STRUCTURAL ANALYSIS OF RADIO-INDUCED LESIONS OF DNA

The structural impact of spontaneous or radio-induced lesions of DNA is being more and more thoroughly studied by means of analytic techniques such as nuclear magnetic resonance and molecular dynamics. These methods can be used to study the conformational changes that occur between the time the lesions are created and their possible repair. The work conducted by the French Atomic Energy Commission (CEA) demonstrates in particular that several structures are generally associated with a given lesion, because of the temporary increase in the local flexibility of the DNA molecule. The equilibrium among the different conformations also depends on the DNA sequences involved and on the physical and chemical conditions of the medium, e.g., temperature, acidity, basicity.

Emmanuel Joly/CEA

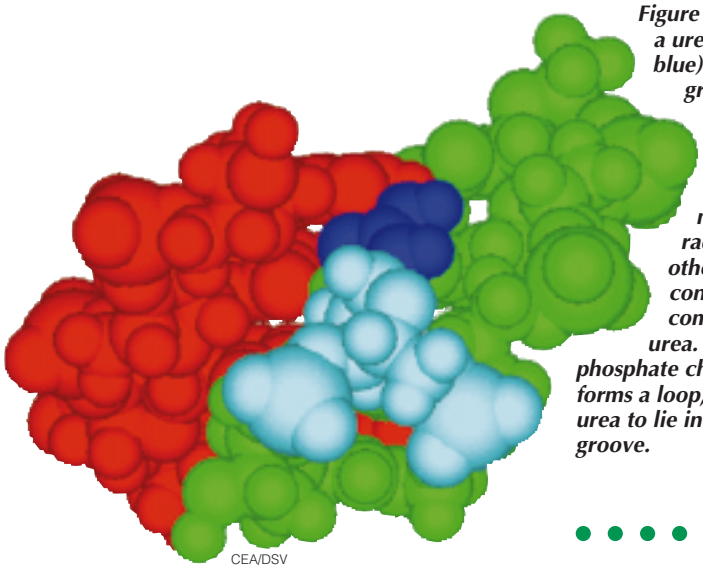
The nuclear magnetic resonance spectrometer set up at CEA/Saclay is used to determine the structures of biological macromolecules such as proteins and DNA fragments. The magnetic field produced by a superconducting coil maintained at the temperature of liquid helium in a cryostat enables us to study atoms with non-zero spin (magnetic momenta associated with nuclei resulting from the rotation of the different charges). The analysis of the interactions between such atoms, especially hydrogen atoms, affords steric information that helps resolve the molecular structure in solution. In the background we see the electronics and the computer system that stores, processes and analyses the raw data.

Two methods for the structural analysis of DNA

Ionizing radiation (X- or gamma rays) is the cause of numerous modifications of the DNA molecule because it leads to the creation of free radicals, which are highly reactive toward **DNA**. The study of the impact of ionizing radiation on the DNA structure, in comparison with its normal structure, helps to gain a better understanding of the mechanisms by which the molecule is recognized by the many proteins that are responsible for its upkeep (box A, ***DNA molecule, heredity vector***). The techniques used are chiefly nuclear magnetic resonance (NMR) and molecular dynamics. NMR is a non-destructive spectroscopic method that affords information on the local environment of atomic nuclei in solution. The information can be structural, such as the distance between two atoms(1), or a torsion angle. This technique also enables us to identify atoms involved in **hydrogen bonding**, and to access the molecule's dynamics. All this information is used to build one or more models that are representative of the average structure of the molecule in aqueous solution, water being the **cell** medium. The models obtained provide a basis for molecular dynamics calculations that simulate the movement of the atoms at a given temperature(2). These calculations validate the stability of the models and help us study the distortions of the DNA molecule.

Base pairs of differing stability

Among the **genotoxic** agents acting on the cell from outside, ionizing radiation is responsible for chemical lesions of **bases** or the total elimination of some bases by cleavage of the **sugar-base** bond. If the lost base is a **purine** (adenine or guanine), this elimination reaction is called *depurination*; if it is a **pyrimidine** (cytosine or thymine), the process is *depyrimidination*. The normal DNA molecule possesses several weak covalent bonds(3), favoring spontaneous modifications of bases. However, the rate of occurrence of these modifications can be increased by the products of water **radiolysis** (see *Radiolysis of water*). When a hydroxyl radical or a free electron comes into contact with DNA, a charge transfer ensues. This charge transfer may, for example, cause stepwise fragmentation of thymine, finally yielding a urea residue. The change in the chemical composition of the thymine to give urea induces a modification in the coding information(4) that was initially borne by the thymine. The presence of the urea in a DNA molecule confers **mutagenic** and **lethal** properties; **DNA-polymerase**, both during DNA replication (box B, ***Replication of DNA: near-perfect fidelity***) and during repair, preferentially places a thymine opposite the lesion and not an adenine as it should normally, because thymine forms a particularly stable structure with urea. The thymine-urea pair, inside the DNA **double helix**, is linked by two hydrogen bonds. Conversely, when an adenine is located opposite the urea the combination is unstable, and no structural study of this system is possible. The DNA-polymerase can also bypass the lesion, causing a jump in the reading of the **gene**, and so modifying the nature of the protein it codes for. In this case the urea is turned out of the helix. Instead of coming into contact with solvent the modified base lies in the major groove of the DNA, causing marked distortion of the **phosphodiester backbone** (Figure 1)..



CEA/DSV

Figure 1. Positioning of a urea residue (dark blue) located in the green strand of a DNA molecule, resulting from the modification of a thymine by ionizing radiation. The other strand (red) contains no base complementary to urea. Part of the sugar-phosphate chain (light blue) forms a loop, allowing the urea to lie in the major groove.

Base loss is potentially mutagenic

Abasic sites are formed spontaneously under normal conditions. The number of depurinations is estimated at $3 \cdot 10^{-11}$ per **nucleotide** per second in cell DNA. This process is therefore a frequent one. Extrapolation of ***in vitro*** evaluations of this frequency indicates about 10,000 depurinations per mammalian cell per day. In addition, external agents are able to accelerate the appearance of these sites by acting on the bases of the DNA molecule. Ionizing radiation will also act directly on the deoxyribose-base bond. A missing base, i.e., missing coding information, creates a potentially mutagenic situation. Genetic studies show that DNA-polymerase preferentially incorporates adenine and then gua-

(1) This distance can range from 2 to 5 Å. One angstrom = 10^{-10} m.
(2) Analysis of this simulation provides various sorts of information, e.g., the curvature of the DNA, the diffusion of water molecules, or the variation in the energy of the system.
(3) Chemical bond between two atoms formed by the sharing of one or more pairs of electrons.
(4) Information carried by a DNA base that orders the choice of an **amino acid** in the synthesis of a protein.

Replication of DNA : near-perfect fidelity

B

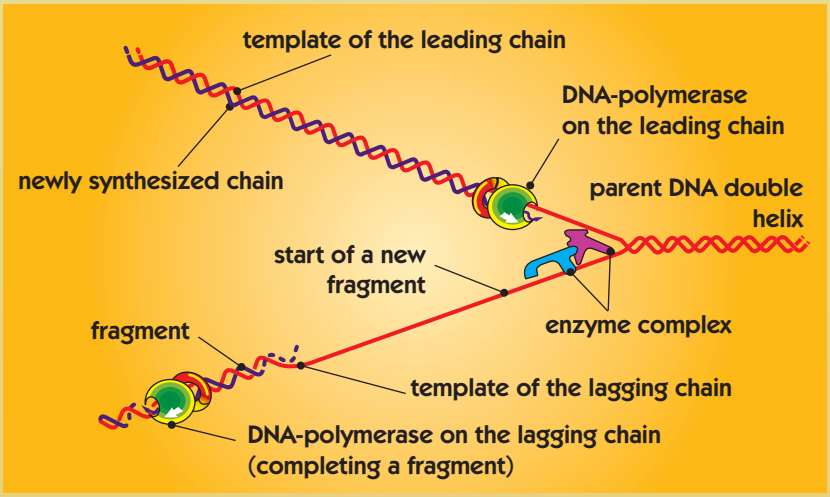
To ensure the transmission of the genetic message the **DNA** molecule has to be replicated a great number of times during the growth of the organism. Before it divides, every cell duplicates its DNA to produce two identical copies, each of them then being transferred to one of the new cells, which thus normally inherits all the genetic information of the parent cell.

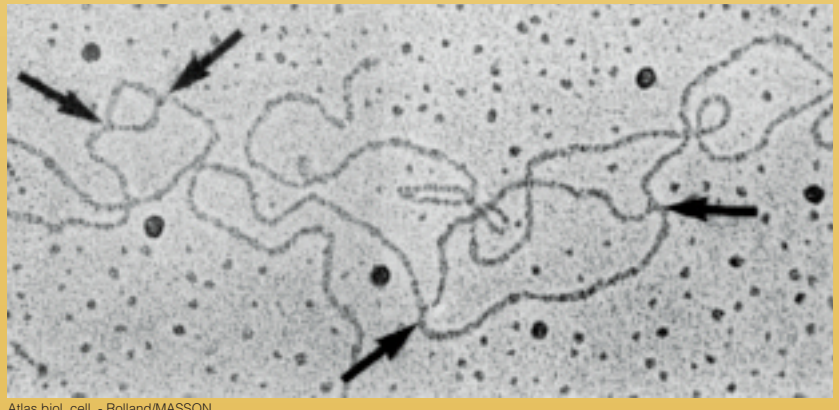
The replication of DNA must therefore be very reliable, and reproduce the **nucleotide** sequence accurately. The integrity of the molecule is safeguarded by a set of **proteins** that will recognize and repair any imperfections it may develop. Among other agents(1) ionizing radiation (especially X- and gamma rays) can be the cause of numerous modifications of the DNA molecule.

The replication of the double helix begins with the unwinding at a particular point of these two complementary chains. An **enzyme** moves along the ladder and severs the linkages that form the rungs. Two separate strands result, forming the two prongs of a replication fork (illustration). At each **replication fork** the DNA of the two new chains is synthesized by a complex composed of several enzymes, of which one is **DNA-polymerase** (diagram). One of the two chains, called the leading chain, is synthesized continuously; the other one, called the lagging chain, is synthesized segment by segment. Each parent strand serves as a template for the construction of a new double helix by stepwise addition of nucleotides that hook onto the half-rungs and build up the missing half of the original ladder. The nucleotide that has to be added at each step is selected so as to form a complementary **base** pair with the nucleotide opposite it on the starting chain (e.g., an adenine is linked to a thymine), thereby forming a new chain, the sequence of which is complementary to that of the parent chain. The chemical reactions brought into play provide the energy necessary for the replication and make it irreversible.

The genetic information is integrally duplicated. Two complete DNA double helix molecules are formed, each composed of an original strand and a newly synthesized strand bearing a nucleotide sequence identical to that of the parent helix, which served as the template. This mechanism of DNA replication is accordingly termed *semi-conservative*.

One of the most impressive features of DNA replication is its fidelity. Complex enzyme processes ensure that the inevitable random errors that arise are identified and corrected. Several correction mechanisms eliminate wrongly positioned nucleotides, enabling replication of each DNA nucleotide sequence with less than one error per billion added nucleotides. Even so, the replication mechanism occasionally skips or adds extra nucleotides, or even mixes them up. Any change of this type in a DNA sequence is a **mutation**, which will subsequently be passed on in all future cell generations, the faulty DNA sequences being copied just as faithfully as the correct ones.





Atlas biol. cell. - Rolland/MASSON

(1) The list of all the assaults that the environment can make on a cell is a long one: oxidative stress, heat shock, chemical stress, **genotoxic** stress, mechanical stress, osmotic shock...

nine opposite an abasic site. Structural studies of oligonucleotides containing an abasic site show that the DNA keeps its overall shape but adopts different conformations according to what base is located opposite the abasic site. If the base is a purine, it remains stacked inside the helix; an **oligonucleotide** containing adenine opposite the baseless site is, however, more stable than one containing guanine. The situation is more complex when the opposing base is a pyrimidine. Cytosine is turned out of the helix, while thymine can be located either inside or outside the helix. These results reveal the importance of the base stacking energy(5) in the structuring of the molecules.

Single-strand breaks and gaps can mostly be repaired

A single-**strand** break, or nick, in a DNA molecule is a frequent event; as many as 150,000 can occur per cell per day (Figure 2). If a strand is nicked in two places not too far apart a segment of the chain is lost (**deletion**), leaving a gap. The organism can also create gaps itself during the DNA repair process. A nucleotide gap is an intermediate in the most frequent repair process in the **bacterium** *Escherichia coli*. This type of repair is called repair by nucleotide excision. A gap is therefore a structure that is frequent, well recognized and easily repaired in living organisms. The study of the structure and the dynamics of oligonucleotides bearing either a single-strand break or a nucleotide gap has shown that a purine opposite the gap is always stacked inside the helix and that a cavity is formed. For a pyrimidine there is an equilibrium between two structures. In one conformation the base is inside the double helix and in the other one the pyrimidine is turned out of the helix and the two adjacent base pairs are stacked normally.

In all cases the presence of a gap increases the flexibility of the oligonucleotide at that position, amplifying its curvature. The hydration of the gap depends on the structure. If the structure is straight the water molecules form two layers; if it is bent they form a single layer. A network of hydrogen bonds is formed between the water molecules, and between these and the bases around the gap. The organizing of the water molecules in the cavity helps stabilize the overall structure. The explanation for these mechanisms is thermodynamic. The equilibria reached are governed by the gain in stacking energy obtained for a structure according to its contact area between bases. A pyrimidine, the stacking energy of which is lower than that of purine, is more easily turned out of the helix (Figure 3).

The mechanism of gap repair requires that the unpaired base lies inside the helix. For those conformations in which the base is located outside it the repair process fails in about 30% of cases.

Complex, resolved by crystallography, of human DNA-polymerase (yellow), an enzyme whose function is to synthesize DNA, and a molecule of DNA (red) with a single-strand break, indicated by the two arrows. The right angle formed by the DNA is an interesting feature.

Structure resolved by Sawaya et al., deposited at the Protein Data Bank (PDB) with reference 1BPZ.pdb

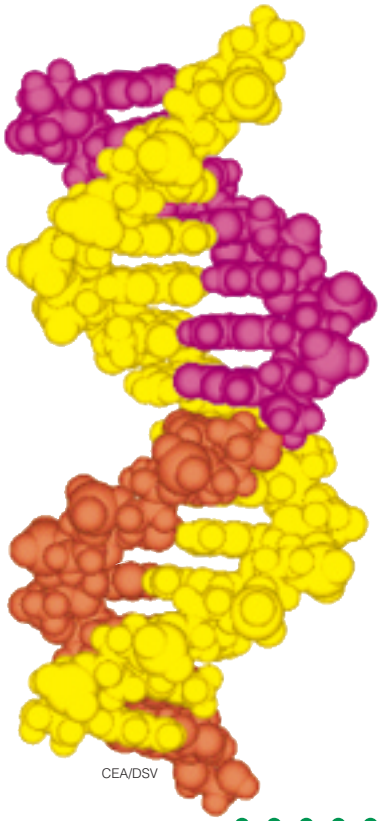
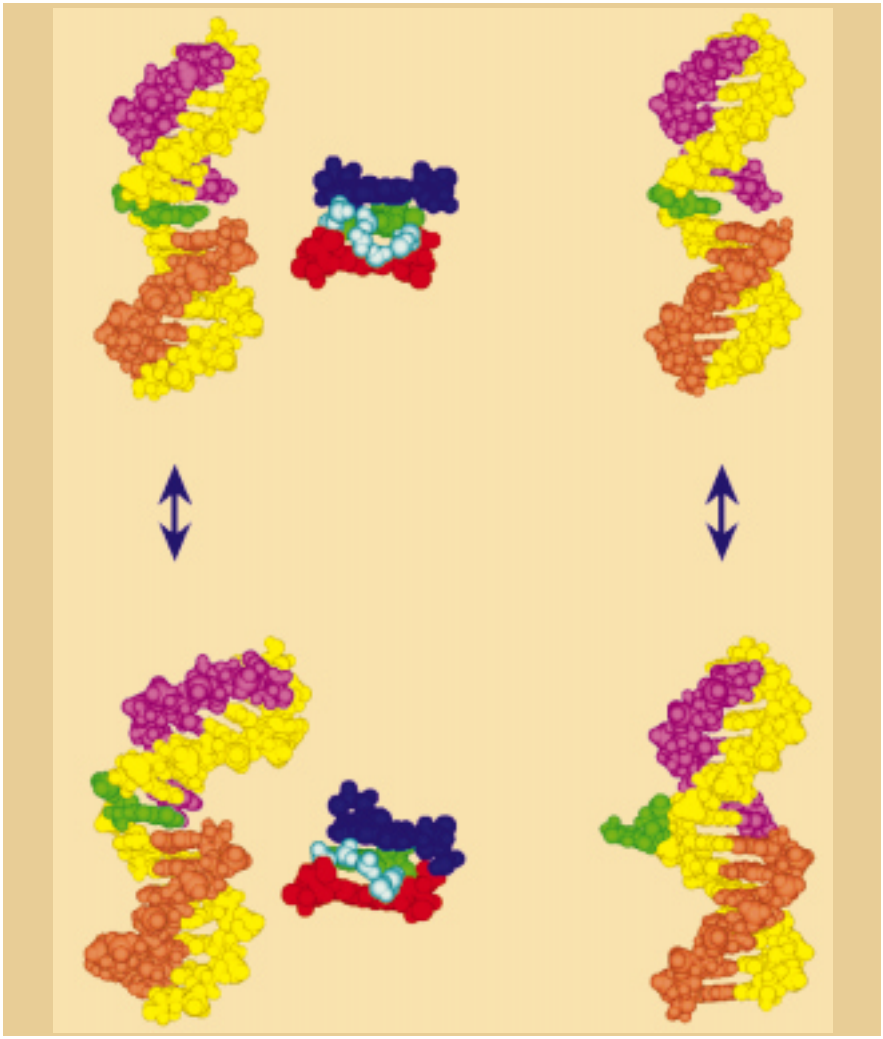


Figure 2. Structure of a DNA molecule displaying a single-strand break. The unbroken strand is yellow; the severed strand is purple on one side of the break and orange on the other.

(5) The architecture of the double helix of DNA is maintained first of all by the hydrogen bonds between the bases in the two strands, and secondly by the stacking energy, i.e., the contact area between bases on the same strand.



CEA/DSV

Figure 3. Structures of DNA molecules containing a nucleotide gap opposite a base (green): a purine (guanine) on the left, a pyrimidine (thymine) on the right. The unbroken strands are yellow, the complementary strands are purple and orange. In the case of purine the straight form (top) and the bent form (bottom) of the oligonucleotide correspond respectively to gap hydration structures with two layers and one layer of water molecules (blue). The two base pairs adjacent to the cavity are dark blue and red. In the case of pyrimidine the top drawing depicts the conformation in which the base is located inside the helix and the bottom one that in which the base is turned outward.

Proteins for specific structures?

There are a great many proteins associated with the DNA repair complex the functions of which are still unclear. Even so their presence is known to be compulsory. Some of them may be involved in the recognition of specific DNA conformations, such as that in which the base lies outside the helix. They may therefore be vitally important for the maintenance of the **genome**. Earlier results together with those of work in progress already demonstrate the remarkable performance of the repair systems, which to be efficient have to recognize and repair all the errors introduced in the genes. An unrepaired defect that does not jeopardize the life of the cell, i.e., the occurrence of a **mutation**, must remain a very rare event.

Yves Boulard
and **Georges Victor Fazakerley**
Department of Cellular and Molecular Biology
Life Sciences Division
CEA/Saclay - France