An Overview of DNA-Protein Interactions

Article *in* Current Chemical Biology · October 2015

DOI: 10.2174/2212796809666151022202255

2 authors:

Nilkanta Chowdhury
Sidho Kanho Birsha University
18 PUBLICATIONS **58** CITATIONS
SEE PROFILE

Angshuman Bagchi
University of Kalyani
156 PUBLICATIONS **703** CITATIONS
SEE PROFILE

Send Orders for Reprints to reprints@benthamscience.ae

Current Chemical Biology, **2015,** *9,* **000-000**

An Overview of DNA-Protein Interactions

Nilkanta Chowdhury and Angshuman Bagchi*

Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, Nadia, India

A. Bagchi

Abstract: DNA Protein interactions play very vital roles in any living cell. It controls various cellular processes which are very essential for living beings, viz. replication, transcription, recombination, DNA repair etc. There are several types of proteins found in a cell. But only those proteins interact with DNA, which have the DNA binding domains. Each DNA binding domain has at least one motif, which is a conserved amino acid sequence of this protein, which can potentially recognize a double stranded or a single stranded DNA. These DNA binding domains possess an affinity to bind to either double stranded or single stranded DNA. There are mainly two broad types of DNA protein interactions: 1) Sequence specific DNA binding and 2) Sequence non-specific DNA binding. In case of sequence specific DNA protein interactions, a DNA binding protein binds to a DNA on a site having a specific nucleotide sequence. But in case of sequence non specific DNA protein interactions, the DNA binding protein can bind to a DNA in a random position on the DNA. As for example, the sequence specific DNA protein interaction is found to occur in case of transcription. The transcription factors are a special kind of DNA binding proteins. They can only recognize a specific DNA sequence. The sequence non-specific DNA protein interaction occurs in replication. During replication the DNA double strand is melted by helicase enzyme, and a replication fork is made. A special kind of protein called single strand binding protein or SSB binds to the melted single strand of DNA and stabilizes the system by preventing them to be re-natured. There are several motifs present, which are involved in DNA binding, for example, helix-turn-helix, leucine zipper, zinc finger, helix-loop-helix etc. In this review, comprehensive analyses of DNA-protein interactions are made. The detailed discussions on the topic would be very much essential for all fields of biochemistry and biophysics.

Keywords: DNA-protein binding, Specificity of binding, Non-specific interaction, DNA binding motifs, Domains, DNA.

1. INTRODUCTION

Nucleic acids and proteins are very two of the most essential macromolecules for any living cell. They are involved in various bio-molecular functions viz., Replication, Transcription, Translation, Mutation, DNA repair and so on and so forth. Most if not all the important biological functions are the outcomes of the complex interplay of DNA protein interactions. The central dogma of molecular biology is entirely dependent on the mode of DNA protein interactions be it DNA-Replication, Transcription, Translation, Recombination or DNA repair. Scientists have therefore been intrigued to decipher the mechanism of DNA-protein interactions. DNA is a negatively charged macro molecule to which protein molecules bind in a number of different ways [1]. The interactions of proteins with DNA are mostly non-covalent interactions like hydrogen bonding (H-bonding), van der Waals interactions and ionic bonds [2]. The DNA binding proteins generally have specific DNA binding domains. DNA binding domains are specific clusters of amino acid sequences which remain conserved in the same family of DNA binding proteins. A DNA binding domain possesses at least one amino acid sequence motif which can recognize a single stranded or a double

*Address correspondence to this author at the Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, Nadia, India; Tel: +919051948843;
E-mail: angshu@klyuniv.ac.in

2212-7968/14 $58.00+.00

stranded DNA. A motif is a typical arrangement of specific secondary structural elements joined together to perform a specific biological task. There are several DNA binding motifs found to be present, viz.- Helix-Turn-Helix (HTH), Leucine Zipper, Zinc Finger to name a few [3-5]. On the basis of the mode of interaction between the DNA and protein, the DNA-protein interactions can be classified as: The Sequence Specific Binding or Sequence Specific DNA-protein interaction and Sequence non-specific Binding or Sequence Non-Specific DNA-protein interaction.

In the sequence specific binding, the DNA-binding protein can bind tightly to the target DNA in a sequence specific manner. The DNA binding domain recognizes and binds to a specific base sequence of the DNA, which is called the recognition sequence. Sometimes a DNA binding protein can randomly bind to a DNA molecule in a sequence independent manner, which is a sequence non-specific DNA-protein interaction [6]. In such cases the DNA binding protein loosely binds to the DNA [7]. However, both the sequence specific and sequence non-specific DNA-protein interactions are essential for life. In the present review, we try to analyze the different modes of DNA-protein interaction schemes. The different types of non-covalent interactions involved in the binding of proteins with DNA will be discussed. An overview of the molecular recognition process will be elucidated. This review is therefore aimed at scientists to have a firsthand knowledge of DNA-protein interactions.

1.1. Non-Covalent Binding Forces in DNA-Protein Interactions

The major binding forces responsible for DNA-protein interactions are the non-covalent binding forces, viz., Hydrogen Bonding (H-bonding), van der walls interactions, hydrophobic forces and salt bridges or ionic interactions.

1.1.1. Hydrogen Bonding

The term 'Hydrogen Bond' was first introduced in the book, named *The Nature of the Chemical Bond,* in 1912 [8]. A hydrogen bond is a weak non-covalent force of attraction. It is an outcome of dipole-dipole interactions between polar molecules. A hydrogen bond is weaker than any kind of covalent or electrostatic forces. But it is much stronger than a van der walls interaction [9-12]. Based on the donor and acceptor atoms, the H-bonds can be classified into two different categories: a) Inter-molecular b) Intra-molecular [13]. A DNA binding protein binds to the DNA by forming H-bonds with the negatively charged sugar phosphate backbone of the DNA. The DNA binding motif generally contains a collection of positively charged basic amino acid residues. The H-bonding is mediated by the side chains of the basic amino acid residues with the negatively charged sugar-phosphate backbone of the DNA [14].

1.1.2. van der Waals Interactions

The van der Waals interactions are another major source of interactions between DNA and proteins. The hydrophobic side chains of the amino acids in DNA binding proteins make van der Waals contact with the hydrophobic C-skeleton of the DNA bases [15].

1.1.3. Hydrophobic Interaction

This hydrophobic contact is used to read the DNA bases. Protein side chains perform hydrophobic interactions to differentiate thymine from cytosine. This hydrophobic interaction with bases also has an essential role in the sequence-specific interactions between bacterial cold shock protein and single-stranded DNA [16, 17].

1.1.4. Salt Bridges or Ionic Interactions

This type of interactions is also mediated by the charged or polar side chains of the amino acids. The amino acid residues like His, Arg having positively charged side chains bind to the negatively charged phosphate backbones of the DNA [8].

Before going into the details of DNA-protein interactions, a few words on DNA structure as well as DNA binding domains and motifs, are worth mentioning.

1.2. DNA or Deoxyribonucleic Acid

DNA or deoxyribonucleic acid is a double stranded, helical, bio-polymer. This polymer is made up of a simple unit, called nucleotide. The nucleotide is made up of nitrogen containing

nucleobases, a monosaccharide sugar, which is called deoxyribose (in case of DNA the sugar molecule is deoxyribose, but in case of RNA it is just a ribose sugar), and a phosphate group. The primary nucleobases are Adenine (A), Guanine (G), Thymine (T), Cytosine (C), and Uracil (U). Among them Uracil is found to be present only in RNA. Adenine (A) and Guanine (G) are made up of double ring structures, and belong to the family of purines. Thymine (T), Cytosine (C) and Uracil (U) are made up of single ringed structures, and belong to the pyrimidines. When a nucleobase forms a glycosidic bond with the 1' carbon atom of the deoxyribose (or ribose in RNA), then the structure formed is called a nucleoside. And when a phosphate group binds to the 5' carbon of the sugar, then the structure formed is called a nucleotide [18, 2].

In a double stranded helix, the two strands are bound together by H-bonding and the two strands of a DNA molecule run anti-parallel. DNA double strand is found to exist in three different conformations, viz: A-DNA, B-DNA and Z-DNA. B form of the DNA is the form which is very commonly found in cellular environments (Fig. **1**) [19].

There are some groovy regions found in a DNA double helix. These groovy regions are simply and distinguishingly termed as major groove and minor groove. The groove portion with a higher width is the major groove. The width of major groove is about 22 Å. And the groove with a comparatively lower width is called the minor groove. The width of the minor groove is about 12 Å [20, 21].

The higher width in the major groove provides more accessibility to read the sequence than the minor groove. As a results proteins like transcription factors which have to bind to a specific sequence in a DNA, interact or access the major groove in the DNA [22].

1.3. DNA Binding Protein Domains & Motifs

In a protein there are some conserved amino acid sequences or conserved tertiary structures which can fold and evolve independently to the rest of the protein body. They are called protein domains. In 1973 the domain concept was first introduced by Wetlaufer [4]. There are several categories of protein domains according to their functions, positions etc. DNA binding domain is one of them. A DNA binding domain must have a DNA binding motif, which has the ability to recognize a single stranded or a double stranded DNA and can bind either on a specific sequence, which is called the recognition sequence or on a random position on the DNA. There are several DNA binding motifs found to be present [5]. They are:

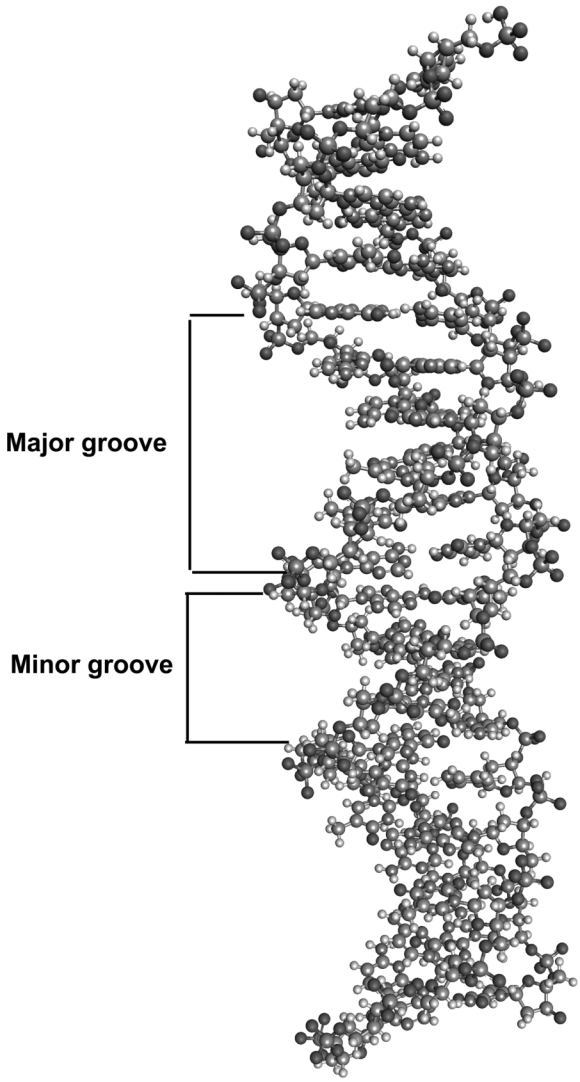

Fig. (1). A DNA double helix of B-form DNA.

1.3.1. Helix-Turn-Helix

Helix-Turn-Helix (HTH) (Fig. **2**) is a DNA binding protein motif, commonly found in several transcription factors, about 20 to 25 amino acids long. It's found in both eukaryotes and prokaryotes. The HTH motif is composed of two α-helices,

connected by a short amino acid chain, which builds the turn. The helix near to the C-terminal is called the recognition helix. This recognition helix interacts with the major groove to play an important role in the specific DNA sequence recognition. However, the amino acid sequences of this recognition helix differ from protein to protein. Transcriptional factor, LasR of *Pseudomonas aeuriginosa* possess a HTH motif in the C-terminal DNA binding domain [23-27].

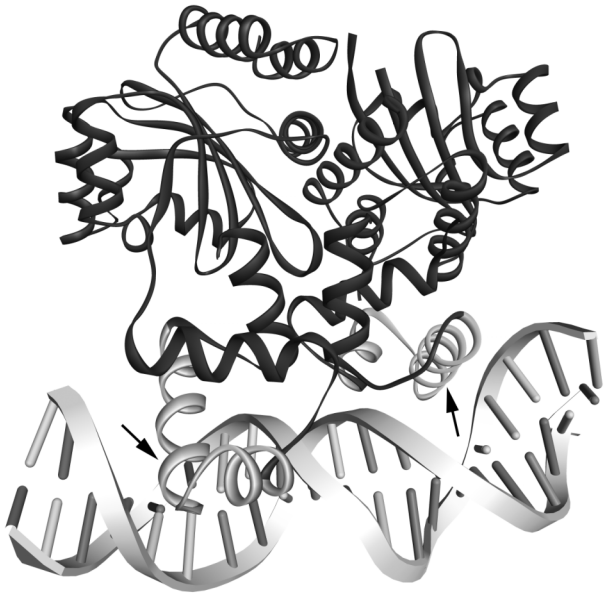

Fig. (2). Helix-turn-helix motif from a X-ray crystallographic structure. PDB ID: 1L3L [23]. The HTH motif of each chain is indicated by an arrow.

1.3.2. Zinc Finger Motif

This is another DNA binding amino acid sequence motif. The specialty of the motif is that the motif binds to the DNA with the help of Zn^{+2} ions. The compositions of the motifs vary greatly depending on their Zn^{+2} ion binding amino acid residues, viz - Cys2His2Zn1, Cys4Zn1, Cys3His1Zn1 and Cys6Zn2. The proteins having the Cys2His2 motif are referred to as zinc finger proteins (Fig. **3**). A characteristic feature of this protein motif is that, the zinc finger motif is coordinated to one or more Zn^{+2} ions. This Zn^{+2} ion stabilizes the three dimensional structure of the motif. The zinc finger domain is generally 23 to 28 amino acids long. The motif generally consists of a single recognition helix and a 2-strand beta-sheet. *Xenopus laevis* TFIIIA contains a zinc finger motif [28-30].

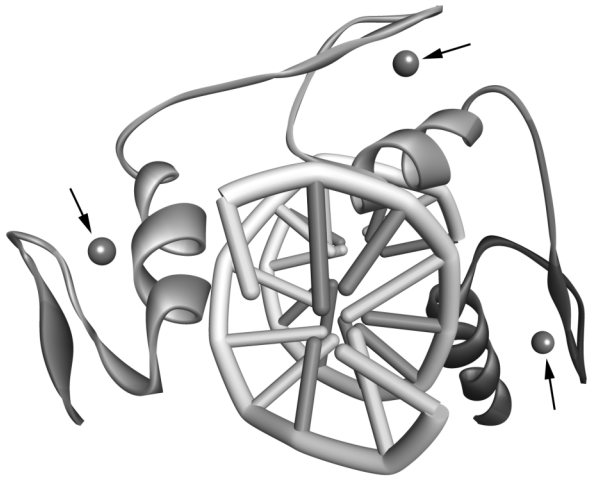

Fig. (3). Zinc finger motif from a X-ray crystallographic structure. PDB ID: 1ZAA [28]. The zinc atoms is indicated by arrows.

1.3.3. Leucine Zipper

The leucine zipper DNA binding motif is 60 to 80 amino acids long. The proteins bearing this motif generally act as a dimer. This motif is so named because two α-helices from each monomer are bound together by the interaction of the two Leucine residues from the two α-helices, forming a coiled-coil structure, and its look like a teeth of a zipper. Just after this dimerization region, these two α-helices form a Y-shaped structure which interacts with the DNA major grooves. c-fos, c-jun, etc are the Leucine zipper proteins [31-33]. It has been observed that several gene regulatory proteins bind to a DNA as homodimers. Each monomer provides a specific interaction with the DNA, so a homodimer can provide a strong specific binding. And very commonly in these types of homodimers, the dimerization interface is distinct from the DNA binding interface (as for example, the HTH motif). But in the Leucine zipper motif containing proteins these two interfaces are combined. A Leucine zipper motif contains an alpha helix with a Leucine residue at every 7^{th} amino acid position (Fig. **4**).

1.3.4. Helix-Loop-Helix

A helix-loop-helix (HLH) is another DNA-binding protein motif found to be present in several transcription factors. The HLH motif is made up of two α-helices, connected by a loop. Proteins with HLH motif forms both homo and hetero-

dimers. Human protein AHR, ARNT, etc possesses HLH motif (Fig. **5**) [34-36].

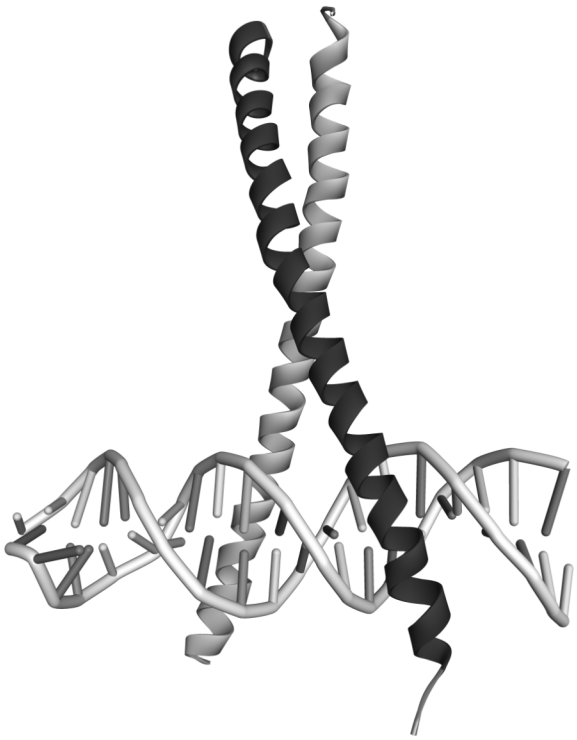

Fig. (4). Leucine zipper motif from a X-ray crystallographic structure. PDB ID: 1YAS [31].

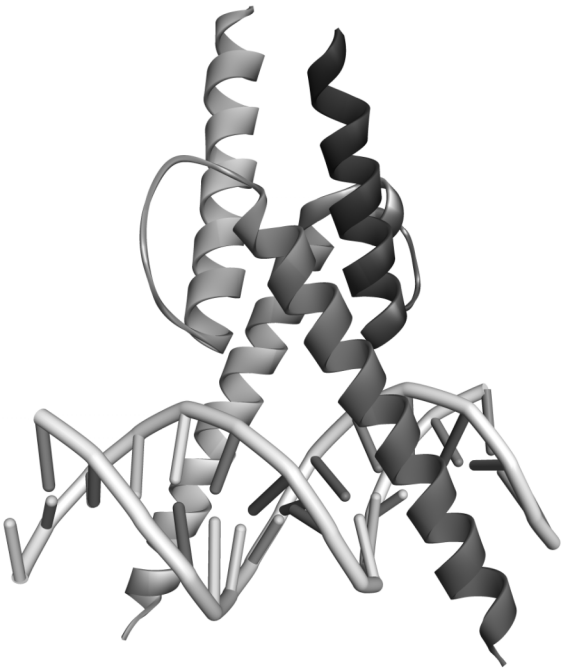

Fig. (5). Helix-loop-helix motif from a X-ray crystallographic structure. PDB ID: 4H10 [34].

1.3.5. HMG-Box

HMG-box (High Mobility Group box) (Fig. **6**) is another class of DNA binding protein motif. The HMG-box is made up of 3 α-helices connected by loops. The DNA-binding protein which has this motif can only bind to any DNA conformation, except B-type conformation (kinked or unwound) with high affinity. These HMG-box motifs are found in proteins with high mobility, which are involved in DNA-transcription, DNA-Replication and DNA-Repair methods. Human LEF1 contains a HMG-box [37-39].

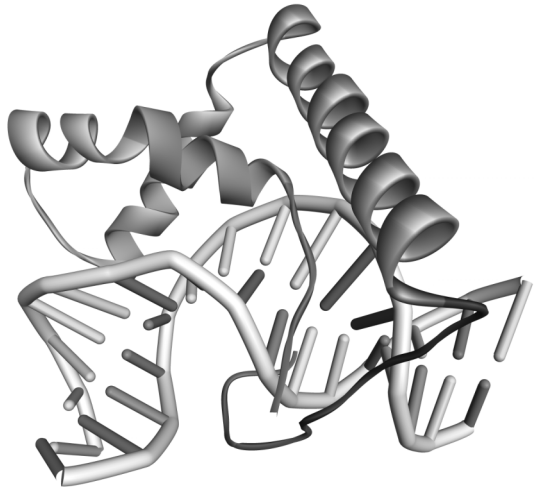

Fig. (6). HMG-Box from a X-ray crystallographic structure. PDB ID: 2lEF [37].

1.4 DNA-Protein Interaction

DNA protein interactions are inevitable for living beings. This interaction carries the central dogma of any life. Such interactions are found in any cells be it a lower prokaryotic organism or a higher eukaryotic organism [18, 22].

As mentioned earlier, that DNA-Protein interactions can be distinguished as:

1) Sequence specific DNA-protein interactions

2) Sequence non-specific DNA-protein interactions

1.4.1. Sequence Specific DNA-Protein Interaction

In Sequence specific DNA-protein interactions, a DNA binding protein, with its DNA binding motif,

interacts with a DNA in the major groove, in a sequence dependent manner. The modes of such interactions are dependent on the specific base sequences of the DNA, and the specific structural conformations of the DNA. The specific base sequences are also referred to as the specific sites.

1.4.1.1. Specific Site

We have discussed earlier in this article, that in a DNA double helix, groovy regions are formed. The wider grooves (22 Å) are called major grooves and the comparatively less wide grooves (12 Å) are called minor grooves. The wideness of a major groove provides a better exposure of the nucleotide bases to a DNA binding protein than the minor groove. In case of sequence specific DNA-protein interactions, the DNA binding proteins interact in the major groove of the DNA. It uses the exposure of DNA bases in the major groove, which helps the DNA binding protein to identify and read the specific sequence. But the minor groove cannot provide a sufficient exposure for a DNA binding protein, where it can distinguish a specific sequence [21, 2] (Fig. **7**). The distribution pattern of hydrogen bond acceptor, hydrogen bond donor, hydrogen atom, and methyl group, in major groove is different for all the four base pairings, A*T, T*A, G*C, C*G. So a specific sequence of base pairing generates an array of unique arrangement patterns necessary for interactions with specific DNA binding sequence motifs in DNA binding proteins. That is how a DNA major groove can provide the specificity to the sequence specific DNA binding proteins. However, such is not the case for the minor groove. The modes of base pairings of A*T and T*A as well as G*C and C*G is the same for the minor groove. So in minor groove the DNA bases are indistinguishable. Thus, the minor groove cannot provide sequence specificity to the sequence specific DNA binding proteins. Some zinc finger proteins and HMG protein form hydrogen bonds in the minor grooves of DNAs [2, 40, 41].

1.4.1.2. Sequence Dependent DNA Bending

DNA bending is an important feature for site specific DNA protein interactions. This property provides the second source for sequence specificity in DNA protein interactions, by providing a structural complementarity for the protein. The bending of the DNA depends on the sequence of the nucleic acids. A specific sequence provides a specific binding site for the DNA binding protein, and the specific bend formed due to the specific sequence makes a specific bent conformation, which helps the protein to bind with a lower free energy expense. The bend is formed due to the distortion in the sugar phosphate backbone. The bending of the DNA is grouped in three categories-

- Local distortion.
- Local distortion at successive base.
- Several severe local distortion.

The local distortion leads to the DNA bending which occurs at a single site. In the second type a collective bending occurs due to distortion in successive bases. And in the third type several severe local distortions produce kinks on DNA [42, 43]. The sequence specific DNA-protein interaction is needed during Transcription, Restriction endonuclease activity etc. During transcriptions, a transcription factor binds to the promoter region containing a specific DNA sequence, called a recognition sequence and controls the rate of the transcription of genetic information from DNA to mRNA [44]. For example LasR is a transcription factor present in *Pseudomonas aeuriginosa*. This transcription factor controls the transcription of *hcnABC* genes [27].

1.4.2. Sequence Non-Specific DNA-Protein Interactions

In case of a sequence non-specific DNA-protein interaction, a DNA binding protein is known to possess a general affinity to bind to a DNA strand. The protein uses ionic interactions to bind with the sugar-phosphate backbone of the DNA. Proteins showing site-specific binding with DNA also have some affinity to non specific DNA sequence. There are several events occurring in a cell where some DNA-binding proteins have to interact with DNA, but not in sequence specific manners. DNA Packaging is one of them [45, 46]. For example, in *Sulfolobus solfataricus* a histone like protein, Sso7d binds with its DNA by a triple-stranded β-sheet in the DNA minor groove and helps in DNA packing [46]. The other examples of DNA-protein sequence non specific interactions include homologous recombination, DNA repair etc. During

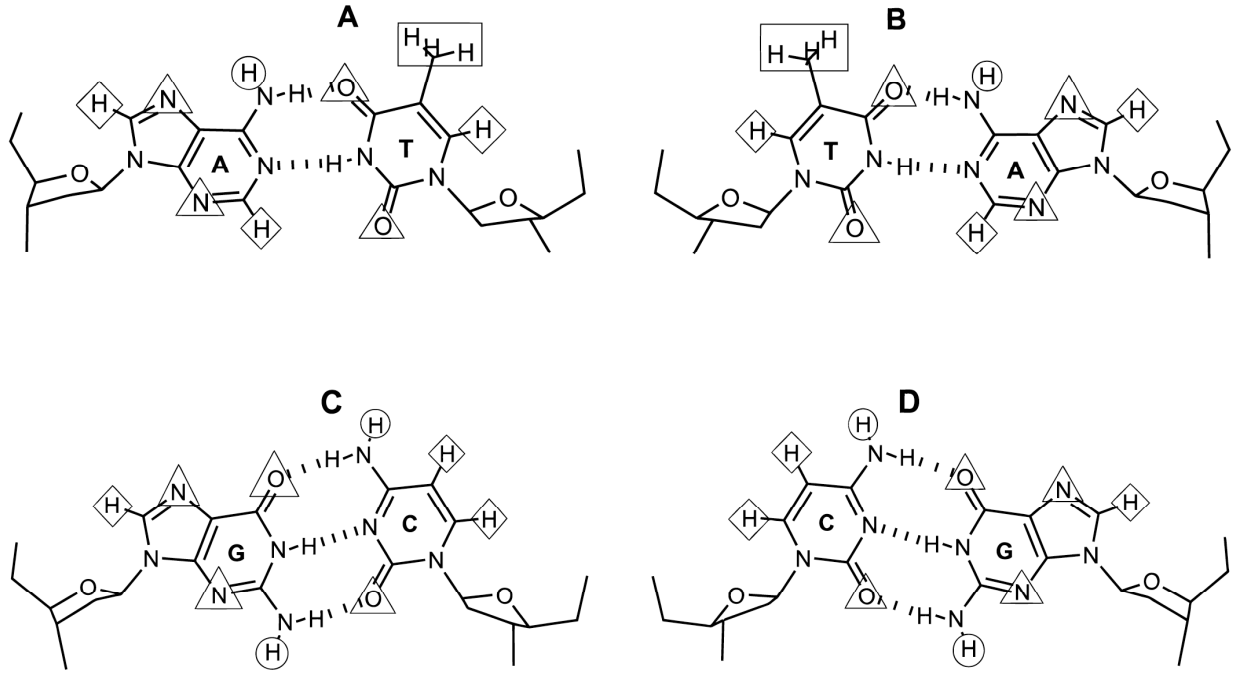

Fig. (7). The above picture is showing the distribution pattern of - Hydrogen bond acceptor, enclosed by triangle sign. Hydrogen bond donor, enclosed by circle. Hydrogen atom, enclosed by rhombus. And methyl group, enclosed by rectangle. For all the four base pairing, A*T, T*A, G*C, C*G, in major and minor groove. The upper face is representing the major groove, and the lower face representing the minor groove in each base pairs (**A**, **B**, **C**, **D**).

replication there are several proteins interacting with the DNA, but not in a sequence specific manner, viz- topoisomerase, helicase, single strand binding protein etc. They all interact with DNA, but not on a specific site. In case of homologous recombination, the RecA protein helps in the pairing between a single-stranded DNA molecule and a homologous sequence from another DNA molecule, a process called synapsis. When a DNA repair is required during a DNA double strand break, it is performed by DNA ligase enzyme. All of these DNA-protein interactions occur in a sequence independent manner [10].

1.5. Biological Processes Involving DNA-Protein Interactions

There are many important biological processes mediated by DNA-protein interactions. Among them the most important ones are: DNA-replication, DNA-transcription and DNA repair.

1.5.1. DNA Replication

DNA replication is the biological process, which is needed to make a replica of the DNA for a new cell which is produced during cell division. There are several enzymes that interact with DNA, during this process. To initiate this process of replication, an initiator protein interacts with the DNA in the region, called origin of replication, which is an AT rich region [47]. The most important enzymes are DnaB, DnaC, SSB protein, DNA Polymerase, RNaseH and ligase. These proteins make direct contacts with DNA in this process [48-50].

1.5.2. Transcription

Transcription is another essential biological process needed for gene expression. In this process DNA is transcribed to messenger RNA or mRNA. This mRNA contains a coded message, which is decoded when it is translated during translation process. In this process also some transcription performing proteins are involved, which interact with the DNA during this process. The transcription process is mainly conducted by two types of proteins, called the transcription factors, and RNA polymerase. The transcription factor binds to a promoter region on the DNA and recruits the RNA polymerase. A transcription factor interacts with the DNA in a sequence specific manner by binding

to a specific promoter sequence. Both the transcription factor and RNA polymerase directly interacts with the DNA during this process [10].

1.5.3. DNA Repair

A cell cannot function properly if DNA damage occurs. This damage can corrupt the information present in the genome. There are mainly two types of DNA repair system is present, based on their functions [51-55].

- Direct reversal repair of DNA damage.
- Excision repair of DNA damage.

In this repair system some proteins are involved to participate in the repair mechanism. The important enzymes are O^6-methylguanine methyltransferase, apurinic/apyrimidinic endonuclease or AP endonuclease, DNA glycosylase.

1.6. Detection of DNA Protein Interaction

There are so many *in vitro* and *in vivo* techniques which are used in the detection of DNA-protein Interactions. The following section gives a brief description of the different techniques.

1.6.1. DNA-Protein Interaction Study X-Ray Crystallography Method

X-ray crystallography is the most efficient tool which is used to identify the atomic level details DNA-protein interactions. The technique requires a high quality crystal of the specimen molecule. The crystallization technique is more art than a science. A crystallographer can make a 3D image of electron density of the specimen under study by the measurement of the diffracted beams and the angel of the beams. The position, chemical bonds of an atom in the crystal and much other information can be determined from this electron density. X-ray crystallography is very useful technique for a macromolecule study, but the limitation of this technique is the difficulties in crystallizing of a macromolecule [56].

1.6.2. DNA-Protein Interaction Study by Nuclear Magnetic Resonance (NMR)

Nuclear Magnetic Resonance (NMR) is a spectroscopic method in which atomic nuclei in a magnetic field absorb and re-emit electromagnetic radiation. This technique is used for characterization of the macromolecules and complexes [57, 58]. NMR studies of large macromolecules and complexes are performed in liquid state. In NMR study of Protein-DNA complex, structural restraints are collected, which are used to perform the structural calculation. Different types of information are collected: ambiguous restraints, which is generally derived by mapping of the interaction surface, and unambiguous restraints, which is provided by mainly three types of data, and paramagnetic relaxation enhancement (PRE), intermolecular nuclear Overhauser effects (NOEs), and residual dipolar couplings experiments. All these combined data are used to compute a high quality DNA-protein complex structure [59].

Besides these techniques, there are other techniques which are widely used for the detection of DNA-protein interactions, viz: Electrophoretic Mobility Shift Assay (EMSA) [60, 61]. It is an electrophoretic separation technique of a protein–RNA or protein–DNA mixture on a agarose gel or poly-acrylamide gel. A specific type of immune-precipitation technique called Chromatin Immuno-precipitation is also used to detect the interaction between DNA and proteins [62]. DNase footprinting assay is a technique used to detect the DNA-protein interaction by using the deoxyribonuclease or DNase enzyme [63]. Bio-molecular interaction analysis by Surface Plasmon Resonance (SPR) spectroscopy is an optical technique used to detect DNA-protein interaction [64]. Fluorescence method, electron microscopy, atomic force microscopy, yeast two-hybrid system, bacterial one hybrid system are some of the techniques which are widely used in the detection of DNA-protein interaction [65, 66].

1.7. DNA-Protein Interaction-Final Words

DNA-protein interactions are necessary for many different biological processes. Without such interactions, life cycle would be stopped. However, the detailed mechanism of DNA-protein interactions at the molecular level is still not very clear. This review is therefore aimed at giving a brief overview of the different modes of DNA-protein interactions. This review may be considered as a first hand guide to the scientists who are interested in this field.

CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

The authors would like to thank the Bioinformatics Infrastructure Facility (BIF) housed in the Department of Biochemistry and Biophysics, University of Kalyani, for providing the necessary support.

REFFERENCES

[1] Neidle S. Principles of Nucleic Acid Structure. The Building-Blocks of DNA and RNA. In: Kirsten Funk, Ed. Principles of Nucleic Acid Structure. 1st ed. London: Academic Press 2008; pp. 20-37.

[2] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. The Chemical Components of a Cell. In: Marjorie Anderson, Sherry Granum, Eds. Molecular Biology of the Cell. 5th ed, New York: Garland science, Taylor & Francis Group; 2002; pp. 45-124.

[3] Jones S, van Heyningen P, Berman HM, Thornton JM. Protein-DNA interactions: A structural analysis. J Mol Biol. 1999; 287(5): 877-96.

[4] Phillips DC. The Three-Dimensional Structure of an Enzyme Molecule. Sci Am. 1966 Nov; 215(5): 78-90.

[5] Chiang YS, Gelfand TI, Kister AE, Gelfand IM. New classification of supersecondary structures of sandwich-like proteins uncovers strict patterns of strand assemblage. Proteins Struct Funct Genet. 2007; 68(4): 915-21.

[6] Cai YH, Huang H. Advances in the study of protein-DNA interaction. Amino Acids. 2012; 43(3): 1141-6.

[7] Thermo Scientific. Protein:DNA Interactions. Nucleic Acid Interaction Handbook. Guidebook. pp. 2.

[8] Pauling L. The nature of the chemical bond—1992. J Chem Educ. 1992; 69(7): 519.

[9] Anslyn E V, Dougherty DA. Solution and Non-Covalent Binding Forces. In: John Murdzek, Ed. Modern Physical Organic Chemistry. 1st ed. United States: Wilsted and Taylor Publishing Services. 2006; 69: pp. 145-205.

[10] Lodish HF, Berk A, Zipursky SL, Matsudaira P, Baltimore D, James D. Chemical Foundations. In: Katherine Ahr Parker, Ed. Molecular Cell Biology. 7th ed. United States: W.H Freeman and company 2012; pp. 23-58.

[11] Arunan E, Desiraju GR, Klein RA, et al. Defining the hydrogen bond: An account (IUPAC Technical Report). Pure and Applied Chemistry. 2011; 83(8): 1619-36.

[12] Nelson DL, Cox MM. Weak Interactions in Aqueous Systems. In: Sara Tenney, Ed. Lehninger Principles of Biochemistry. 5th ed. United States: W. H. Freeman and Company 2008; pp. 43-53.

[13] McNaught AD, Wilkinson A. IUPAC Compendium of Chemical Terminology. In: Alan D. McNaught and Andrew Wilkinson, Eds. IUPAC Compendium of Chemical Terminology(The "Gold Book"). 2nd ed. UK: Blackwell Scientific Publications, Oxford. 1997; pp. 1-1670.

[14] Tsai CS. Biochemistry: Studies of Life at the Molecular Level. In: John Wiley & Sons, Inc Ed. An Introduction to Computational Biochemistry. United States 2002; pp. 1-4.

[15] Van Oss CJ, Absolom DR, Neumann AW. Applications of net repulsive van der Waals forces between different particles, macromolecules, or biological cells in liquids. Colloids and Surfaces. 1980; 1(1): 45-56.

[16] Harrison SC, Aggarwal AK. DNA recognition by proteins with the helix-turn-helix motif. Annu Rev Biochem. 1990; 59: 933-69.

[17] Chandler D. Interfaces and the driving force of hydrophobic assembly. Nature. 2005; 437(7059): 640-7.

[18] Watson JD, Crick FHC. Molecular structure of nucleic acids. Nature. 1953; 171(4356): 737-8.

[19] Clausen-Schaumann H, Rief M, Tolksdorf C, Gaub HE. Mechanical stability of single DNA molecules. Biophys J. 2000; 78(4): 1997-2007.

[20] Ghosh A, Bansal M. A glossary of DNA structures from A to Z. Acta Crystallographica - Section D Biological Crystallography. 2003; 59(4): 620-6.

[21] Wing R, Drew H, Takano T, et al. Crystal structure analysis of a complete turn of B-DNA. Nature. 1980; 287(5784): 755-8.

[22] Pabo CO, Sauer RT. Protein-DNA recognition. Annu Rev Biochem. 1984; 53: 293-321.

[23] Zhang R, Pappas KM, Brace JL, et al. Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. Nature. 2002; 417(6892): 971-4.

[24] Brennan RG, Matthews W. The Helix-Turn-Helix DNA Binding Motif. J Biol Chem. 1989; 264(4): 22-5.

[25] Wintjens R, Rooman M. Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. J Mol Biol. 1996; 262(2): 294-313.

[26] Matthews BW, Ohlendorf DH, Anderson WF, Takeda Y. Structure of the DNA-binding region of lac repressor inferred from its homology with cro repressor. Proc Natl Acad Sci U S A. 1982; 79(5): 1428-32.

[27] Laville J, Blumer C, Von Schroetter C, et al. Characterization of the hcnABC gene cluster encoding hydrogen cyanide synthase and anaerobic regulation by ANR in the strictly aerobic biocontrol agent Pseudomonas fluorescens CHA0. J Bacteriol. 1998; 180(12): 3187-96.

[28] Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science. 1991; 252(5007): 809-17.

[29] Klug A, Rhodes D. Zinc fingers: A novel protein fold for nucleic acid recognition. Cold Spring Harb Symp Quant Biol. 1987; 52: 473-82.

[30] Hanas JS, Hazuda DJ, Bogenhagen DF, Wu FY, Wu CW. Xenopus transcription factor A requires zinc for binding to the 5 S RNA gene. J Biol Chem. 1983; 258 (23): 14120-5.

[31] Ellenberger TE, Brandl CJ, Struhl K, Harrison SC. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted ? helices: Crystal structure of the protein-DNA complex. Cell. 1992; 71: 1223-37.

[32] Vinson CR, Sigler PB, McKnight SL. Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science. 1989; 246(4932): 911-6.

[33] Turner R, Tjian R. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. Science. 1989; 243(4899): 1689-94.

[34] Wang Z, Wu Y, Li L, Su X-D. Intermolecular recognition revealed by the complex structure of human CLOCK-BMAL1 basic helix-loop-helix domains with E-box DNA. Cell Res. 2013; 23(2): 213-24.

[35] Anthony-Cahill SJ, Benfield PA, Fairman R, Wasserman ZR, Brenner SL, Stafford WF, et al. Molecular characterization of helix-loop-helix peptides. Science. 1992; 255(5047): 979-83.

[36] Massari ME, Murre C. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. Mol Cell Biol. 2000; 20(2): 429-40.

[37] Love JJ, Li X, Case DA, Giese K, Grosschedl R, Wright PE. Structural basis for DNA bending by the architectural transcription factor LEF-1. Nature. 1995; 376 (6543): 791-5.

[38] Štros M, Launholt D, Grasser KD. The HMG-box: A versatile protein domain occurring in a wide variety of DNA-binding proteins. Cellular and Molecular Life Sciences. 2007; 64(19-20): 2590-606.

[39] Thomas JO. HMG1 and 2: architectural DNA-binding proteins. Biochem Soc Trans. 2001; 29(Pt 4): 395-401.

[40] Boutonnet N, Hui X, Zakrzewska K. Looking into the grooves of DNA. Biopolymers. 1993; 33(3): 479-90.

[41] Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. Origins of specificity in protein-DNA recognition. Annu Rev Biochem. 2010; 79: 233-69.

[42] Nadassy K, Wodak SJ, Janin J. Structural features of protein-nucleic acid recognition sites. Biochemistry. 1999; 38(7): 1999-2017.

[43] Pabo CO, Nekludova L. Geometric analysis and comparison of protein-DNA interfaces: why is there no simple code for recognition? J Mol Biol. 2000; 301(3): 597-624.

[44] Kessler C, Manta V. Specificity of restriction endonucleases and DNA modification methyltransferases a review (Edition 3). Gene. 1990; 92(1-2): 1-248.

[45] Ganguly A, Rajdev P, Williams SM, Chatterji D. Nonspecific interaction between DNA and protein allows for cooperativity: a case study with mycobacterium DNA binding protein. J Phys Chem B. 2012; 116(1): 621-32.

[46] Agback P, Baumann H, Knapp S, Ladenstein R, Härd T. Architecture of nonspecific protein-DNA interactions in the Sso7d-DNA complex. Nat Struct Biol. 1998; 5(7): 579-84.

[47] Weigel C, Schmidt A, Rückert B, Lurz R, Messer W. DnaA protein binding to individual DnaA boxes in the Escherichia coli replication origin, oriC. EMBO J. 1997; 16(21): 6574-83.

[48] Fass D, Bogden CE, Berger JM. Crystal structure of the N-terminal domain of the DnaB hexameric helicase. Structure. 1999 Jun; 7(6): 691-8.

[49] Augustin MA, Huber R, Kaiser JT. Crystal structure of a DNA-dependent RNA polymerase (DNA primase). Nat Struct Biol. 2001; 8(1): 57-61.

[50] Reyes-Lamothe R, Sherratt DJ, Leake MC. Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. Science. 2010; 328(5977): 498-501.

[51] Of R, Dna O, In D, et al. Serial Review : Oxidative DNA Damage and Repair of oxidative DNA damage in nuclear and mitochondrial DNA, and some changes with aging in Science (80-). 2002; 32(9): 804-12.

[52] Ma W, Westmoreland JW, Gordenin DA, Resnick MA. Alkylation base damage is converted into repairable double-strand breaks and complex intermediates in g2 cells lacking ap endonuclease. PLoS Genet. 2011; 7(4).

[53] Zharkov DO. Base excision DNA repair. Cell Mol Life Sci. 2008; 65(10): 1544-65.

[54] Kunkel TA, Erie DA. DNA mismatch repair. Annu Rev Biochem. 2005; 74: 681-710.

[55] Karanjawala ZE, Lieber MR. DNA damage and aging. Mech Ageing Dev. 2004; 125(6): 405-16.

[56] Smyth MS, Martin JH. x ray crystallography. Mol Pathol. 2000; 53(1): 8-14.

[57] Pervushin K, Riek R, Wider G, Wuthrich K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci. 1997 Nov 11; 94(23): 12366-71.

[58] Salzmann M, Pervushin K, Wider G, Senn H, Wuthrich K. TROSY in triple- resonance experiments: New perspectives for sequential NMR assignment of large proteins. Proc Natl Acad Sci. 1998 Nov 10; 95(23): 13585-90.

[59] Campagne S, Gervais V, Milon A. Nuclear magnetic resonance analysis of protein-DNA interactions. J R Soc Interface. 2011; 8(61): 1065-78.

[60] Garner MM, Revzin A. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: Application to components of the Escherichia coli lactose operon regulatory system. Nucleic Acids Res. 1981; 9(13): 3047-60.

[61] Fried M, Crothers DM. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. 1981; 9(23): 6505-25.

[62] Collas P. The current state of chromatin immunoprecipitation. Mol Biotechnol. 2010; 45(1): 87-100.

[63] Brenowitz M, Senear DF, Shea MA, Ackers GK.; Quantitative DNase footprint titration: A method for

studying protein-DNA interactions. Methods Enzymol. 1986; 130: 132-81.

[64] Tanious F a., Nguyen B, Wilson WD. Biosensor-Surface Plasmon Resonance Methods for Quantitative Analysis of Biomolecular Interactions. Methods Cell Biol. 2008; 84(07): 53-77.

[65] Joung JK, Ramm EI, Pabo CO. A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. Proc Natl Acad Sci U S A. 2000; 97(13): 7382-7.

[66] Bulyk ML. Discovering DNA regulatory elements with bacteria. Nat Biotechnol. 2005; 23(8): 942-4.

Received: August 10, 2015 Revised: October 18, 2015 Accepted: October 20, 2015