Improvement of fidelity of molecular DNA computing using laser spectroscopy

2015 Laser Phys. 25 035202

(http://iopscience.iop.org/1555-6611/25/3/035202)

IOP Publishing | Astro Ltd Laser Physics

Laser Phys. **25** (2015) 035202 (10pp) doi:10.1088/1054-660X/25/3/035202

Improvement of fidelity of molecular DNA computing using laser spectroscopy

T A Dolenko[1], S A Burikov[1], K A Laptinskiy[1], A A Moskovtsev[2,3], M V Mesitov[2] and A A Kubatiev[2,3]

[1] Moscow State University, Department of Physics, 119991, Leninskie Gory 1/2, Moscow, Russia
[2] Research Institute of General Pathology and Pathophysiology RAMS, 125315, Baltiyskaya street, 8, Moscow, Russia
[3] Russian Medical Academy of Postgraduate Education, 125315, Baltiyskaya street, 8, Moscow, Russia

E-mail: tdolenko@mail.ru and tdolenko@lid.phys.msu.ru

Received 15 December 2014
Accepted for publication 23 December 2014
Published 28 January 2015

Abstract

This work is devoted to application of laser Raman spectroscopy to improve the fidelity of molecular computing by DNA strands. The developed method provides determination of concentrations of specific nitrogenous bases not lower than 0.03 g l^{-1} and the accuracy of determining the total concentration of DNA in solutions 0.02–0.04 g l^{-1} by the Raman spectra of DNA solutions. The method allows us to control the temperature of DNA solutions in the process of molecular computing with an accuracy better than 0.2 °C. The obtained results provide required control of the parameters of DNA solutions to increase the speed and accuracy of solving the problems in the results of molecular computing.

Keywords: DNA strand, molecular computations, nitrogenous bases, Raman spectroscopy

(Some figures may appear in colour only in the online journal)

1. Introduction

Natural means of storage, delivery and processing of information in biological systems have always stimulated natural interest because of their high efficiency and mechanisms of functioning that are still not understood completely. As a result of discoveries in the last decades in the field of biochemistry and the development of new information and molecular technologies it is now possible to transfer information to biomolecular media and to read it back. This has raised the question of elaboration of the principles of operation and practical construction of biocomputers. Development of calculations on biological structures is very important not only for progress of new information technologies, but for understanding the functioning of living systems on the molecular level. Some advantages of molecular computing are immanent large-scale parallelism of operations, high density of information per unit volume, high energetic efficiency and availability of experimental realizations of solution of simple problems.

In spite of the fact that the strategic direction of development of conventional calculations on the basis of silicon computers is the use of multicore and multithreaded calculations, the degree of parallelism achieved here is incomparable with the degree of parallelism inherent in biological and quantum systems. However, quantum and molecular computational paradigms have also many common problems, most notably fidelity of calculations. This problem is very complex and versatile. For molecular DNA computing, the solution of this problem lies in elaboration of principles for estimation of the state of biological structures during DNA computing and in increasing their fidelity.

1.1. Molecular DNA computing

The main idea of calculations using biological structures is construction of new models and new algorithms on the basis of knowledge of the structure and functions of molecules of these structures and on the basis of operations performed with them in living cells with the help of different enzymes.

In living cells, genetic information is encoded in the molecule of deoxyribonucleic acid—DNA. DNA is one of the two types of nucleic acids, providing storage, delivery from generation to generation and realization of a genetic program of development and functioning of living organisms. DNA is a

long polymeric molecule consisting of repeating elements—nucleotides [1–3]. Each nucleotide consists of a nitrogenous base (adenine (A), guanine (G), cytosine (C), thymine (T)), deoxyribose, and phosphate group. Nucleotides in the strand are bonded with each other by a phosphodiester bond between the deoxyribose and phosphate group. Thus, in these supermolecules, phosphate and deoxyribose play the role of supporting structure—backbone (they alternate each other in the strand) and the sequences of nitrogenous bases encode information.

Use of DNA molecules as the element base of a new generation of computers became possible due to DNA replication—self-reproduction of nucleic acids providing accurate copying of genetic information and its delivery from generation to generation [4]. During replication, denaturation of the DNA double helix occurs and synthesis of new polynucleotide strands begins simultaneously. The synthesis occurs with the help of a special enzyme—DNA-polymerase. It slides along DNA and synthesizes a new strand on its basis. In this strand, all nitrogenous bases are replaced by their complementary bases (PCR—polymerase chain reaction). The enzyme starts the reaction only in the presence of a DNA-primer forming a double-stranded segment of DNA. Free nucleotides are attached to this primer.

The new strand synthesized on each of the initial strands, is identical to another initial strand. As the result of the synthesis, two identical double helixes are formed. Each of these helixes consists of one initial strand and one new strand [1, 2]. Thus, only one of the two strands composing the initial DNA molecule is transferred from one generation to another. This is the semi-conservative mechanism of DNA replication [5]. In fact, synthesis of a nucleic acid strand on a DNA-matrix is a realization of a Turing machine, which consists of two tapes and programmable control desk. The control desk reads the data from one tape, processes these data according to a certain algorithm and records the data on another tape. Polymerase also consecutively reads the initial data from the initial DNA tape and forms a new strand on this basis.

For construction of computational structures on the basis of DNA, specialists in theory of computing and mathematical modeling describe the properties of DNA strands and operations with them with the help of formal programming languages [6]. One can almost completely formalize operations with DNA—synthesis, elongation, shortening and cutting, cross linking, modification, interlacing, sequencing etc.

The possibility of DNA-based computing was demonstrated for the first time in 1994 by Adleman [7]. In his study, DNA strands were used for solution of the problem of proving the existence of a Hamiltonian path (passing each vertex of the graph exactly once) in a directed graph G with selected starting and final vertexes.

Adleman assigned a unique code to each of the 7 vertexes of the graph—a strand consisting of 20 nucleotides, a certain sequence of nitrogenous bases (A), (T), (G), (C). The arcs of the graph were also encoded by strands, beginning with the second half of the code of vertex of departure and ending with the first half of the code of the vertex of arrival.

Then for sequences encoding vertexes of the graph, their complementary strands were synthesized. Using these complementary strands and strands encoding arcs of the graph, all possible paths were created randomly in the test tube. The desired solution of the stated problem should have been encoded by a DNA sequence consisting of 140 nucleotides (7 * 20).

For synthesis and separation of such a molecule, about 100 trillion DNA molecules containing all possible 20-nucleotide sequences encoding graph vertexes and paths between them, were placed in the test tube. Because of the mutual attraction of nitrogenous bases A-T, G-C, single DNA strands were randomly bonded, enzyme ligase joining short molecules into large-scale complexes (oligonucleotides—into threads).

At that, DNA molecules simulating all possible paths between vertexes were synthesized. It was only necessary to isolate molecules corresponding to the desired solution. Adleman has solved this problem by biochemical methods, consecutively removing strands that did not include first vertex (starting point) and last vertex (endpoint), then by removing strands containing more than 7 vertexes and no vertexes. After such selection, any of the remaining DNA molecules represented a solution of the problem. Solution of the problem took 7 d [8]. Note that most of the effort was spent not on synthesis of the initial data and elaboration of the algorithm, but on filtration and extraction of the final result in the form suitable for human perception.

Later on, the approach suggested by Adleman was developed in the solution of a number of other problems with the help of DNA computing—breaking the DES encryption system (data encryption standard system) [9, 10], solution of some chess problems [11], matrix multiplication [12] etc. In the studies of the group of Shapiro [13, 14], it was suggested the research not be restricted by a special algorithm for solution of a concrete problem but to create the technology of a multi-purpose computer on the basis of known properties of biomolecules. Within this group, finite molecular automation was realized—a system consisting of initial state, a variety of finite states, and functions of transitions. At the present time, algorithms of the solution of a number of NP-complete problems related to logic, theory of graphs, computer networks, set theory, mathematical programming, optimization of processes and programs, have been elaborated theoretically on the base of DNA computations [15].

1.2. Fidelity of molecular computing

Molecular computing on DNA strands is basically a set of biochemical reactions. All operations performed with DNA molecules are rather complicated and laborious and this is why on the stage of every reaction, errors on the molecular level can take place [16]. The most vexing problem is the last stage of reading out the result—isolation, assortment and analysis of DNA strands. Contemporary methods are unable to provide sequencing of strands even with the length of several thousands of bases. Solutions of the problem can be lost during intermediate operations (molecules gather at the walls of the test tube), there is no guarantee that there will be no damage to

DNA molecules, DNA molecules tend to disintegrate eventually etc. In real experiments, these losses can reach 10% from the initial amount of molecules, whereas the elimination of even 1% of molecules from the process of calculations leads to an incorrect solution of the problem.

During molecular computations, when implementing different algorithms, one has to deal not only with different DNA concentrations and different length of strands but with different types and concentrations of buffer solution, different temperature profiles of reaction, i.e. with the change of numerous parameters of media influencing the accuracy of the solution of the problem. Authors of many studies [14–20] note considerable dependence of the accuracy of calculations on the concentration of biological molecules, temperature and the temporal reaction regime, type of buffer, values of pH of mixtures, on properties and composition of oligonucleotides. Thus, one of the main problems of the development of molecular DNA computing is minimization of errors during biochemical reactions, i.e. increasing the fidelity of molecular computing.

This situation causes necessity of elaboration of evaluation methods for the state of biological structures and methods of increasing fidelity of DNA computing, i.e. minimization of experimental errors during DNA computing.

Obviously these methods should satisfy the following requirements:

- they must be noninvasive, i.e. non-destructive (*in vivo*)
- they must work in real-time mode for permanent control of changes during calculations;
- they should simultaneously measure as many parameters of the media as possible.

Laser Raman spectroscopy satisfies these requirements. As it is known, this method is widely used for solution of problems of express monitoring of natural waters and water solutions [21, 22], simultaneous determination of several parameters of water media [23], including solution of inverse multi-parameter problems of laser spectroscopy of water media using modern methods of pattern recognition—artificial neural networks [24, 25].

With the help of laser Raman spectroscopy, structure and properties of DNA molecules were studied in detail [26, 27]. Bands corresponding to vibrations of bonds in nitrogenous bases and DNA backbone were identified; behavior of DNA vibrational spectra in dependence on change of temperature, type of DNA conformation, etc was studied [28, 29]. This gives the premise for use of this method for the solution of problems of control of the state of reaction mixture during molecular computing.

This study concerns use of Raman spectroscopy for increase of fidelity of molecular computing on DNA strands, namely for express and remote control and measurement of parameters of media where biochemical reactions with DNA molecules take place. The results of the use of elaborated methods for determination of concentration of individual nitrogenous bases and total concentration of DNA in solutions using their Raman spectra and the results of control of the temperature regime of biochemical reactions by Raman spectra, are presented.

2. Materials and methods

2.1. Objects of research

A peculiarity of molecular DNA computing is the fact that for these calculations one needs a large amount of 'actuation medium' in the solution—from units of $g\,l^{-1}$ of DNA molecules ($3.32\,g\,l^{-1}$ [30]) up to tens of $g\,l^{-1}$ ($32.5\,g\,l^{-1}$ [31])—depending on the current task. During molecular computing, the length of DNA strands in the solution and DNA concentration may vary in a wide range. It depends both on specifics of the concrete problem solved by DNA molecules and on the stage of solution of this problem. This is why samples containing strands with length and DNA concentration differing by more than two orders were chosen as objects of the research.

The following samples were used as objects of the research.

(a) Sample 1: sodium deoxyribonucleate containing linear double-stranded DNA excreted from salmon sperm and consisting of various nucleotides. DNA is placed in 3x SSC buffer (water solution of 0.45 M NaCl + 0.045 M $Na_3C_6H_5O_7$*5.5H_2O). Maximal concentration of sample 1 in the solution was $25\,g\,l^{-1}$.

(b) Sample 2 containing specially synthesized and sequenced linear DNA strands consisting of 100 prescribed nitrogenous bases (86 of them adenine, the rest—cytosine) in 3x SSC buffer. Maximal concentration of sample 2 was $9\,g\,l^{-1}$. Sequence of nucleotides in the DNA strand was the following:

AAA-AAA-AAA-AAA-AAA-AAA-AAA-AAA-ACC-CAA-AAA-AAA-AAA-AAA-ACC-CCA-AAA-AAA-AAA-AAA-AAA-AAA-AAA-AAA-CCC-AAA-AAA-AAA-AAA-AAA-AAA-CCC-C,

(A—adenine, C—cytosine).

(c) Sample 3: λ-DNA—linear double-stranded DNA (48510 known bases) excreted from bacteriophage lambda in buffer (water solution with 10 mM TRIS ($C_4H_{11}NO_3$)-HCl and 1 mM EDTA ($C_{10}H_{16}N_2O_8$)). Maximal concentration of sample 3 was $0.3\,g\,l^{-1}$. This sample was produced by Thermo Scientific.

2.2. Experiment

The specified content of DNA in the solution (see 2.1) provides the possibility of use of Raman spectroscopy in its conventional variant—collecting signals from the bulk of solution. This allows control of practically the whole bulk of 'actuation medium'. Another advantage is very high sensitivity of the Raman spectra to many parameters of molecules and solutions—molecular connectivity (at the first place, strength of hydrogen bonds), temperature and saline composition of solutions. This allows simultaneous monitoring of change of these parameters by Raman spectra. Furthermore, excitation of the Raman signal was performed by laser radiation with 488 nm wavelength, so it is a nondestructive method for DNA molecules (it does not cause damage or modification of DNA molecules).

The Raman signal was excited by an argon laser (wavelength 488 nm, output power 200 mW). Registration of spectra

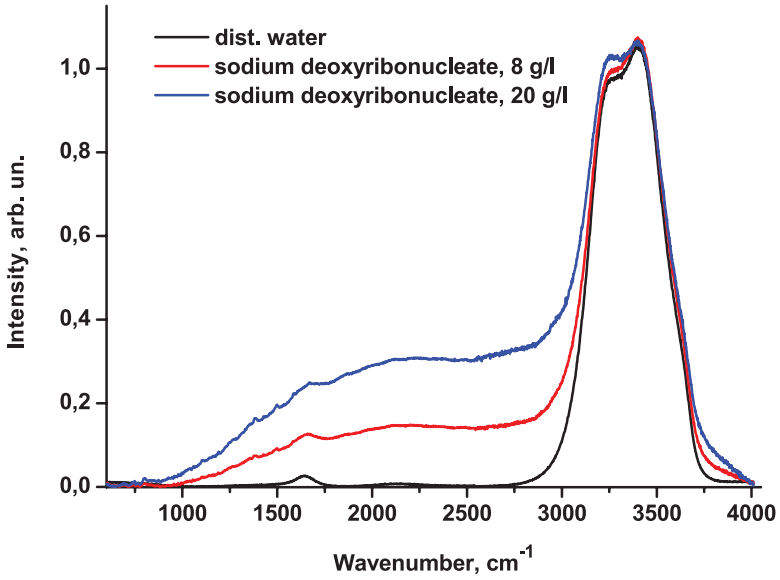

Figure 1. Raman spectra of the solutions of Sample 1 with different concentrations.

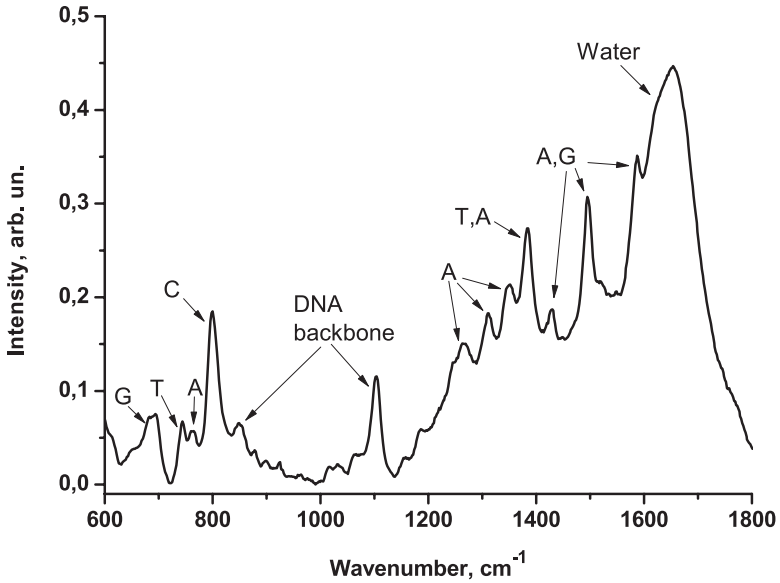

Figure 2. Low-frequency Raman spectrum of solution of sodium deoxyribonucleate (20 g l^{-1}). A—adenine marker, C—cytosine marker, G—guanine marker, T—thymine marker.

was performed in 90° geometry by a monochromator (Acton 2500i, grade 900 grooves mm^{-1}, focal length 500 mm) and a CCD camera (Horiba Jobin Yvon, model Synapse BIUV). In order to remove the strong signal of elastic scattering, an edge filter (Semrock) was used. Spectra were measured in the range 500–4000 cm^{-1} with practical resolution 2 cm^{-1}. System of thermostabilization (on the base of thermostat KRIO-VT-01) allowed fixing the temperature of the samples with 0.1 °C accuracy in the temperature range from −30 to 100 °C.

2.3. Determination of DNA concentration by Raman spectra

Figure 1 presents Raman spectra of the solution of sample 1 with different concentrations. DNA Raman bands are mainly situated in the low-frequency part of the spectrum—from 500 to 1600 cm^{-1}. The intense band with maximum near 3400 cm^{-1} is the valence band of water, the band with maximum near 1650 cm^{-1} is the bending band of water.

The most important information about DNA strands in the solutions is contained in the low-frequency part of the spectrum—from 500 to 1600 cm^{-1} (figure 2): bands of vibrations of DNA backbone (vibrations of phosphate groups near 1100 cm^{-1}, phosphodiester bonds near 840 cm^{-1}) and band-markers of nitrogenous bases (adenine near 760, 1260, 1300, 1350 cm^{-1}, cytosine near 800 cm^{-1}, guanine near 685 cm^{-1}, thymine near 740 cm^{-1}). The band near 840 cm^{-1} corresponds to the vibrations of a phosphodiester bond between deoxyribose of one nucleotide and phosphate group of another nucleotide. The presence of this band in the spectrum attests to the existence of DNA threads in the solution. Band-markers of

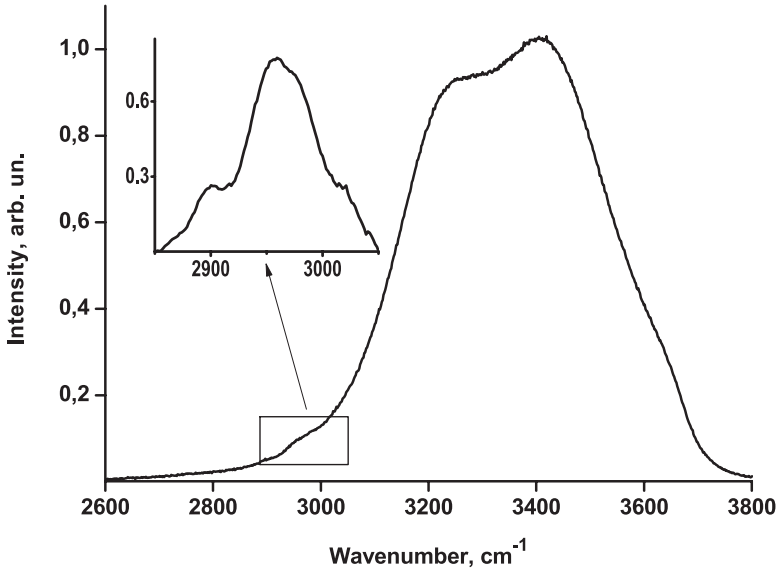

Figure 3. Raman spectrum of solution of sodium deoxyribonucleate (25 g l^{-1}), the band of CH-groups near 2950 cm^{-1} (difference spectrum) is shown in the inset.

bases provide the possibility to determine of concentration of nitrogenous bases in the solutions. By the position of a band it is possible to identify the bases and the intensity of bands allows determining the concentrations of individual bases.

The total concentration of DNA molecules can be determined by the intensity of the band near 1100 cm^{-1} corresponding to the vibrations of the phosphate groups of DNA backbone. Moreover, total concentration of DNA can be determined by the intensity of the band of CH-groups of deoxyribose contained in the DNA backbone. This band is situated near 2900 cm^{-1} and overlaps with the water Raman valence band (figure 3). However, this band can be extracted by use of difference spectra (after subtraction of spectrum of solvent from spectrum of solution, see inset in figure 3).

Accuracy of the determination of substance concentration by Raman spectra is defined in particular by the methods of spectra processing and by taking into account instrument factors. Such factors as instability of laser power and influence of elastic scattering worsen the quality of the obtained data and increase the error of determination of concentration.

In order to eliminate these factors, obtained spectra were processed in the following manner: the weak fluorescence background was subtracted, then the spectra were normalized by the area of water Raman valence band. Such normalization allows one to eliminate the influence of instability of the laser power on the result.

In order to eliminate the influence of elastic scattering (i.e. to extract the useful Raman signal against the background of a much stronger elastic scattering signal at the non-shifted frequency), the method of so-called *R*-representation was used [32, 33]. This method takes into account the Bose–Einstein distribution of molecules over quantum states at a given temperature. The following equation (1) gives the expression for the $R(\nu)$-representation [32, 33]:

$$R(\nu) = \frac{\nu * (1 - e^{-\frac{h\nu}{kT}})}{(\nu_0 - \nu)^4} * I(\nu). \qquad (1)$$

Here ν is the frequency shift of a Raman line, ν_0 is the frequency of exciting radiation, k is the Boltzmann constant, T is temperature, $I(\nu)$ is the experimental intensity. The useful signal of Raman scattering in the low-frequency region was extracted from the spectra of all solutions by the method of *R*-representation.

3. Results

3.1. Determination of concentration of individual oligonucleotides

For elaboration of the method of determination of concentration of individual nitrogenous bases, the dependence of Raman spectra of sample 2 on concentration was obtained in the range 0–9 g l^{-1} with increment 0.3 g l^{-1}. In figure 4, experimental Raman spectra of sample 2 with different concentrations in the range 500–1500 cm^{-1} after *R*-representation are presented.

For determination of the adenine concentration, bands of its vibrations near 1300–1400 cm^{-1} were used. For more correct extraction of these bands, the method of difference spectra was used (spectrum of solvent was subtracted from spectra of solutions).

In figure 5, the dependence of the integral intensity of bands-markers of adenine in the region 1300–1400 cm^{-1} on concentration is presented. This dependence can be approximated by a straight line. Accuracy of determination of the concentration of adenine by this linear dependence is 0.02 g l^{-1}, which is 0.2% of the maximum concentration of DNA in the solution.

Determination of the concentration of cytosine is more complicated. Its marker (a band near 800 cm^{-1}) strongly overlaps with the band of vibrations of the phosphodiester bond between nucleotides near 840 cm^{-1}. For extraction of the band of cytosine, spectra in this area (770–870 cm^{-1}) were decomposed into two Gaussian contours (figure 4) by the least-squares method.

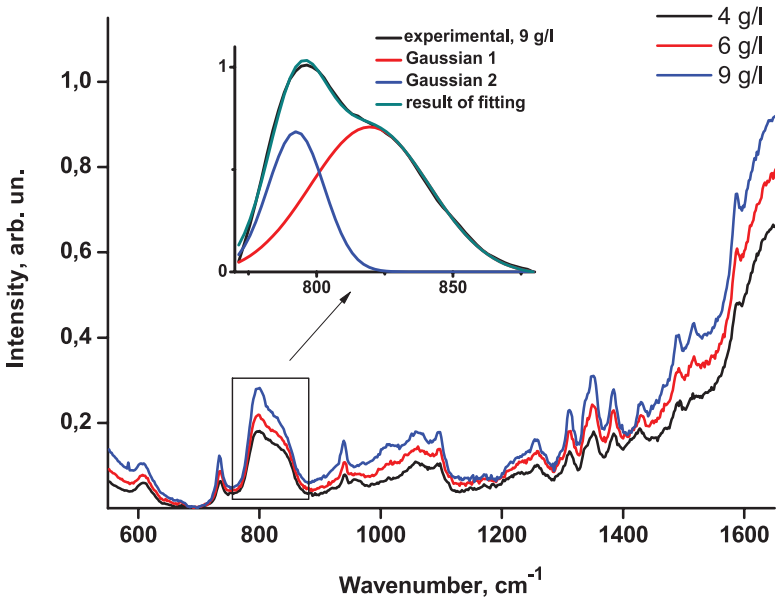

Figure 4. Raman spectra of sample 2 at different concentrations. Decomposition of spectra near 800 cm^{-1} is shown in the inset.

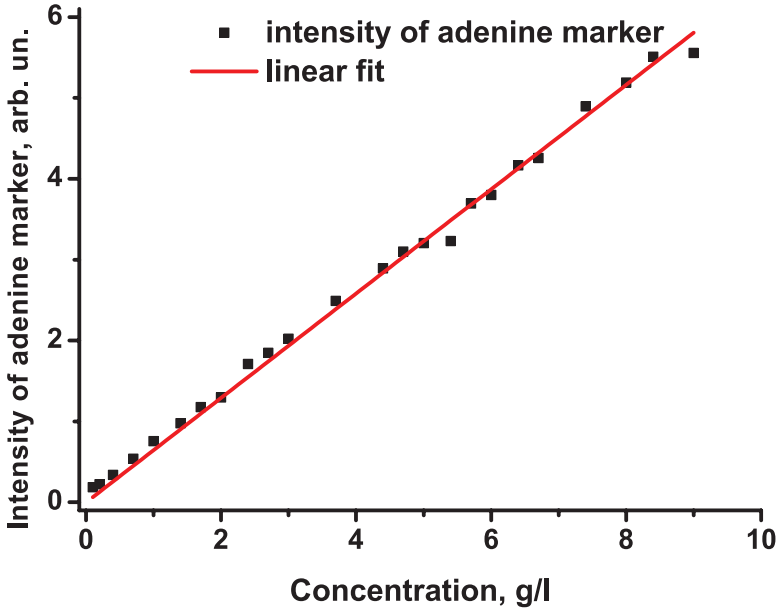

Figure 5. Dependence of intensity of adenine marker on concentration of sample 2.

For determination of the concentration of cytosine, the area under the contour with maximum at 795 cm^{-1} was used. The obtained dependence of integral intensity on concentration can be also approximated by a straight line. It allows determining concentration with accuracy 0.03 $g\,l^{-1}$, which is 0.3% of the maximum DNA concentration in the solution.

As was mentioned above, the losses of substance during molecular computing can reach 10%. So, the obtained accuracy of determination of concentration of adenine and cytosine is certainly sufficient for monitoring of concentration of the bases.

3.2. Determination of total DNA concentration

Total concentration of DNA in the solution can be determined by the integral intensity of bands corresponding to vibrations of phosphate groups near 1100 cm^{-1} and by the intensity of bands of CH-groups near 2900 cm^{-1}. To solve this problem, Raman spectra of DNA solutions were measured in the concentration range 0–9 $g\,l^{-1}$ with increment 0.4 $g\,l^{-1}$ for sample 2 (figure 4) and in concentration range 0–0.3 $g\,l^{-1}$ with increment 0.05 $g\,l^{-1}$ for sample 3.

Table 1. Accuracy of determination of DNA concentration using linear dependencies of integral intensity of different bands of the Raman spectrum on DNA concentration.

DNA sample	By band of phosphate groups		By valence band of CH-groups	
	($g\,l^{-1}$)	(%, rel. max. conc.)	($g\,l^{-1}$)	(%, rel. max. conc.)
Sample 2	0.044	0.5	0.04	0.4
Sample 3	0.015	5	0.01	3.3

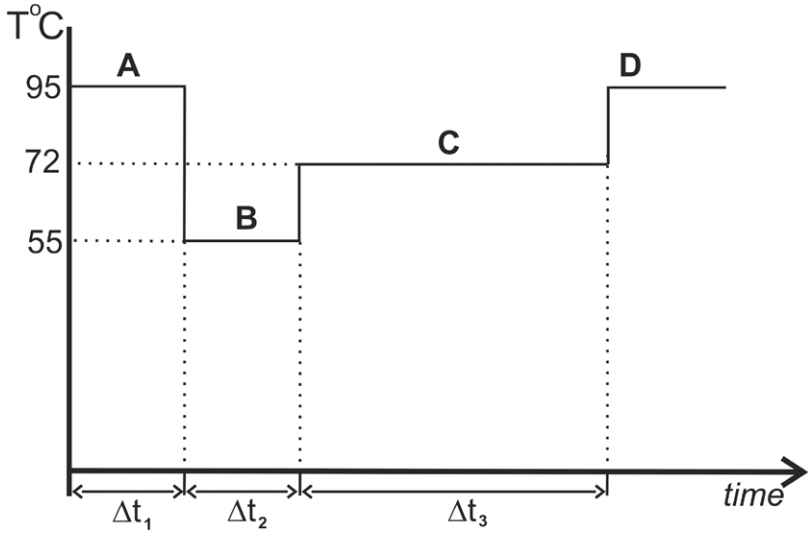

Figure 6. Typical temperature profile of the polymerase chain reaction.

For Raman bands of phosphate groups of both samples, the dependencies of integral intensity of difference spectra in the region 1050–1130 cm^{-1} on concentration of DNA were obtained. These dependencies were approximated with good accuracy by straight lines.

The obtained linear dependencies provided good accuracy of determination of total concentration of DNA for solutions of both samples. The error values are given in table 1.

It turned out that integral intensities of difference spectra of bands of vibrations of CH-groups in the range 2850–3000 cm^{-1} are also linearly dependent on the concentration of DNA for solutions of all samples. Accuracies of determination of DNA concentration using these linear dependencies are also presented in table 1.

As it can be seen from the presented results, better accuracy of determination of total concentration of DNA in the solutions is provided by using valence bands of vibrations of CH-groups, as they are more intense in comparison with bands of phosphate groups.

Our results obtained for solutions of sample 3 demonstrate that even at maximal concentration of 'actuation medium' in a molecular computer, 0.3 $g\,l^{-1}$ accuracy of determination of its concentration during molecular computing makes units of percent. It provides reliable control of losses of solutions of DNA calculations. In real experiments, concentration of DNA varies from units to tens of $g\,l^{-1}$ (see 2.1). Thus even the upper-bound estimate of errors demonstrates that the problem of determination of total DNA concentration in the solutions using Raman spectroscopy was solved with good accuracy.

3.3. Measurement and control of temperature of the DNA solutions during molecular computing

During molecular computing, measurement and control of temperature play a special role. The hydrogen bonding between complementary nitrogenous bases is the mechanism of connection of linear strands of DNA into the double helix. The hydrogen bonds are very sensitive to changes in temperature. This fact is used in molecular computations. Selecting the mode of temperature change, one can achieve the increase in efficiency of the reactions between DNA molecules, management of the processes of denaturation and renaturation. The temperature may vary from 0 to 95 °C.

As an example, the PCR (polymerase chain reaction) that is multiple selective copying of a particular stretch of DNA can be mentioned [34]. Figure 6 shows the typical temperature profile of this reaction, which has three stages. At the first stage (denaturation, plot A in figure 6), a solution with double strands of DNA is heated to the temperature of denaturation (usually to 94–96 °C during time period Δt_1 = 0.5–2 min); as the result, the DNA strands diverge. At the second step (annealing, plot B in figure 6), the temperature is reduced (usually, to 60–70 °C, depending on the length and nucleotide composition of the primer), so that the primers can get connected with a single strand. The time of this stage Δt_2 is about 0.5 min. At the third stage (elongation, plot C in figure 6), replication occurs in the strand and the primer acts as a seed. The temperature of elongation depends on the type of polymerase. The most commonly used polymerases *Taq* and *Pfu* have the maximum of their activity at the temperatures of 72 °C. The time of this stage of the reaction Δt_3 depends on the type of polymerase and the length of the copied object. Usually, the estimation of the time of elongation is 1 min per fragment length of 1000 pairs of bases. Thus, one cycle of polymerase chain reaction, consisting of three stages, may take several minutes, during which the temperature is changed by several tens of degrees. Then the cycle repeats (plot D in figure 6). Usually several tens of cycles are conducted. Thus, for all the reaction time (over an hour), the temperature is changed by tens of degrees multiple times. The violation of the temperature regime adversely affects the yield of the reaction.

In this regard, it becomes evident that the problem of temperature control in operations with DNA is extremely important. From the technical point of view, the situation is complicated by the fact that in some cases the volumes of material are very small—up to tens of microliters. This significantly limits the use of standard submersible temperature meters. The remote methods of sensing, in particular, Raman

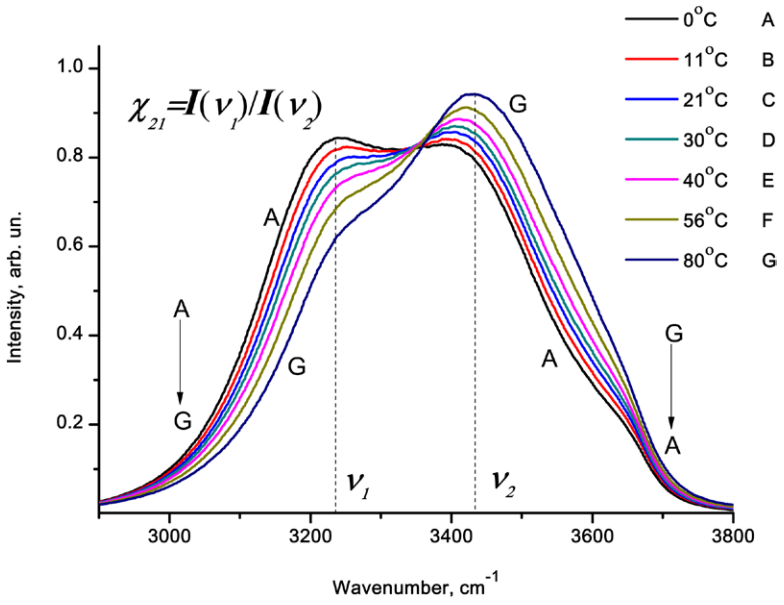

Figure 7. Temperature dependence of Raman valence band in the solutions of sample 1 of DNA.

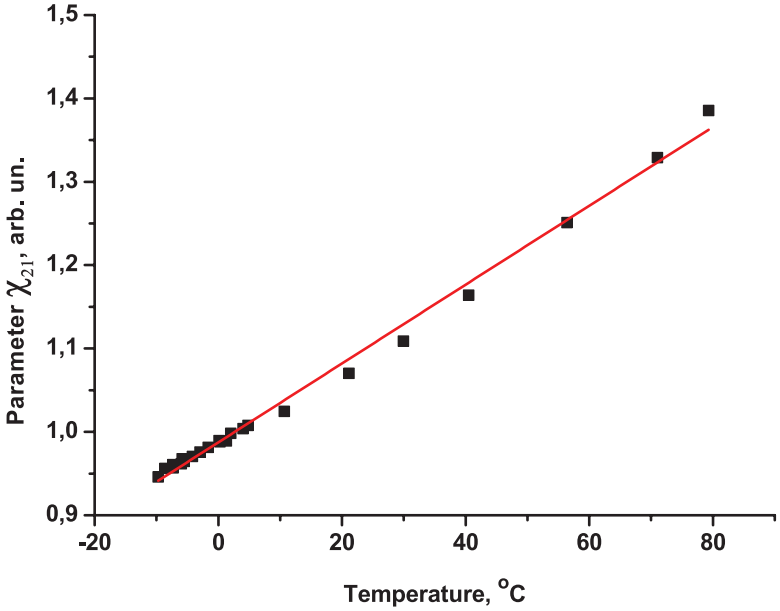

Figure 8. Dependence of the parameter χ_{21} on temperature.

spectroscopy, come to the foreground. The fact that in the process of molecular computing, in most cases, DNA molecules are in the aqueous solutions–buffers, allows one to use the dependence of Raman spectra of water (primarily the valence band of the spectrum) on temperature [21, 22]. This allows monitoring of the temperature of the solution in the mode of non-contact express sensing.

The method of measuring the temperature of water and aqueous solutions by the valence Raman band was developed by the authors a long time ago [21]. In this study, this method was adapted to measure the temperature of the 'actuation medium' of a molecular DNA computer. Figure 7 presents the temperature dependence of the Raman valence band of water in the solutions of sample 1 of DNA on temperature in the range of temperatures from 0 to 80 °C.

As can be seen, the Raman valence band is very sensitive to the change of temperature. With increasing temperature, the intensity of its high-frequency region increases and that of the low-frequency region decreases. This is caused by the changes of the structure of the hydrogen bond network under change of temperature and by reflection of these changes in vibrational spectra [21]. As a parameter characterizing the deformation of the valence band, the parameter $\chi_{21} = I_2/I_1$, which is equal to the ratio of intensities of high-frequency and low-frequency regions of the water valence band, was used [21]. The selection of specific values of the wavenumbers for determination of I_2 and I_1 is illustrated in figure 7. The values of the wavenumbers ν_2 and ν_1 are defined by specific points of the derivative of the Raman valence band. Figure 8 shows the dependence of the parameter χ_{21} on temperature.

Approximation of this dependence by a straight line allows one to determine the temperature of the DNA solution with an accuracy of 0.2 °C. This accuracy is satisfactory for solving the problem of temperature control in the process of molecular computing. It should be noted that the method of determining the temperature using the parameter χ_{21} does not require complex and time-consuming processing of the spectra and it does not require any special processing programs. So, this method can be easily used to monitor the 'actuation medium' of molecular DNA computing in real-time.

4. Conclusions

In the present study, it has been shown that the method of laser Raman spectroscopy provides the principal opportunity to determine the concentration of individual nitrogenous bases of DNA and the total concentration of DNA molecules in solutions by Raman spectra, and also to control the temperature of biochemical reactions in the process of molecular DNA computing.

(a) The upper-bound estimates show that the developed method provides determination of concentrations of specific nitrogenous bases with the accuracy not lower than 0.03 g l^{-1} (0.3 %) and the accuracy of determining the total concentration of DNA in solutions not lower than 0.02–0.04 g l^{-1} (3–5%). The obtained values of the accuracy of measuring the concentration of single-strand and double-strand DNA molecules of various lengths in the solution provide a precision monitoring of the concentration of DNA in the solutions during molecular computing.

(b) The method allows one to control the temperature of flow of chemical reactions in DNA solutions in the process of molecular computing with an accuracy better than 0.2 °C. This provides due control of the temperature, necessary to increase the speed and fidelity of solving the problems by molecular computing.

(c) The proposed method is non-invasive, remote and works in real time mode. It allows determining several parameters of 'actuation medium' in computational molecular devices simultaneously.

The obtained results allow one to use the developed method to monitor DNA solution states and to control errors in the process of molecular computing by DNA strands.

Acknowledgments

This work was supported by grant no.14-02-00710-a of the Russian Foundation for Basic Research.

All authors want to express their deep appreciation to Professor of the Physical Department of MSU S P Kulik for initiating research on this topic, for his permanent interest in this study, and for discussion of the results.

References

[1] Watson J and Crick F 1953 Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid *Nature* **171** 737–8

[2] Watson J 2003 Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid *Am. J. Psychiatry* **160** 623–4

[3] Bruce A 2007 *Molecular Biology of the Cell* 5th edn (New York: Garland Science)

[4] Alba M 2001 Replicative DNA polymerases *Genome Biol.* **2** 1

[5] Ogawa T and Okazaki T 1980 Discontinuous DNA replication *Annu. Rev. Biochem.* **49** 421–57

[6] Paun G, Rozenberg G and Salomaa A 1998 *DNA Computing: New Computing Paradigms* (Berlin: Springer)

[7] Adleman L M 1994 Molecular computation of solutions to combinatorial problems *Science* **266** 1021–4

[8] Adleman L 1998 Computing with DNA *Sci. Am.* **279** 54–61

[9] Lipton R J 1995 DNA solution of hard computational problems *Science* **268** 542–5

[10] Boneh D, Dunworth C and Lipton R 1995 Breaking DES using a molecular computer *Technical Report* CS-TR-489–95 (Princeton, NJ: Princeton University)

[11] Faulhammer D *et al* 2000 Molecular computation: RNA solutions to chess problems *Proc. Natl Acad. Sci.* **97** 1385

[12] Oliver J S 1999 Computation with DNA: matrix multiplication DIMACS series in discrete mathematics and theoretical computer science *Am. Math. Soc.* **44** 113–22

[13] Benenson Y *et al* 2001 Programmable and autonomous computing machine made of biomolecules *Nature* **414** 430–4

[14] Benenson Y *et al* 2004 An autonomous molecular computer for logical control of gene expression *Nature* **429** 423–9

[15] Cormen T, Leiserson C, Rivest R and Stein C 2001 *Introduction to Algorithms* 2nd edn (New Delhi: Prentice Hall of India)

[16] Kari L, Losseva E and Sosik P 2005 DNA computing and errors: a computer science perspective *Molecular Computational Models: Unconventional Approaches* (Hershey, PA: Idea Group Publishing)

[17] Benenson Y *et al* 2003 DNA molecule provides a computing machine with both data and fuel *Proc. Natl Acad. Sci. USA* **100** 2191–6

[18] Boneh D and Lipton R 1995 Making DNA computers error resistant *Technical Report* CS-TR-491–5 (Princeton, NJ: Princeton University)

[19] James K *et al* 1998 The fidelity of template-directed oligonucleotide ligation and its relevance to DNA computation *Nucl. Acid Res.* **26** 5203–11

[20] Faulhammer D, Lipton R and Landweber L 2000 Fidelity of enzymatic ligation for DNA computing *J. Comput. Biol.* **7** 839–48

[21] Dolenko T, Churina I, Fadeev V and Glushkov S 2000 Valence band of liquid water Raman scattering: some peculiarities and applications in the diagnostics of water media *J. Raman Spectrosc.* **31** 863–70

[22] Nikitin S, Manka C, Grun J and Bowles J 2012 A technique for contactless measurement of water temperature using Stokes and anti-Stokes comparative Raman spectroscopy *Rev. Sci. Instrum.* **83** 033105

[23] Burikov S, Dolenko T, Fadeev V and Sugonyaev A 2005 A new opportunities in the determination of inorganic compounds in water by the method of laser raman spectroscopy *Laser Phys.* **15** 1–5

[24] Gerdova I, Dolenko S, Dolenko T, Churina I and Fadeev V 2002 New opportunity solutions to inverse problems in laser spectroscopy involving artificial neural networks *Izvest. Akad. Nauk Ser. Fiz.* **66** 1116–24

[25] Dolenko S, Burikov S, Dolenko T, Efitorov A, Gushchin K and Persiantsev I 2014 Neural network approaches to solution of the inverse problem of identification and determination of partial concentrations of salts in multi-component water solutions *Artificial Neural Networks and Machine Learning* (*Lecture Notes in Computer Science* vol 8681) (Berlin: Springer) pp 805–12

[26] Thomas G 1999 Raman spectroscopy of protein and nucleic acid assembles *Annu. Rev. Biophys. Biomol. Struct*. **28** 1–27
[27] Prescott B, Steinmetz W and Thomas G 1984 Characterization of DNA structures by laser raman spectroscopy *Biopolymers* **23** 235–56
[28] Movileanu L, Benevides J and Thomas G 2002 Temperature dependence of the raman spectrum of DNA II raman signatures of premelting and melting transitions of Poly (dA)Poly(dT) and comparison with Poly(dA-dT)Poly(dA-dT) *Biopolymers* **63** 181–94
[29] Martin J and Wartell R 1982 Changes in raman vibrational bands of calf thymus DNA during the B-to-A transition *Biopolymers* **21** 499–12
[30] Lee J, Shin S, Park T and Zhang B 2004 Solving traveling salesman problems with DNA molecules encoding numerical values *Biosystems* **78** 39–47
[31] James K *et al* 1998 The fidelity of template-directed oligonucleotide ligation and its relevance to DNA computation *Nucl. Acid Res.* **26** 5203–11
[32] Brooker M, Nielsen O and Praestgaard E 1988 Assessment of correction procedures for reduction of Raman spectra *J. Raman Spectrosc.* **19** 71–8
[33] Murphy W *et al* 1989 Further assessment of reduction procedures for Raman spectra *J. Raman Spectrosc.* **20** 695–9
[34] Rudnickaya G and Evstrapov A 2008 Microarray device for polymerase chain reaction. Part 1: the basic principles of PCR; construction materials and microchips (review) *Nauchnoe Priborostroenie* **18** 3–20