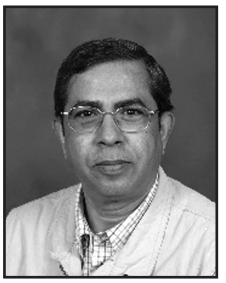
MUTATIONAL HOT SPOTS IN DNA: WHERE BIOLOGY MEETS PHYSICS

BY JULIA A. BERASHEVICH AND TAPASH CHAKRABORTY



SUMMARY

A lot of issues in modern science are interdisciplinary in nature and require coordinated efforts from physics, chemistry, biology and many other disciplines. The phenomena of mutation of the DNA code is one of those problems. The oxidative damage of DNA is a significant source for mutations in living cells that results in aging, carcinogenesis and many types of human cancers, plus it is linked to diseases such as AIDS, Parkinson's disease, Alzheimer's disease, and rheumatoid arthritis. Originally, mutation of the hereditary code was considered to be primarily a biological problem, but later investigations have clearly revealed that it requires contributions from physicists as well.

The human body is made up of many organs which are built from different types of cells performing various functions. The cells are the basic units of life. Several hundred thousand cells are reproduced every day to replace the aged and damaged cells. Most cells have a similar structure, they contain several functional components surrounded by a membrane protecting them from accidental damage. One of the most important components of the human cell is the nucleus, where necessary information for cell functionality and reproduction is stored. These instructions are coded in the genes - a part of the deoxyribonucleic acid (DNA) molecule.

DNA is the key molecule for cell reproduction and functionality. The structure of the DNA molecule is schematically presented in Figure 1 (a). This molecule is a linear polymer with a double-helical secondary structure where two strands are twisted around each other. Each strand contains a sugar-phosphate backbone and the attached bases connected to an opposite strand by hydrogen bonds. The hereditary information is coded by four unit bases: adenine (A), guanine (G), cytosine (C) and thymine (T). The structure of these bases allows formation of proper stable base pairs through the hydrogen bonds. The general rule for base pairing is that, thymine can only bond with adenine (T-A), while cytosine can only bond with guanine (C-G) (see Figure 1 (a)). This property is called complementarity [1].

How does our body control the processes of cell reproduction and functionality? The mechanism of cell reproduction consists in division of an undamaged cell - the parent cell which splits into two daughter cells. Prior to cell division nuclear division occurs when the parent DNA is duplicated. The duplication process consists of replication of the parent DNA into two new DNA molecules that provide transmission of hereditary information to the next cell generation. Schematically, replication of the DNA molecule is shown in Figure 1 (b). For the cell functionality, reading of instructions is produced by the transcription of DNA into RNA (Ribonucleic acid) molecules, using DNA as a template. The transcription process differs from replication only by its products, where the parent DNA causes the formation of a new RNA strand complimentary to one of the DNA strands. This RNA is then transported out of the nucleus where it is translated into specific protein molecules that make our bodies function. The groups of three bases within the RNA strand are used for coding a single amino acid within the resulting protein structure. It is similar to the way we store information in our everyday life using the binary code where a set of 0 and 1 can code any data.

DNA replication is crucial for cell functionality and reproduction and can be treated as the most important process in cell life. Infringement of replication or damage to the DNA molecule can induce mutations of the hereditary information thereby diminishing or altering cell functionality. DNA mutations result in aging, carcinogenesis and

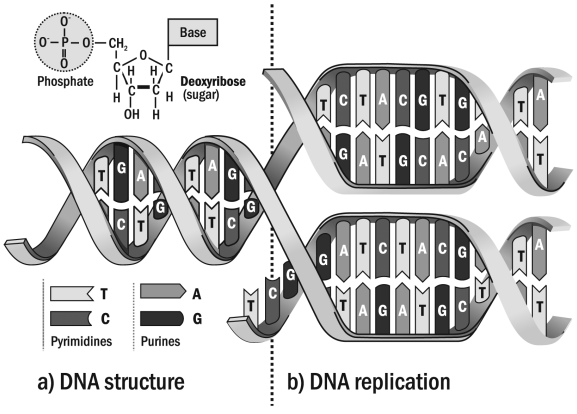


Fig 1 (a) The structure of the DNA molecule and (b) DNA replication process (schematic).

Julia A. Berashevich and Tapash Chakraborty [tapash@physics.umanitoba.ca], Department of Physics and Astronomy, University of Manitoba, Winnipeg, MB, R3T 2N2

many types of human cancers, plus it can directly cause diseases such as AIDS, Parkinson's disease, Alzheimer's disease, and rheumatoid arthritis. Not surprisingly, understanding the damage and mutation mechanisms in the DNA molecule has become an important topic in medicine and biology.

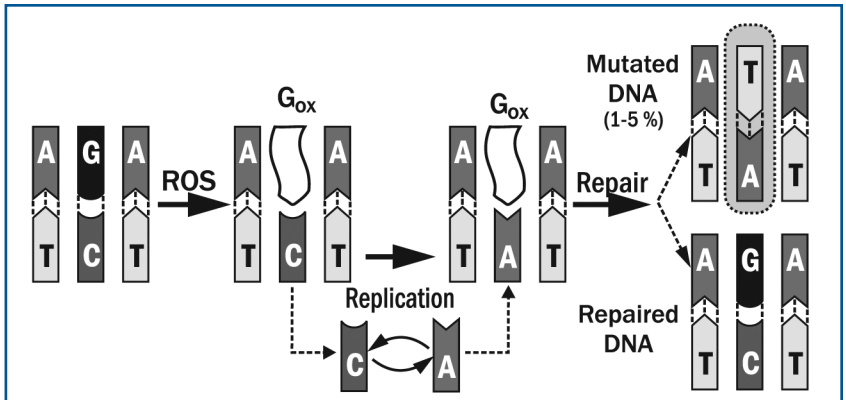


Fig. 2 The scheme for mutation of the G-C base pair into T-A base pair in the DNA molecule through oxidation of guanine.

DNA DAMAGE

More than several thousand DNA damage per human cell occur every day due to several factors, some of which are often mentioned in the popular press, such as the oxidants, ionizing radiation, ultraviolet radiation and toxic chemicals. In addition, there are less publicized sources as well, such as hydrolysis, metabolism and alkylation. Cells have mechanisms for repairing the DNA damage which can be divided into three general categories: base excision repair (repair of a single damaged nucleotide), nucleotide excision repair (repair damage of 2 - 30 bases) and mismatch repair (for repairing mispaired bases). The existing repair mechanisms cover most of the DNA damage in nature and operate by the rule that the undamaged strands of the double helical DNA structure can be a template to repair the damaged strand. However, accidental failures in the repair process result in irreversible mutations where the DNA structure is not damaged but the hereditary code is changed. The frequency of DNA mutations directly depends on the regularity of the mutagenetic damage. Only some types of damage are highly mutagenetic.

The most common type of damage is the single-strand break. This break arises in the cells spontaneously [2]. The repair mechanism for this damage is highly efficient. Sometimes this damage can turn into a DNA double-strand break which is then the most difficult damage to repair. However, the double-strand break is the most unusual DNA damage and is potentially lethal for the cells. The other common damage of the DNA molecule is the loss of one base due to hydrolysis or thermal disruption. This damage most often can be successfully repaired by the base excision repair mechanism. Therefore, the most common damage is mutagenetically poor. The most frequent mutagenetic damage is caused by deamination (loss of amino group) and oxidative damage both of which can result in irreversible mutation.

It is now widely accepted that among all types of DNA damage, oxidative damage is the most cancer causing transversion mutation. Moreover, oxidative damage was also found to increase the frequency of mutations induced by deamination damage [3]. The mechanism of oxidative damage occurs as follows: The metabolism, ultraviolet radiation, ionizing radiations, environmental pollutions and several intracellular sources can cause the formation of the reactive oxygen species (ROS) in our system which are highly reactive molecules that can damage all sorts of cellular components. The DNA molecule within the nuclei is surrounded by many types of proteins which can protect the DNA from the oxidative damage. The protection is based on the positive charge of these proteins at physiological pH that can successfully bind a ROS. However, oxidative stress in the nucleus may result in the oxidation of proteins, their modification into hydroperoxides, and cross link formation between the proteins and DNA. These processes may rapidly decrease the protection susceptibility of the proteins [4]. The degeneration of physiological functionality of our body with age and morphological age-changes are associated with the oxidative stress as well [5]. The age-related progressive accumulation of oxidative damage also causes corruption of cell functionality and reproduction.

DNA MUTATIONS

Why does oxidative damage lead to DNA mutations? The process of oxidative damage and consequent mutations are sketched in Figure 2. The contact of the ROS with the DNA molecule can occur directly with the resultant base oxidation or with participation of proteins and protein-hydroperoxides, when a radical cation is inserted into DNA. The induced radical reacts with water and also leads to base oxidation. Experimental investigations have shown that the *hot spots* for oxidative damage in the DNA molecule are the guanine bases, either single or stacked [4]. The guanine base has the lowest ionization potential among all bases in the DNA molecule and is the preferred spot for the oxidative damage. However, the oxidative damage of guanine can occur even if the ROS incorporation is far from the guanine location [5]. Therefore, it has been proposed [6,7] that a radical cation can migrate over the DNA chain until it reaches a guanine base which acts as a trap for the positive charge. The oxidative damage accompanied by a long-range charge transfer is called the *long-range oxidative damage*. The product of the oxidative damage - oxidized guanine G_{ox} - is a premutagenic lesion that can pair with adenine as well as cytosine during DNA replication. Replacement of the cytosine by the adenine, i.e., appearance of a G_{ox}-A pair, is known as mispaired damage. The mismatch repair mechanism corrects a mispaired A-C and T-G more efficiently than G-A and T-C [8]. Failure in the mismatch repair results in mutation of the genetic code (Figure 2). The mutation frequencies of

G_{ox}-A mispaired damage lies in the range 1-5% depending on the pair bonding and its stability [9]. This value is considerably large if we consider the number of DNA damage events in a cell per day (more than 100,000). An example of the lethality of DNA mutation and hence oxidative damage can be found in the famous cloning procedure. Here the source of genes for cloning is only from one parent. With age, the parent cells accumulate ROS and in particular, the DNA accumulates mutations caused by the ROS. Therefore, using genes of adult animals plus the additional damage of the embryonic cell that occurs during culturing in the laboratory, cause the irreversible gene mutations that may lead to early death of cloned animals [10]. In nature, the problem of DNA mutations is mitigated by participation of the gene set from both parents in embryonic cell division. In that case, identification of the defective genes leads to switching of the gene duplication process from one gene set to another.

Decreasing DNA oxidative damage and preventing mutation in the hereditary code has recently become an important challenge for researchers. The widely applied method of decreasing oxidative damage is by using antioxidants such as vitamin C, which can efficiently counteract DNA mutations [11]. Another method is based on blockading long-range oxidative damage. It is known that some people have a *genetic predisposition* to cancer because they are born with DNA mutations in particular genes. Examination of their gene sequence, detection of the hot spots for oxidative damage and blockade of the radical migration to these spots, could help prevent the crucial mutations. The blockade could be performed by the creation of a transfer channel around the hot spot absorbing the free radicals and leading to eventual neutralization. Clearly, charge migration in DNA is indeed a subject where biological interest overlaps with the physics aspect of the problem to be discussed below.

Since the 1990s when it was discovered that the mutational hot spots occur at the stacked guanine bases [12], investigations of charge migration in the DNA molecule and related topics became important for biophysicists, chemists and physicists. Of particular interest are the mechanisms of charge migration and the influence of the DNA sequences and the surrounding environment on the efficiency of the charge migration.

CHARGE MIGRATION THROUGH DNA

How does the DNA molecule conduct charge? The overlapping of the π-electrons of the stacked bases and formation of the π clouds above and below the bases result in a π pathway for charge migration along the DNA. For the hole migration, the energies of the highest occupied orbital of the bases form a potential profile on a π pathway, regulating the rate of charge migration. Guanine is known to be the base with the lowest ionization potential, and the stacking of several guanines provides even lower potential. Therefore, guanine (G), duplex guanine (GG) and triplex guanine (GGG) are *traps* for a hole migrating through the DNA. Adenine, cytosine and thymine, on the other hand, mostly play the role of potential barriers. DNA molecules can then be described as an assembly of potential barriers and wells.

The charge transfer in condensed matter is a pure physical problem. In the past century this problem was extensively studied, starting from the charge transfer description within classical mechanics to the quantum theory. From the quantum theory it is well known that there are two possible mechanisms for charge migration through the potential barrier - quantum mechanical tunneling and incoherent hopping [13]. Domination by one of these mechanisms over the other mostly depends on the height and width of the potential barrier as well as the temperature. In a DNA duplex, when the number of A:Ts is three or less, the hole propagation occurs by tunneling. For additional A:Ts, the commonly accepted idea is that, in the time it takes to tunnel beyond three As, a few holes acquire enough thermal energy to jump onto the bridge made out of As and then move by incoherent, nearest-neighbor hopping between the As, and eventually reach the hole trap. Such a behavior has indeed been observed experimentally in Refs. [14] (Figure 3). Here, for a thin barrier formed by stacked T-A pairs (≤ 2-3 base pairs), tunneling from the donor (Sa - stilbenedicarboxamide or a single guanine) to the acceptor (Sd - stilbenediether or the triple guanine), characterized by the exponential dependence on the barrier width, has been found to be the dominant mechanism (open symbols). For the thick barrier, the prevalence of the incoherent hopping was observed - the weak distance dependence introduced by the closed symbols in Figure 3. In solid state physics, this situation can be easily simulated by a tight-binding Hamiltonian or a system of kinetic equations that includes the temperature effects. These two models have been widely used for simulation of the charge transfer in the DNA molecule [13,15,16].

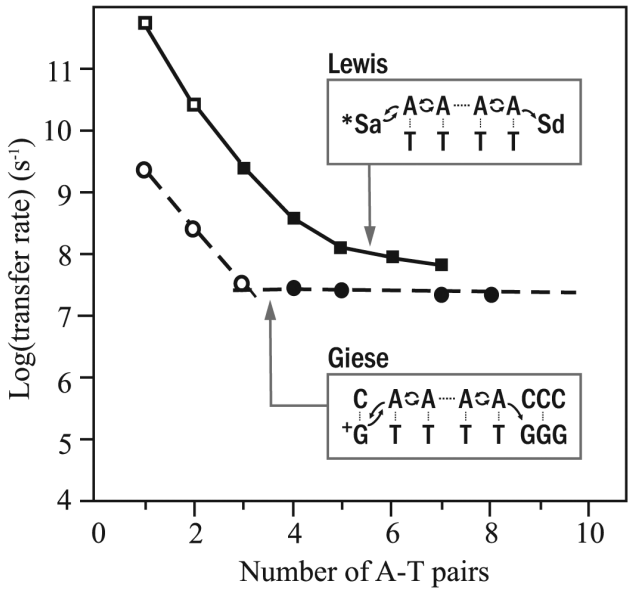


Fig. 3 The experimental data for dependence of the charge transfer rate through DNA molecule on the number of A-T pairs between the donor and the acceptor [14]. Open symbols are for tunneling and filled symbols are for incoherent hopping.

However, the charge transfer in a biological system is rather an interdisciplinary problem combining the physical as well as the biological and the chemical aspects and can not be adequately solved by pure physical approaches developed for the rigid lattice. The picture of charge migration in the molecular structure is not so simple because of high flexibility of the geometry and interactions with the solvent environment. Any change of the charge density at the DNA sites causes geometry distortion of the states and change of electrostatic interactions with the surrounding environment. Therefore, the electronic structure of the active sites and their energy are not constant during the charge transfer along the molecular chain.

During the years 1956 to 1965, a successful approach to describe the electron transfer between molecules in solution was developed by Rudolph A. Marcus, who incorporated within his theory the quantum mechanical and chemical aspects of the problem [17]. However, because of an important assumption made that the transfer reaction occurs between two active groups separated by the solvent molecules, its application to the investigation of charge transfer in complex molecules is not fully correct. The active sites participating in the charge transfer in large molecules most often are not separated by highly mobile water molecules, but by not so mobile molecular objects. In the case of DNA, the distance between the neighboring nucleobases is so small (3.4 Å) that presence of any other molecules is very unlikely. Taking into account all these issues, a more accurate *polaron model* for charge migration in the polymer and DNA molecules has been developed [18]. This theory is based on coupling of the charge density at the active states and the geometry distortion of these states during the polaron propagation. Moreover, this model has been found to be highly effective for investigation of charge transfer in polymer molecules [19]. However, application of the polaron model to DNA remains questionable. For example, important issues that can completely change the predicted behavior of the polaron, are the polaron size and its dependence on the DNA structure and the environment [20,21].

Because of intense experimental investigations of the DNA molecule the charge migration has been extensively studied for different environmental conditions and for a wide range of DNA sequences. For example, variation of the DNA conductance from an insulator to a metallic behavior and the contradictory transport characteristics with decreasing temperature have been reported [20,22]. There are also intriguing magnetic properties of DNA observed in experiments [23]. However, because of the complexity of the charge transfer phenomena in molecular structures, the widely used models, such as the tight-binding approach, the system of kinetic equations and the polaron model, are all inadequate to explain these observed phenomena. In fact, it is safe to conclude that for charge migration in DNA only one problem has been solved successfully so far, albeit only for the simplest types of DNA sequence, i.e., the description of the charge transfer through potential barriers by competition of the quantum tunneling and incoherent hopping (Fig. 3). An important problem now is to find an adequate theory that includes the temperature effect and also the contribution of the solvent environment.

The problem of charge migration in the DNA molecule is important not only for issues associated with the DNA mutations discussed above, but are also important for application of DNA in molecular electronics. The miniaturization of the components of the electronic circuits in semiconductor microelectronics has reached a saturation level and any further decrease of the device size is limited mostly by the lithographic technology. This problem has generated a growing interest in the scientific world in molecular electronics. The interest is motivated by several factors: charge migration, molecular recognition and self-assembly properties. Self-assembly of the molecular building blocks into well-structured systems allows us to solve a problem of the limitation of physical manipulation during fabrication of the nanosize devices. Moreover, any damage to the structure of the molecular and electrical elements can be repaired as well by the molecular recognition property without any physical contact. An example of the application of the self-assembly properties of the DNA molecule is presented in Figure 4 [24]. The DNA molecules are incorporated into the supramolecular arrays and programmable self-assembled two-dimensional nanogrids are constructed. This grid is then used as templates for neatly organizing gold nanoparticles. The structure thus obtained with clusters of gold atoms has many useful potential applications, from drug delivery or cancer-treatment directly at the damaged site in the human body to digital data storage. Because of these important possibilities, it is now widely believed in the scientific community that molecular electronics is perhaps the most promising technology in the near future.

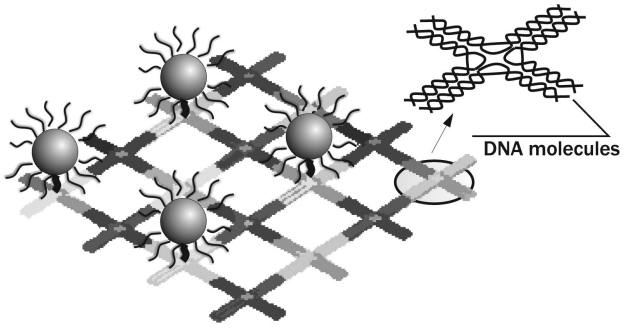


Fig. 4 The self-assembled DNA nanogrid with attached gold nanoparticles. The cross structure of the DNA template is composed of four DNA branch junctions.

In conclusion, charge transfer in the DNA molecule is the focal point for developing an anti-mutation medical treatment, for example for people with the *genetic predisposition*, and for application of the DNA molecule in molecular electronics. Our

present understanding of charge migration in the DNA molecule is far from satisfactory for the above-mentioned purposes and therefore necessarily requires further investigations in this fascinating field [15].

ACKNOWLEDGEMENTS

This work has been supported by the Canada Research Chair program and a Canadian Foundation for Innovation (CFI) fund.

REFERENCES

1. The structure of DNA was deciphered in 1953 by James Watson and Francis Crick (Nobel prize in 1962) in the physics department (Cavendish Laboratory) of the University of Cambridge. Numerous accounts of their work leading to that discovery have been published in the literature. See for example, J.D. Watson, *The Double Helix: A Personal Account of the Discovery of the Structure of DNA* (Penguine, 1968); H.F. Judson, *The Eighth Day of Creation* (Simon and Schuster, New York, 1979).
2. S.F. El-Khamisya, and K.W. Caldecott, *Neuroscience*, **145**, 1260 (2007).
3. D.-H. Lee, T. R. O'Connor, G.P. Pfeifer, *Nucleic Acids Research* , **30**, 3566 (2002).
4. A. Furukawa, Y. Hiraku, S. Oikawa, C. Luxford, M.J. Davies, S. Kawanishi, *Biochem J.*, **388**, 813 (2005).
5. E.R. Stadtman, *Science*, **257**, 1220 (1992).
6. M. Nunez, D.B. Hall, J.K. Barton, *Chem. Biol.*, **6**, 85 (1999).
7. B. Giese, *Acc. Chem. Res*., **33**, 631 (2000); G.B. Shuster, *Acc. Chem. Res.*, **33**, 253 (2000).
8. G. Marra, P. Schär, *Biochem. J*., **338**, 1 (1999).
9. F. Seela, S. Budow, K.I. Shaikh, A.M. Jawalekar, *Org. Biomol. Chem.*, **3**, 4221 (2005).
10. D. Humpherys, K. Eggan, H. Akutsu, K. Hochedlinger, W.M. Rideout III, D. Biniszkiewicz, R. Yanagimachi, and R. Jaenisch, *Science*, **293**, 95 (2001).
11. A. Besaratinia, S. Kim, S.E. Bates, G.P. Preifer, *Proceedings of the National Academy of Sciences*, **104**, 5953 (2007).
12. D.G. Lowy and J.M.C. Gutteridge, *Ann. Rev. Biochem*., **62**, 851 (1993). S.P. Hussain *et al*., *Oncogene*, **9**, 2277 (1994).
13. J. Jortner *et al*., *Proceedings of the National Academy of Sciences*, **95**, 12759 (1998).
14. F.D. Lewis *et al*., *Angew. Chem. Int. Ed*., **45**, 7982 (2006); B. Giese, J. Amaudrut, A.-K. Khler, M. Spormann, S. Wessely, *Nature*, **412**, 318 (2001).
15. T. Chakraborty (Ed.), *Charge Migration in DNA: Perspectives from Physics, Chemistry, and Biology* (Springer, 2007).
16. Li XQ, Y. Yan Y, *J. Chem. Phys*., **115**, 4169 (2001).
17. For a review, see R. Marcus, *Rev. Mod. Phys*., **65**, 599 (1993).
18. T. Dauxois, M. Peyrard, A. R. Bishop, *Phys. Rev. E*, **47**, 684 (1993).
19. S.V. Rakhmanova and E.M. Conwell, *Appl. Phys. Lett.*, **75**, 1518 (1999).
20. K.-H. Yoo *et al., Phys. Rev. Lett*., **87**, 198102 (2001).
21. G.B. Schuster (Ed.), *Long range charge transfer in DNA* (Springer, 2004).
22. D. Porath, A. Bezryadin, S. de Vries, C. Dekker, *Nature*, **403**, 635 (2000).
23. S. Nakamae, M. Cazayous, A. Sacuto, P. Monod, and H. Bouchiat, *Phys. Rev. Lett*., **94**, 248102 (2005).
24. Hao Yan, Sung Ha Park, Gleb Finkelstein, John H. Reif and Thomas H. LaBean, *Science*, **301**, 1882 (2003).