©IDOSR PUBLICATIONS
International Digital Organization for Scientific Research ISSN: 2579-0730
IDOSR JOURNAL OF BIOLOGY,CHEMISTRY AND PHARMACY 1(1)1-6, 2017.

Overview of Deoxyribonucleic Acid (DNA)

Havey M. James

Biological Science University of Leicester England

Email: haveyjam6@gmail.com

ABSTRACT

This article is the overview of deoxyribonucleic acid (DNA). DNA, or deoxyribonucleic acid, is the hereditary material in humans and almost all other organisms. Nearly every cell in a person's body has the same DNA. Most DNA is located in the cell nucleus (where it is called nuclear DNA), but a small amount of DNA can also be found in the mitochondria (where it is called mitochondrial DNA or mtDNA). Mitochondria are structures within cells that convert the energy from food into a form that cells can use. The main role of DNA in the cell is the long-term storage of information. It is often compared to a blueprint, since it contains the instructions to construct other components of the cell, such as proteins and RNA molecules. The DNA segments that carry genetic information are called genes, but other DNA sequences have structural purposes, or are involved in regulating the expression of genetic information. In eukaryotes such as animals and plants, DNA is stored inside the cell nucleus, while in prokaryotes such as bacteria and archaea, the DNA is in the cell's cytoplasm. Unlike enzymes, DNA does not act directly on other molecules; rather, various enzymes act on DNA and copy its information into either more DNA, in DNA replication, or transcribe it into protein. DNA usually occurs as linear chromosomes in eukaryotes, and circular chromosomes in prokaryotes. DNA is central to biotechnology and medicine by virtue of the fact that it not only provides the basic blueprint for all life, it is a fundamental determinant of how the body functions and the disease process. Understanding the structure and function of DNA has helped revolutionise the investigation of disease pathways, assess an individual's genetic susceptibility to specific diseases, diagnose genetic disorders, and formulate new drugs. It is also critical to the identification of pathogens.

Keywords: Overview, Deoxyribonucleic, DNA, acid.

INTRODUCTION

Deoxyribonucleic acid (DNA) is a molecule composed of two chains that coil around each other to form a double helix carrying genetic instructions for the development, functioning, growth and reproduction of all known organisms and many viruses. DNA and ribonucleic acid (RNA) are nucleic acids; alongside proteins, lipids and complex carbohydrates (polysaccharides), nucleic acids are one of the four major types of macromolecules that are essential for all known forms of life [1].

DNA, or deoxyribonucleic acid, is the hereditary material in humans and almost all other organisms. Nearly every cell in a person's body has the same DNA. Most DNA is located in the cell nucleus (where it is called nuclear DNA), but a small amount of DNA can also be found in the mitochondria (where it is called mitochondrial DNA or mtDNA). Mitochondria are structures within cells that convert the energy from food into a form that cells can use

The two DNA strands are also known as polynucleotides as they are composed of simpler monomeric units called nucleotides. Each nucleotide is composed of one of four nitrogen-containing nucleobases (cytosine [C], guanine [G], adenine [A] or thymine [T]), a sugar called

deoxyribose, and a phosphate group. The nucleotides are joined to one another in a chain by covalent bonds between the sugar of one nucleotide and the phosphate of the next, resulting in an alternating sugar-phosphate backbone [2]. The nitrogenous bases of the two separate polynucleotide strands are bound together, according to base pairing rules (A with T and C with G), with hydrogen bonds to make double-stranded DNA. The complementary nitrogenous bases are divided into two groups, pyrimidines and purines. In DNA, the pyrimidines are thymine and cytosine; the purines are adenine and guanine [3] [4].

Both strands of double-stranded DNA store the same biological information. This information is replicated as and when the two strands separate [5]. A large part of DNA (more than 98% for humans) is non-coding, meaning that these sections do not serve as patterns for protein sequences. The two strands of DNA run in opposite directions to each other and are thus antiparallel. Attached to each sugar is one of four types of nucleobases (informally, bases). It is the sequence of these four nucleobases along the backbone that encodes genetic information. RNA strands are created using DNA strands as a template in a process called transcription, where DNA bases are exchanged for their corresponding bases except in the case of thymine (T), which RNA substitutes for uracil (U). Under the genetic code, these RNA strands specify the sequence of amino acids within proteins in a process called translation [6].

Within eukaryotic cells, DNA is organized into long structures called chromosomes. Before typical cell division, these chromosomes are duplicated in the process of DNA replication, providing a complete set of chromosomes for each daughter cell. Eukaryotic organisms (animals, plants, fungi and protists) store most of their DNA inside the cell nucleus as nuclear DNA, and some in the mitochondria as mitochondrial DNA or in chloroplasts as chloroplast DNA [7]. In contrast, prokaryotes (bacteria and archaea) store their DNA only in the cytoplasm, in circular chromosomes. Within eukaryotic chromosomes, chromatin proteins, such as histones, compact and organize DNA. These compacting structures guide the interactions between DNA and other proteins, helping control which parts of the DNA are transcribed [8].

DNA was first isolated by Friedrich Miescher in 1869. Its molecular structure was first identified by Francis Crick and James Watson at the Cavendish Laboratory within the University of Cambridge in 1953, whose model-building efforts were guided by X-ray diffraction data acquired by Raymond Gosling, who was a post-graduate student of Rosalind Franklin. DNA is used by researchers as a molecular tool to explore physical laws and theories, such as the ergodic theorem and the theory of elasticity [9] [10]. The unique material properties of DNA have made it an attractive molecule for material scientists and engineers interested in micro- and nano-fabrication. Among notable advances in this field are DNA origami and DNA-based hybrid materials.

The information in DNA is stored as a code made up of four chemical bases:

- Adenine (A),
- Guanine (G),
- Cytosine (C),
- Thymine (T).

Human DNA consists of about 3 billion bases, and more than 99 percent of those bases are the same in all people. The order, or sequence, of these bases determines the information available for building and maintaining an organism, similar to the way in which letters of the alphabet appear in a certain order to form words and sentences [11].

Functions of DNA

All known cellular life and some viruses contain DNA. The main role of DNA in the cell is the long-term storage of information. It is often compared to a blueprint, since it contains the instructions to construct other components of the cell, such as proteins and RNA molecules [12]. The DNA segments that carry genetic information are called genes, but other DNA sequences have structural purposes, or

are involved in regulating the expression of genetic information. In eukaryotes such as animals and plants, DNA is stored inside the cell nucleus, while in prokaryotes such as bacteria and archaea, the DNA is in the cell's cytoplasm. Unlike enzymes, DNA does not act directly on other molecules; rather, various enzymes act on DNA and copy its information into either more DNA, in DNA replication, or transcribe it into protein. Other proteins such as histones are involved in the packaging of DNA or repairing the damage to DNA that causes mutations [13] [14]. DNA is a long polymer of simple units called nucleotides, which are held together by a backbone made of sugars and phosphate groups. This backbone carries four types of molecules called bases and it is the sequence of these four bases that encodes information. The major function of DNA is to encode the sequence of amino acid residues in proteins, using the genetic code. To read the genetic code, cells make a copy of a stretch of DNA in the nucleic acid RNA. These RNA copies can then used to direct protein synthesis, but they can also be used directly as parts of ribosomes or spliceosomes.

Biological functions

DNA usually occurs as linear chromosomes in eukaryotes, and circular chromosomes in prokaryotes. The set of chromosomes in a cell makes up its genome; the human genome has approximately 3 billion base pairs of DNA arranged into 46 chromosomes. The information carried by DNA is held in the sequence of pieces of DNA called genes. Transmission of genetic information in genes is achieved via complementary base pairing. For example, in transcription, when a cell uses the information in a gene, the DNA sequence is copied into a complementary RNA sequence through the attraction between the DNA and the correct RNA nucleotides [15]. Usually, this RNA copy is then used to make a matching protein sequence in a process called translation, which depends on the same interaction between RNA nucleotides. In alternative fashion, a cell may simply copy its genetic information in a process called DNA replication. The details of these functions are covered in other articles; here the focus is on the interactions between DNA and other molecules that mediate the function of the genome [16].

Genes and genomes

Each length of DNA that codes for a specific protein is called a gene. For instance, one gene codes for the protein insulin, the hormone that helps control levels of sugar in the blood. Humans have around 20,000–30,000 genes, although estimates varyTrusted Source.

Our genes only account for around 3 percent of our DNA, the remaining 97 percent is less well understood [17]. The outstanding DNA is thought to be involved in regulating transcription and translation.

Genomic DNA is tightly and orderly packed in the process called DNA condensation, to fit the small available volumes of the cell. In eukaryotes, DNA is located in the cell nucleus, with small amounts in mitochondria and chloroplasts. In prokaryotes, the DNA is held within an irregularly shaped body in the cytoplasm called the nucleoid. The genetic information in a genome is held within genes, and the complete set of this information in an organism is called its genotype. A gene is a unit of heredity and is a region of DNA that influences a particular characteristic in an organism. Genes contain an open reading frame that can be transcribed, and regulatory sequences such as promoters and enhancers, which control transcription of the open reading frame [18].

In many species, only a small fraction of the total sequence of the genome encodes protein. For example, only about 1.5% of the human genome consists of protein-coding exons, with over 50% of human DNA consisting of non-coding repetitive sequences. The reasons for the presence of so much noncoding DNA in eukaryotic genomes and the extraordinary differences in genome size, or C-value, among species, represent a long-standing puzzle known as the "C-value enigma". However, some DNA sequences that do not code protein may still encode functional non-coding RNA molecules,

which are involved in the regulation of gene expression [19].

How does DNA create proteins?

For genes to create a protein, there are two main steps:

Transcription: The DNA code is copied to create messenger RNA (mRNA). RNA is a copy of DNA, but it is normally single-stranded. Another difference is that RNA does not contain the base thymine (T), which is replaced by uracil (U).

Translation: The mRNA is translated into amino acids by transfer RNA (tRNA).

mRNA is read in three-letter sections called codons. Each codon codes for a specific amino acid or building block of a protein. For instance, the codon GUG codes for the amino acid valine.

Base pairing

In a DNA double helix, each type of nucleobase on one strand bonds with just one type of nucleobase on the other strand. This is called complementary base pairing. Here, purines form hydrogen bonds to pyrimidines, with adenine bonding only to thymine in two hydrogen bonds, and cytosine bonding only to guanine in three hydrogen bonds. This arrangement of two nucleotides binding together across the double helix is called a Watson-Crick base pair. Another type of base pairing is Hoogsteen base pairing where two hydrogen bonds form between guanine and cytosine. As hydrogen bonds are not covalent, they can be broken and rejoined relatively easily. The two strands of DNA in a double helix can thus be pulled apart like a zipper, either by a mechanical force or high temperature [20]. As a result of this base pair complementarity, all the information in the double-stranded sequence of a DNA helix is duplicated on each strand, which is vital in DNA replication. This reversible and specific interaction between complementary base pairs is critical for all the functions of DNA in organisms.

Top, a GC base pair with three hydrogen bonds [21]. Bottom, an AT base pair with two hydrogen bonds. Non-covalent hydrogen bonds between the pairs are shown as dashed lines. The two types of base pairs form different numbers of hydrogen bonds, AT forming two hydrogen bonds, and GC forming three hydrogen bonds (see figures, right). DNA with high GC-content is more stable than DNA with low GC-content.

As noted above, most DNA molecules are actually two polymer strands, bound together in a helical fashion by noncovalent bonds; this double-stranded (dsDNA) structure is maintained largely by the intrastrand base stacking interactions, which are strongest for G,C stacks. The two strands can come apart—a process known as melting—to form two single-stranded DNA (ssDNA) molecules. Melting occurs at high temperature, low salt and high pH (low pH also melts DNA, but since DNA is unstable due to acid depurination, low pH is rarely used).

The stability of the dsDNA form depends not only on the GC-content (% G,C basepairs) but also on sequence (since stacking is sequence specific) and also length (longer molecules are more stable) [22]. The stability can be measured in various ways; a common way is the "melting temperature", which is the temperature at which 50% of the ds molecules are converted to ss molecules; melting temperature is dependent on ionic strength and the concentration of DNA. As a result, it is both the percentage of GC base pairs and the overall length of a DNA double helix that determines the strength of the association between the two strands of DNA. Long DNA helices with a high GC-content have stronger-interacting strands, while short helices with high AT content have weaker-interacting strands. In biology, parts of the DNA double helix that need to separate easily, such as the TATAAT Pribnow box in some promoters, tend to have a high AT content, making the strands easier to pull apart.

In the laboratory, the strength of this interaction can be measured by finding the temperature necessary to break the hydrogen bonds, their melting temperature (also called Tm value). When all the base pairs in a DNA double helix melt, the strands separate and exist in solution as two entirely independent molecules. These single-stranded DNA molecules have no single common shape, but some conformations are more stable than others.

Application

The analysis of DNA is pivotal to understanding both the biological mechanisms of life and diseases that arise when this process goes wrong. Many different applications have been developed to understand this process. Today scientists can analyse the molecule through a range of techniques, including DNA sequencing which helps work out its structure, through to PCR, which rapidly amplifies tiny quantities of DNA into billions of copies. Such techniques underpin all tests carried out today to for example identify a genetic mutation that causes cancer, or to determine whether a person carries a gene for a hereditary disease that can be passed on to their offspring. In addition, scientists have found ways to manipulate and construct new forms of DNA, known as recombinant DNA or gene cloning. Such technology is crucial to the mass production of many drugs, such as interferon, and the development of gene therapy.

CONCLUSION

DNA is central to biotechnology and medicine by virtue of the fact that it not only provides the basic blueprint for all life, it is a fundamental determinant of how the body functions and the disease process. Understanding the structure and function of DNA has helped revolutionise the investigation of disease pathways, assess an individual's genetic susceptibility to specific diseases, diagnose genetic disorders, and formulate new drugs. It is also critical to the identification of pathogens. Aside from its medical uses, the fact that DNA is unique to each individual makes it a vital forensic tool identifying criminals, the remains of a missing person, and determining the biological parent of a child. Within agriculture DNA is also used to help improve animal livestock and plants.

REFERENCES

1. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Peter W (2002). Molecular Biology of the Cell (Fourth ed.). New York and London: Garland Science. ISBN 0-8153-3218-1.
2. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2014). Molecular Biology of the Cell (6th ed.). Garland. p. Chapter 4: DNA, Chromosomes and Genomes. ISBN 978-0-8153-4432-2.
3. Berg J, Tymoczko J, Stryer L (2002). Biochemistry. W.H. Freeman and Company. ISBN 0-7167-4955-6.
4. Carr S (2012). "Watson-Crick Structure of DNA". Memorial University of Newfoundland. Archived from the original on 19 July 2016. Retrieved 13 July 2016.
5. Casella E, Markewych O, Dosmar M, Heman W (1999) Production and expression of dTMP-enriched DNA of bacteriophage SP15. J Virology 28 (3) 753–66
6. Ghosh A, Bansal M (2003). "A glossary of DNA structures from A to Z". Acta Crystallographica Section D. 59 (Pt 4): 620–26.
7. Gregory SG, Barlow KF, McLay KE, Kaul R, Swarbreck D, Dunham A, et al. (2006). "The DNA sequence and biological annotation of human chromosome 1". Nature. 441 (7091): 315–21.
8. Irobalieva RN, Fogg JM, Catanese DJ, Catanese DJ, Sutthibutpong T, Chen M, Barker AK, Ludtke SJ, Harris SA, Schmid MF, Chiu W, Zechiedrich L (2015). "Structural diversity of supercoiled DNA". Nature Communications. 6: 8440.
9. IUPAC-IUB Commission on Biochemical Nomenclature (CBN) (December 1970). "Abbreviations and Symbols for Nucleic Acids, Polynucleotides and their Constituents. Recommendations 1970". The Biochemical Journal. 120 (3): 449–54.
10. Kiljunen S, Hakala K, Pinta E, Huttunen S, Pluta P, Gador A, Lönnberg H, Skurnik M (2005). "Yersiniophage phiR1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine

replaced by deoxyuridine". Microbiology. 151 (Pt 12): 4093–102.
11. Mandelkern M, Elias JG, Eden D, Crothers DM (2001). "The dimensions of DNA in solution". Journal of Molecular Biology. 152 (1): 153–61.
12. Mashaghi A, Katan A (2013). "A physicist's view of DNA". De Physicus. 24e (3): 59–61.
13. Molnár P, Marton L, Izrael R, Pálinkás HL, Vértessy BG (2018) Uracil moieties in Plasmodium falciparum genomic DNA. FEBS Open Bio 8(11):1763–1772
14. Purcell A. "DNA". Basic Biology. Archived from the original on 5 January 2017.
15. Russell P (2001). iGenetics. New York: Benjamin Cummings. ISBN 0-8053-4553-1.
16. Saenger W (1984). Principles of Nucleic Acid Structure. New York: Springer-Verlag. ISBN 0-387-90762-9.
17. Simpson L (March 1998). "A base called J". Proceedings of the National Academy of Sciences of the United States of America. 95 (5): 2037–38.
18. Tropp BE (2012). Molecular Biology (4th ed.). Sudbury, Mass.: Jones and Barlett Learning. ISBN 978-0-7637-8663-2.
19. Uchiyama J, Takemura-Uchiyama I, Sakaguchi Y, Gamoh K, Kato S, Daibata M, Ujihara T, Misawa N, Matsuzaki S (2014). "Intragenus generalized transduction in Staphylococcus spp. by a novel giant phage". The ISME Journal. 8 (9): 1949–52.
20. Verma S, Eckstein F (1998). "Modified oligonucleotides: synthesis and strategy for users". Annual Review of Biochemistry. 67: 99–134.
21. Watson JD, Crick FH (1993). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid" (PDF). Nature. 171 (4356): 737–38.
22. Yakovchuk P, Protozanova E, Frank-Kamenetskii MD (2006). "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix". Nucleic Acids Research. 34 (2): 564–74.