ISSN 0096-3925, Moscow University Biological Sciences Bulletin, 2021, Vol. 76, No. 2, pp. 52–58. © Allerton Press, Inc., 2021.
Russian Text © The Author(s), 2021, published in Vestnik Moskovskogo Universiteta, Seriya 16: Biologiya, 2021, Vol. 76, No. 2, pp. 67–75.

RESEARCH ARTICLE

Design of Nucleic Acid Biosensors Based on CRISPR/Cas Systems and Reporter Split Proteins

R. V. Novikov[a,b,*,#], A. K. Gribkova[a,b,,#], J. G. Kacher[a,b,***,#], P. A. Zaytsev[a,b,****], G. A. Armeev[a,b,*****], G. S. Gluhov[a,b,******], and A. K. Shaytan[a,b,c,*******] (ORCID: 0000-0003-0312-938X)**

[a] *Department of Biology, Moscow State University, Moscow, 119234 Russia*
[b] *Sirius University of Science and Technology, Sochi, 354340 Russia*
[c] *Higher School of Economics University, Moscow, 109028 Russia*
**e-mail: novikov.roman.2017@post.bio.msu.ru*
***e-mail: gribkova.anna.2013@post.bio.msu.ru*
****e-mail: kacher.julia.2015@post.bio.msu.ru*
*****e-mail: zaitsev@mail.bio.msu.ru*
******e-mail: armeev@mail.bio.msu.ru*
*******e-mail: gluhovg@gmail.com*
********e-mail: shaytan_ak@mail.bio.msu.ru*

Received February 11, 2021; revised March 28, 2021; accepted April 12, 2021

Abstract—Highly sensitive, specific, rapid, and easy-to-use diagnostic methods for the detection of nucleic acids of pathogens are required for the diagnosis of many human, animal, and plant diseases and environmental monitoring. The approaches based on the use of the natural ability of bacterial CRISPR/Cas9 systems to recognize DNA sequences with a high specificity under isothermal conditions are an alternative to the polymerase chain reaction method, which requires expensive laboratory equipment. The development of the methods for signal registration with the formation of a DNA/RNA/Cas9 protein complex is a separate bioengineering task. In this work, a design was developed and the applicability of a biosensor system based on the binding of two dCas9 proteins with target DNA sequences (without their cutting) and detection of their colocalization using reporter systems based on split enzymes was studied. Using the methods of molecular modeling, possible mutual positions of two dCas9 proteins at a detectable locus of genomic DNA, allowing the split enzyme domains attached to them to interact in an optimal way, were determined. The optimal distances on DNA between binding sites of dCas9 proteins in different orientations were determined, and the dependence of the complex structure on the distance between the binding sites of dCas9 proteins was modeled. Using the methods of bioinformatics, the genomes of a number of viruses (including SARS-CoV-2) were analyzed, and the presence of genomic loci unique to the species, allowing the possibility of landing pairs of dCas9 proteins in optimal positions, was demonstrated. The possibility of a combined use of dCas9 proteins from different bacteria to expand the spectrum of detected loci was analyzed. The results of the work indicate a fundamental possibility of the creation of highly specific nucleic acid biosensors based on a combination of CRISPR/Cas9 technologies and split enzymes.

Keywords: biosensors, CRISPR/Cas, nucleic acids, split proteins, beta-lactamase, molecular modeling, genomic analysis

DOI: 10.3103/S0096392521020036

INTRODUCTION

The methods for the detection of nucleic acids with a given nucleotide sequence play a crucial role in the diagnosis of human and animal diseases, environmental monitoring, genotyping, establishment of kinship, etc. An ideal test system for detecting nucleic acids should have a high sensitivity, specificity, and low cost of analysis and allow for detection as soon as possible without the need for expensive equipment. To date, diagnostic test systems based on the use of polymerase chain reaction (PCR) are most widespread. Theoretically, PCR allows one to register individual molecules in the reaction mixture; however, the sensitivity limit of PCR-based test systems used in practice strongly depends on the protocol used and can reach hundreds of thousands of DNA molecules per milliliter [1]. The PCR method requires expensive equipment, which allows one to cyclically change the temperature of a small volume of the reaction mixture at clearly defined intervals as well as to register the fluorescence signals.

[#] Equal contribution.

Thus, PCR analyses are possible only in specialized laboratories, which leads to a significant time interval between the collection of biomaterial and getting the result (usually at least a day). These limitations are a stimulus for the development and creation of more advanced diagnostic test systems that can, among other things, be used for a rapid diagnosis directly at the points of medical care and in other situations with limited access to laboratory infrastructure or requiring a fast receipt of the results of the analysis. The need for such systems has become especially relevant in connection with the pandemic of SARS-CoV-2 coronavirus [1]. The key directions in solving this problem are the following: (1) transition to "isothermal" approaches, when it is just enough to maintain the reaction mixture, in which a key stage of nucleic acid detection is carried out, at a constant temperature; (2) the use of simple visual signal detection methods (for example, by a change in the color of the solution or the appearance of colored stripes similar to immunochromatographic test systems). The loop-mediated isothermal amplification (LAMP) method, based on the use of Bst polymerase, which is able to conduct amplification at a constant temperature, can be given as a promising isothermal approach already used in practice. This approach usually allows one to obtain a result within 60 min, does not require the use of expensive equipment, and it is also more resistant to reaction inhibitors than PCR; however, it requires a thermostat with a sufficiently high temperature (~65°C) and involves a rather complicated primer design procedure, which makes it difficult to select them for a specific task [2].

Along with detection methods based only on nucleic acid amplification, fundamentally different approaches, based on the use of CRISPR-Cas systems and their ability to bind specifically to DNA/RNA sequences in a programmable manner, began to be developed not so long ago. Such approaches were called CRISPR-based diagnostics (CRISPR-Dx). CRISPR-Cas systems are the adaptive immune system of bacteria destroying alien genetic elements entering the cell [3]. With a primary infection, a number of proteins of the system captures the genetic information of the pathogen and writes it to the CRISPR locus of the bacterial genome in the form of so-called spacer sequences. The type II class 2 CRISPR-Cas systems, in which the effector Cas9 nuclease (particularly, SpCas9 of *Streptococcus pyogenes* bacterium) plays a key role, are the most studied. Cas9 forms a complex with two RNA, including crRNA, which carries information about the spacer sequence, and small auxiliary tracrRNA. This complex binds specifically to DNA sequences due to the formation of heteroduplex between the pathogen DNA region (called protospacer) and the complementary region of crRNA spacer. The presence of a fixed DNA motif (PAM sequence) recognized by Cas9 protein itself is also critical for binding to the target. PAM sequences in Cas9 systems from different bacteria vary in the composition and length (approximately from three to eight nucleotides). After binding, the nuclease cuts the target DNA. For editing genomes, artificial variants of Cas9 systems, where crRNA and tracrRNA are combined into a single guide RNA (gRNA), are widely used. Variations of the class 2 CRISPR-Cas systems, for example, based on Cas12, Cas13, or Cas14 type nucleases can have the ability to bind single-stranded DNA or RNA as well as have so-called collateral activity, when the enzyme begins to cut any RNA or DNA in the solution upon binding to the target. The latter property underlies the currently developed nucleic acid detection systems, such as SHERLOCK, DETECTR, and those related to them. The Cas13 nuclease, which recognizes single-stranded RNA fragments, is used in the SHERLOCK system. Thus activated collateral enzyme activity in relation to RNA is measured using a cleavage of signaling molecules that consist of a fluorophore and a quencher bound by a synthetic RNA molecule [4]. There are also variants for the adaptation of a signal reading system for the use of immunochromatographic strips [5]. The DETECTR system uses DNA-specific Cas12 and cas14 nucleases that bind to double-stranded or single-stranded DNA molecules, respectively. The above-described approaches and those similar to them usually assume a step of preamplification using one of isothermal methods, for example, LAMP or RPA [2, 6]. The sensitivity of resulting methods is comparable to the best variants of PCR analysis (detection of single nucleic acid molecules per milliliter) [4]. A short detection time (from 10 min) can be attributed to the advantages of the method. At the moment, the complexity of obtaining the system components, the absence of ready-made commercially available test systems, and the need to use RNA/DNA modified with fluorophores are limitations of the method.

An alternative approach to the use of CRISPR-Cas systems for the detection consists in the use of their properties to bind to target DNA sequences (binding-based biosensing) [7]. By introducing point mutations into the Cas9 protein, a dCas9 protein (D10A and H841A substitutions for SpCas9), which still binds to DNA but does not cut it, can be obtained [8]. Binding of Cas9-gRNA complexes to DNA is highly sensitive to the presence of even single nucleotide mismatches between the sequences of the spacer and protospacer [9]. Such sensitivity opens up prospects for the creation of sensors capable of genotyping organisms and viruses and detection of somatic mutations.

In this work, we studied the question about the optimal construction of test systems based on the binding of dCas9-gRNA complexes with the target DNA locus. As a principle of signal detection, a scheme based on the binding of the fragments of split enzymes linked to dCas9 proteins was considered. Using modeling methods, optimal arrangement of the pairs of dCas9-proteins and lengths of peptide linkers

for the attachment of split fragments were detected. Using the methods of genomic analysis, we demonstrated that the proposed construction of the biosensor allows it to bind specifically to genomic loci of a number of viruses, including to distinguish between the haplotypes of viruses that differ in single mutations.

MATERIALS AND METHODS

Molecular modeling. Atomistic molecular models of binding of two dCas9 proteins to DNA were created using Python scripts for the UCSF Chimera program [10]. The structure SpCas9/gRNA/DNA with PDB ID 5Y36 was selected for modeling the dCas9/gRNA/DNA complex [11]. The structure of ideal DNA in B-form was used for modeling the additional DNA between two dCas9 complexes. To connect it with the models of dCas9 complexes, a structural alignment of two terminal nucleotides of connective DNA with two terminal nucleotides of DNA duplex in the structure of dCas9 complex was carried out. The alignment was carried out using the method of minimizing the standard deviation of the positions of the following nucleotide atoms: O5', C5', C4', O4', C3', O3', C2', C1'. To measure the distances and angles in the resulting structures, MDAnalysis and NumPy packages were used [12, 13]. For measuring the optimal location for the ends of the polypeptide chain of dCas9 proteins relative to DNA axis, we calculated a dihedral angle α formed by the following points: geometric centers of N- and C-terminal amino acid residues of two dCas9 proteins and geometric centers of two central C1' atoms of DNA separating dCas9 proteins. The construction of graphs was carried out using a Matplotlib library [14].

Genomic analysis. A Python program code was realized to search for detectable loci (targets) in selected genomes or genes for different combinations of dCas9 proteins from different organisms and different combinations of mutual orientation of dCas9 proteins. dCas9 proteins of *Streptococcus pyogenes* (SpdCas9), *Staphylococcus aureus* (SadCas9), *Campylobacter jejuni* (CjdCas9), and *Streptococcus thermophilus* (StdCas9) were used; they differ in the length and sequence of bound PAM sites (NGG, NNGRRT, NNNNAYAC, NNAGAAW, respectively, where N is any nucleotide, R is A or G, Y is C or T, W is A or T) as well as in the optimal length of a protospacer recognized by dCas9 protein (20, 21, 22 or 20 nucleotides, respectively). Nucleotide sequences were downloaded from the NCBI RefSeq database [15]. A search for targets was carried out in annotated protein-coding regions (including the regions annotated as hypothetical proteins) using regular expressions. To analyze RNA viruses, a complementary double-stranded DNA was used. To search for targets, restrictions were imposed on possible distances between the binding sites of dCas9 proteins along DNA corresponding to optimal variants of mutual location of dCas9 proteins determined during molecular modeling. For each found pair of PAM sites, the appropriate sequences of protospacers were determined. In order to analyze the number of potential targets and their specificity for the discussed test systems, the complete genomes of the following virus species were selected: SARS-CoV-2 isolate Wuhan-Hu-1 (NC_045512.2, size 29903 bp), SARS-CoV (NC_004718.3, size 29751 bp), and MERS-CoV (NC_019843.3, size 30119 bp). A search for possible targets in the genes was carried out in the genes N (1259 bp) and E (227 bp) of SARS-CoV-2 virus. Demonstration of a biosensor's applicability for the detection of single nucleotide polymorphisms was carried out on the example of 13 nonsynonymous point mutations that distinguish the haplotypes of SARS-CoV-2 virus (UK variant VOC-202012/01 and isolate Wuhan-Hu-1) in the genes ORF1ab (C3267T, C5388A, T6954C), Spike (A23063T, C23271A, C23604A, C23709T, T24506G, G24914C), ORF8 (C27972T, G28048T, A28111G), and N (C28977T). For a global alignment of genomes, MAFFT [16] with basic settings was used.

RESULTS AND DISCUSSION

General principle of biosensor operation. When developing biosensors based on specific binding of proteins to nucleic acids, a conjugation of this binding with the generation of some detectable signal from the test system is a key question. Such signal generation can be realized by a simultaneous binding of two CRISPR-dCas9 complexes with adjacent, spatially close regions of a DNA locus. In this case, dCas9 proteins can be used as carriers for a spatial convergence of different pairs of reporter protein fragments or domains prelinked to dCas9 proteins. So-called split enzymes are of a special interest [17]. Split enzymes are the enzymes artificially divided into two protein fragments that can be spontaneously combined into a whole functional enzyme. For example, firefly luciferase and bacterial β-lactamase were adapted to work as split enzymes [18, 19]. β-lactamase, as a split enzyme, has a number of advantages: small size (21 kDa), independence from minor fluctuations of pH, possibility of colorimetric signal detection as a result of cleavage of a chromogenic substrate (for example, nitrocefin), and signal amplification due to the accumulation of reaction products over time. Taking into account the above, a principal scheme of the biosensor based on dCas9 proteins bound to the fragments of β-lactamase is given in Fig. 1a. When designing a specific biosensor construction, a number of fundamental questions arise. First, it is necessary to decide on the optimal mutual orientation of dCas9 complexes relative to each other. This location is determined by (1) the variants of choosing DNA strands with which the complexes will bind (PAM-in, PAM-out, or PAM-direct, Fig. 1a), (2) the distance along DNA between the binding sites of the complexes, (3) mutual orientation

of split enzymes' attachment points relative to the DNA axis. Second, it is necessary to decide on the optimal attachment of the split system domains (to C- or N-terminus of each dCas9 protein). Third, it is necessary to decide on the optimal length of a peptide linker connecting split domains with dCas9 proteins. Finally, since binding of dCas9 proteins with specific loci is limited by the presence of the appropriate PAM sites at the binding sites, the question of studying the presence of pairs of sites suitable for landing in the genomes of detected organisms is important. A difference in the sequence and length of PAM sites in dCas9 proteins from different species of bacteria opens up the possibilities for combining different dCas9 proteins in order to optimize the number of potential landing sites, on the one hand, and specificity of the biosensor, on the other hand. The system design, taking into account the above considerations and limitations, is a multiparameter problem, the study of which the following sections are devoted to.

Molecular modeling of a biosensor based on two dCas9 proteins. When designing a biosensor based on the fusion system of dCas9 proteins with the parts of the split enzyme, determination of the mutual orientation of dCas9 proteins as a part of the nucleoprotein complex and the distance between them is an important step since these parameters directly affect the efficiency of the biosensor. We carried out such an analysis by molecular modeling of the systems of two dCas9 complexes in three mutual orientations (Fig. 1a). It is important to note that there is a change in the distance between terminal amino acid residues of dCas9 proteins (to which split domains are attached), as well as a change in their orientation location relative to DNA characterized by the angle α (see Materials and Methods), when changing the distance between dCas9 complexes. Along with the mutual orientation of the proteins, these parameters affect the efficiency of the biosensor: when choosing a complex with a large value of the distance between the attachment points of split fragments, there is a possibility that the parts of the split enzyme will not combine, and such a combination can be prevented by DNA separating the complexes if the selected complex has a large value of the angle between the residues. Based on modeling, a dependence of the distance between N- and C-terminal residues of dCas9 proteins (to which the parts of the split enzyme can be attached), as well as α angle, on the distance between PAM sites was obtained (Fig. 1b). Based on these data, optimal variants of the attachment of the split enzyme parts were determined, namely, C termini of dCas9 proteins for PAM-in, N termini of dCas9 proteins for PAM-out orientation, and C terminus of dCas9 protein located at 5'-terminus of recognizable locus and N-terminus of dCas9 protein located at 3'-terminus of recognizable locus (the locus orientation is determined along the DNA strand forming a heteroduplex) for PAM-direct orientation. The optimal models with the lowest values of distances between the attachment points of split domains and α angles without steric overlap were also determined for each of the above systems. The following variants are the most optimal: a model with the distance of 41 nucleotides between PAM sites for the orientation PAM-direct, 29 nucleotides for PAM-in, and 52 nucleotides for PAM-out (Fig. 1c). An extended set of seven optimal distances between PAM sites corresponding to minimal distances between the attachment points of split domains (for PAM-direct orientation, the cut-off threshold was 70 Å) at the values of α angle less than 72 degrees was also selected for the genomic analysis. With subsequent genomic analysis, the results obtained were generalized for dCas9 proteins from different organisms assuming that changes in the system depending on the variant of dCas9 protein will be insignificant.

To create fusion systems, the use of a linker connecting dCas9 proteins with the parts of the split enzyme is a prerequisite. A glycine–serine linker $(GGGGS)_n$ is widely used to create such systems [20]. There is a bilateral limitation when choosing the linker length: a too short linker can interfere with the combination of split domains; at the same time, a too long linker will lead to the fact that the probability of the interaction of domains will decrease. Based on minimal distances between the attachment points of linkers obtained in modeling (R_{min}: PAM-direct, 37 Å; PAM-in, 26 Å; PAM-out, 26 Å), we estimated the optimal length of linkers. The distance between terminal amino acid residues of beta-lactamase is insignificant (8 Å); therefore, it was not taken into account in the calculations. According to the simplest models of polymer physics, the probability distribution for the distance between N- and C-terminal amino acid residues of the linker has a Gaussian form $P_N(R) = (2\pi\langle R^2\rangle/3)^{-3/2}\exp(-3R^2/2\langle R^2\rangle)$, where the width of this distribution $\langle R^2\rangle = ll_k$ (l, contour length of amino acid residue; l_k, Kuhn segment length). The contour length of one amino acid is 3.4 Å; Kuhn segment for peptides, 8.4 Å [21]. Based on these estimations, the linker length of ten amino acids is reasonable, since it provides the probability of its being in a state with $R > R_{min}/2$ in more than 50% of conformations. With an increase in the distance between the attachment points of split domains to 70 Å, the probability is still significant (approximately 10%).

Analysis of the presence of potential targets in virus genomes. We developed a program code to check the presence of potential targets in detectable genomes, genes, and mutation sites for the developed biosensor out of two dCas9 proteins (see Materials and Methods). The variants of a molecular test system consisting of the combinations of dCas9 proteins were considered: SpdCas9, StdCas9, CjdCas9, and SadCas9 located on the target in the orientations PAM-out, PAM-in, and PAM-direct (a total of 64 possible variants). A search for possible targets was carried out in

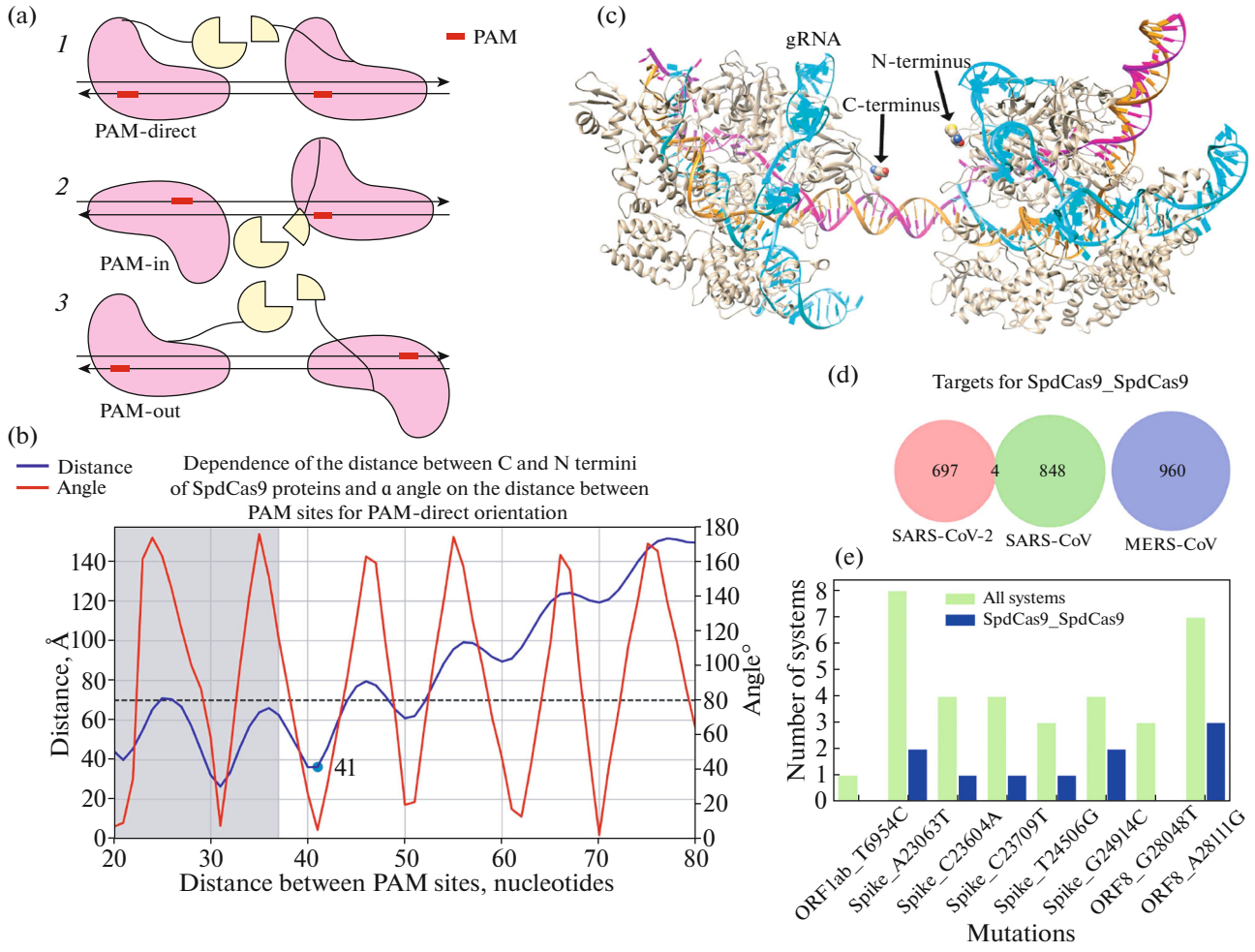


Fig. 1. (a) Scheme of possible variants of mutual location of two dCas9 proteins with beta-lactamase fragments linked to them on DNA locus as a part of the considered biosensor; PAM motifs are highlighted in red; 5'-3' direction of DNA strands is indicated by arrows. (b) Graph of the dependence of the distance between C and N termini of SpdCas9 proteins in PAM-direct orientation, as well as α angle, on the distance between PAM sites; sterically inaccessible region is highlighted in gray; optimal configuration is marked with a dot; threshold distance of 70 Å is demonstrated by a dotted line. (c) Structure of two SpdCas9 proteins as part of a nucleoprotein complex in PAM-direct orientation with the distance between PAM sites of 41 nucleotides; C and N termini of SpdCas9 proteins are highlighted by spheres. (d) Venn diagram for found potential targets in the genomes of coronaviruses for the systems out of two SpdCas9 proteins. (e) Histogram of the number of possible systems for the detection of nonsynonymous single nucleotide substitutions differing in SARS-CoV-2 haplotypes (UK variant VOC-202012/01 and SARS-CoV-2 isolate Wuhan-Hu-1); dark color, number of systems of two SpdCas9 proteins; light color, number of all possible systems.

the genomes of coronaviruses. The number of targets was 697, 848, and 960 for SARS-CoV-2, SARS-CoV, and MERS-CoV. Out of them, only four targets were nonunique and were found both in the genome of SARS-CoV-2 and in the genome of SARS-CoV (Fig. 1d), which indicates a possibility of designing highly specific biosensors for specific viruses. The revealed diversity of targets unique for the virus provides opportunities for a choice of the most optimal and well-binding gRNA when developing specific test systems. We also studied the question about the possibility of using only a number of genes that are already used in the existing test systems as targets. A search for viral targets was carried out in the genes E and N of SARS-CoV-2. A maximal number of targets in the gene N was found for the system realized on the basis of two SpdCas9 proteins and is 59. For the whole genome of SARS-CoV-2, there are 704 possible targets in this system. Two potential targets were revealed for the same system for relatively small gene E. However, it is possible to increase the number of targets by expanding the spectrum of used dCas9 proteins, and four targets are already possible in the gene E for the systems from SpdCas9 and StdCas9. A prevailing number of targets was found for the systems with SpdCas9, SadCas9, and CjdCas9, which is due to less limitations on the sequence of PAM sites. No possible targets were found for the system StdCas9-StdCas9 in the orientation PAM-in in the genome of SARS-CoV-2.

The emergence of new variants of viruses (haplotypes) with single mutations emphasizes the need to develop molecular instrumentation with a single nucleotide specificity able to distinguish one variant from another. We carried out a search for specific targets in mutation sites of the variant SARS-CoV-2 UK VOC-202012/01. The results of the analysis of the number of possible variants of test systems (consisting only of SpdCas9 proteins or all possible pairs of dCas9 proteins) for landing on mutation sites are given in Fig. 1e. For eight of 13 sites of single nucleotide nonsynonymous substitutions, it is possible to select targets for the detection using at least one of the considered test systems, and four possible variants of the detection system were found for the substitution A23063T (N501Y) in the spike protein, which leads to better protein binding to cellular receptors [22].

Thus, as a result of modeling and bioinformatic analysis, we demonstrated a fundamental possibility of the design of biosensors for the detection of nucleic acids of pathogens based on simultaneous binding of two CRISPR-dCas9 complexes to target DNA locus. At the same time, signal detection in the proposed construction occurs due to the interaction of domains of the reporter system linked to dCas9 proteins. This kind of biosensors seem to be promising, since they are easily programmed to bind to given DNA loci by specifying gRNA of the required sequence. The proposed approach refers to a developing wide spectrum of different approaches to the detection of nucleic acids using the ability of Cas proteins to recognize DNA or RNA sequences in a programmable way [23]. At the moment, SHERLOCK and DETECTR systems that use a collateral activity of Cas12, Cas13, and Cas14 enzymes reached the highest degree of practical elaboration [5, 24]. However, there are also developments associated with the use of enzymatic activity of Cas9 proteins or their variations, for example, CRISDA [25], RACE [26], and CASLFA [27]. In comparison with these systems, the approach studied in this article has a number of functional peculiarities. Thus, the enzymatic activity of Cas proteins is usually more sensitive to the presence of discrepancies between the spacer sequence and bound DNA region than their ability to bind to a given locus. Thus, in our variant of the detection system, possible nontarget effects at the level of binding of one dCas9 protein will be more pronounced. However, this decrease in the specificity can be compensated by the need for simultaneous binding of two dCas9 proteins in neighboring places. Their interaction through reporter split domains will lead to the presence of a cooperative effect in the binding of two dCas9 proteins, which, in turn, will improve the signal to noise ratio when registering the final signal. In theory, the use of pairs of high-precision Cas9 proteins [28] will also allow for additionally increasing the accuracy of this system. The use of colorimetric reaction in the proposed system worsens the detection limit as compared with the systems in which fluorescent probes are used but provides a number of advantages by reducing the cost of the system due to the absence of probes and the need to use a complex equipment. The absence of synthetic components in the biosensor construction also opens up the opportunities for its integration into regulatory genetic schemes realized inside living organisms. A high specificity of binding of Cas9 proteins opens up prospects for the discovery of single nucleotide substitutions, which is an urgent task in the detection of viral haplotypes (for example, SARS-Cov-2 variants).

ACKNOWLEDGMENTS

This work was carried out using the equipment of the Center for Collective Use of Ultra-High Performance Computing Resources of Moscow State University.

FUNDING

This work was supported by the Russian Foundation for Basic Research (project no. 19-34-51053).

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflict of interests.

This article does not contain any studies involving human participants performed by any of the authors.

REFERENCES

1. MacKay, M.J., Hooker, A.C., Afshinnekoo, E., et al., The COVID-19 XPRIZE and the need for scalable, fast, and widespread testing: 9, *Nat. Biotechnol.,* 2020, vol. 38, no. 9, pp. 1021–1024.
2. Notomi, T., Loop-mediated isothermal amplification of DNA, *Nucleic Acid Res.,* 2000, vol. 28, no. 12, e63.
3. Makarova, K.S., Wolf, Y.I., Iranzo, J., et al., Evolutionary classification of CRISPR-Cas systems: A burst of class 2 and derived variants, *Nat. Rev. Microbiol.,* 2020, vol. 18, no. 2, pp. 67–83.
4. Kellner, M.J., Koob, J., Gootenberg, J.S., Abudayyeh, O.O., and Zhang, F., SHERLOCK: Nucleic acid detection with CRISPR nucleases, *Nat. Protoc.,* 2019, vol. 14, no. 10, pp. 2986–3012.
5. Gootenberg, J.S., Abudayyeh, O.O., Kellner, M.J., Joung, J., Collins, J.J., and Zhang, F., Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6, *Science,* 2018, vol. 360, no. 6387, pp. 439–444.
6. Lobato, I.M. and O'Sullivan, C.K., Recombinase polymerase amplification: Basics, applications and recent advances, *TrAC Trends Anal. Chem.,* 2018, vol. 98, pp. 19–35.
7. Li, Y., Li, S., Wang, J., and Liu, G., CRISPR/Cas systems towards next-generation biosensing, *Trends Biotechnol.,* 2019, vol. 37, no. 7, pp. 730–743.
8. Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A., Repur-

posing CRISPR as an RNA-guided platform for sequence-specific control of gene expression, *Cell,* 2013, vol. 152, no. 5, pp. 1173–1183.

9. Zhang, L., Rube, H.T., Bussemaker, H.J., and Pufall, M.A., The effect of sequence mismatches on binding affinity and endonuclease activity are decoupled throughout the Cas9 binding site, *bioRxiv,* 2017. https://doi.org/10.1101/176255
10. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E., UCSF Chimera—A visualization system for exploratory research and analysis, *J. Comput. Chem.,* 2004, vol. 25, no. 13, pp. 1605–1612.
11. Huai, C., Li, G., Yao, R., Zhang, Y., Cao, M., Kong, L., Jia, C., Yuan, H., Chen, H., Lu, D., and Huang, Q., Structural insights into DNA cleavage activation of CRISPR-Cas9 system: 1, *Nat. Commun.,* 2017, vol. 8, no. 1, 1375.
12. Gowers, R., Linke, M., Barnoud, J., Reddy, T., Melo, M., Seyler, S., Domanski, J., Dotson, D., Buchoux, S., Kenney, I., and Beckstein, O., MDAnalysis: A Python package for the rapid analysis of molecular dynamics simulations, *Proceedings of the 15th Python in Science Conference,* Austin: SciPy, 2016, pp. 98–105.
13. Harris, C.R., Millman, K.J., van der Walt, S.J., et al., Array programming with NumPy: 7825, *Nature,* 2020, vol. 585, no. 7825, pp. 357–362.
14. Hunter, J.D., Matplotlib: A 2d graphics environment, *Comput. Sci. Eng.,* 2007, vol. 9, no. 3, pp. 90–95.
15. Pruitt, K.D., Tatusova, T., and Maglott, D.R., NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins, *Nucleic Acid Res.,* 2007, vol. 35, pp. D61–D65.
16. Katoh, K., Rozewicki, J., and Yamada, K.D., MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization, *Brief Bioinf.,* 2019, vol. 20, no. 4, pp. 1160–1166.
17. Wehr, M.C. and Rossner, M.J., Split protein biosensor assays in molecular pharmacological studies, *Drug Discov. Today,* 2016, vol. 21, no. 3, pp. 415–429.
18. Azad, T., Tashakor, A., and Hosseinkhani, S., Split-luciferase complementary assay: Applications, recent developments, and future perspectives, *Anal. Bioanal. Chem.,* 2014, vol. 406, no. 23, pp. 5541–5560.
19. Porter, J.R., Stains, C.I., Segal, D.J., and Ghosh, I., Split beta-lactamase sensor for the sequence-specific detection of DNA methylation, *Anal. Chem.,* 2007, vol. 79, no. 17, pp. 6702–6708.
20. Chen, X., Zaro, J., and Shen, W.-C., Fusion protein linkers: Property, design and functionality, *Adv. Drug Deliv. Rev.,* 2013, vol. 65, no. 10, pp. 1357–1369.
21. Soranno, A., Longhi, R., Bellini, T., and Buscaglia, M., Kinetics of contact formation and end-to-end distance distributions of swollen disordered peptides, *Biophys. J.,* 2009, vol. 96, no. 4, pp. 1515–1528.
22. Zhao, S., Lou, J., Cao, L., Zheng, H., Chong, M.K.C., Chen, Z., Chan, R.W.Y., Zee, B.C.Y., Chan, P.K.S., and Wang, M.H., Quantifying the transmission advantage associated with N501Y substitution of SARS-CoV-2 in the United Kingdom: An early data-driven analysis, *J. Travel Med.,* 2021, vol. 28, no. 2, taab011.
23. Wang, M., Zhang, R., and Li, J., CRISPR/Cas systems redefine nucleic acid detection: Principles and methods, *Biosens. Bioelectron.,* 2020, vol. 165, 112430.
24. Broughton, J.P., Deng, X., Yu, G., et al., CRISPR-Cas12-based detection of SARS-CoV-2, *Nat. Biotechnol.,* 2020, vol. 38, no. 7, pp. 870–874.
25. Zhou, W., Hu, L., Ying, L., Zhao, Z., Chu, P.K., and Yu, X.-F., A CRISPR-Cas9-triggered strand displacement amplification method for ultrasensitive DNA detection: 1, *Nat. Commun.,* 2018, vol. 9, no. 1, 5012.
26. Wang, R., Zhao, X., Chen, X., Qiu, X., Qing, G., Zhang, H., Zhang, L., Hu, X., He, Z., Zhong, D., Wang, Y., and Luo, Y., Rolling circular amplification (RCA)-assisted CRISPR/Cas9 cleavage (RACE) for highly specific detection of multiple extracellular vesicle MicroRNAs, *Anal. Chem.,* 2020, vol. 92, no. 2, pp. 2176–2185.
27. Wang, X., Xiong, E., Tian, T., Cheng, M., Lin, W., Wang, H., Zhang, G., Sun, J., and Zhou, X., Clustered regularly interspaced short palindromic repeats/Cas9-mediated lateral flow nucleic acid assay, *ACS Nano,* 2020, vol. 14, no. 2, pp. 2497–2508.
28. Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F., Rationally engineered Cas9 nucleases with improved specificity, *Science,* 2016, vol. 351, no. 6268, pp. 84–88.

Translated by A. Barkhash