Each of the cells in your body carries about 1.5 gigabytes of genetic information, an amount of information that would fill two CD ROMs or a small hard disk drive. Surprisingly, when placed in an appropriate egg cell, this amount of information is enough to build an entire living, breathing, thinking human being. Through the efforts of the international human genome sequencing projects, you can now read this information. Along with most of the biological research community, you can marvel at the complexity of this information and try to understand what it means. At the same time, you can wonder at the simplicity of this information when compared to the intricacy of the human body.

Read-Only Memory
DNA is read-only memory, archived safely inside cells. Genetic information is stored in an orderly manner in strands of DNA. DNA is composed of a long linear strand of millions of nucleotides, and is most often found paired with a partner strand. These strands wrap around one another in the familiar double helix, as shown here. The code is quite easy to read: you simply step down the strand of DNA one nucleotide at a time and read off the bases: A, T, C or G. This is exactly what your cells do: they scan down a messenger RNA (copied from the DNA), and use ribosomes to build proteins based on the code that is read. This is also how researchers determine the sequence of a DNA strand: they clip off one nucleotide at a time to see what it is.

Your Inheritance
Your genetic information, inherited from your parents, is your most precious possession. It guided the construction of your body in the first nine months of your life and it continues to control all of the basic functions of living. Each of your cells is constantly using this information, asking questions about how to control blood sugar levels and body temperature, how to digest different foods and how to deal with new environmental challenges, and thousands of other important questions. The answers are held in the DNA. Hundreds of different proteins are built to interact with this information: to read it and use it to build new proteins, to copy it when the cell divides, to store and protect it when it is not actively being used, and to repair the information when it becomes corrupted by chemicals or radiation.

A Central Icon
DNA is arguably one of the most beautiful molecules in living cells. Its graceful helix is pleasing to the eye. DNA is also one of the most familiar molecules, the central icon of molecular biology, easily recognized by everyone. To some, it may carry a negative connotation, being a pervasive symbol for activists against genetically engineered produce. To others, it may bring to mind advances in forensics such as the DNA fingerprinting used in many recent high-profile trials. Some may have seen it in science fiction, modified to build dinosaurs or store cryptic messages from aliens. To all it is a pervasive symbol of our growing understanding of the human body and our close kinship with the rest of the biosphere, and the moral and ethical issues that must be addressed in the face of that knowledge.

Molecular Information
DNA is perfect for the storage and readout of information. It is laden with information. Every surface and edge of the molecule carries information. The basic mechanism by which DNA stores and transmits genetic information was discovered in the 1950’s by Watson and Crick. This basic information is stored in the way that the bases match one another on opposite sides of the double helix—adenine with thymine, guanine with cytosine—forming a set of complementary hydrogen bonds. These are shown in the diagram (on the reverse) with red arrows. Additional ‘extragenetic’ information is read from the surfaces that are left exposed in the double helix. In the major groove (the wider of

About the RCSB PDB Molecule of the Month
Using selected molecules from the PDB archive, each feature includes an introduction to the structure and function of the molecule, a discussion of its relevance to human health and welfare, and suggestions for viewing and accessing further details.

The RCSB PDB Molecule of the Month is read by students, teachers, and scientists worldwide at www.pdb.org.

This November 2001 edition was written and illustrated by David S. Goodsell (RCSB PDB and The Scripps Research Institute).
DNA

Exploring the Structure

We often think of DNA as a perfect, smooth double helix. In reality, DNA has a lot of local structure. The small piece of DNA shown here, from PDB entry 1bna, shows some of the common variations. At the top, the helix is bent to the left, distorted by the way that the helices are packed into the crystal. At the bottom, two of the bases are strongly propeller twisted—they are not in one perfect plane. This improves the way that the bases stack on top of one another along each strand, stabilizing the whole double helix. As more and more structures of DNA are studied, it is becoming clear that DNA is a dynamic molecule, quite flexible on its own, which is bent, kinked, knotted and unknotted, unwound and rewound by the proteins that interact with it.

Variations on a Theme

DNA adopts the familiar smooth double helix, termed a B-helix, under the typical conditions found in living cells. An example is shown in the center, exemplified by the crystal structure in PDB entry 1bna, shown at the top superimposed over the idealized version of the B-helix. Under other conditions, however, DNA can form other structures, as revealed in two early crystal structures: PDB entries 1ana on the left and 2dcg on the right. The one on the left, with tipped bases and a deep major groove, is termed A-DNA. It is formed under dehydrating conditions. Also, RNA most often shows this form, because its extra hydroxyl group on the sugar gets in the way, making the B-form unstable (look, for instance, at the A-helical structure of transfer RNA shown in the October 2001 issue of *Molecule of the Month*). The form on the right, which winds in the opposite direction from A-DNA and B-DNA, is termed Z-DNA. It is found under high salt conditions and requires a special type of base sequence, with many alternating cytosine-guanine and guanine-cytosine base pairs.

Further information about DNA


The Nucleic Acid Database, ndbserver.rutgers.edu

References:


The Protein Data Bank (PDB) is a worldwide repository of the structures of large molecules of biological interest. The PDB was founded in 1971 with a deposit from William S. Hammes of the X-ray crystal structure of myoglobin. Since then, the PDB has expanded to contain over 90,000 entries describing structures of ribonucleic acid (RNA), deoxyribonucleic acid (DNA), proteins, viruses, organelles, and whole cells. The PDB is a service of the Rutgers, The State University of New Jersey, and is funded by the National Institutes of Health and the National Science Foundation.