nucleic acid that stores genetic information. If all the DNA in a typical mammalian cell were stretched out end to end, it would extend more than 2 m. Ribonucleic acid (RNA) is the nucleic acid responsible for using the genetic information encoded in DNA to produce the thousands of proteins found in living organisms.

Primary Structure of Nucleic Acids

Nucleotides are joined together through the phosphate group of one nucleotide connecting in an ester linkage to the OH group on the third carbon atom of the sugar unit of a second nucleotide. This unit joins to a third nucleotide, and the process is repeated to produce a long nucleic acid chain (Figure \(\PageIndex{1}\)). The backbone of the chain consists of alternating phosphate and sugar units (2-deoxyribose in DNA and ribose in RNA). The purine and pyrimidine bases branch off this backbone.

Each phosphate group has one acidic hydrogen atom that is ionized at physiological pH. This is why these compounds are known as nucleic acids.
Like proteins, nucleic acids have a primary structure that is defined as the sequence of their nucleotides. Unlike proteins, which have 20 different kinds of amino acids, there are only 4 different kinds of nucleotides in nucleic acids. For amino acid sequences in proteins, the convention is to write the amino acids in order starting with the N-terminal amino acid. In writing nucleotide sequences for nucleic acids, the convention is to write the nucleotides (usually using the one-letter abbreviations for the bases, shown in Figure \(\PageIndex{1}\)) starting with the nucleotide having a free phosphate group, which is known as the 5′ end, and indicate the nucleotides in order. For DNA, a lowercase \(d\) is often written in front of the sequence to indicate that the monomers are deoxyribonucleotides. The final nucleotide has a free OH group on the 3′ carbon atom and is called the 3′ end. The sequence of nucleotides in the DNA segment shown in Figure \(\PageIndex{1}\) would be written 5′-dG-dT-dA-dC-3′, which is often further abbreviated to dGTAC or just GTAC.

**Secondary Structure of DNA**

The three-dimensional structure of DNA was the subject of an intensive research effort in the late 1940s to early 1950s. Initial work revealed that the polymer had a regular repeating structure. In 1950, Erwin Chargaff of Columbia University showed that the molar amount of adenine (A) in DNA was always equal to that of thymine (T). Similarly, he showed that the molar amount of guanine (G) was the same as that of cytosine (C). Chargaff drew no conclusions from his work, but others soon did.

At Cambridge University in 1953, James D. Watson and Francis Crick announced that they had a model for the
secondary structure of DNA. Using the information from Chargaff’s experiments (as well as other experiments) and data from the X-ray studies of Rosalind Franklin (which involved sophisticated chemistry, physics, and mathematics), Watson and Crick worked with models that were not unlike a child’s construction set and finally concluded that DNA is composed of two nucleic acid chains running antiparallel to one another—that is, side-by-side with the 5’ end of one chain next to the 3’ end of the other. Moreover, as their model showed, the two chains are twisted to form a double helix—a structure that can be compared to a spiral staircase, with the phosphate and sugar groups (the backbone of the nucleic acid polymer) representing the outside edges of the staircase. The purine and pyrimidine bases face the inside of the helix, with guanine always opposite cytosine and adenine always opposite thymine. These specific base pairs, referred to as complementary bases, are the steps, or treads, in our staircase analogy (Figure \(\PageIndex{2}\)).

![Figure \(\PageIndex{2}\) DNA Double Helix. (a) This represents a computer-generated model of the DNA double helix. (b) This represents a schematic representation of the double helix, showing the complementary bases.](image)

The structure proposed by Watson and Crick provided clues to the mechanisms by which cells are able to divide into two identical, functioning daughter cells; how genetic data are passed to new generations; and even how proteins are built to required specifications. All these abilities depend on the pairing of complementary bases. Figure \(\PageIndex{3}\) shows the two sets of base pairs and illustrates two things. First, a pyrimidine is paired with a purine in each case, so that the long dimensions of both pairs are identical (1.08 nm).
Figure 1 Complementary Base Pairing. Complementary bases engage in hydrogen bonding with one another: (a) thymine and adenine; (b) cytosine and guanine.

If two pyrimidines were paired or two purines were paired, the two pyrimidines would take up less space than a purine and a pyrimidine, and the two purines would take up more space, as illustrated in Figure 2. If these pairings were ever to occur, the structure of DNA would be like a staircase made with stairs of different widths. For the two strands of the double helix to fit neatly, a pyrimidine must always be paired with a purine. The second thing you should notice in Figure 1 is that the correct pairing enables formation of three instances of hydrogen bonding between guanine and cytosine and two between adenine and thymine. The additive contribution of this hydrogen bonding imparts great stability to the DNA double helix.

Figure 2 Difference in Widths of Possible Base Pairs
Summary

• DNA is the nucleic acid that stores genetic information. RNA is the nucleic acid responsible for using the genetic information in DNA to produce proteins.

• Nucleotides are joined together to form nucleic acids through the phosphate group of one nucleotide connecting in an ester linkage to the OH group on the third carbon atom of the sugar unit of a second nucleotide.

• Nucleic acid sequences are written starting with the nucleotide having a free phosphate group (the 5′ end).

• Two DNA strands link together in an antiparallel direction and are twisted to form a double helix. The nitrogenous bases face the inside of the helix. Guanine is always opposite cytosine, and adenine is always opposite thymine.
Concept Review Exercises

1. a. Name the two kinds of nucleic acids.
   b. Which type of nucleic acid stores genetic information in the cell?

2. What are complementary bases?

3. Why is it structurally important that a purine base always pair with a pyrimidine base in the DNA double helix?

Answers

1. a. deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)
   b. DNA

2. the specific base pairings in the DNA double helix in which guanine is paired with cytosine and adenine is paired with thymine

3. The width of the DNA double helix is kept at a constant width, rather than narrowing (if two pyrimidines were across from each other) or widening (if two purines were across from each other).

Exercises

1. For this short RNA segment,
   a. identify the 5’ end and the 3’ end of the molecule.
   b. circle the atoms that comprise the backbone of the nucleic acid chain.
   c. write the nucleotide sequence of this RNA segment.
2. For this short DNA segment,
   a. identify the 5’ end and the 3’ end of the molecule.
   b. circle the atoms that comprise the backbone of the nucleic acid chain.
   c. write the nucleotide sequence of this DNA segment.

3. Which nitrogenous base in DNA pairs with each nitrogenous base?
   a. cytosine
b. adenine

c. guanine

d. thymine

4. Which nitrogenous base in RNA pairs with each nitrogenous base?
   a. cytosine
   b. adenine
   c. guanine
   d. thymine

5. How many hydrogen bonds can form between the two strands in the short DNA segment shown below?
   5' ATGCGACTA 3' 3' TACGCTGAT 5'

6. How many hydrogen bonds can form between the two strands in the short DNA segment shown below?
   5' CGATGAGCC 3' 3' GCTACTCGG 5'

---

Answers

1. c. ACU

3. a. guanine
   b. thymine
c. cytosine  
d. adenine

5. 22 (2 between each AT base pair and 3 between each GC base pair)

Learning Objectives

- Describe how a new copy of DNA is synthesized.
- Describe how RNA is synthesized from DNA.
- Identify the different types of RNA and the function of each type of RNA.

We previously stated that deoxyribonucleic acid (DNA) stores genetic information, while ribonucleic acid (RNA) is responsible for transmitting or expressing genetic information by directing the synthesis of thousands of proteins found in living organisms. But how do the nucleic acids perform these functions? Three processes are required: (1) replication, in which new copies of DNA are made; (2) transcription, in which a segment of DNA is used to produce RNA; and (3) translation, in which the information in RNA is translated into a protein sequence.

Replication

New cells are continuously forming in the body through the process of cell division. For this to happen, the DNA in a dividing cell must be copied in a process known as replication. The complementary base pairing of the double helix provides a ready model for how genetic replication occurs. If the two chains of the double helix are pulled apart, disrupting the hydrogen bonding between base pairs, each chain can act as a template, or pattern, for the synthesis of a new complementary DNA chain.

The nucleus contains all the necessary enzymes, proteins, and nucleotides required for this synthesis. A short segment of DNA is “unzipped,” so that the two strands in the segment are separated to serve as templates for new DNA. DNA polymerase, an enzyme, recognizes each base in a template strand and matches it to the complementary base in a free nucleotide. The enzyme then catalyzes the formation of an ester bond between the 5′ phosphate group of the nucleotide and the 3′ OH end of the new, growing DNA chain. In this way, each strand of the original DNA molecule is used to produce a duplicate of its former partner (Figure 9). Whatever information was encoded in the original DNA double helix is now contained in each replicate helix. When the cell divides, each daughter cell gets one of these replicates and thus all of the information that was originally possessed by the parent cell.

![Figure 9: A Schematic Diagram of DNA Replication. DNA replication occurs by the sequential unzipping of segments of the double helix. Each new nucleotide is brought into position by DNA polymerase and is added to the growing strand by the formation of a phosphate ester bond. Thus, two double helixes form from one, and each consists of new DNA strands that are identical to the original DNA strands.](image)
of one old strand and one new strand, an outcome called *semiconservative replications*. (This representation is simplified; many more proteins are involved in replication.)
Example {\PageIndex{1}}

A segment of one strand from a DNA molecule has the sequence 5′-TCCATGAGTTGA-3′. What is the sequence of nucleotides in the opposite, or complementary, DNA chain?

**SOLUTION**

Knowing that the two strands are antiparallel and that T base pairs with A, while C base pairs with G, the sequence of the complementary strand will be 3′-AGGTACTCAACT-5′ (can also be written as TCAACTCATGGA).

Exercise {\PageIndex{1}}

A segment of one strand from a DNA molecule has the sequence 5′-CCAGTGAATTGCCTAT-3′. What is the sequence of nucleotides in the opposite, or complementary, DNA chain?

What do we mean when we say information is encoded in the DNA molecule? An organism’s DNA can be compared to a book containing directions for assembling a model airplane or for knitting a sweater. Letters of the alphabet are arranged into words, and these words direct the individual to perform certain operations with specific materials. If all the directions are followed correctly, a model airplane or sweater is produced.

In DNA, the particular sequences of nucleotides along the chains encode the directions for building an organism. Just as saw means one thing in English and was means another, the sequence of bases CGT means one thing, and TGC means something different. Although there are only four letters—the four nucleotides—in the genetic code of DNA, their sequencing along the DNA strands can vary so widely that information storage is essentially unlimited.

---

**Transcription**

For the hereditary information in DNA to be useful, it must be “expressed,” that is, used to direct the growth and functioning of an organism. The first step in the processes that constitute DNA expression is the synthesis of RNA, by a template mechanism that is in many ways analogous to DNA replication. Because the RNA that is synthesized is a complementary copy of information contained in DNA, RNA synthesis is referred to as transcription. There are three key differences between replication and transcription:

1. RNA molecules are much shorter than DNA molecules; only a portion of one DNA strand is copied or transcribed to make an RNA molecule.
2. RNA is built from ribonucleotides rather than deoxyribonucleotides.
3. The newly synthesized RNA strand does not remain associated with the DNA sequence it was transcribed from.

The DNA sequence that is transcribed to make RNA is called the **template strand**, while the complementary sequence on the other DNA strand is called the **coding or informational strand**. To initiate RNA synthesis, the two DNA strands unwind at specific sites along the DNA molecule. Ribonucleotides are attracted to the uncoiling region of the DNA molecule, beginning at the 3′ end of the template strand, according to the rules of base pairing. Thymine in DNA calls for adenine in RNA, cytosine specifies guanine, guanine calls for cytosine, and adenine requires uracil. RNA polymerase—an enzyme—binds the complementary ribonucleotide and catalyzes the formation of the ester linkage
between ribonucleotides, a reaction very similar to that catalyzed by DNA polymerase (Figure \(\PageIndex{2}\)).

Synthesis of the RNA strand takes place in the 5’ to 3’ direction, antiparallel to the template strand. Only a short segment of the RNA molecule is hydrogen-bonded to the template strand at any time during transcription. When transcription is completed, the RNA is released, and the DNA helix reforms. The nucleotide sequence of the RNA strand formed during transcription is identical to that of the corresponding coding strand of the DNA, except that U replaces T.

Figure \(\PageIndex{2}\): A Schematic Diagram of RNA Transcription from a DNA Template. The representation of RNA polymerase is proportionately much smaller than the actual molecule, which encompasses about 50 nucleotides at a time.

Example \(\PageIndex{2}\)

A portion of the template strand of a gene has the sequence 5’-TCCATGAGTTGA-3’. What is the sequence of nucleotides in the RNA that is formed from this template?

**SOLUTION**

Four things must be remembered in answering this question: (1) the DNA strand and the RNA strand being synthesized are antiparallel; (2) RNA is synthesized in a 5’ to 3’ direction, so transcription begins at the 3’ end of the template strand; (3) ribonucleotides are used in place of deoxyribonucleotides; and (4) thymine (T) base pairs with adenine (A), A base pairs with uracil (U; in RNA), and cytosine (C) base pairs with guanine (G). The sequence is determined to be 3’-AGGUACUCAACU-5’ (can also be written as 5’-UCAACUCAUGGA-3’).

Exercise \(\PageIndex{2}\)

A portion of the template strand of a gene has the sequence 5’-CCAGTGAATTGCCTAT-3’. What is the sequence of nucleotides in the RNA that is formed from this template?

Three types of RNA are formed during transcription: messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). These three types of RNA differ in function, size, and percentage of the total cell RNA (Table \(\PageIndex{1}\)). mRNA makes up only a small percent of the total amount of RNA within the cell, primarily because each molecule of mRNA exists for a relatively short time; it is continuously being degraded and resynthesized. The molecular dimensions of the mRNA molecule vary according to the amount of genetic information a given molecule contains. After transcription, which takes place in the nucleus, the mRNA passes into the cytoplasm, carrying the genetic message from DNA to the ribosomes, the sites of protein synthesis. Elsewhere, we shall see how mRNA directly determines the sequence of amino acids during protein synthesis.
Table \PageIndex{1}: Properties of Cellular RNA in Escherichia coli

<table>
<thead>
<tr>
<th>Type</th>
<th>Function</th>
<th>Approximate Number of Nucleotides</th>
<th>Percentage of Total Cell RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRNA</td>
<td>codes for proteins</td>
<td>100–6,000</td>
<td>~3</td>
</tr>
<tr>
<td>rRNA</td>
<td>component of ribosomes</td>
<td>120–2900</td>
<td>83</td>
</tr>
<tr>
<td>tRNA</td>
<td>adapter molecule that brings the amino acid to the ribosome</td>
<td>75–90</td>
<td>14</td>
</tr>
</tbody>
</table>

Ribosomes are cellular substructures where proteins are synthesized. They contain about 65% rRNA and 35% protein, held together by numerous noncovalent interactions, such as hydrogen bonding, in an overall structure consisting of two globular particles of unequal size.

Molecules of tRNA, which bring amino acids (one at a time) to the ribosomes for the construction of proteins, differ from one another in the kinds of amino acid each is specifically designed to carry. A set of three nucleotides, known as a codon, on the mRNA determines which kind of tRNA will add its amino acid to the growing chain. Each of the 20 amino acids found in proteins has at least one corresponding kind of tRNA, and most amino acids have more than one.

Figure \PageIndex{3}: Transfer RNA. (a) In the two-dimensional structure of a yeast tRNA molecule for phenylalanine, the amino acid binds to the acceptor stem located at the 3′ end of the tRNA primary sequence. (The nucleotides that are not specifically identified here are slightly altered analogs of the four common ribonucleotides A, U, C, and G.) (b) In the three-dimensional structure of yeast phenylalanine tRNA, note that the anticodon loop is at the bottom and the acceptor stem is at the top right. (c) This shows a space-filling model of the tRNA.

The two-dimensional structure of a tRNA molecule has three distinctive loops, reminiscent of a cloverleaf (Figure \PageIndex{3}). On one loop is a sequence of three nucleotides that varies for each kind of tRNA. This triplet, called
the anticodon, is complementary to and pairs with the codon on the mRNA. At the opposite end of the molecule is the acceptor stem, where the amino acid is attached.

**Summary**

- In DNA replication, each strand of the original DNA serves as a template for the synthesis of a complementary strand.
- DNA polymerase is the primary enzyme needed for replication.
- In transcription, a segment of DNA serves as a template for the synthesis of an RNA sequence.
- RNA polymerase is the primary enzyme needed for transcription.
- Three types of RNA are formed during transcription: mRNA, rRNA, and tRNA.

**Concept Review Exercises**

1. In DNA replication, a parent DNA molecule produces two daughter molecules. What is the fate of each strand of the parent DNA double helix?
2. What is the role of DNA in transcription? What is produced in transcription?
3. Which type of RNA contains the codon? Which type of RNA contains the anticodon?

**Answers**

1. Each strand of the parent DNA double helix remains associated with the newly synthesized DNA strand.
2. DNA serves as a template for the synthesis of an RNA strand (the product of transcription).
3. codon: mRNA; anticodon: tRNA

**Exercises**

1. Describe how replication and transcription are similar.

2. Describe how replication and transcription differ.

3. A portion of the coding strand for a given gene has the sequence 5′-ATGAGCGACTTTGCGGGATTA-3′.
   a. What is the sequence of complementary template strand?
   b. What is the sequence of the mRNA that would be produced during transcription from this segment of DNA?

4. A portion of the coding strand for a given gene has the sequence 5′-ATGGCAATCCTAAACGCTGT-3′.
   a. What is the sequence of complementary template strand?
   b. What is the sequence of the mRNA that would be produced during transcription from this segment of DNA?
Learning Objectives

• describe the characteristics of the genetic code.
• describe how a protein is synthesized from mRNA.

One of the definitions of a gene is as follows: a segment of deoxyribonucleic acid (DNA) carrying the code for a specific polypeptide. Each molecule of messenger RNA (mRNA) is a transcribed copy of a gene that is used by a cell for synthesizing a polypeptide chain. If a protein contains two or more different polypeptide chains, each chain is coded by a different gene. We turn now to the question of how the sequence of nucleotides in a molecule of ribonucleic acid (RNA) is translated into an amino acid sequence.

How can a molecule containing just 4 different nucleotides specify the sequence of the 20 amino acids that occur in proteins? If each nucleotide coded for 1 amino acid, then obviously the nucleic acids could code for only 4 amino acids. What if amino acids were coded for by groups of 2 nucleotides? There are $4^2$, or 16, different combinations of 2 nucleotides (AA, AU, AC, AG, UU, and so on). Such a code is more extensive but still not adequate to code for 20 amino acids. However, if the nucleotides are arranged in groups of 3, the number of different possible combinations is $4^3$, or 64. Here we have a code that is extensive enough to direct the synthesis of the primary structure of a protein molecule.
The genetic code can therefore be described as the identification of each group of three nucleotides and its particular amino acid. The sequence of these triplet groups in the mRNA dictates the sequence of the amino acids in the protein. Each individual three-nucleotide coding unit, as we have seen, is called a codon.

Protein synthesis is accomplished by orderly interactions between mRNA and the other ribonucleic acids (transfer RNA [tRNA] and ribosomal RNA [rRNA]), the ribosome, and more than 100 enzymes. The mRNA formed in the nucleus during transcription is transported across the nuclear membrane into the cytoplasm to the ribosomes—carrying with it the genetic instructions. The process in which the information encoded in the mRNA is used to direct the sequencing of amino acids and thus ultimately to synthesize a protein is referred to as translation.

Before an amino acid can be incorporated into a polypeptide chain, it must be attached to its unique tRNA. This crucial process requires an enzyme known as aminoacyl-tRNA synthetase (Figure \(\PageIndex{1}\)). There is a specific aminoacyl-tRNA synthetase for each amino acid. This high degree of specificity is vital to the incorporation of the correct
amino acid into a protein. After the amino acid molecule has been bound to its tRNA carrier, protein synthesis can take place. Figure 2 depicts a schematic stepwise representation of this all-important process.

Figure 2: The Elongation Steps in Protein Synthesis

Early experimenters were faced with the task of determining which of the 64 possible codons stood for each of the 20 amino acids. The cracking of the genetic code was the joint accomplishment of several well-known geneticists—notably Har Khorana, Marshall Nirenberg, Philip Leder, and Severo Ochoa—from 1961 to 1964. The genetic dictionary they compiled, summarized in Figure 3, shows that 61 codons code for amino acids, and 3 codons serve as signals for the termination of polypeptide synthesis (much like the period at the end of a sentence). Notice that only methionine (AUG) and tryptophan (UGG) have single codons. All other amino acids have two or more codons.
Example \(\PageIndex{1}\): Using the Genetic Code

A portion of an mRNA molecule has the sequence 5′-AUGCCACGAGUUGAC-3′. What amino acid sequence does this code for?

SOLUTION

Use Figure \(\PageIndex{3}\) to determine what amino acid each set of three nucleotides (codon) codes for. Remember that the sequence is read starting from the 5′ end and that a protein is synthesized starting with the N-terminal amino acid. The sequence 5′-AUGCCACGAGUUGAC-3′ codes for met-pro-arg-val-asp.

Exercise \(\PageIndex{4}\)

A portion of an RNA molecule has the sequence 5′-AUGCUGAAUUGCGUAGGA-3′. What amino acid sequence does this code for?

Further experimentation threw much light on the nature of the genetic code, as follows:

1. The code is virtually universal; animal, plant, and bacterial cells use the same codons to specify each amino acid (with a few exceptions).
2. The code is “degenerate”; in all but two cases (methionine and tryptophan), more than one triplet codes for a given amino acid.
3. The first two bases of each codon are most significant; the third base often varies. This suggests that a change in the third base by a mutation may still permit the correct incorporation of a given amino acid into a protein. The third base is sometimes called the “wobble” base.

4. The code is continuous and nonoverlapping; there are no nucleotides between codons, and adjacent codons do not overlap.

5. The three termination codons are read by special proteins called release factors, which signal the end of the translation process.

6. The codon AUG codes for methionine and is also the initiation codon. Thus methionine is the first amino acid in each newly synthesized polypeptide. This first amino acid is usually removed enzymatically before the polypeptide chain is completed; the vast majority of polypeptides do not begin with methionine.

**Summary**

In translation, the information in mRNA directs the order of amino acids in protein synthesis. A set of three nucleotides (codon) codes for a specific amino acid.

**Concept Review Exercises**

1. What are the roles of mRNA and tRNA in protein synthesis?
2. What is the initiation codon?
3. What are the termination codons and how are they recognized?

**Answers**

1. mRNA provides the code that determines the order of amino acids in the protein; tRNA transports the amino acids to the ribosome to incorporate into the growing protein chain.

2. AUG

3. UAA, UAG, and UGA; they are recognized by special proteins called release factors, which signal the end of the translation process.

**Exercises**

1. Write the anticodon on tRNA that would pair with each mRNA codon.
   a. 5′-UUU-3′
   b. 5′-CAU-3′
   c. 5′-AGC-3′
   d. 5′-CCG-3′

2. Write the codon on mRNA that would pair with each tRNA anticodon.
   a. 5′-UUG-3′
b. 5′-GAA-3′

c. 5′-UCC-3′

d. 5′-CAC-3′

3. The peptide hormone oxytocin contains 9 amino acid units. What is the minimum number of nucleotides needed to code for this peptide?

4. Myoglobin, a protein that stores oxygen in muscle cells, has been purified from a number of organisms. The protein from a sperm whale is composed of 153 amino acid units. What is the minimum number of nucleotides that must be present in the mRNA that codes for this protein?

5. Use Figure \(\PageIndex{3}\) to identify the amino acids carried by each tRNA molecule in Exercise 1.

6. Use Figure \(\PageIndex{3}\) to identify the amino acids carried by each tRNA molecule in Exercise 2.

7. Use Figure \(\PageIndex{3}\) to determine the amino acid sequence produced from this mRNA sequence: 5′-AUGAGCGACUUUGCGGGAUUA-3′.

8. Use Figure \(\PageIndex{3}\) to determine the amino acid sequence produced from this mRNA sequence: 5′-AUGGCAAUCCUAAACGCUGU-3′

**Answers**

1. a. 3′-AAA-5′
   
b. 3′-GUA-5′
   
c. 3′-UCG-5′
   
d. 3′-GGC-5′

3. 27 nucleotides (3 nucleotides/codon)

5. 1a: phenyalanine; 1b: histidine; 1c: serine; 1d: proline

7. met-ser-asp-phe-ala-gly-leu