

UNIT 8

DNA BINDING PROTEIN MOTIFS

Structure

8.1	Introduction	Leucine zipper
	Expected Learning Outcomes	Helix-turn-Helix
8.2	Overview of DNA Binding Proteins	Helix- loop-Helix and Homeodomain
	Role of DNA Double Helix in DNA-Protein Complex	8.4 Summary
	Types of Binding Interactions	8.5 Terminal Questions
8.3	DNA Binding Protein Motifs	8.6 Answer
	Zinc Finger	

8.1 INTRODUCTION

DNA is the central molecule of life. It interacts with other nucleic acids (RNA) with various proteins to regulate a wide range of biological functions. The interaction between Gene(DNA) and DNA binding proteins are required for the numerous biological activities including DNA packaging, transcription, replication, and repair. Therefore, DNA-protein interactions are the fundamental to all life processes. Gene activity is mostly regulated at the level of transcription. This unit explores the interactions of DNA with DNA binding proteins. DNA binding proteins play a crucial role in cellular function by closely interacting with DNA, which helps maintain genomic integrity and regulate complex physiological responses. They control essential cellular functions like as gene expression, replication, repair, and recombination by binding to certain DNA regions. DNA binding proteins play a crucial role in accurately interpreting DNA information and ensuring the proper functioning of DNA molecules. Before transcription initiation, specific transcription factors must bind on the specific sequence of gene to repress and activate gene expression. The question of interest is that how DNA binding proteins bind to the specific sequence to the gene.

Therefore, in this unit you will learn about the DNA binding motifs/domain, their types, binding mechanism and functions. In next block-III you will study about transcription process (Unit 9)

Expected Learning Outcomes

After studying this unit, you should be able to:

- ❖ define DNA binding protein;
- ❖ explain mechanism how proteins bind with major groove of DNA;
- ❖ explain DNA binding motifs Zinc finger, leucine zipper, helix turn helix, helix loop helix and homeodomain; and
- ❖ discuss the function of DNA binding motifs.

8.2 OVERVIEW OF DNA BINDING PROTEINS

DNA-binding proteins are proteins that possess DNA-binding domains capable of binding to single or double-stranded DNA. These proteins have unique sequences with functional groups capable of recognizing base pairs and interacting with the minor and major groove of DNA. Some DNA-binding ligands interact with the minor groove of nucleic acid. DNA-binding proteins can be categorized into four types based on the roles they perform.

- i) **DNA replication factors** perform the synthesis of whole genomes or DNA fragments in living organisms and in laboratory settings.
- ii) **Histones** have a role in transcription and the packaging of chromosomes within the cell nucleus.
- iii) **Transcription factors** are proteins that play a role in regulating gene expression, acting as activators during transcription.
- iv) **Repair factors** play a role in excising single base pairs or particular oligonucleotides and replacing them with appropriate nucleotides. An example is nucleases that cut DNA molecules.

You have learnt in the previous unit that DNA-histone interactions are crucial for chromatin structure and gene regulation, aiding in the condensation and organization of DNA. Recall that tertiary DNA structure, chromatin is bound with the transcription factor and other DNA binding proteins and folded into looped domain. The DNA within a chromosome is a lengthy double helix structure. Initially, it was believed that these proteins could need direct contact with the hydrogen bonds within the base pairs inside the double helix to differentiate between different DNA sequences. The outside of the double helix contains DNA sequence information that gene regulatory proteins may identify without needing to unwind the double helix.

8.2.1 Role of DNA Double Helix in DNA-Protein Complex

DNA-binding proteins can identify and attach to particular DNA sequences through various motifs which are discussed in next section. The recognition process entails the interaction between binding proteins' structural motifs and DNA nucleotide bases. You know that DNA has a double-stranded form with **two grooves – minor and major** (Fig. 8.1a). DNA possesses negative charges because of the phosphate residues, allowing a basic protein region with positive charges to attract. The DNA can potentially create a stable combination with the molecule. Hydrophobic interactions and hydrogen bonding between amino acid side-chains and DNA bases are significant to form the DNA-protein complex. These interactions establish the affinity between proteins and DNA.

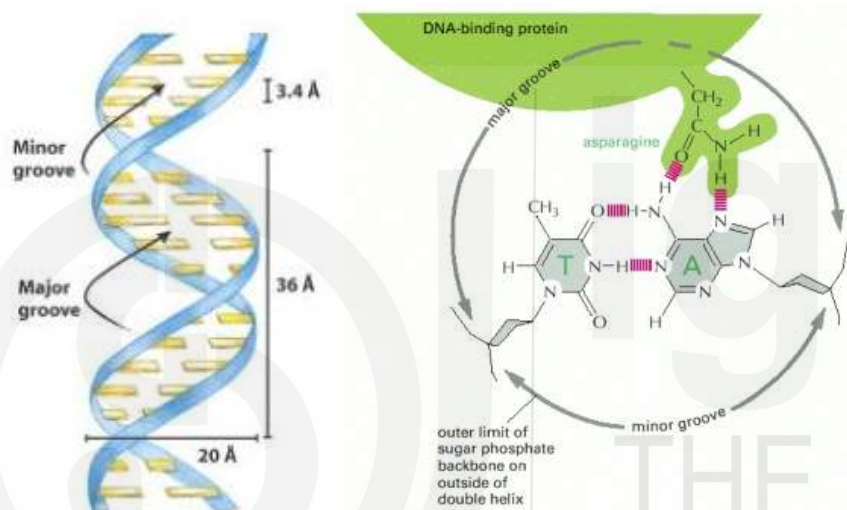


Fig. 8.1: a) Minor and major groove in DNA b) The binding of a gene regulatory protein to the major groove of DNA.

Source: *Molecular Biology of the Cell*. 4th edition. Alberts B, Johnson A, Lewis J, et al. New York: Garland Science; 2002.

The DNA base pairs are exposed in two grooves on its surface. The major groove is wider while the minor groove is narrower. Most DNA-binding protein motifs have structural features that fit into these grooves. Have you thought where the DNA binding protein bind the DNA double helix. DNA is at the surface of the double helix, the edges of each base pair are allowed to forming hydrophobic patches, hydrogen bond donors/acceptors with the DNA binding proteins. This binding complex form a unique pattern by which proteins can identify and interact with the major and minor grooves. Specific DNA-binding proteins often bind to the major groove of B-DNA due to the exposure of additional functional groups that can recognise a base pair. As a result, this interaction alters the base-pair configurations of DNA (Fig.8.1b). In this way, DNA binding proteins binds with the DNA without to open the double helix structure and regulate the gene expression. Different kind DNA binding proteins with metal ions can form different type of domains with the DNA about which you will study in the section.

Therefore, the protein usually binds to the major groove of DNA, with minimum involvement of the minor groove in the interactions. However, protein-DNA interactions trigger a sequence of biochemical changes that control essential biological processes in cells. One example of these binding interaction is transcription factors, which are a specific to binds to the DNA during transcription process. Such factors regulate gene expression by interacting with specific DNA motifs and histones associated with the DNA structure. Thus, DNA binding proteins play a crucial role in DNA replication, repair, and RNA synthesis regulation during transcription. They ensure accurate gene expression by selectively binding to specific DNA regions like promoters, enhancers, and regulatory elements. Proteins (Enzymes) can interact with single-stranded DNA to separate the double helix during DNA replication, repair, and recombination. An example of this is replication protein A which specifically bind to single-stranded DNA, stabilizing it and protecting it from degradation by nucleases and initiate the formation of stem-loop structures.

8.2.2 Types of Binding Interactions between DNA-Binding Proteins and Nucleic acid

DNA-protein interactions occur through **direct contact** between DNA base pairs and particular amino acids in the protein structure, or **indirect contact** enabled mainly by water molecules and conformational changes in the DNA structure. Proteins interact with DNA through non-covalent bonding such as electrostatic interactions (salt bridges), dipolar interactions (hydrogen bonding), hydrophobic interactions, and dispersion forces (base stacking). These interactions guide whether the binding between a protein and DNA is specific and tight or non-specific and loose. Such types of interactions are known as the **non-specific DNA-protein interactions**. These interactions aid in the organization of DNA into a condensed form known as chromatin, which is essential for controlling transcription.

The binding mechanism is lies on the interaction between aromatic and positively charged residues and phosphate groups of the DNA backbone. Because these interactions are based on electrostatic forces, it is believed that they constitute the basis for the specificity of the DNA-Protein complex formation. Hydrogen bonds between polar amino acids and DNA atoms increase binding affinity during complex formation. Additionally, hydrogen bonds form between DNA main chain atoms and amino acid residues, while hydrophobic interactions enhance complex stability. DNA plays a role in the complexation process, and its conformational flexibility influences binding specificity. There is a link between DNA stiffness and the binding selectivity of protein-DNA complexes. The conformational changes of DNA allow structural rearrangements to take place that are essential in mediating complex formation. It is well established that DNA binding proteins induced a conformational change in the DNA structure by which functional role of DNA is regulated.

SAQ 1

- a) In the following statements choose the correct alternative word given in parentheses
- i) DNA binding protein often bind to the (minor/major grooves) of DNA.
 - ii) DNA protein complex forms via (non covalent binding/covalent)
 - iii) The binding interaction between (aromatic and positively charged residues and phosphate groups/ deoxyribose sugar of the DNA backbone
- b) List the binding interaction occurs in DNA-protein complex?
-

Let us learn DNA binding protein motifs

8.3 DNA BINDING PROTIEN MOTIFS

The DNA binding protein motifs are short, conserved patterns in the amino acid sequences of DNA binding proteins; facilitating the interaction of these proteins with the DNA double helical structure selectively. The structures exhibit diversity through forms such as alpha-helices, beta-sheets, and other secondary components that enable effective and precise binding to DNA.

These motifs are distinguished by their functions, since they serve diverse purposes. Transcription factors are DNA-binding proteins that include particular patterns which regulate their binding to certain promoter or enhancer regions, hence affecting the activation or inhibition of transcription. Different types of DNA binding protein motifs play a crucial role in controlling DNA replication and repair mechanisms. Various DNA binding motifs that are evolutionarily conserved can be found in comparable forms across numerous animals.

A single DNA binding protein contains multiple DNA binding motifs, to bind and interact with different DNA sequences and regulate various cellular processes. The DNA binding protein motifs represent just a subset of the diverse structural elements that DNA binding proteins utilize to interact specifically with DNA. The variations in these motifs contribute to the wide array of functions performed by DNA binding proteins in the intricate landscape of cellular processes. There are some motifs classified on the basis of their structure as:

- **Zinc Finger**
- **Leucine zippers**
- **Helix-Turn-Helix (HTH)**
- **Homeodomain motifs**

- **Helix-Loop-Helix (HLH)**
- **Winged Helix-Turn-Helix (wHTH)**

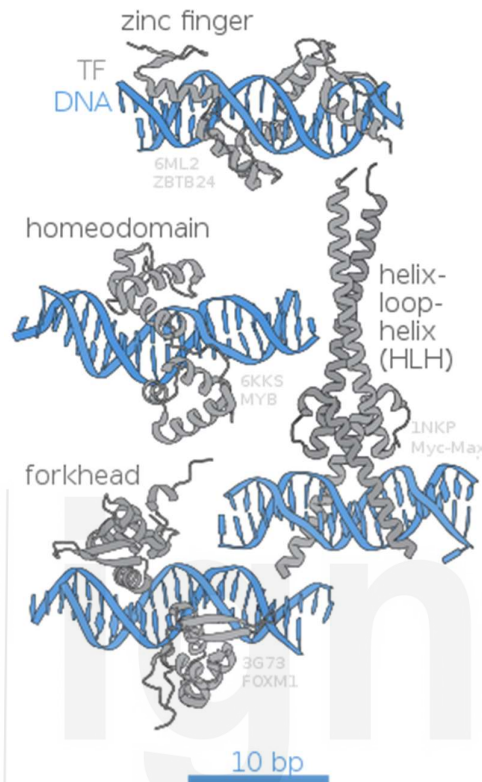


Fig. 8.2: Overview of different DNA binding motifs, in grey: the binding domain of zinc finger, homeodomain, helix-loop-helix, forkehead) and in blue is the DNA double helix.

https://en.wikipedia.org/wiki/DNA-binding_domain#/media/File:Transcription_factors_DNA_binding_sites.svg under attribution [CC BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/)

Let us discuss the most common DNA binding motifs:

8.3.1 Zinc Finger Motif

The zinc finger was discovered in 1985 as a result of our biochemical investigations into the interaction between the *Xenopus* protein transcription factor IIIA (TFIIIA) and 5S RNA. It is among the most common and versatile small structural motif found in eukaryotic in transcriptional factors of eukaryotes. The zinc-finger domain targets the DNA sequence and aligns with the three base pairs on DNA with a small number of amino acids in the α -helix structure. It plays wide range of function such as transcriptional regulation, DNA repair, development and differentiation of several tissues. It is generally present in eukaryotes. It exhibits a finger-like structure with the stable coordination of a zinc ion (Zn^{2+}) that bind to cysteine and histidine residues thus named as Zinc fingers (Fig.8.3). The Sp1-like proteins and Krüppel-like factors (KLFs) are the zinc finger motifs involved in the eukaryotic transcription.

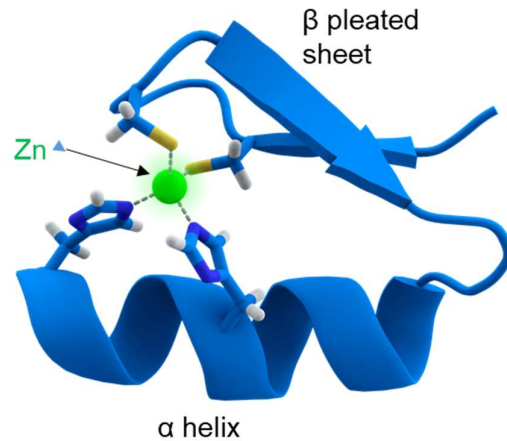


Fig. 8.3: Shows the cartoon structure of the Cys2His2 zinc finger motif, consisting of an α helix and an antiparallel β sheet. The zinc ion (green) is coordinated by two histidine residues and two cysteine residues. (Image source: https://commons.wikimedia.org/wiki/File:Zinc_finger_rendered.png).

The characteristic features of Zinc Finger Motif:

- It consists of a small, independently folded domain with 25 to 30 amino acids organized around a zinc ion.
- The structural core consists of the tetrahedrally coordination of a zinc ion with two cysteines and two histidines (Cys2His2 motif), maintaining stability of the zinc finger structure.
- It occurs in tandem repeats within a protein allowing recognition of longer DNA sequences with high specificity.
- The zinc finger protein has reserved sequence pattern of X3-Cys-X2-4-Cys-X12-His-X3-4-His-X4-, where X represents any amino acid. The Zn^{2+} ion is coordinated in a tetrahedral manner with the side chains of Cys and His. These side chains are located on one of two antiparallel beta strands and an alpha helix, respectively. The zinc finger, which is made stable by the presence of zinc, attaches to the major groove of DNA.

Function of Zinc Finger Proteins:

1. This motif act as transcription factors, regulating the expression of genes by binding to specific DNA sequences in the promoter regions.
2. These motifs are also involved in DNA repair processes, recognizing damaged DNA and facilitating repair mechanisms.
3. These zinc finger proteins are critical for embryonic development and tissue differentiation regulating genes expression.

8.3.2 Leucine Zipper Motif

The leucine zipper motif is characterised by a unique arrangement of amino acids, with leucine occurring at every seventh position in a repeated pattern known as heptad repetition (Fig. 8.4). This structural motif is mostly seen in

transcription factors to aid in protein dimerization. This motif is essential for establishing stable connections between proteins, facilitating their coordinated binding to particular DNA sequences. The leucine zipper is formed by the binding of two separate alpha helix monomers that are connected to DNA. It has amphipathic interaction between two ZIP domains which is situated in the alpha-helix of each monomer and is composed of leucines or leucine-like amino acids.

The characteristic features of Leucine Zipper Motif:

- It consists of a repeating pattern of seven amino acids, with hydrophobic amino acid, leucine as seventh residue.
- The hydrophobic core stabilizes the coiled-coil structure where two alpha helices twist around each other and promotes dimerization.
- The dimerization is essential for the functional activity of DNA binding proteins.

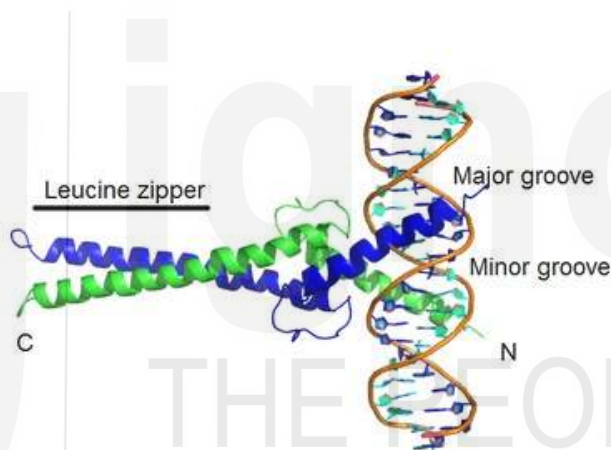


Fig. 8.4: Leucine Zipper the Helix-loop-helix (HLH) dimer, is shown bound to DNA fragment — each alpha helix represents a monomer. Attribution—the [Creative Commons Attribution-Share Alike 4.0 International](#).

Function of Leucine Zipper Proteins:

1. The leucine zipper facilitates the stable binding of two protein subunits to DNA.
2. Transcription factors containing leucine zippers play key roles in the regulation of gene expression.
3. Leucine zipper proteins are involved in various cell signaling pathways.
4. Some leucine zipper proteins are critical for embryonic development and tissue differentiation.

8.3.4 Helix-Turn-Helix (HTH) Motif

The Helix-Turn-Helix (HTH) motif consists of two alpha helices linked by a little turn, where the second helix is usually responsible for identifying and attaching to DNA (Fig. 8.5). It plays a vital role in the precise identification of DNA

sequences, enabling the control of gene expression and various cellular activities. The λ repressor of bacteriophage lambda makes two helix-turn-helix motifs.

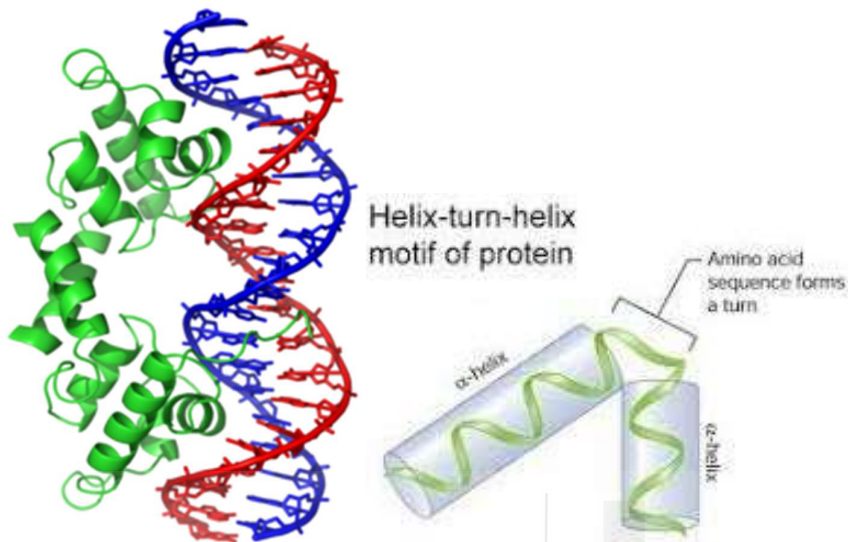


Fig. 8.5: (a) Helix turn helix motif shows the green protein binds with DNA. b) amino acid sequence turns to form helix turn helix. [CC BY-SA 3.0](#).

The characteristic features of Helix-Turn-Helix Motif are as follows:

- The HTH motif contains two alpha helices, the first for stability and the second for DNA interaction.
- The second helix, also referred to as the recognition helix, aids in the particular recognition of DNA sequences.
- The two helices are joined by a short turn that contains three to five amino acid residues.
- The recognition helix fits into the major groove of the DNA, where it makes specific contacts with the exposed bases allowing the protein to recognize and bind to a particular DNA sequence with high specificity.

Function of Helix-Turn-Helix Proteins:

1. It functions as transcription factors that regulate gene expression.
2. It enables DNA binding proteins to bind to specific DNA sequences to activate or repress transcription.
3. It versatile, and of different forms exist in different proteins.
4. It can mediate protein-DNA interactions through the specific recognition of DNA sequences critical for the precise targeting of transcription factors to their respective binding sites in the genome.

8.3.5 Helix-Loop-Helix (HLH) Motif

The Helix-Loop-Helix (HLH) consists of two helices that are linked by loop structure (Fig. 8.6). Transcription factors often exhibit this type of DNA binding

motifs. It is required for the assembly of homo- or heterodimeric complexes, enabling proteins to bind to specific DNA regions and control gene expression. Myc and Max serve as transcription factor that regulates the gene expression.

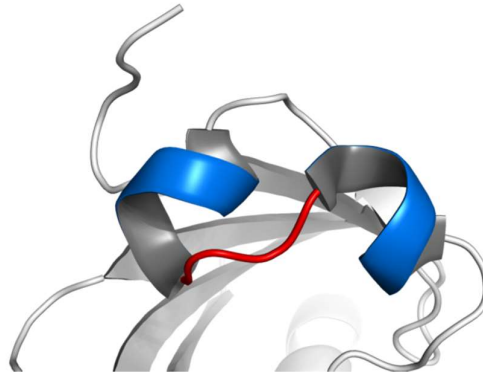


Fig. 8.6: Basic-helix-loop-helix structural motif of proteins (PDB ID 1X0o). Two helices (blue) are connected by a short loop region (red). CC BY-SA 3.0.

The characteristic features of Helix-Loop-Helix Motif are as follows:

- The HLH motif is small motifs with two alpha helices of typically about 20 amino acids each contains. These two helices are connected by a loop region.
- It contains basic amino acids (positively charged amino acids) which facilitates the DNA binding interactions.
- It can form homo and heteodimerization.
- HLH proteins play a role in a range of biological activities, such as differentiation, development, and regulation of the cell cycle.

8.3.6 Homeodomain

The homeodomain was first identified in the homeotic genes of the fruit fly *Drosophila melanogaster* (Fig.8.7). It is a conserved protein domain that is involved in gene expression by binding of proteins to specific DNA sequences, influencing the activation or repression of target genes.

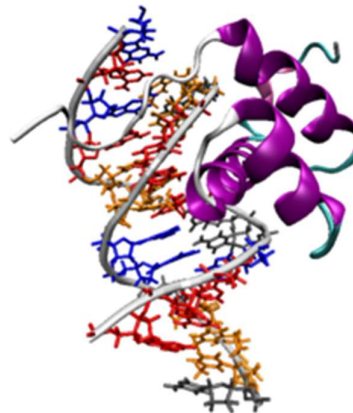


Fig. 8.7: The structure of antennapedia homeodomain protein of *Drosophila melanogaster* bound to a major and minor groove of DNA. Attribution-CC BY.

Image source

<https://en.wikipedia.org/wiki/Homeobox#/media/File:Homeodomain-dna-1ahd.png>

The characteristic features of the homeodomain:

- It is typically about 60 amino acids in length and forms a characteristic three-dimensional structure.
- It has three alpha helices in which the second and third helices arranged in an antiparallel orientation.
- The third helix helps in the DNA binding.
- Homeodomains recognize and bind to specific DNA sequences known as homeoboxes or Hox genes which are 180-base pair DNA segment with a core 60-base pair sequence that is essential for binding by the homeodomain-containing proteins.

Function of the homeodomain:

1. The homeodomain establishes hydrogen bonding as well as Vander-Waals interactions with the DNA bases, allowing for precise recognition.
2. Homeodomain binding proteins are transcription factors that regulate the expression of genes involved in development and differentiation.
3. This is evolutionarily conserved motif across a wide range of species from insects to mammals and plays a critical role in establishing the body plan and segment identity during embryonic development.

SAQ 2**Fill in the blanks:**

- a) Which protein motif shows the zip like motifs?.....
- b) Which DNA protein motifs recognize the homeobox of gene?.....
- c) HTH motif contains two alpha helices, the first for and the second for
- d) The distinctive pattern of amino acids with a heptad repetition where every seventh residue is a leucine is called.....
- e) Which DNA binding motif contains tetrahedrally coordination of a zinc ion with two cysteines and two histidines?.....
- f) The two alpha helices linked by a little turn is referred to the
- g) The two helices are connected by a loop region is called.....

8.4 SUMMARY

So far, you have learned that:

- DNA-binding proteins are those proteins that bind to the specific DNA sequence and form distinct domains. The DNA binding motifs generally contain helix segment. These domains have a vital function in identifying and attaching to particular DNA sequences, enabling a range of biological activities including transcription, replication, repair, and recombination and cell division.
- Proteins with DNA-binding domains can bind single or double-stranded DNA. These proteins have unique sequences with functional groups that recognise base pairs and interact with DNA's minor and major grooves.
- DNA double helix contains minor groove and major groove which allow to expose DNA base pairs for binding with other proteins. Major groove is wider than minor groove and provides stable binding interactions between DNA and DNA-binding protein
- The margins of each base pair on the double helix can produce hydrophobic patches and hydrogen bond donors/acceptors with DNA binding proteins.
- Zinc Finger motif is independently folded structural motifs which coordinate a zinc ion. It is responsible for interacting with a few nucleotides, and variations exist, such as C_2H_2 zinc fingers and C_4 zinc fingers.
- Leucine zippers motifs are dimerization motifs consisting of alpha helices with a characteristic heptad repeat, where every seventh amino acid is a leucine.
- Helix-Turn-Helix (HTH) motif contains two alpha helices connected by a short turn and found in prokaryotes.
- Homeodomain motifs are approximately 60 amino acid residues long and form a helix-turn-helix structure. They are found in many transcription factors involved in developmental processes.
- **Helix-Loop-Helix (HLH) motifs** are composed of two alpha helices connected by a loop. HLH proteins are often involved in the regulation of transcription.

8.5 TERMINAL QUESTIONS

1. What is DNA binding motif?
2. Explain how DNA binding proteins binds with DNA.
3. Explain the features and role of the following DNA binding motifs
 - (a) Helix-turn-helix

- (b) Helix loop helix
 - (c) Homeodomain
4. Differentiate between Zinc finger motif and Leucine zipper motif
 5. Draw the labeled diagram of DNA with major and minor grooves.
-

8.6 ANSWER

Self Assessment Questions

1.
 - a)
 - i) Major grooves of DNA.
 - ii) Non covalent binding The binding interaction
 - iii) Between (aromatic and positively charged residues and phosphate groups
 - b) Non-covalent interactions such as electrostatic interactions (salt bridges), dipolar interactions (hydrogen bonding), hydrophobic interactions, and dispersion forces (base stacking).
2.
 - a) Leucine Zipper Motif
 - b) Homeodomain
 - c) Stability and the DNA interaction
 - d) The distinctive pattern of amino acids with a heptad repetition where every seventh residue is a leucine is called Leucine Zipper Motif.
 - e) Zinc finger motif
 - f) HTH motif
 - g) HLH motifs

Terminal Questions

1. Please refer to section 8.2
2. Please refer to section 8.2.1
3. Please refer to section 8.3
 - a) Subsection 8.3.4,
 - b) Subsection 8.3.5 &
 - c) Subsection 8.3.6
4. Please refer to subsection 8.3.1 and 8.3.2
5. Please refer to Fig. 8.1



ignou
THE PEOPLE'S
UNIVERSITY