UDC 575.174:599

Analysis of Mutation Mechanisms in Human Mitochondrial DNA

I. V. Kornienko[1] and B. A. Malyarchuk[2]

[1]*Central Laboratory for Medical Forensic Identification 124,*
Department of Defense of the Russian Federation, Rostov-on-Don, 344010 Russia
[2]*Institute of the Biological Problems of the North, Far East Division,*
Russian Academy of Sciences, Magadan, 685000 Russia
Received February 4, 2005

Abstract—The cause of the high variability of human mitochondrial DNA (mtDNA) remains largely unknown. Three mechanisms of mutagenesis that might account for the generation of nucleotide substitutions in mtDNA have been analyzed: deamination of DNA nitrous bases caused by deamination agents, tautomeric proton migration in nitrous bases, and the hydrolysis of the glycoside bond between the nitrous base and carbohydrate residue in nucleotides against the background of the free-radical damage of DNA polymerase γ. Quantum chemical calculations demonstrated that the hydrolysis of the N-glycoside bond is the most probable mechanism; it is especially prominent in the H strand, which remains free during mtDNA replication for a relatively long time. It has also been found that hydrolytic deamination of adenine in single-stranded regions of the H strand is a possible cause of the high frequency of T $\longrightarrow$ C transitions in the mutation spectra of the L-chain of the major mtDNA noncoding region.

Key words: human mitochondrial DNA, major noncoding region, hypervariable segments 1 and 2, nucleotide substitutions, hot spots, mechanisms of mutagenesis

INTRODUCTION

The human mitochondrial genome is characterized by considerable variation, its mechanisms remaining largely unknown [1]. The nucleotide sequence variation of the major mitochondrial DNA (mtDNA) noncoding region (or control region), which contains hypervariable segments (HVSs) termed HVS1 and HVS2, has been studied in most detail. The mutation spectra of HVS1 and HVS2 are characterized by a considerably larger number of transitions compared to transversions, as well as a larger number of pyrimidine transitions, than those of purine [2, 3]. Analysis of the HVS1 and HVS2 mutation spectra reconstructed with the use of phylogenetic analysis of population data has demonstrated that the mutation process in human mtDNA depends on the nucleotide context, and the high frequency of mutations (23.4 and 49.4% in HVS1 and HVS2, respectively) is explained in terms of the model of mtDNA chain dislocation during replication [3]. Analysis of mtDNA mutation spectra has shown that oxidative damage (e.g., the formation of 8-OH-dG) leading mainly to transversions is not the main cause of mitochondrial mutations. However, the mutagenesis mechanisms that result in AP-sites, such as spontaneous loss of bases, base deamination, and some types of nucleotide oxidative damage, may be among the factors determining the high proportion of transitions in mtDNA mutation spectra. To date, promising approaches to revealing the mechanisms of the high mtDNA variation rate has been proposed [4]; however, further studies in this important field are necessary.

We analyzed three molecular mechanisms of mutagenesis that might be involved in the formation of mtDNA mutations: deamination of DNA nitrous bases caused by deamination agents, tautomeric proton migration in nitrous bases, and the hydrolysis of the glycoside bond between the nitrous base and carbohydrate residue in nucleotides against the background of the free-radical damage of DNA polymerase γ.

EXPERIMENTAL

The rate of oxidative deamination of nitrous bases was determined as described previously [5]. The degree of completeness of the reaction

$$RNH_2 + HNO_2 \longrightarrow ROH + N_2 + H_2O$$
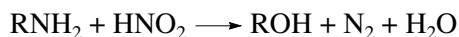
was estimated by $\alpha(RNH_2)$, the molar proportion of the compound RNH_2: the more RNH_2 has been converted into ROH, the lower the $\alpha(R\text{-}NH_2)$.

All nonempirical calculations included complete optimization of geometry by the Hartree–Fock method in the RHF/6-31G** basis [6]. Semiempirical calculations were performed by the PM3 method

Table 1. Distribution of point substitutions in the L strand of the human mtDNA control region

mtDNA region, nucleotide substitutions	HVS1		HVS2	
	number of mutations	percentage of mutations	number of mutations	percentage of mutations
Transitions				
C ⟶ T	944	42.7	110	22.0
T ⟶ C	638	28.8	213	42.6
G ⟶ A	188	8.5	71	14.2
A ⟶ G	294	13.3	94	18.8
Transversions				
A ⟶ C	32	1.4	5	1.0
C ⟶ A	45	2.0	0	0
T ⟶ G	5	0.2	1	0.2
G ⟶ T	0	0	1	0.2
A ⟶ T	25	1.1	0	0
T ⟶ A	3	0.1	4	0.8
G ⟶ C	7	0.3	1	0.2
C ⟶ G	31	1.4	0	0

Note: The distribution of mutations is estimated from previously published data [3].

based on neglecting two-atom differential overlapping [7, 8]. The calculations were performed by means of a Pentium III-450 PC with the use of the GAMESS software [9].

Complete optimization of the geometry of molecular structures corresponding to stationary points (minimums, $\lambda = 0$, where λ is the number of negative eigenvalues of Hesse's matrix at the given stationary point) on the surface of potential energy was performed to a gradient of 10^{-6} Hartry–Bohr. The calculation in the 6-31G** basis of the matrix of power constants was performed analytically using the GAMESS three-point method [9].

RESULTS AND DISCUSSION

Polymorphism of the mtDNA control region was mainly accounted for by point nucleotide substitutions (Table 1), transition sites being considerably more numerous than transversion sites. This may have resulted from, first, selective effects of mutagens on DNA bases; second, spontaneous ketone–enol and amine–imine tautomerism of nitrous bases; and third, replication errors. Asynchronous replication characteristic of the mitochondrial genomes of vertebrates leads to the asymmetry of nucleotide frequencies observed in mtDNA chains [10]. The rate of accumulation of point substitutions in the mtDNA control region is an order of magnitude higher than in the

nuclear genome, because the mitochondrial genome contains an "open" single-stranded region (D loop) [11]. Earlier studies using *Naemophilus influenzae* transforming DNA demonstrated that denaturation of the polynucleotide chain increased the mutagenic effect of nitrous acid [12, 13]. Indeed, the H-bond rupture during denaturing makes nitrous bases more accessible to mutagenic factors.

Modifiers of nitrous bases not causing deletions or base crosslinking are of special interest. These are, e.g., deamination agents, which transform purines and pyrimidines with an exocyclic amino group according to the formula

$$R-NH_2 \longrightarrow R-OH.$$
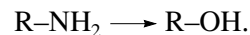
This may result in point substitutions via the formation of unusual H-bonds by modified bases during DNA replication. One of the mechanisms of point substitutions in DNA is cytosine and adenine deamination caused by NO-containing agents. Nitrous acid (HNO_2) is a classic example of a mutagen of this type. It is formed from organic precursors, e.g., nitrosoamines, as well as from nitrites and nitrates. Nitrous acid effectively removes amino groups from cytosine, adenine, and guanine.

Nitrous acid transforms guanine into xanthine, which still forms H-bonds (two instead of three) with cytosine. Therefore, guanine deamination does not result in point mutations but may cause inactivation because of weakening the bonds between the nitrous base and deoxyribose. Uracil, the product of cytosine deamination, can form two H-bonds with adenine. Therefore, deamination causes the substitution G–C ⟶ A–T. Finally, adenine deamination leads to the formation of hypoxanthine, which, after tautomeric transformation, forms two H-bonds with cytosine, and the A–T pair is substituted by the G–C pair.

There are abundant but contradictory data on deamination gene mutations, mostly on the comparative HNO_2-induced mutability of adenine and cytosine. This mutability is mainly determined by the adenine and cytosine deamination rates. For example, some researchers believe that the deamination rate is higher for cytosine than for adenine [14]; therefore, the probability of substitutions of G–C by A–T is higher than that of the reverse substitution. Other researchers report the rate constants, indicating that the rate increases in the order cytidine < adenosine < guanosine [15]. Our data indicate that the adenine deamination rate is higher than the cytosine deamination rate [5]. Table 2 shows the results of the interaction of adenine and cytosine with nitrous acid. According to these data, adenine is deaminated more intensely than cytosine.

Our experimental data were confirmed by the results of *ab initio* quantum chemical calculations in the RHF/6-31G** basis. Indeed, electrophilic substi-

tution reactions preferably occur at the site of the highest electron density and the largest negative charge. The small difference in energy between the border orbitals of the electrophile NO^+ (–0.270 eV) and the substrates (adenine and cytosine) indicates that orbital control of the process is more preferable. The highest occupied molecular orbital of adenine (–0.307 eV) is lower than the highest occupied orbital of cytosine (–0.337 eV), which affiliates electron transfer from the highest occupied orbital of the substrate to the lowest free orbital of the reagent (in the case of adenine). Therefore, adenine must be more prone to interact with nitrous acid.

Hydrolytic base deamination in single-stranded DNA regions is more intense than in double-stranded ones [16]. Therefore, adenine deamination in the mtDNA H strand is expected to lead to substitution of thymine by cytosine in the L strand [10]. Note that experimental data only partly agree with these predictions: indeed, many L spectra of the mtDNA HVS1 and HVS2 contain only the T $\longrightarrow$ C substitutions (Table 1). Thus, regarding single-stranded regions on the H strand, the mechanism of HNO_2-induced mutagenesis (as well as the fact that adenine is deaminated more than cytosine under these conditions) may account for the preference of T $\longrightarrow$ C point substitutions compared to G $\longrightarrow$ A transitions in the L strand of the mtDNA control region. Note that the results of our study agree with the data on the frequency distribution of nucleotide substitutions as dependent on the period of time when segments of the mtDNA H strand remain single-stranded (D_{ssH}) [17]. Calculations performed in [17] demonstrate that the frequency of only A $\longrightarrow$ G transitions linearly increase with increasing D_{ssH}, whereas the frequency of C $\longrightarrow$ T transitions rapidly reaches a plateau. As can be seen in Table 1, the proportion of T $\longrightarrow$ C transitions in the mutation L spectrum of HVS2, which remains single-stranded during replication for a longer time than does HVS1, was substantially higher than in the HVS1 mutation spectrum. However, the formation of point mutations via nitrous base deamination cannot account for transversions or the larger proportion of pyrimidine than purine transitions.

Let us consider another mechanism of point substitutions, which is based on the spontaneous tautomerism of nitrous bases. The transitions of protons between nitrogen and oxygen atoms in nucleotides determine the amine–imine and ketone–enol tautomerisms (Fig. 1). Available crystallographic and spectroscopic data conclusively demonstrate that ketone–amine tautomers are more prevalent in water solutions [18, 19].

The change in the position of the hydrogen atom alters the pattern of H-bonds formed by it: the ketone group with proton-acceptor properties in the enol form becomes a proton donor, and the amino group

Table 2. Molar proportion of RNH_2 in the system after the reaction $RNH_2 + HNO_2 \longrightarrow ROH + N_2 + H_2O$ in 0.1 N HCl

HNO_2 concentration, M	Adenine, $5 \cdot 10^{-4}$ M		Cytosine, $5 \cdot 10^{-4}$ M	
	A_{270}	$\alpha(R\text{-}NH_2)$	A_{285}	$\alpha(R\text{-}NH_2)$
$5 \cdot 10^{-5}$	0.66	0.82	0.53	0.91
$5 \cdot 10^{-4}$	0.64	0.78	0.52	0.90
$5 \cdot 10^{-3}$	0.62	0.70	0.51	0.87
$5 \cdot 10^{-2}$	0.60	0.64	0.50	0.83

Note: A_{270} and A_{285} are the optical densities at 270 and 285 nm, respectively.

Table 3. Energy characteristics of the classic and tautomeric forms of nitrous bases calculated for the gas phase by the semiempirical PM3 method [7, 8]

Nitrous base	Heat of the formation, kcal/mol		Energy of the transition state, kcal/mol	Potential barrier, kcal/mol
	classic form	tautomeric form		
Adenine	+57.3	+66.4	+104.3	+47.0
Cytosin	–12.3	–10.7	+30.1	+42.4
Guanine	+10.1	+9.9	+49.5	+39.4
Thymine	–75.8	–66.4	–29.5	+46.3

giving up a proton when transiting into the substituted imine form becomes a proton acceptor. Because of tautomeric transformations, the enol forms of uracil and guanine can mimic cytosine and adenine when pairing, and the imine forms of cytosine and adenine can mimic uracil and guanine (Fig. 1). In Watson and Crick's self-replicating system based on specific pairing, these transformations may lead to nucleotide substitutions [20]. Since the DNA double helix consists of base pairs, the tautomeric transition of one of the components of this pair induces the tautomeric transition of the other component. Thus, base pairs rather than individual nitrous bases undergo tautomeric transformations in a DNA double helix. There are several hypotheses concerning the mechanisms of tautomeric transformations leading to nucleotide substitutions. In our opinion, P.O. Levdin's hypothesis is the most plausible (cited from [21]). According to it, it is not necessary that a proton should change place within a base (according to Watson and Crick's model); instead, proton exchange within a base pair may occur through movement along an H-bond (Fig. 2).

The results of our quantum chemical calculations demonstrated that, if a proton is transferred from one atom to another within a base molecule, it must overcome a very high potential barrier (Table 3). As we will show below, in contrast to Watson and Crick's model, the potential barrier is substantially lower if
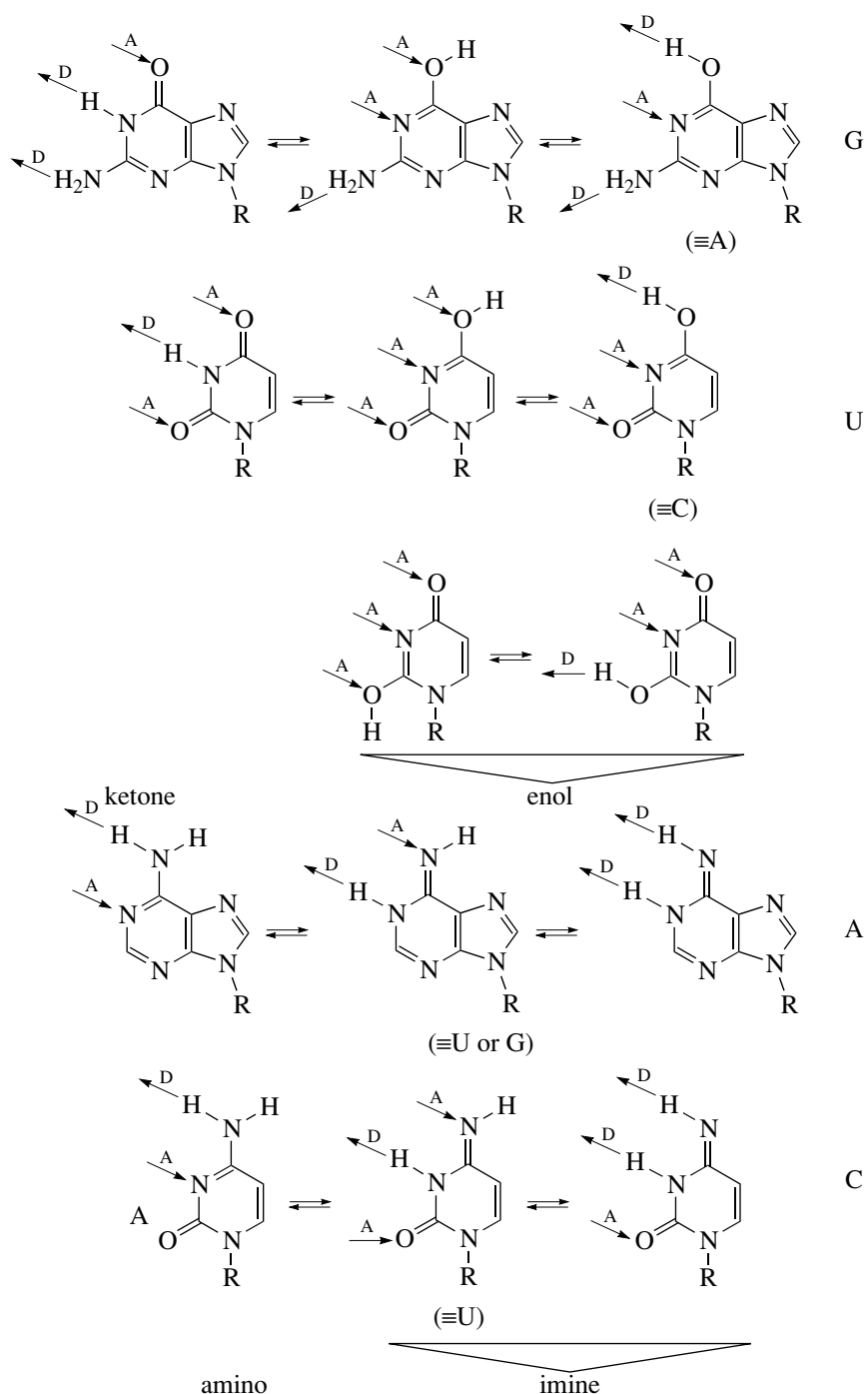
Fig. 1. Ketone–enol and amine–imine tautomerisms of bases in nucleotides. Arrows with symbols A and D indicate the acceptor and donor groups capable of forming H-bonds. G and U in the enol form become similar to A and C, respectively; A and C in the imine form are equivalent to U or G and to U, respectively.

proton exchange occurs between two nitrous bases in a pair according to Levdin (Fig. 2). Thus, Levdin's mechanism seems more plausible.

It has not been demonstrated conclusively that point substitutions occur by Levdin's mechanism. In addition, it is unknown whether the frequency of point substitutions depends on the type of the base pair. To theoretically estimate the mutation frequency, we calculated the height of the potential barrier for the proton exchange between two nitrogen atoms and between a nitrogen and an oxygen atoms in pairs A–T and G–C. For this purpose, we first determined the mechanism of proton exchange in base pairs A–T and
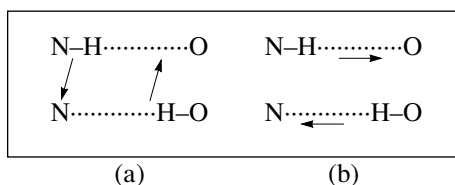
Fig. 2. Tautomeric transition of the A–T base pair [21]: (a) according to Watson and Crick's model; (b) according to Levdin's model.

G–C and compared the potential barriers for these pairs. We calculated the heat of the formation of the A–T and G–C pairs and then compared the potential barriers for different cases of proton transition between two bases in both pairs (Fig. 3, Table 4).

Three mechanisms of proton exchange between two bases in the A–T and G–C pairs are possible (Fig. 3).

(1) First, a proton moves from a nitrogen atom (of adenine and cytosine) to an oxygen atom (of thymine and guanine). Then, another proton moves from a nitrogen atom (of thymine and guanine) to a nitrogen atom (of adenine and cytosine).

(2) First, one proton moves from a nitrogen atom (of thymine and guanine) to a nitrogen atom (of adenine and cytosine). Then, another proton moves from a nitrogen atom (of adenine and cytosine) to an oxygen atom (of thymine and guanine).

(3) Both protons move simultaneously.

The comparison of potential barriers for these mechanisms of proton transitions within each pair showed that the second mechanism was the most probable (Table 4); i.e., it is most likely that a proton first moves from a nitrogen atom (of thymine and guanine) to a nitrogen atom (of adenine and cytosine). After that, another proton moves from a nitrogen atom (of adenine and cytosine) to an oxygen atom (of thymine and guanine) (Fig. 3). As follows from Table 4, the energy barriers of proton transition in pairs A–T and G–C are practically the same. Thus, this mechanism does not allow us to explain the higher frequency of point substitutions between pyrimidine bases than between purine bases in the mtDNA HVS1 and HVS2. In addition, this mechanism of the formation of point mutations cannot explain the appearance of transversions.

The mechanism of point mutations in single-stranded nucleotide fragments of the D loop is based on the hydrolysis of the glycoside bond between nitrous bases and the carbohydrate residue in nucleotides against the background of the free-radical damage of DNA polymerase γ. As we will show below, this mechanism can account for the formation of transversions and predominant accumulation of transitions.

Under the conditions of local accumulation of protons in the mitochondrial matrix, nucleotides are hydrolyzed according to the following scheme:
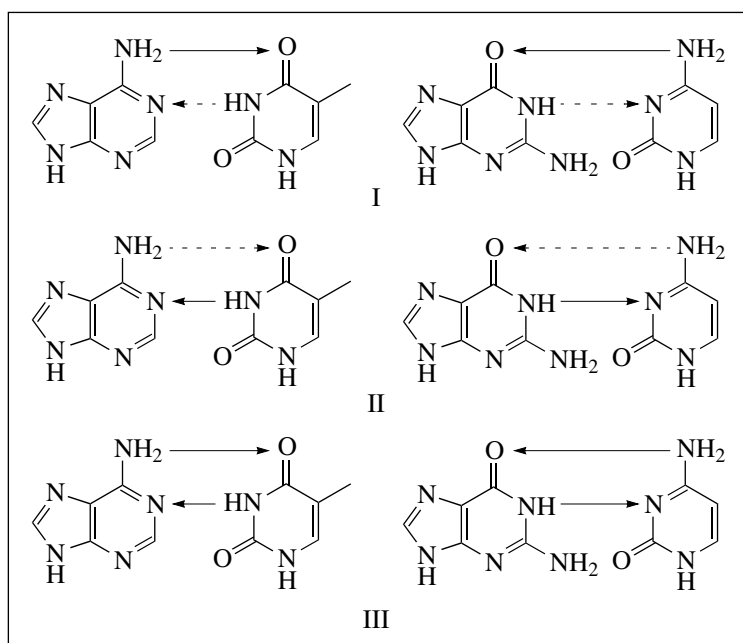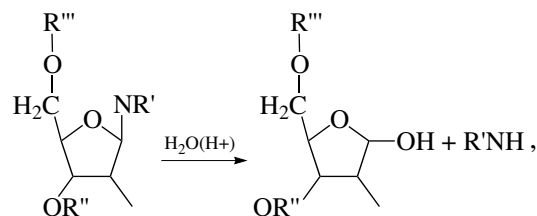

Fig. 3. Three possible mechanisms of proton exchange between two bases in pairs A–T and G–C.

Table 4. Potential barriers of proton transitions between two bases in pairs A–T and G–C calculated for the gas phase by the PM3 method [7, 8]

Base pair	Mechanism of proton transition	Energy of the potential barrier to overcome, kcal/mol
A-T	I	+33.4
A-T	II	+17.4
A-T	III	+27.9
G-C	I	+25.1
G-C	II	+19.3
G-C	III	+24.9

Table 5. Summary positive charges of the nitrogen atoms involved in the glycoside bonds of nucleosides calculated for the gas phase by the semiempirical PM3 method [7, 8]
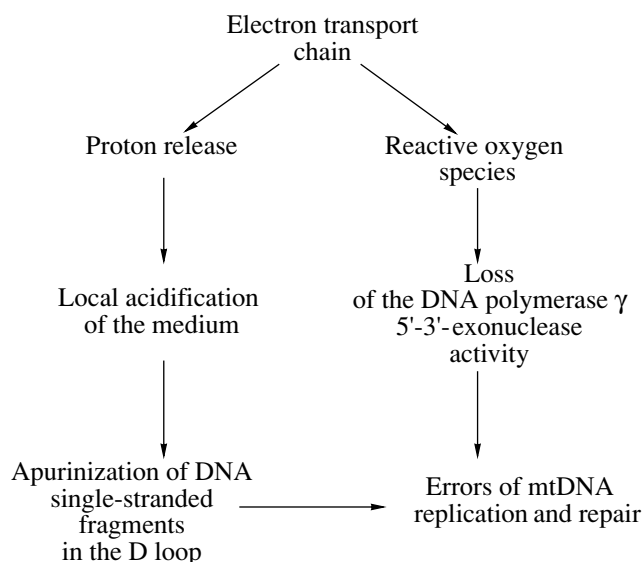
Nucleoside	Charge of the nitrogen atom involved in the glycoside bond
Deoxyriboadenosine	+0.196
Deoxyriboguanosine	+0.242
Deoxyribothymidine	+0.043
Deoxyribocytidine	+0.020

where R' is a nitrous base and R" and R'" are hydrogen atoms or unsubstituted and substituted phosphate groups.

Regarding the reaction mechanism, it should be remembered that the rate of hydrolysis of these compounds is determined by the positive charge of the nitrogen atom involved in the nucleophile attack in the glycoside bond: the higher the charge, the easier the hydrolysis.

The results of quantum chemical calculations demonstrated that the hydrolysis of the N-glycoside bond in purine bases was easier than in pyrimidine ones (Table 5), which agrees with experimental data on the hydrolysis of nucleosides and their monophosphates [13, 15]. As can be seen from Table 5, purines rather than pyrimidines are expected to preferably undergo acid hydrolysis. The hydrolysis forms a gap that will be filled by the corresponding nitrous base according to the Watson–Crick complementary interaction. The specificity of the formation of H-bonds between bases

is largely determined by the 5'–3' exonuclease activity of DNA polymerase.

The utilization of molecular oxygen leads to the accumulation of reactive oxygen species in the mitochondria; in addition to single- and double-strand breaks of DNA resulting from the cleavage of deoxyribose, they may disturb chemical bonds within nucleotides and lead to the formation of apurine and apyrimidine sites, DNA–DNA and DNA–protein crosslinks, as well as DNA adducts [4]. Oxidation may also lead to complete or partial loss of the DNA polymerase exonuclease activity, which has been demonstrated in the cases of sequenase (a modified bacteriophage T7 DNA polymerase) [22] and human mitochondrial DNA polymerase γ [23].

Thus, point mutations in the mtDNA major noncoding region may occur as follows (Fig. 4). Single-stranded DNA regions represented by the H strand appear during mtDNA replication [24]. In the absence of H-bonds between nucleotides that are characteristic of the DNA double helix, open purines and pyrimidines in these single-stranded regions become more sensitive to acidification of the surrounding medium [16]. DNA depurinization (loss of A and G considered above) results in a gap that may be filled with other nitrous bases during replication.

E.B. Freese was the first to suggest a mechanism of the mutagenic effect of acidification [13, 25]. It followed from the model of equiprobable filling of gaps that two-thirds of all mutations induced by the decrease in pH had to be transversions. However, the analysis of mtDNA nucleotide sequences (Table 1) demonstrated that most point substitutions were transitions rather than transversions. Let all gaps be filled with any nucleotide with the same probability. In this case, the DNA secondary structure may be distorted. For example, if a purine substitutes a pyrimidine, the geometry of the double helix will be inevitably disturbed. However, substitution of a purine by a pyrimidine does not disturb the spatial structure of DNA (Fig. 5). Thus, substitutions of three types are the most probable: pyrimidine $\longrightarrow$ pyrimidine, purine $\longrightarrow$

Electron transport chain

Proton release

Reactive oxygen species

Local acidification of the medium

Loss of the DNA polymerase γ 5'-3'-exonuclease activity

Apurinization of DNA single-stranded fragments in the D loop $\longrightarrow$ Errors of mtDNA replication and repair

Fig. 4. Mechanism of point mutation in the major mtDNA noncoding region.
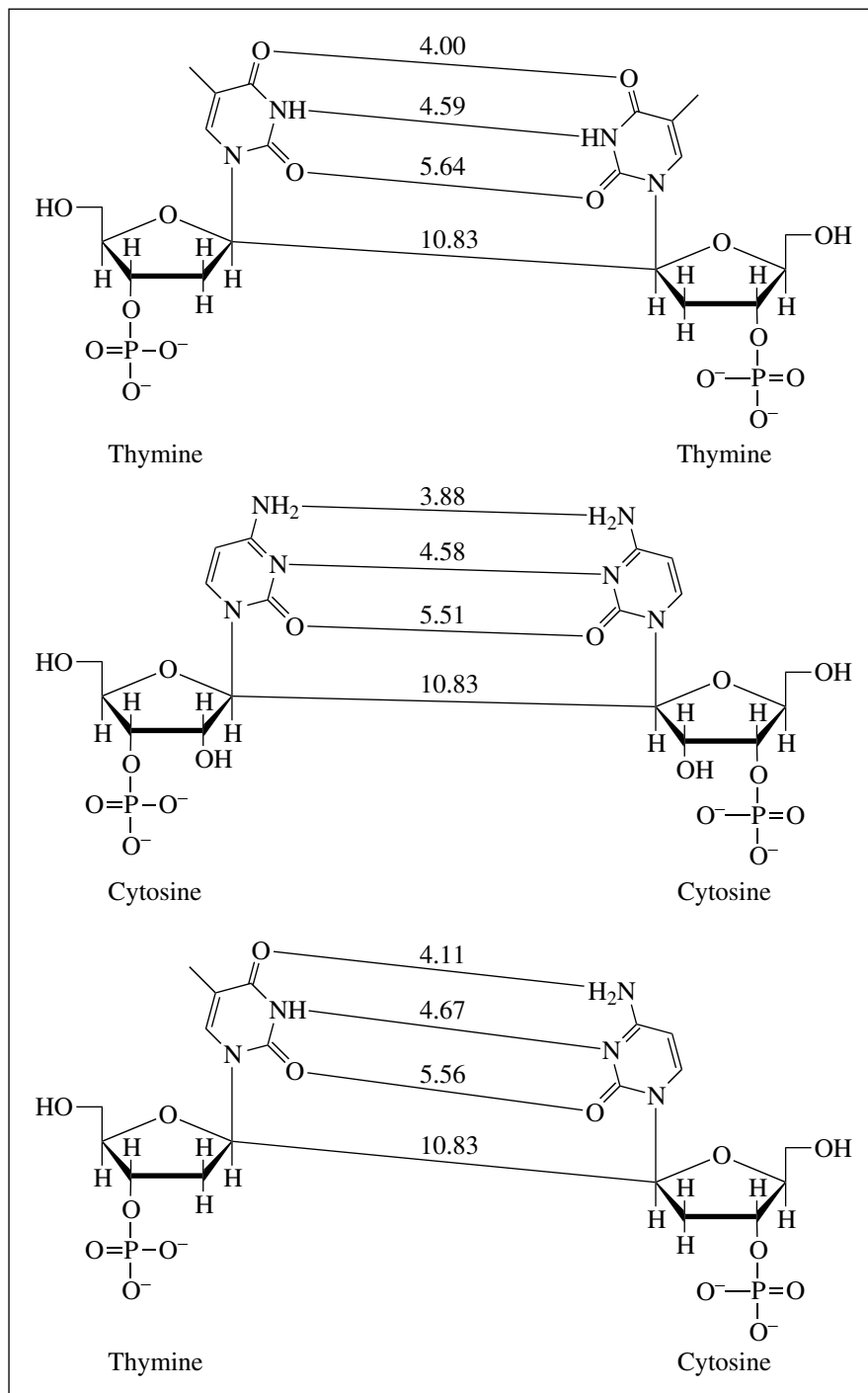
Fig. 5. Lengths of the bonds between pyrimidines (Å) in the classic B structure of the DNA double helix calculated by the PM3 method [7, 8] at fixed distances between deoxyribose C1' atoms.

purine, and purine ⟶ pyrimidine. Only these substitutions will not distort the DNA secondary structure. This mechanism may account for the preference of the transitions T ⟶ C and C ⟶ T and the formation of transversions. Indeed, the preferable hydrolysis of the N-glycoside bond in deoxyadenosine and deoxyguanosine in the H strand of the mtDNA noncoding region leads predominantly to T ⟶ C and C ⟶ T substitutions in the complementary L strand, which agrees with the pattern observed when analyzing the nucleotide sequences of the HVSs of the mtDNA control region.

REFERENCES

1. Wallace D.C. 1999. Mitochondrial diseases in man and mouse. *Science*. **283**, 1482–1488.

2. Budowle B., Wilson M.R., DiZinno J.A., Stauffer C., *et al*. 1999. Mitochondrial DNA regions HV I and HV II population data. *Forensic Sci. Int*. **103**, 23–35.

3. Malyarchuk B.A., Rogozin I.B. 2004. Mutagenesis by transient misalignment in human mitochondrial DNA control region. *Ann. Hum. Genet*. **68**, 324–339.

4. Marcelino L.A., Thilly W.G. 1999. Mitochondrial mutagenesis in human cells and tissues. *Mutat. Res*. **434**, 177–203.

5. Gaeva E.B., Narezhnaya E.V., Kornienko I.V. 2004. Determining mutagenic activity using an example of interaction between DNA bases and nitrous acid. *Klin. Lab. Diagn*. **1**, 16–18.

6. Clark T. 1990. *Computer Chemistry* (Russian translation). Moscow: Mir.

7. Stewart J.J.P. 1989. Optimization of parameters for semiempirical methods: 1. Method. *J. Comput. Chem*. **10**, 209.

8. Stewart J.J.P. 1989. Optimization of parameters for semiempirical methods: 2. Applications. *J. Comput. Chem*. **10**, 221.

9. Schmidt M., Baldridge K.K., Boatz J.A., *et al*. 1993. *J. Comput. Chem*. **14**, 1347 (package of *ab initio* programs GAMESS, Version 5.0, 1998).

10. Tanaka M., Ozawa T. 1994. Strand asymmetry in human mitochondrial DNA mutations. *Genomics*. **22**, 327–335.

11. Brown W.M., George M., Wilson A.C. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA*. **76**, 1967–1971.

12. Strack H.B., Freese E.B., Freese E. 1964. Comparison of mutation and inactivation rates induced in bacteriophage and transforming DNA by various mutagens. *Mutat. Res*. **1**, 10–21.

13. Soyfer V.N. 1969. *Molekulyarnye mekhanizmy mutageneza* (Molecular Mechanisms of Mutagenesis). Moscow: Nauka.

14. Ashmarin I.P. 1974. *Molekulyarnaya biologiya* (Molecular Biology). Leningrad: Meditsina, pp. 186–191.

15. Kochetkov N.K., Budovskii E.I., Sverdlov E.D., *et al*. 1970. *Organicheskaya khimiya nukleinovykh kislot* (Organic Chemistry of Nucleic Acids). Moscow: Khimiya.

16. Lindahl T. 1993. Instability and decay of the primary structure of DNA. *Nature*. **362**, 709–715.

17. Krishnan N.M., Raina S.Z., Pollock D.D. 2004. Analysis of among-site variation in substitution patterns. *Biol. Proceed. Online*. **6**, 180–188.

18. Roberts J., Casserio M. 1978. *Fundamentals of Organic Chemistry* [Russian translation]. Moscow: Mir, vol. 2.

19. Cantor C.R., Schimmel P.R. 1980. *Biophysical Chemistry*. San Francisco: W. H. Freeman.

20. Watson J.D., Baker T.A., Bell S.P., Gann A., Levine M., Losick R. 2004. *Molecular Biology of the Gene*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Lab. Press.

21. Ladik J. 1975. *Quantum Biochemistry for Chemists and Biologists* [Russian translation]. Moscow: Mir.

22. Brown T.A. 1994. *DNA Sequencing*. Oxford: IRL Press.

23. Graziewicz M.A., Day B.J., Copeland W.C. 2002. The mitochondrial DNA polymerase as a target of oxidative damage. *Nucleic Acids Res*. **30**, 2817–2824.

24. Clayton D.A. 1991. Replication and transcription of vertebrate mitochondrial DNA. *Annu. Rev. Cell Biol*. **7**, 453–478.

25. Freese E.B. 1961. Transitions and transversions induced by depurination agents. *Proc. Natl. Acad. Sci. USA*. **47**, 540–545.