ISSN 1607-6729, Doklady Biochemistry and Biophysics, 2012, Vol. 447, pp. 267–272. © Pleiades Publishing, Ltd., 2012.
Original Russian Text © N.N. Vtyurina, S.L. Grohovsky, A.B. Vasiliev, I.I. Titov, P.M. Ponomarenko, M.P. Ponomarenko, S.E. Peltek, Yu.D. Nechipurenko, N.A. Kolchanov, 2012, published in Doklady Akademii Nauk, 2012, Vol. 447, No. 2, pp. 217–222.

BIOCHEMISTRY, BIOPHYSICS AND MOLECULAR BIOLOGY

Contextual DNA Features Significant for the DNA Damage by the 193-nm Ultraviolet Laser Beam

N. N. Vtyurina[a, b], S. L. Grohovsky[a], A. B. Vasiliev[b], I. I. Titov[c], P. M. Ponomarenko[c], M. P. Ponomarenko[c], S. E. Peltek[c], Yu. D. Nechipurenko[a, b], and Academician N. A. Kolchanov[c, d, e]

Received May 20, 2012

DOI: 10.1134/S1607672912060038

Based on original measurements of the frequencies of guanine (G) damage in DNA by ultraviolet (UV) laser light with a wavelength of $\nu^{-1} = 193$ nm, we found that these frequencies depended on the local nucleotide environment of G. This is the first study to demonstrate the reliability of forecasting the frequency of UV damage of G by the sequence of a known DNA on the basis of this sequence using an independent control.

UV lasers are used in cancer therapy [1], cosmetics [2], bioluminescence [3], microsurgery [4], footprinting, and many other fields of biology. In numerous experiments it was established that the laser-induced UV radiation can damage DNA molecules [5–7]. The most common result of UV-induced DNA damage at wavelengths $\nu^{-1} \le 290$ nm is the formation of 7,8-dihydroxy-8-oxoguanine, 2,2-diamino-oxazolone, and other oxidation products of the guanine cation (G^+), resulting in DNA strand breaks [5–7] or nucleotide substitutions [8]. Such damage is fundamentally different from the DNA damage caused by other agents (e.g., ultrasound [9]). These most frequent UV-induced guanine damages are generally associated with the sufficiency of $h\nu$ energy of absorbed photon (h is the Planck's constant) for the electron (e^-) release from DNA to form a "hole" (DNA^+) [5–7], whose migration along the DNA strand often ends by the appearance of G^+ due to the lower ionization threshold of G compared to A, T, and C [7] and an immediate attack of this G^+ by free radical anions. It was shown that the frequency of UV-induced guanine damages in DNA sequences can vary widely, indicating its dependence on the DNA nucleotide context around guanine residues [5–7]. However, this dependence still remains obscure. The question about a systematic approach to in silico analysis of the characteristics of the DNA nucleotide context that affect the frequency of guanine UV damage has not yet been raised.

This is the first study to perform a systematic in silico analysis of the contextual DNA features under UV laser irradiation. The features of local environment of G that significantly affect the frequency of UV-induced damage of G include (1) consensus, (2) the position-weight matrix, (3) tetranucleotides YNVW (15-letter code of the IUPAC-IUB nomenclature) located in the 5' direction from G [10], and (4) contextual features of DNA reflecting the frequency of its contacts with the histone-like proteins [11]. Based on these four characteristics, which were first revealed by us, we constructed a linearly additive estimate of the frequency of UV-induced damage of guanine on its local environment in DNA and for the first time showed that the predictions of this estimate significantly agreed with the measurements of an independent control experiment [6].

We analyzed the experimental data [12] on the frequency of damage f_{G_n} of 43 guanine residues in a DNA region 316 bp long (positions 146 to 461 in the *Escherichia coli* pGEM7 (f+) plasmid under irradiation with 193-nm ultraviolet laser beam (table: $1 \le n \le 43$ in the descending order of f_{G_n} values from 1.59 to 0.00 logarithmic units, ln). A set of DNA sequences $S_{G_n} = \{s^n_{-10} \dots s^n_0 = G_n \dots s^n_j \dots s^n_{10}\}$, with a center in each G_n and a length of 21 bp each, was created (table).

At first, we built the simplest of the common contextual DNA characteristics—the consensus of UV damage of guanine. To do this, all 43 sequences S_{G_n} were grouped into six groups. In the control set $S_{\text{cont}\,G_n}$, a representative of each of the six groups with the f_{G_n}

[a] *Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, ul. Vavilova 32, Moscow, 119991 Russia*
[b] *Moscow State University, Moscow, 119992 Russia*
[c] *Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, pr. Akademika Lavrent'eva 10, Novosibirsk, 630090 Russia*
[d] *Novosibirsk State University, ul. Pirogova 2, Novosibirsk, 630090 Russia*
[e] *Kurchatov Institute National Research Center, pl. Akademika Kurchatova 1, Moscow, 123182 Russia*

Frequencies of damage f_{G_n} of 43 guanine residues (G_n) at positions (146–461) in the *E. coli* pGEM7 (f+) plasmid by 193-nm ultraviolet laser beam [12], DNA characteristics found on their basis, and predictions made using Eq. (6)

n	Local environment of G_n in DNA, sequence, S_{G_n}	Frequency f_{G_n}, ln, exp. data [12]	DNA parameters important for UV damage of G_n				Prediction by Eq. (6)
			matches with consensus	similarity-PWM_{G_n}	tetranucle-otide YNVW	B-DNA helix, $P_{23[-7;\,2]}$	
1	cagaacat t t **G**at accaaacc	1.59	2	−2.02	1.40	14.41	1.44
2	t t aaaccct g**G**gaaccgcaag	1.57	0	−1.95	2.06	12.45	1.61
3	t t t aaaccct **G**ggaaccgcaa	1.55	2	−2.76	1.55	13.48	1.28
4	cgcaaggt t g**G**gcaaat aaag	1.53	2	−2.69	1.20	13.48	1.21
5	cccct t cct t **G**gt at ggaaaa	1.43	1	−2.95	1.16	14.15	1.27
6	agaacact aa**G**agct cagat c	1.43	1	−2.74	0.51	12.63	1.07
7	ggcaaat aaa**G**gct aat cat a	1.39	4	−2.78	0.93	13.96	1.04
8	ct t gt t t t aa**G**aacagt t t gt	1.38	2	−3.45	1.13	14.48	1.13
9	at aagt t t t t **G**cagaat aat g	1.33	3	−2.88	0.30	14.90	1.01
10	accaact cag**G**aaaccact t g	1.29	0	−2.59	1.06	13.26	1.33
11	aaccaact ca**G**gaaaccact t	1.28	2	−2.56	0.73	13.81	1.16
12	gct cagat ca**G**aacat t t gat	1.19	3	−3.21	1.42	12.69	1.05
13	ccgcaaggt t **G**ggcaaat aaa	1.15	3	−3.37	1.02	14.16	1.03
14	t t at at t at g**G**t t t acat aag	1.13	4	−2.87	1.34	10.16	0.87
15	aagaacat ag**G**aaaat agaac	1.12	1	−2.35	0.10	11.71	0.99
16	agt t t t t gca**G**aat aat gt t c	1.10	3	−3.85	0.32	14.70	0.82
17	ggaaccgcaa**G**gt t gggcaaa	1.10	0	−2.78	0.24	10.69	0.95
18	cct gggaacc**G**caaggt t ggg	1.02	5	−4.34	0.10	9.60	0.22
19	agaacagt t t **G**t aaccat aaa	0.95	3	−2.89	0.50	14.16	1.00
20	gcagaat aat **G**t t ct at cagt	0.93	4	−3.48	1.23	11.83	0.84
21	t ccagccact **G**cccct t cct t	0.86	3	−3.52	1.32	11.40	0.89
22	cat accat aa**G**t t t t t gcaga	0.85	2	−3.01	0.39	11.58	0.86
23	t cacat cct t **G**t t t t aagaac	0.85	2	−3.08	1.16	12.90	1.10
24	aacact aaga**G**ct cagat cag	0.82	5	−3.41	0.91	11.85	0.72
25	t aaaccct gg**G**aaccgcaagg	0.81	2	−2.87	1.29	11.56	1.08
26	gact t ggat a**G**at t ccaaaag	0.80	3	−3.25	0.60	11.85	0.81
27	t at cagt cca**G**ccact gcccc	0.78	3	−3.98	0.52	12.16	0.68
28	cat aagact t **G**gat agat t cc	0.77	3	−3.12	0.20	14.15	0.90
29	gaaaccact t **G**t ct cacat cc	0.73	6	−3.89	0.44	13.14	0.54
30	gcaaat aaag**G**ct aat cat aa	0.72	4	−3.25	1.26	12.31	0.92
31	caaaaaat ca**G**cact ct t t t a	0.71	4	−2.91	1.20	13.64	1.06
32	at t t acat aa**G**act t ggat ag	0.71	3	−3.19	0.40	11.21	0.74
33	t t t aagaaca**G**t t t gt aacca	0.69	4	−3.63	1.02	14.15	0.91
34	t aagagct ca**G**at cagaacat	0.67	0	−2.62	0.32	12.29	1.10
35	gat t ccaaaa**G**aacat aggaa	0.67	4	−3.48	0.63	14.78	0.90
36	at aagact t g**G**at agat t cca	0.64	4	−3.31	0.30	13.26	0.76
37	gcaaggt t gg**G**caaat aaagg	0.51	2	−3.25	0.20	11.83	0.79
38	t gt t ct at ca**G**t ccagccact	0.51	6	−3.52	1.01	11.36	0.62
39	gt t t acat aa**G**cat t t acat a	0.50	3	−2.86	0.40	10.45	0.75
40	t t t at at t at **G**gt t t acat aa	0.43	4	−3.19	1.13	9.79	0.74
41	t aggaaaat a**G**aacact aaga	0.41	3	−2.81	0.30	13.21	0.92
42	gaaccgcaag**G**t t gggcaaat	0.37	4	−3.81	0.29	10.69	0.50
43	aaagaacat a**G**gaaaat agaa	0.00	4	−2.82	0.10	11.71	0.71
Linear correlation coefficient (significance), $r(\alpha)$		training	−0.41 (0.05)	0.35 (0.05)	0.45 (0.05)	0.44 (0.05)	0.67 (10^{-6})
		control	−0.91 (0.05)	0.86 (0.05)	0.48 (0.05)	0.44 (0.05)	

Note: 43 guanine residues (G_n) are shown in bold, control is underlined, the last column shows the prediction for controls.

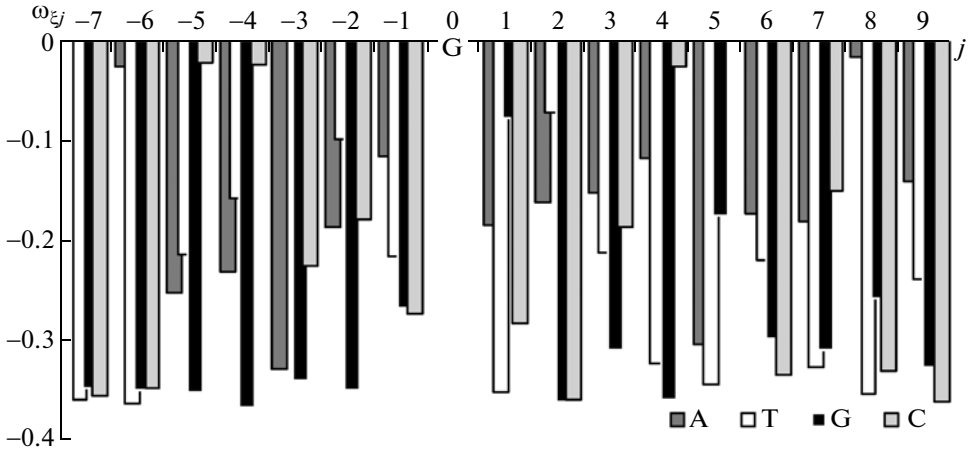

Fig. 1. Graphical representation of the position-weight matrix $\omega_{\xi j}$ of nucleotides ξ at positions j in the local environment of UV damaged guanine residues at $j = 0$. For $j \in \{-10, -9, -8, 10\}$, $\omega_j \equiv 0$ was set due to the lack of significantly frequent or rare nucleotides here. Numbers above—position in the vicinity of G.

value closest to its mean intra-group estimate (Table 1: no. 4, 12, 19, 28, 35, 42) was chosen. The remaining 37 sequences S_{G_n} formed the training set $S_{train'G_n}$.

On the basis of the training set $S_{train'G_n}$, the frequency $f_{\xi j}$, weighted with allowance of the measured f_{G_n} frequencies of UV damage of G_n (the conventional Boltzmann's population), was estimated for each nucleotide $\xi \in \{A, T, G, C\}$ at each position j from −10 to 10:

$$f_{\xi j} = \frac{\sum\limits_{S_{G_n} \in \{S_{train G_n}\}} \delta(\xi = s_j^n) f_{G_n}}{\sum\limits_{S_{G_n} \in \{S_{train G_n}\}} f_{G_n}}, \quad (1)$$

where $\delta(\xi = s_j^n) = 1$ for a match of s_j^n with a given ξ and 0 in the case of a mismatch, $\xi \neq s_j^n$.

On the basis of $f_{\xi j}$ values, found using Eq. (1), for each ξ we estimated the mean frequency and its standard error $f_\xi \pm \sigma_\xi$ for the $S_{train\,G_n}$ sample in general. Using the Student's t test, the most significantly frequent nucleotide was found for each position j, and the consensus of the environment of UV damages in guanine was constructed on this basis. However, the number of matches S_{G_n} with this consensus and the f_{G_n} frequencies correlated with each other neither in the training nor in the control sets (data not shown). In view of this, the consensus of the most significantly rare nucleotides in the vicinity of UV damages of G_n was constructed, which took the form ttaaagcHtcg-actgc (H is the code "not guanine" and "–" is any nucleotide in the IUPAC-IUB nomenclature [10]). In the control sample $S_{cont\,G_n}$, the number of matches to this consensus was significantly correlated with the f_{G_n} frequencies (Table 1: $r = -0.91$, $\alpha < 0.05$). Thus, the better the local environment of G_n coincides with the ttaaagcHtcg-actgc consensus, the lower the frequency of UV damaged guanine residues.

Then, the next (in terms of complexity) characteristic of the nucleotide context of DNA sites—the position-weight matrix $\omega_{\xi j}$ in the vicinity of UV damaged G_n—was determined:

$$\omega_{\xi j} = \frac{f_{\xi j}}{\max_\xi(f_{\xi j}) + \max_j(f_{\xi j})} \ln\left(\frac{f_{\xi j}^2}{\max_\xi(f_{\xi j})\max_j(f_{\xi j})}\right), \quad (2)$$

where $\max_j(f_{\xi j})$ is the weight $\omega_{\xi j}$ of nucleotide ξ at each position j in the G_n environment normalized to the maximum frequency of this nucleotide, added to the conventional Shannon's information content formula. The graphical representation of $\omega_{\xi j}$ is shown in Fig. 1. At positions $j \in \{-10, -9, -8, 10\}$, it was found that $\omega_{\xi j} \equiv 0$, which means the absence of significantly frequent or significantly rare nucleotides at these positions.

The table shows the values of similarity of the control sequences $S_{G_n} \in \{S_{cont\,G_n}\}$ with the weight matrix $\omega_{\xi j}$, calculated using the formula

$$PWM(S_{G_n}) = \sum_{j=-10}^{10} \omega_{s_j^n j}. \quad (3)$$

They are significantly correlated with f_{G_n} in the control ($r = 0.86$, $\alpha < 0.05$) (see the table). Therefore, the more the local environment of guanine residues satisfies the weight matrix shown in Fig. 1, the higher the frequency of damage of these guanine residues.

Next, using the ACTIVITY software, which was developed by us earlier [13], we analyzed 43 S_{G_n} sequences to identify the contextual DNA characteristics that, firstly, significantly affect the frequencies f_{G_n} of UV-induced guanine damages and, secondly,

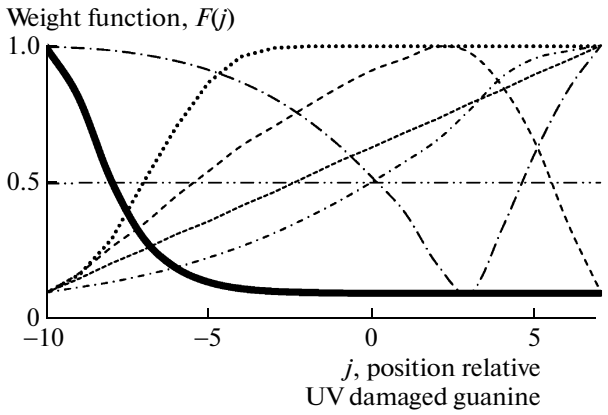

Fig. 2. Examples of seven of the 360 weight functions $F(j)$ vs positions j relative to UV damaged guanine residues at $j = 0$. $F(j)$ functions (Eq. (4)) determine the effect of tetranucleotides Z on f_G—the frequency of UV damage at localization Z at position j on the basis of the heuristic rule "the higher $F(j)$, the greater the effect of Z at position j in sequence S_{G_n} on the frequency f_{G_n} of UV damaged G_n." The bold line shows the weight function $F(j)$, found by the ACTIVITY program [13] as a result of analysis of the training set data for the tetranucleotide YNWV.

meet the requirements for the application of regression analysis to data. All possible tetranucleotides $Z = z_1z_2z_3z_4$ (in the 15-letter code of the IUPAC-IUB CBN nomenclature [10]: $z \in \{A, T, G, C, W = A + T, R = A + G, M = A + C, K = T + G, Y = T + C, S = G + C, B = T + G + C, V = A + G + C, H = A + T + C, D = A + T + G, N = A + T + G + C\}$ were considered as contextual characteristics. In total, we considered $15^4 = 50\,625$ tetranucleotides (i.e., an exhaustive search was performed). In the ACTIVITY software [13], the contribution of oligonucleotide Z to the guanine damage frequency f_{G_n} was set by the formula

$$Z_F(S_{G_n}) = \sum_{j=-10}^{7} \delta(s_j^n = z_1)\delta(s_{j+1}^n = z_2) \times \delta(s_{j+2}^n = z_3)\delta(s_{j+3}^n = z_4)F(j). \quad (4)$$

Here, $0 \le F(j) \le 1$ is the weight function, which has the following meaning: the higher the $F(j)$, the greater the contribution of Z at position $F(j)$ to the frequency f_{G_n} of UV-induced guanine damages. For each of the 15^4 tetranucleotides $Z = z_1z_2z_3z_4$, 360 weight functions F (a total of $360 \times 15^4 \approx 10^7$ variants ZF) were considered (Fig. 2). The necessity of consideration of ten million of contextual characteristics Z_F was determined by the absence of any information on the effect of oligonucleotides on the frequency of UV damage of guanine residues.

A set of 43 sequences S_{G_n} was divided into two subsets: even and odd sequences formed the training and control subsets, respectively (table). In the training subset, the ACTIVITY software [13] tested each Z_F characteristic with respect to five types of correlations and six requirements of the regression analysis. As a result, the program calculated the U value $(Z_F; f)$, which reflected, firstly, the significance of the Z_F contribution to the f_{G_n} frequency of UV-induced guanine damages and, secondly, the applicability of the regression analysis to $\{Z_F; f_{G_n}\}$ data. Of the 10^7 considered characteristics of Z_F, we selected the one that had the greatest positive value $U(Z_F; f)$. The ACTIVITY software [13] was set up so that the Bonferroni correction for multiple tests gave the estimate $p < 10^{-20}$ of the probability of accidental appearance for random causes of the final positive value $U(Z_F; f)$, calculated by this program.

The largest value $U(Z_F; f) = 0.25$ was found for the tetranucleotide YNVW, characterized by the weight function F (shown with a bold line in Fig. 2). This feature decreased from $F = 1$ at the 5' end of the sequence to the minimum $F = 0.1$ at its center. The $YNVW_F$ values calculated using formula (4) are summarized in the table. They are significantly correlated with the f_{G_n} values in the control ($r = 0.48$, $\alpha < 0.05$). Thus, the larger the number of tetranucleotides YNVW in the 5' direction from the guanine residue, the higher the frequency of guanine damages by UV radiation. Additionally, the tetranucleotide YNVW fits the regression analysis better than any other tetranucleotide, which allowed it to be used it at the final stage of the study when constructing the regression equation (Eq. (6)) to predict the frequency of UV damages of guanine residues in random DNA sequences.

Finally, we used the ACTIVITY software [13] to analyzed the contribution of 38 physicochemical and conformational properties of the B-DNA helix to the frequency of UV damages of guanine residues. For each of these properties, which were taken from the database [14], the mean value in all possible regions $[a; b]$ of analyzed DNAs was estimated by equation

$$P_{k[a;\,b]}(S_{G_n}) = \frac{1}{b-a}\sum_{j=a}^{b-1} P_k(s_j^n s_{j+1}^n), \quad (5)$$

where $-10 \le a < b \le 10$ and $P_k(\xi\zeta)$ is the value of k-th property of the dinucleotide $\xi\zeta$ step of B-DNA helix, $1 \le k \le 38$ [14]. In total, $38 \times (20 \times 19)/2 = 7220$ variants of $P_{k[a;\,b]}$ were analyzed. Because of the lack of any information on the effect of conformational and physicochemical properties of B-DNA helix on the frequency of UV damages of guanine residues, it was necessary to test each of the 7220 $P_{k[a;\,b]}$ variants.

The analysis of the training subset of even DNA sequences from the table showed $U(P_{23[-7;\,2]}; f) = 0.15$; i.e., the greatest contribution to the guanine damage frequency is made by the physicochemical property "the frequency of dinucleotide contact with a histone-like protein" [11] in the range $[-7, 2]$ relative to guanine. The $P_k(\xi\zeta)$ values used to calculate this property

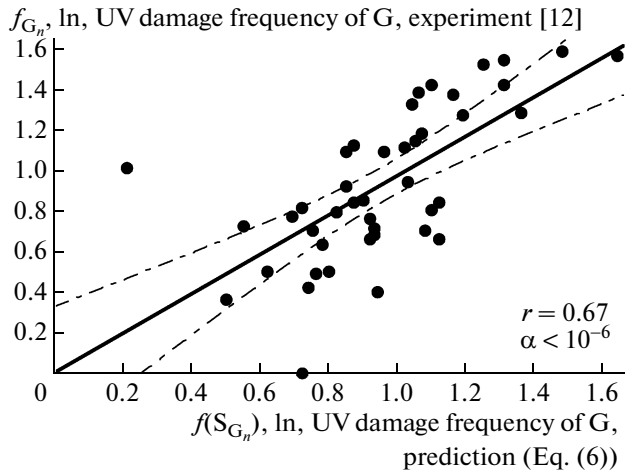

Fig. 3. Correlations of the prediction (Eq. (6)) of the frequencies fS_{G_n} of UV damage of G_n in the region [146, 461] of the *Escherichia coli* pGEM7 (f+) plasmid by 193-nm UV laser beam with the f_{G_n} values measured in the experiment [12], which were analyzed in this work. The dashed line line shows 5% confidence limits (STATISTICA package).

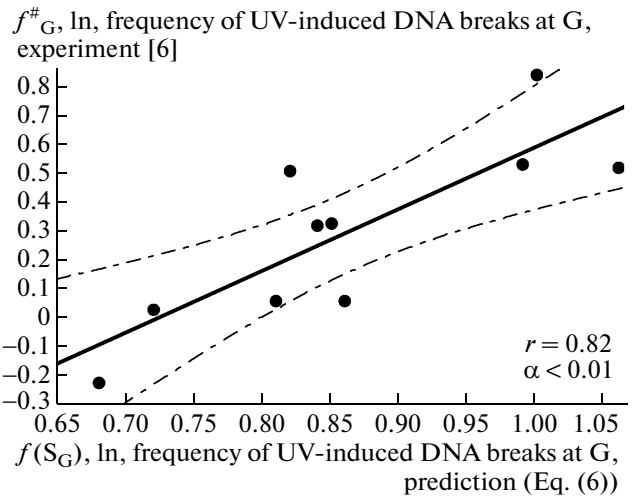

Fig. 4. Correlation between the frequencies $f^{\#}_{G_n}$ of DNA breaks at guanine residues induced by 193-nm UV laser beam in the DNA region 100 bp long in the mouse gene MIP-1α (macrophage inflammatory protein 1α, GenBank AC = X12531: 201–300) and the in silico prediction (Eq. (6)) of the frequencies fS_{G_n} of UV damages in guanine residues [6]. The dashed line shows 5% confidence limits (STATISTICA package).

(Eq. (5)) were taken from the database [14]. They were experimentally determined earlier as a result of analysis of 177 experimentally established DNA sites 145 bp long that were in contact with the core particle of hen nucleosome [11] and, therefore, were taken as a basis for estimating the frequency of contacts with histones for each of the 16 possible dinucleotides [11].

The values calculated using Eq. (5) are summarized in the table. They are significantly correlated with the frequency of guanine damage in the control ($r = 0.44$ ($\alpha < 0.05$)). Thus, the more readily the local environment of guanines comes in contact with the histone-like proteins, the higher the UV radiation-induced damage of these guanines under experimental conditions in vitro [12]. In addition, this contextual characteristic meets the requirements of the regression analysis, which allows it to be used in the regression equation (Eq. (6)) to predict the frequency of UV damages of guanine residues in random DNA sequences.

It should be noted that *E. coli* contains histone-like proteins HU [15], which are evolutionary homologous of hen histones from the experiment described in [11] and which pack the bacterial DNA in nucleosome-like structures [15]. *E. coli* strains defective for HU proteins are characterized by an abnormally high sensitivity of DNA to UV radiation [15]. This allowed us to make a conjecture that a mechanism for protecting the DNA regions that are most easily damaged by UV radiation could be selected in the course of evolution of the bacterial genomes by the binding of these regions to histone-like proteins.

Thus, as a result of analysis performed in this study, we, firstly, have identified four contextual DNA features that are important for estimating the frequencies of UV damages of guanine residues and, secondly, found that these characteristics are not correlated with one another. For this reason, they were used without optimization to construct a linearly additive formula to estimate the frequency of UV damages of guanine residues:

$$f(S_{G_n}) = 0.69 - 0.07N_{cons}(S_{G_n}) + 0.19PMW(S_{G_n}) + 0.22YNWV_F(S_{G_n}) + 0.07P_{23[-7;\,2]}(S_{G_n}), \quad (6)$$

where $N_{cons}S_{G_n}$ is the number of matches of S_{G_n} with the consensus ttaaagcHtcg-actgc in the local environment of UV damaged G_n.

The frequencies of UV damages of guanine residues for 43 S_{G_n} sequences, calculated according to Eq. (6), were significantly correlated with the experimental data (table, Fig. 3). In addition, using Eq. (6), we predicted the frequencies of UV damages of guanine residues in an independent experiment [6] (Fig. 4). In this experiment, $f^{\#}_G$ frequencies of UV-induced DNA breaks at guanine residues in the mouse gene MIP-1α under irradiation with 193-nm laser beam were measured in vitro. As can be seen from Fig. 4, the results of our prediction was significantly correlated with the experimental data ($r = 0.82$, $\alpha < 0.01$).

To conclude, it should be noted that all the four identified contextual features of the local environment of guanine residues make partial linearly additive contributions to the frequencies of UV damage of these guanine residues with indistinguishably equal minimum levels of significance $\alpha < 0.05$, whereas their total contribution (Eq. (6)) has a 1000-fold higher level of significance ($\alpha < 10^{-6}$). This finding points out

to the high complexity of the multistage cooperative molecular mechanism of UV radiation-induced damage of guanine residues in DNA molecules.

ACKNOWLEDGMENTS

The study was supported by the Russian Foundation for Basic Research (project nos. 11-04-02001, 11-04-01888, and 12-04-01584), the program of the President of the Russian Federation "Leading Scientific Schools" (project NSh-5278.2012.4), the Integratsiya Foundation for Special Research Programs (project nos. 6.8 and 30.29), and the programs of the Presidium of the Russian Academy of Sciences "Biodiversity," "Molecular and Cellular Biology," and "The Origin and Evolution of the Biosphere."

REFERENCES

1. Thogersen, J., Knudsen, C.S., Maetzke, A., et al., *Int. J. Radiat. Biol.*, 2007, vol. 83, pp. 289–299.
2. Do, J.E., Shin, J.Y., Kim, D.Y., et al., *Photodermatol. Photoimmunol. Photomed.*, 2011, vol. 27, pp. 147–151.
3. Maloney, M.D., Goater, J.J., Parsons, R., et al., *J. Bone Joint Surg. Am.,* 2006, vol. 88, pp. 753–761.
4. Schastak, S., Yafai, Y., Yasukawa, T., et al., *Lasers Surg. Med.,* 2007, vol. 39, pp. 353–357.
5. Gurzadyan, G.G., Gorner, H., and Schulte-Frohlinde, D., *Photochem. Photobiol.*, 1993, vol. 58, pp. 477–485.
6. Melvin, T., Cunniffe, S.M., O'Neill, P., et al., *Nucleic Acids Res.*, 1998, vol. 26, pp. 4935–4942.
7. Saito, I., Takayama, M., Sugiyama, H., et al., *J. Am. Chem. Soc.,* 1995, vol. 117, pp. 6406–6407.
8. Kirpota, O.O., Endutkin, A.V., Ponomarenko, M.P., et al., *Nucleic Acids Res.,* 2011, vol. 39, pp. 4836–4850.
9. Grokhovsky, S.L., Il'icheva, I.A., Nechipurenko, D.Yu., et al., *Biophys. J.*, 2011, vol. 100, pp. 117–125.
10. IUPAC-IUB CBN, *J. Mol. Biol.*, 1971, vol. 55, pp. 299–310.
11. Satchwell, S.C., Drew, R., and Travers, A.A., *J. Mol. Biol.*, 1986, vol. 191, pp. 659–675.
12. Vtyurina, N.N., Grokhovskii, S.L., Filimonov, I.V., et al., *Biophysics,* 2011, vol. 56, no. 3, pp. 399–402.
13. Omelyanchuk, N.A., Ponomarenko, P.M., and Ponomarenko, M.P., *Mol. Biol.* (Moscow), 2011, vol. 45, no. 2, pp. 366–375.
14. Ponomarenko, J.V., Ponomarenko, M.P., Frolov, A.S., et al., *Bioinformatics,* 1999, vol. 15, pp. 654–668.
15. Li, S. and Waters, R., *J. Bacteriol.*, 1998, vol. 180, pp. 3750–3756.