SUPRAMOLECULAR BIOORGANIC CHEMISTRY: DNA BINDING AGENTS

Abstract

Noncovalent DNA recognition by pyrrole oligopeptides agents have been illustrated in this chapter, and is contrasted with biomolecular DNA recognition. Although the size, surface dimensions, and structure of DNA present novel opportunities and challenges, the principles and forces involved in DNA recognition are comparable to those observed elsewhere in the larger field of supramolecular chemistry.

Keywords: Noncovalent interaction, molecular recognition, DNA binding agents

Authors

Netra Pal Singh
Department of Chemistry
D.D.U. Gorakhpur University
U. P. India
npsmcm.in@gmail.com

Maitreyi Singh
Department of Chemistry
D.D.U. Gorakhpur University
U. P. India

I. INTRODUCTION

Supramolecular chemistry is one of the most popular and faster growing areas of experimental chemistry. It is highly interdisciplinary in nature and attract not only chemist but biochemist, biologist, environmental scientist engineers and physicist. It is closely related to the bio-organic and bioinorganic chemistry. Its foundation was laid down less than fifty years ago. In 1987, the founding fathers of supramolecular chemistry were Charles J. Pederson, Donald J. Cram, Jean-Marie Lehn. They were awarded Noble prize in chemistry "*for their development and use of molecules with structure-specific interactions of high selectivity"*. The concept of supramolecular chemistry deals with the introduction of Host-Guest interaction. In this the host molecule is a molecular entity consisting of convergent binding site where as the guest molecule plays a role in the form of divergent binding site. The host guest relationship is based on the backbone of non-covalent interaction. The enhancement of this branch of science needs more attention for its growth and development where it requires the resources of molecular chemistry involving the chemistry of covalent bond and the different sorts of frame-ups with those of non-covalent bonds which leads to the formation of supramolecular entities comprising fruitful features such as molecular recognition, self- replication and self-assembly. In the field of supramolecular chemistry various designs and structures have been proposed by the scientist which reflects the apprehension of structural and functional relationship between the non-covalent bonds. Supramolecular chemistry holds well-established discipline of chemical science. As the supramolecular chemistry revolves around the concept of non-covalent interactions in between the building blocks, it exhibits stimuli-responsive actions. The chief criteria for molecular recognition, self-assembly, catalysis and self-replication is specifically headed by the intra and intermolecular non-covalent interactions. [1-3] Molecular recognition is a binding of guest molecule to complementary host molecules to form host-guest complex. The non-covalent interactions including hydrogen bonding, л-л interactions, van der Waals and electrostatic interactions, completely elaborates its utility in different biological systems of chemistry and thus possessing "special nature". Supramolecular chemistry focuses on understanding the fundamentals of noncovalent interactions and how to apply them to artificial systems. The planned and purposeful linking of various chemical entities is the subject of supramolecular chemistry.

1. Noncovalent interactions in bioorganic chemistry

- **Electrostatic interactions:** The noncovalent interaction can be studied well in electrostatic interactions where there is an interaction between large molecules such as proteins and DNA bearing net electrical charges which gets engage to small ions like Na^+, Mg^{2+} and Cl^-, solvent and other charged molecules in the cell. Among all non-covalent interactions, electrostatic interaction has longest range and is non-directional. Most of the relevant example of biomolecules pertaining electrostatic interaction includes salt bridges or ion pairs in protein.

- **Hydrogen bonding interactions:** The other important type of interaction is the hydrogen bonding interaction constitute vital role in the formation of organic molecular host-guest complexes. The hydrogen bonding interaction is found between the hydrogen atom covalently bonded to the donor atom (such as NH, NH_2, or -OH) and are also created when two non-bonding electrons act as the acceptor group (such

as C=O or =N-). [4]. The hydrogen bond possesses both kind of properties of covalent and non-covalent interactions. The bond angle and bond length are nearly fixed and are generally shorter than would be predicted from the sum of the van der Waals radii, which serves as an indicator of covalent character. Its relatively low bond energy when compared to covalent bonds serves as an example of its noncovalent nature. Edge-to-face aromatic interactions lead to a unique sort of hydrogen bond [5]. Here, the electron-rich face of one aromatic molecule attracts the proton of another. This kind of contact increases the rigidity of the involved molecules' linear orientation, which significantly affects the tertiary structure of proteins. Edge-to-face interactions have energies that fall midway between electron donor acceptor interactions and conventional hydrogen bond. Among the noncovalent interactions so far studied also includes the electron donor-acceptor interaction that is the aggregation of electrostatic interaction and charge transfer effects.

- **Van der Waals interactions*:*** The other kind of most common interaction is Van der Waal interaction. Here, atoms or molecules are smoothly driven to one another before the outermost electron orbitals begin to overlap. Repulsion increases significantly as the distance between two centers become small. The optimum radius for the closest molecular packing is known as the van der Waals radius. There exists an attraction between two different polarizable aromatic rings which configures a specific type of van der Waals interaction, often known as the n-electron stacking force. They are crucial in the construction of the DNA double helix.

- **Hydrophobic interactions:** Hydrophobic interaction is a kind of noncovalent interaction which deals with the affiliation of hydrophobic molecule in aqueous solution, this occurs by the loss of water molecules from the solvation shell and an accompanying increase in entropy [6,7]. It serves as the primary force of attraction among hydrophobic groups in host-guest molecular combinations in water. It lacks direction and specificity. The van der Waals interactions are maximized by the internal arrangement of hydrophobic side chains, which also leads to a rise in the entropy of the entire system from the release of solvent molecules. As a result, the hydrophobic effect becomes the main element of stabilization for protein globular shapes. The primary force behind the self-assembly of membranes and liposomes is also the binding of substrates to enzymes and the interactions between antigens and antibodies.[8]

2. **Molecular recognition in bioorganic chemistry:** Molecular recognition was defined by Lehn as "a process including both binding and selection of substrates by a certain receptor molecule, as well as potentially a specialized function." This gives a vivid description of the occurrence of a framework of intermolecular (noncovalent) interactions with clearly defined structural characteristics. The morphology of the interaction site must be compatible with the interaction surface of the substrates in terms of shape, charge distribution, size, and nature. The terms "host" and "guest" molecules are frequently used to describe receptors and substrates in synthetic systems. Antibody action, enzymatic catalysis, as well as the application of the essential tenet of molecular biology all rest on the idea of molecular recognition. In 1894, Fischer proposed his "lock-and-key model," which states the fundamental molecular recognition concept to explain the geometric properties required for enzyme catalysis [9]. This paradigm proposes that a given

substrate fits a single enzyme like a key fits a single lock. However, Jorgensen, who cited multiple cases in which either the host, the guest, or even both experience significant structural changes on binding, has reexamined the idea of a rigid lock and key compatibility in his study on "Rusting of the Lock and Key Model for Protein-Ligand Binding."[10].

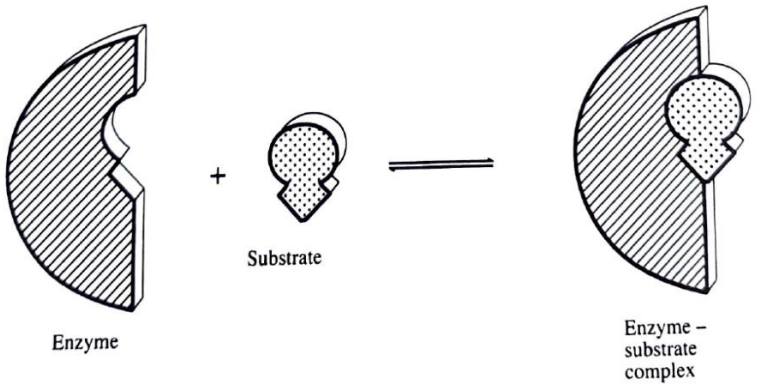


Figure 1: Fisher's rigid lock-and-key model

- **Chemical receptors:** Receptors are certain proteins that engage stereotypically with particular substances. The affinities by which the molecules are attached and the resultant effect are two different fundamental interactions together make up the receptor activity. The process of binding involves noncovalent interactions with the receptor's binding site, and the result is typically a change in the receptor's and possibly the substrate's conformation. The equilibrium orientation of the receptor is perturbed by substrate binding, and it is compensated via conformational change. In order to generate a physiological or biochemical reaction, such change can be transmitted via the receptor through noncovalent interaction and the peptide backbone. It is possible to determine the selectivity of substrates binding receptors, and it has been demonstrated for various systems that substrate structure affects binding constants. In terms of quality, the receptor binding site's size, shape, and charge distribution must complement with those of the substrate recognition domains. The best affinity results from the best fit. It happens frequently in nature for large proteins to recognize and bind tiny compounds. Similar to enzyme-substrate binding and antibody-antigen interactions, factors affecting molecular recognition that results in substrate binding to receptors are largely the same. However, the outcome following the binding phase is different. When an antibody interacts with a substrate, the substrate doesn't change (at least constitutionally), but the receptor does. The structural changes of the substrate after binding are of primary interest in catalytic antibody-substrate and enzyme-substrate interactions.

3. **DNA and nucleic acid-protein interaction:** The creation of double-stranded DNA is the classic instance of molecular recognition in biological systems. The backbone to which the bases are linked is made up of the phosphodiester linkages that connect the deoxyribose sugars in DNA. Through steric and electrostatic forces, the anionic phosphodiester moieties and the more inflexible sugars enforce a regular helical shape on the backbone.[11] Changes in the phosphodiester linkages and sugar moiety have an impact on the helical pattern. The double-helical structure, which was first hypothesized by Watson and Crick, is formed when two strands are coiled in an antiparallel direction

around a common axis. [12] The planes of the bases are perpendicular to the helix axis and extend into the middle of the helix. The pyrimidine bases and the purine bases (Adenine or Guanine) form hydrogen bonds to hold the two strands together (Thymine or Cytosine). As a result, every A is coupled with a T and every C with a G, creating two and three hydrogen bonds, respectively. This is because the bases are complimentary to one another. This extremely selective binding interaction is an excellent illustration of molecular recognition. The order of the matching base pair sequence has no bearing on the double-helical structure. The ability of DNA to function depends on this independence. Every type of DNA has a different base sequence because the specific order encodes genetic information. Several helical structures are observed in the nucleic acid and one of the most common DNA is B-DNA. It has a diameter of 2200 pm, a rise per residue of about 340 pm, and it has 10 residues per turn. Two separate grooves, a big major groove and a smaller minor groove, may be seen on the helix.

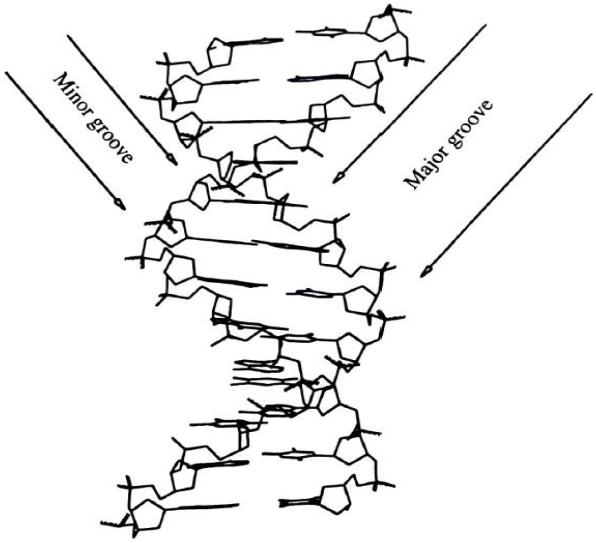


Figure 2: B-DNA showing major and minor groove

Independent of a thorough understanding of the structure of the interacting molecules, the fundamental idea of double complementarity as a need for molecular recognition imposes many predictions regarding the nature of the interactions. A complementary motif in the interaction protein is necessary to match the precise geometric features and dimensions of the helical motif in a DNA molecule. Protein α-helices and β-pleated sheet structures are similar to the B DNA helix in this regard, with α-helices closely fitting the major groove and an antiparallel B-ribbon (equivalent to a two-strand sheet) fitting the minor groove.[13]

4. DNA binding agents

Structure of DNA*:* Since the three-dimensional structure of DNA is well understood, DNA agent complexes can now be meaningfully interpreted in terms of their structural and mechanistic properties. There are two generally known right-handed antiparallel double helix types of DNA. A and B-form DNA as well as the Z-form DNA, a third left-handed antiparallel duplex, have all been discovered [14,15]. The dimensions of the major and minor grooves and the arrangement of the base pairs with respect to the helix axis are the most important differences between A and B-form DNA. A-DNA is distinguished by a deep and

narrow major groove and a wider, shallower minor groove, whereas the more prevalent B-DNA has a wide major groove and a narrow minor groove of almost the same depth. The A-DNA base pairs are slanted and displaced from the helix axis, which is located in the primary groove, whereas the B-DNA base pairs are orthogonal to the helix axis and are located in the middle of the structure. Notably, the major and minor grooves of A- and B-form DNA expose various sections of the base pairs. The sugar-phosphate backbone of the left-handed Z-DNA is zigzag, and it has essentially no major groove and just a very deep, narrow minor groove.[16]

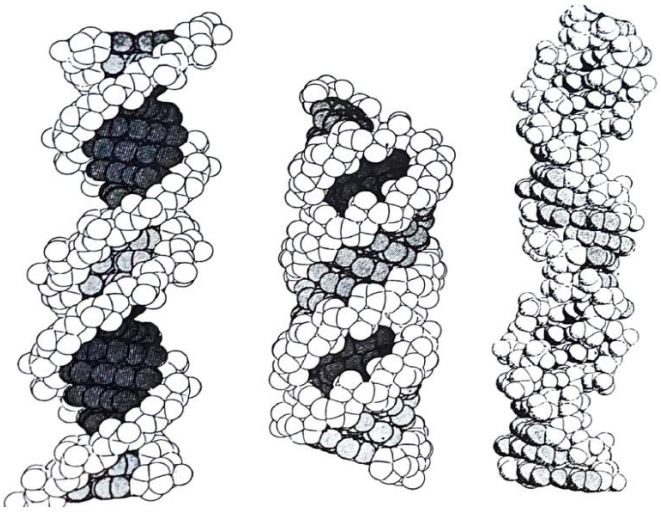


Figure 3: B-DNA A-DNA Z-DNA

As more genetic information about many creatures, including the human genome, becomes available, the emphasis has shifted to finding techniques to regulate the expression of certain genes. For the sake of this convenience, there is a development of various agents that can selectively bind to particular genes as well as regulate their expression. The capacity to turn a gene on or off is crucial for medical treatment as well as for elucidating the intricate and interconnected biological pathways in the cell. As more genetic information about many creatures, including the human genome, becomes available, the emphasis has shifted to finding techniques to regulate the expression of certain genes. For the sake of this convenience, there is a development of various agents that can selectively bind to particular genes as well as regulate their expression. The capacity to turn a gene on or off is crucial for medical treatment as well as for elucidating the intricate and interconnected biological pathways in the cell [17].

- **Noncovalent Minor Groove Binding Agents**

Distamycin and Netropsin (Pyrrole Oligopeptides): Distamycin and Neptropsin are the antitumor antibiotic which are mostly in the form of crescent shaped structure and consist of two and three N-methyl pyrrole carboxamides, respectively. Netropsin possess two cationic charges where each charge is placed at each end of its structure whereas distamycin have only one positive charge. It was discovered that netropsin binds in the minor groove and that it is only specific for AT-rich sequences. They strongly prefer AT-rich sequences and bind in the minor groove of B-DNA, extending either five (distamycin) or four (netropsin) base pairs. The special agents have been demonstrated to be duplex DNA-specific, and s-stranded DNA does not bind to them. One has gathered

comprehensive data on their sequence selectivity. X-ray diffraction studies and NMR studies have been done on their complexes with synthetic oligodeoxynucleotides in order to understand the structural features of their interaction with duplex DNA. Hydrogen-bonding, hydrophobic, van der Waals, and electrostatic interaction are all thought to play significant roles in how they bind to and recognize duplex DNA. According to the x-ray crystal structure of netropsin bound to the core AT-rich region of the Dickerson dodecamer, 5-d(CGCGAATTCGCG), the agents are kept in the minor groove by particular hydrogen bonds with the bases and van der Waals contacts with the deoxyribose backbone creating the minor groove walls, and adopt a bound conformation comparable to the natural helical shape of the minor groove [18]. The outside N-3 adenines of the 5'-d$(AATT)_2$ binding site are in contact with the two cationic ends of netropsin. Four intermolecular van der Waals contacts between the DNA bases and agent and three intermolecular hydrogen bonds between the netropsin's amide hydrogen atoms and the adenine N-3 or thymine C-2 carbonyl that face the minor groove were also noted. The adenine C-2 hydrogen forms two of these Van der Waals interactions with the pyrrole C-3 hydrogen, while the other two are formed in between adenine C-2 hydrogen and the methylene groups close towards the distal amide linkage. The 1:1complexes have showed that the carboxamide hydrogens either create bifurcated hydrogen bonds with two near thymine C-2 carbonyls on opposite strands of the helix or make hydrogen bonds to bridge the adenine N-3 and the thymine C-2 carbonyl on contiguous base and opposite helix strands. However, tight van der Waals interactions rather than hydrogen bonds are the forces driving one to attach to four or more A-T base pairs preferentially. The shorter width of the minor groove in an AT-rich sequence compared to a GC-rich sequence and the frictional projection of the guanine C-2 amine into the minor groove both increase the dominating stabilizing van der Waals interactions within the AT-rich sites. Furthermore, binding is made easier by the agents' complimentary shapes, that acquire coupled helical conformations which suit the minor groove's curvature and pitch.

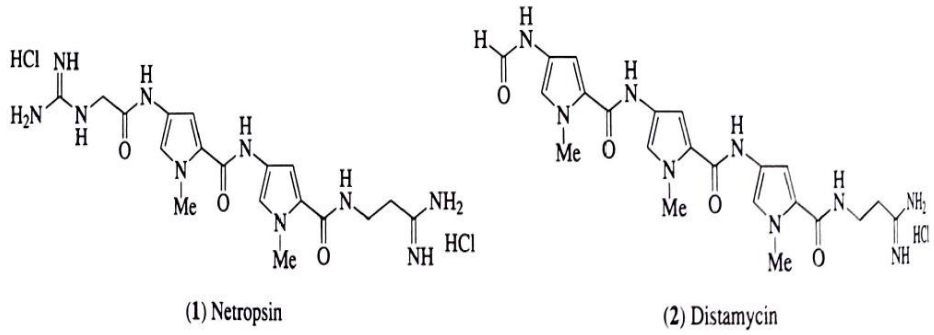
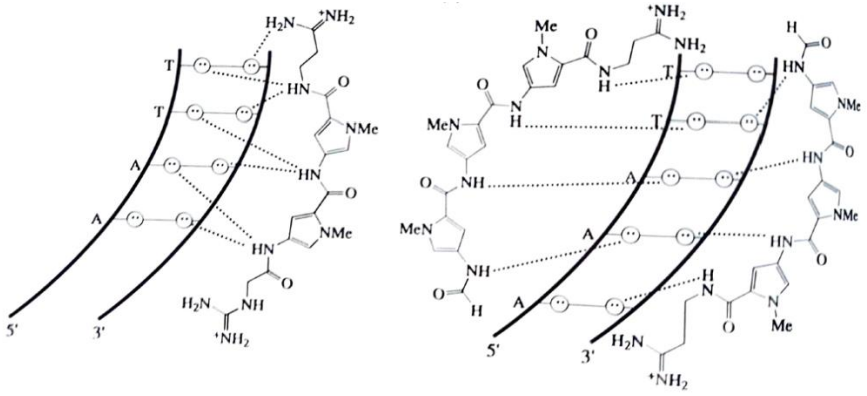


Figure 4: Structure of Netropsin and Distamycin

Distamycin has the ability to covalently bind to DNA to create 2:1 complex in which it binds parallel to one another in the minor groove. A 2:1 distamycin-DNA complex's initial X-ray crystal structure has been determined using the alternating B-DNA duplex 5'-d(ICICICIC)2 [19]. In an effort to improve the DNA binding affinity, enhanced versions of netropsin and distamycin have been created, [20,21] Oligo(N-methylpyrrolecarboxamides)$_n$ where n=2-9 comprise three to ten pyrrole carboxamides and may bind four to nine A-T base pairs, respectively. More linked tri- or tetra-N-

methylpyrrole carboxamides dimers and trimers that were joined head-to-head or tail-to-tail by fumaric acid or B-alanine linkers have also been created. It has been demonstrated that these synthetic compounds can bind duplex DNA up to 16 A-T base pairs [22,23]. More than five N-methyl pyrrole units cause extended analogues to shift out of phase in the study, and not all pyrrole carboxamide subunits can produce favorable results.



a. **b.**

b.

Figure 5: (a) 1:1 binding model for the complex formed between netropsin and a 5'-AATT-3' sequence (b) 2:1 binding model for the complex formed between distamycin and a 5'-TGTCA-3' sequence

II. CONCLUSION

The demand for individualized treatments and the requirement for probes that affect gene expression present the field of DNA identification with timely opportunities. This has traditionally indicated recognition that is specific to a sequence. Specifically, biomolecular motifs—modified peptides or oligonucleotides have been used to accomplish that. Selecting a single DNA structure type binding to simple duplex DNA and the wide range of other potential DNA structures will be the productive part. Understanding the distinct grooves and л-surfaces is probably crucial for binding to duplex DNA.

REFERENCE

[1] J.M. Lehn, *Angew. Chem., Int. Ed. Engl.,* 1988, **27**, 89.
[2] C. J. Pedersen, *Angew. Chem., Int. Ed. Engl.,* 1988, **27**, 1021.
[3] D. J. Cram, *Angew. Chem., Int. Ed. Engl.,* 1988, **27**, 1009.
[4] . M. H. Abraham, *Chem. Soc. Rev.,* 1993, **22**, 73.
[5] C. A. Hunter, *Chem. Soc. Rev.,* 1994, **23**, 101.
[6] C. K. Matthews and K. E. van Holde, "Biochemistry." Benjamin-Cummings, Menlo Park, CA, 1990, chap 2.
[7] H.-J. Schneider, *Angew. Chem., Int. Ed. Engl.,* 1991, **30**, 1417.
[8] R. Breslow, in "Supramolecular Chemistry." ed. V. Balzani and L. de Cola, Kluwer, Dordrecht, 1992, pp. 411-28
[9] E. Fischer, *Ber. Dtsch. Chem. Ges.*. 1894, **27**, 2985.
[10] W. L. Jorgensen, *Science,* 1991**, 254**, 954.
[11] L. Stryer, "Buchemistry, 3rd edn, Freeman. San Francisco, 1988, chap 35.
[12] J. D. Watson and F.H.C. Crick, *Nature (London),* 1953, **171**, 737,
[13] J. W. Lown, *Chemtracts Org. Chem.,* 1993**, 6**, 205.

[14] R. E. Dickerson, *Methods Enzymol*., 1992, **211,** 67.
[15] R. E. Dickerson, H. R. Drew, B. N. Conner, R. M. Wing, A. Y. Fratini, and M. L. Kopka, *Science*, 1982, **216**, 475.
[16] A. Rich, A. Nordheim, and A. H.-J. Wang, *Annu. Rev. Biochem.*, 1984, **53**, 791.
[17] Michael J. Hannon *Chem. Soc. Rev.*, 2007, **36**, 280–295.
[18] M. L. Kopka and T. A. Larsen, in "Nucleic Acid Targeted Drug Design, eds. C. L. Propst and T. J. Perun, Dekker, New York,1992, pp 303-74.
[19] X. Chen, B. Ramakrishnan, S. T. Rao, and M. Sundaralingam, *Nature Struct. Biol.,* 1994, **1,** 169.
[20] P.G. Schultz and P. B. *Dervan, Proc. Natl. Acad. Sci. USA*, 1983, **80**, 6834.
[21] R.S. Youngquist and P. B. Dervan, *Proc. Natl. Acad Sci. USA*, 1985, **82**, 2565.
[22] R.S. Youngquist and P. B. Dervan, *J. Am. Chem. Soc.,* 1987, **109**, 7564.
[23] W.S. Wade, M. Mrksich, and P. B. Dervan, *J. Am. Chem. Soc.,* 1992, **114,** 8783.

