Can molecular dynamics explain decreased pathogenicity in mutant camphecene-resistant influenza virus?

Sophia S. Borisevicha, Maxim A. Gureevb, Olga I. Yarovaya, Vladimir V. Zarubaev, Gennadiy A. Kostine, Yurii B. Porozov, and Nariman F. Salakhutdinov

aLaboratory of Physical Chemistry, Ufa Institute of Chemistry UFRS RAS, Ufa, Russia; bLaboratory of Bioinformatics, I.M. Sechenov First Moscow State Medical University, Moscow, Russia; cDepartment of Medicinal Chemistry, N.N. Vorozhtsov Novosibirsk Institute of Organic Chemistry, Siberian Branch RAS, Novosibirsk, Russia; dDepartment of Virology, Saint Petersburg Pasteur Institute, St. Petersburg, Russia; eNikolaev Institute of Inorganic Chemistry SB RAS, Novosibirsk, Russia; fDepartment of Food Biotechnology and Engineering, ITMO University, St. Petersburg, Russia; gDepartment of Computational Biology, Sirius University of Science and Technology, Sochi, Russia

Communicated by Ramaswamy H. Sarma

Abbreviations: CLC: camphecene-like compound; CPH: camphecene; HA: hemagglutinin; MD: molecular dynamics; QC: quantum-chemical

INTRODUCTION

The development of new anti-influenza drugs remains an active area, and efforts in this direction will likely continue far into the future. In this paper, we present the results of a theoretical study explaining the mechanisms behind the antiviral activity of camphor derivatives. These include camphecene and a number of its analogues. The compounds tested can inhibit hemagglutinin (HA) by binding to it at two possible sites. Moreover, the binding site located at the site of proteolysis is the most important. Serial passaging of influenza in the presence of camphecene leads to the formation of mutation-associated resistance. Specifically, camphecene causes a significant mutation in HA (V615L). This substitution likely reduces the affinity of the compound for the binding site due to steric restriction of the positioning of camphecene in the binding cavity. Molecular dynamics (MD) simulation results show that the mutant HA is a more stable structure in terms of thermodynamics. In other words, launching conformational rearrangements preceding the transition from pre- to post-fusion requires more energy than in wild type HA. This may well explain the lower virulence seen with the camphecene-resistant strain.

Abbreviations: CLC: camphecene-like compound; CPH: camphecene; HA: hemagglutinin; MD: molecular dynamics; QC: quantum-chemical

INTRODUCTION

Over the last decade, there has been a significant increase in the incidence of viral infections, including those categorized as ‘especially dangerous’. Analysis of the current 2020 pandemic, caused by the new SARS-CoV-2 coronavirus, provides a number of important conclusions. Firstly, increased population density and mobility have led to increases in the speed with which a viral outbreak spreads. Secondly, geographical distance is not a guarantee of protection against infection; imported cases can occur anywhere in the world. Thirdly, a rare disease with isolated cases can suddenly become a pandemic. Influenza, an acute respiratory infection caused by influenza viruses, is a globally widespread infectious disease (Taubenberger & Morens, 2007). According to the World Health Organization (WHO), annual mortality from seasonal influenza has reached as high as 650,000 individuals (Fiore et al., 2009b). Currently, two strategies are used to fight the influenza virus: vaccination (Fiore et al., 2009a) and chemotherapy (Lee & Yen, 2012; van Poelvoorde et al., 2020). In the former, timely production of vaccines requires taking into
account the antigenic properties of currently circulating strains. In addition, vaccination of some individuals may not be feasible due to a number of medical contraindications (Demicheli et al., 2018). In such cases, chemotherapy is the only way to fight the infection.

There are several types of anti-influenza drugs (Lee & Yen, 2012): adamantane derivatives, which are M2 viral ion channel blockers (amantadine, remantadine); and neuraminidase inhibitors (oseltamivir, zanamivir, peramivir, laninamivir). In 2018, the FDA approved the use of baloxavir marboxil, an inhibitor of viral endonuclease, as an anti-influenza drug (Hayden et al., 2018). All of these groups of drugs have their advantages and disadvantages. The main problem with the use of antiviral drugs is the emergence of drug resistance among influenza viruses. Due to random mutations, as well as gene reassortment, the diversity of the virus genome uncontrollably increases (van Poelvoorde et al., 2020). Hence, almost all modern influenza virus isolates (including the pandemic A/H1N1pdm09 virus) are resistant to remantadine due to mutation (S31N) in the M2 proton channel. Moreover, more recently, a new neuraminidase mutation (I223R) has been identified in this virus subtype (A/H1N1pdm09), which leads to the simultaneous resistance of the virus to neuraminidase inhibitors and rimantadine (van der Vries et al., 2012). Therefore, the development of anti-influenza drugs, unfortunately, remains a necessary effort. At the same time, there is an increasing need to search for new compounds whose antiviral effects are aimed at additional alternative viral targets, for example the surface viral protein hemagglutinin (HA).

Viral fusion proteins are divided into three classes based on structural features. Class 1 proteins are homotrimeric formations consisting of three identical subunits. They contain α-helical structures and a fusion peptide, located closer to the N-terminus, which is hidden in the middle of the protein trimer. Such proteins are present in several virus families: Orthomyxoviridae (influenza virus, HA protein); Retroviridae (human immunodeficiency virus, gp41 protein); Paramyxoviridae (respiratory syncytial virus, F protein); Filoviridae (Ebola fever virus, GP2 protein); and Coronaviridae (coronavirus, spike (S) protein). These viruses have a common fusion mechanism, implemented using so-called ‘heptad repeats’ (HR1–HR2); these are sections of the surface protein located before and after the fusion peptide (Bruno et al., 2012; Kielian & Rey, 2006; Porotto et al., 2010).

The function of influenza virus hemagglutinin is to ensure that the viral genome penetrates the cytoplasm of the host cell. This occurs during successive events: binding to the receptor and subsequent fusion of the viral and cell membranes (Behera et al., 2016; Hamilton et al., 2012; Russell et al., 2008; Xu & Wilson, 2011). The virion enters the cell as a result of endocytosis, and the acidic environment of the endosome triggers conformational rearrangement processes in HA. The HA ectodomain is a homotrimer. Each subunit contains two polypeptides: HA1, responsible for the process of receptor binding; and HA2, a fusion domain that undergoes significant conformational changes at low pH values. Following these changes, the so-called ‘fusion peptide’ gains the ability to bind to the cell membrane and fuse the viral and endosomal membranes. Following fusion, the viral genome and its associated proteins penetrate the cytosol (Hamilton et al., 2012). The process of membrane fusion is favorable from the point of view of thermodynamics, but proceeds rather slowly due to kinetic difficulties (Chernomordik & Kozlov, 2003; Xu & Wilson, 2011). It is assumed that as pH decreases, the B-loop (amino acid site 58–75, Figure 1) undergoes conformational changes, followed by a 180° turn, which leads to divergence of the α-helices. The fusion peptide at the N-terminus (shown in purple) moves to become the N-terminus of the new α-helix, formed from an inverted short (shown in red) and central α-helix. Further, the HA1 and HA2 subunits connected by a disulfide bridge diverge and the membrane anchor is exposed.

In other words, the process by which hemagglutinins change their conformational state is associated with a wide range of displacements of different molecular domains relative to each other. Associated changes in the structure of the receptor, at the secondary and tertiary levels, occur. In this regard, the following question becomes quite relevant: Is it possible to use molecular modeling to simulate mechanisms behind the transition of hemagglutinin from inactive to active (post-fusion) conformations? Analysis of literature data shows that this type of modeling work has been active since 1990. Early work was carried out, related to simulated annealing, in order to clarify hemagglutinin’s structural geometry (Weis et al., 1990). Work related to this subject was very rare. Until 2006, the published activity of such studies (examining the dynamics of different hemagglutinin types) was low (1–4 works per year). This was due, first of all, to the resource-intensive nature of such calculations. In 2007, the situation radically changed with the introduction of CUDA (Compute Unified Device Architecture) architecture. Molecular dynamics calculations have become a key application of ‘general-purpose computing on graphics processing units’ (GPGPU). Following the application of general-purpose video cards, the threshold of computing resources needed for calculations has decreased significantly. As a result, a growth in published activity in these areas has been seen, including studies examining hemagglutinin dynamics relative to its conformational rearrangements.

In 2009, a paper was published describing the prospects of studying the molecular dynamics of different types of influenza hemagglutinins. In that work, the researchers postulated that, based on a detailed analysis of receptor-binding domain dynamics, as well as their interaction energy profiles, it is possible to establish sites responsible for the development of a network of contacts and that it is possible to differentiate general and individual receptor characteristics (Newhouse et al., 2009). Contact frequencies, energy components, and analysis of hydrogen bonds made it possible to identify species-specific differences that cannot be established by X-ray diffraction analysis. Thus, influenza hemagglutinin structural dynamics studies are creating new opportunities in the development of promising antiviral agents (Bruno et al., 2015; Newhouse et al., 2009).
Several works, describing dynamic changes in hemagglutinin structure, are currently the most relevant. One of them, published in 2018, describes the complete process of transition from pre-fusion to post-fusion conformation (for influenza H3N2 hemagglutinin) using the smFRET molecular dynamics method (Das et al., 2018). The calculations performed aimed to identifying the system state with the maximum FRET effect. Moreover, a full registration of all the dynamic states of the trimeric hemagglutinin system was carried out. That work allowed us to create a series of reference points for validation of future molecular dynamics simulations. In recent publication (Benton et al., 2020), HA fusion intermediates, which may bridge between the viral and endosomal membranes, were depicted.

Another work analyzed conformational changes within the monomeric structure of H1N1 hemagglutinin in combination with the (HA) inhibitor MBX2329 (Basu et al., 2014) (Figure 1). Dynamics studies were based on the results of molecular docking of this compound at the tert-butyl-hydroquinone site. After a series of simulations in the 200 ns range, the authors studied ligand-protein interaction details within the complexes with decreasing pH. In particular, an analysis of the main components of the estimated free energy was carried out. The authors found that, when its pH was lowered, the ‘INT’ compound interferes with changes in secondary structure elements in hemagglutinin’s HA2 chain due to the stabilization of hydrophobic contacts inside the protein. The authors thus indicated that a specific site is conserved and is a promising site for the design of low-molecular-weight antiviral agents (Guan et al., 2017). It is formed by amino acid positions: HA1 K278; HA1 S289; HA1 L290; HA1 P291; HA1 P304; HA2 I55; HA2 E56; and HA2 T60 (numbering corresponds to PDB code 1RU7). Such publications show that molecular dynamics analysis of influenza hemagglutinin is relevant to work examining key mechanisms (behind HA’s biological activity) and work focused on the design of low-molecular-weight antiviral agents. At the same time, a question remains: ‘Are modern molecular dynamics methods (or their variations) suitable for simulation of full-scale H1N1 HA conformational rearrangements?’ It remains an open question.

We have described previously a new group of camphor-based, antiviral cage compounds (Artyushin et al., 2017; Sokolova et al., 2015; Yarovaya et al., 2019; Zarubaev et al., 2015). Among these, camphecene (1,7,7-trimethylbicyclo [2.2.1] heptane-2-ylidene-aminoethanol) shows high inhibitory activity against a number of influenza strains, including rimantadine-resistant strains. Numerous biological experiments have shown a wide range of anti-influenza activity for camphecene at low concentrations and with very low toxicity. It showed the highest antiviral activity in the first hours of infection (Zarubaev et al., 2015, 2018). We hypothesized that camphecene (CPH) inhibits hemagglutinin function by binding in the hydrophobic region located between two monomers of the HA trimer. According to some studies, potential HA inhibitors (Figure 1) bind precisely in this hydrophobic region (Basu et al., 2014; Guan et al., 2017; Leiva et al., 2018; Xu & Wilson, 2011; Yusuf et al., 2013). In order to further examine the mechanism of antiviral action, a camphecene-resistant influenza virus was obtained (Zarubaev et al., 2018). Passage of influenza A/H1N1 virus in the presence of increasing concentrations of the drug (six passages) yielded the resistant strain.

Sequencing of the camphecene-resistant strains HA gene showed the presence of a significant mutation (V615L) in the stem region of hemagglutinin’s HA2 subunit. Molecular modeling showed that at HA2’s proteolytic site there is a small,
hydrophobic cavity in which camphecene can fit, forming hydrogen bridges with V615 and I9. Replacing valine with leucine leads to a smaller cavity, thereby decreasing the affinity of camphecene for this binding site. Moreover, in vitro experimental data shows that drug resistance can occur. It is extremely important to note that the resulting mutants are characterized by a significant decrease in virulence and pathogenicity in animals. Such a remarkable result is apparently associated with HA structural features. Perhaps the mutation described (Zarubaev et al., 2015) significantly affects protein conformational mobility, thereby limiting flexibility. This may affect the process of conformational transition from pre-fusion to post-fusion and subsequent HA-mediated membrane fusion (Benton et al., 2018; Leiva et al., 2018).

Like camphecene, a number of its analogues (Artyushin et al., 2017; Sokolova et al., 2015; Zarubaev et al., 2015; Yarovaya et al., 2019) exhibit antiviral activity against the influenza A/Puerto Rico/8/34 (H1N1) virus; they also show the greatest effects in the first hours after infection. The assumption (Yarovaya et al., 2019) that these compounds inhibit viral HA also seems quite logical. In addition, structural similarities among analogues suggest that they bind to binding sites of HA in a manner similar to camphecene itself. Our goal here is to explain the mechanism behind the antiviral effect of camphecene and its analogues. We also address whether modern, molecular modeling methods can sufficiently explain the decreased pathogenicity of camphecene-resistant influenza strains. In the first part of this text, we evaluate the affinity of camphecene analogues for HA using molecular docking methods, along with an assessment of the statistical significance of the theoretical results in conjunction with biological experimental data. The second part is devoted to molecular simulation in time intervals; these enable description of HA domain movements in the presence of camphecene. We paid special attention to quantitative assessment of camphecene affinity for wild-type and mutant (V615L) HA. In addition, we suggest that camphecene-resistant influenza's significantly decreased pathogenicity and virulence (Zarubaev et al., 2018) are associated with a change in the spatial parameters of the protein that occur as a result of V615L substitution. Analysis of conformational transition energy barriers allows us to answer the questions posed.

**Experimental methods**

**Compounds**

In this work, camphecene analogues were studied (Figure 2). Their synthesis and biological activities have been described (Artyushin et al., 2017; Sokolova et al., 2015; Yarovaya et al., 2019; Zarubaev et al., 2015). Further in the text, they are abbreviated as a ‘camphecene-like compound’ (CLC2, CLC3, etc.).

**Acidity assessment of camphecene and its analogues**

Titration in aqueous and aqueous-alcoholic media determined the acidity constants for camphecene (1) and its three derivatives (2, 4, and 5). A combined glass electrode with an internal AgCl/Ag reference electrode was used for measurement. When titrated in an aqueous-alcoholic medium, the solvent was prepared as follows: an exact volume of ethyl alcohol (40 ml) was transferred into a 100 ml volumetric flask and brought to the mark with water. A titrant solution (0.04 M sulfuric acid) was prepared in a similar manner. Samples of the study compounds were selected such that the total amount of the substance in the volume of the titrated solution (25–30 ml) was 0.3–0.4 mmol, while the pH jump at the equivalence point was 3–4 units. Titration curves were processed in the Origin 6.0 package (Yusuf et al., 2013), and equivalence points were determined from the differential titration curves. The pH values were used to estimate the acid constants (pKa) with the addition of half the equivalent. Average pKa values were calculated from a minimum of three parallel determinations, with error characterized by the range of variation. Low solubility of some compounds (CLC2, CLC4, CLC5) make determination of their acidity constants in an aqueous medium impossible. However, the obtained data with aqueous-alcoholic media can be extrapolated, taking into account the correction experimentally determined for camphecene.

**Calculation methods**

**Receptor and ligand preparation**

The crystallographic structure of complete hemagglutinin (PDB code 1RU7 (Gamblin et al., 2004) was downloaded from the Protein Data Bank database (Berman et al., 2000). Model protein structures were prepared using the Schrodinger Protein Preparation Wizard tool: hydrogen atoms were added and minimized; missing amino acid side chains were added; bond multiplicities were restored; solvent molecules were removed; and the entire structure was restrained and optimized in the OPLS3e force field (Shivakumar et al., 2012) at physiological and low pH values (5.0 ± 0.2).

The geometric parameters of potential ligands (camphecene analogues) were also optimized, taking into account all permissible conformations. The probability of nitrogen protonation was taken into account by evaluating the acidity constants of a number of analogues (indicated by red asterisks in Figure 2). Protonation of the nitrogen atom creates an additional binding site in the molecule, preferred for the formation of hydrogen and/or salt bridges. Quantum chemical calculations were performed in the Jaguar program environment (Bochevarov et al., 2013), using previously described methodology (Bochevarov et al., 2016).

**Binding site analysis**

The binding site of tert-butyl hydroquinone (TBQH) in the hemagglutinin structure has been described (Russell et al., 2008). By binding in the hydrophobic pocket of the stem portion of the protein, TBQH likely inhibits the conformational changes necessary for indirect fusion. The molecule becomes located between two polypeptide chains of the protein (Russell et al., 2008) and is surrounded by a number
of amino acids, among which the most significant are: Y594; K558; E557; E597; and S554 (PDB code 1RU7 numbering). These are shown in Figure 3. This site is considered a potential camphecene (CPH) binding site and will be referred to as the ’TBQH site’ hereafter. Based on experimental data and theoretical calculations (Zarubaev et al., 2018), an alternative camphecene binding site (CPH site) may be located at the site of protein proteolysis. The molecule is located in close proximity to valine at position 615 (Figure 3).

**Ligand parametrization before docking and dynamics**

Ligand parameters (geometry, charges) were calculated with optimized OPLS3e forcefield. Forcefield optimization processed with use of forcefield builder on basis of calculated QM parameters for our set of ligands (basis set B3LYP 6-311 G++). All ligands before MD simulations were positioned in active site with use of molecular docking method. Best-fitting pose selected by RMSD in cluster less than 1.5Å.

**Molecular docking and dynamics/metadynamics**

Molecular docking procedures and dynamics were performed using Schrodinger Suite (Release 2018-4) software (Schrodinger Release 2018-4, 2018). For calculations, the monomeric form of hemagglutinin was considered. Studies of the effect of camphecene on protein domain movement (wt and mutant) were carried out on the trimeric form of the receptor. Ligand-protein complexes (CPH and HA) with a variable binding site, as described (Zarubaev et al., 2018), were used as sources for the preparation of molecular systems. Camphecene-hemagglutinin complexes (Zarubaev et al., 2018) underwent a procedure of limited minimization of ligand-induced changes within the binding site, taking into account surrounding amino acid charge characteristics using quantum chemical calculations.

Camphene analogues were docked using the forced ligand positioning protocol (induced fit docking) with the following conditions: flexible protein and ligand; grid matrix size of 15 Å; and amino acids (within a radius of 5 Å from the...
ligand) restrained and optimized, taking into account the influence of the ligand. QM (quantum-polarized docking) docking protocol was used for docking procedure only camphecene in to binding sites of HA. Docking solutions were ranked by evaluating the following calculation parameters: docking score (based on GlideScore minus penalties); ligand efficiency (LE, which takes into account atomic distribution of scoring function); and parameter of model energy value (Emodel), including GlideScore value, energy unrelated interactions and the parameters of the energy spent on the formation of the laying of the compound in the binding site.

Complexes of the series of camphecene analogues with hemagglutinin, obtained by molecular docking, were used for subsequent simulations of molecular dynamics. Among the studied structures (Figure 2), two subgroups were chosen for computational analysis: active group (featuring low cytotoxicity and pronounced antiviral activity); and low-activity group (structures featuring low antiviral activity).

Hemagglutinin monomer was used to build molecular systems. Ligand-protein complexes were placed in the orthorhombic system, with a buffer zone 26 Å from the surface of the protein. Systems were filled with 0.15 M aqueous NaCl (TIP3P Solvent Model; NVT environment). The recorded dynamics simulation periods were 300 ns (at 310 K, 37 °C). The protocol for simulation system preparation included preliminary minimization and balancing of system components.

To assess the degree of camphecene’s influence on domain movements in wild-type and mutant HA (V615L) (Zarubaev et al., 2018), a series of metadynamics simulations was performed (ranging from 300 to 600 ns) with physiological and low pH values. For these, the complete receptor structure in trimeric form was considered. As collective variables, the distances between the peptide chains and the dihedral angles of the disulfide bridge (Cys637-Cys8), which connects the HA1 and HA2 hemagglutinin subunits, were chosen. For further information see Figure 4 in supporting materials. The influence of the pH of the medium was taken into account by calculating: the protonated forms of amino acid residues; their tautomeric forms; as well as the distribution of partial charges on the surface of the receptor protein structure. The conditions for simulation system preparation are identical to those previously described for monomeric hemagglutinin (the orthorhombic system, with a buffer zone of 26 Å from protein surface. Systems were filled with 0.15 M aqueous NaCl, water model TIP3P; NVT environment). Metadynamics simulation at 310 K, with Gibbs free energy recording interval 0.01 ns. Height 0.02 ps (gives better accuracy). Resulting free energy surfaces is shown in supporting materials, Figure 3.

Results

Assessment of ligand protonation

Both at low and physiological pH values, some analogues of camphecene, like camphecene itself, can be protonated at the nitrogen atom. As estimated by quantum chemistry, the acidity of camphecene is 7.2. This value correlates with the
experimental pKa value of 7.5, thereby validating the suitability of the chosen calculation methodology. According to the Henderson–Hasselbalch equation (Henderson, 1908), at low pH, the number of deprotonated camphecene molecules is negligible. According to quantum-chemical calculation, a number of camphecene analogues (Figure 2: CLC1, 2, 4, 5, 16) are also characterized by an acidity value pKa > 7.2 (details in Supplementary Materials). Experimental data for CLC2, CLC4, and CLC6 show pKa values above 7.4. For structures in which nitrogen is bonded to the phenyl ring (Bruno et al., 2015; Newhouse et al., 2009), pKa values are below 5.3 (exp. data and theoretical calculations are detailed in the Suppl. Materials). For study compounds, the ratios of protonated and deprotonated forms were subsequently taken into account when evaluating the energy characteristics of ligand/binding-site affinity and the nature of surrounding amino acid interactions.

**Molecular docking**

We suggest that camphecene inhibits HA by binding in at least two of the likely binding sites. According to QM docking results, the energy values characterizing ligand affinity are approximately the same (Figure 4). However, the location of camphecene in the CPH-site is characterized by higher spatial (orientational) stability (Figure 4B). In the TBQH binding site (Figure 4A), the orientation of camphecene may be different, with the formation of hydrogen bridges between the hydroxyl group of the ligand and amino acids E597 or S554. In some cases, a cation–π stacking interaction is possible between the protonated nitrogen atom and the tyrosine aromatic ring at position 594. At the proteolysis site, camphecene is located closer to the viral membrane, with formation of hydrogen bridges: between the protonated nitrogen atom and oxygen of V615; and between the hydroxyl group and G10.

The presence of camphecene in culture medium leads to selection of a virus carrying the V615L mutation at the proteolysis site (CPH-site), thereby providing resistance to this low-molecular-weight agent (Zarubaev et al., 2018). In this work, however, we carried out docking of camphecene and its analogues at both binding sites. We assert that CPH and related compounds can bind at both binding sites and that their inhibitory effects will be cumulative. A combined analysis of computational and experimental data allowed us to evaluate the statistical significance of HA-ligand affinity data. For both sites, the correlation coefficient (experimental and theoretical data correlation) exceeded 0.7 (Figure 5). In other words, the statistical significance of the docking results for the TBQH-site and the CPH-site were almost identical. This indicates that a number of compounds can bind at both binding sites.

Energy characteristics, describing ligands’ interactions with surrounding amino acids, are given in the Supplementary Materials. While not excluding the possibility that CPH or its analogues may bind in a hydrophobic region like TBQH, we consider the CPH binding site to be a priority.

**Molecular dynamics**

Molecular dynamics simulation was carried out for a group of four compounds (CLC2, CLC4, CLC5, CLC16 in Figure 2) featuring various biological activity indexes (IC50) and low cytotoxicities (CC50). They were selected in order to differentiate compounds according to biological activity level, in relation to (comparative) molecular dynamics parameters. Ideally, system molecular dynamics trajectories should foremost describe ligand-retention efficiency (in the HA binding site). In other words, maximum retention time (τ, ns) reflects compound presence at the CPH site (Table 1). To do this, a set of ‘frames’ was selected from each simulation, describing the state of the system for up to 300 ns.

As expected, camphecene’s retention time at the binding site (close to the site of proteolysis) was the longest. CLC4 is retained at hemagglutinin’s CPH binding site for 113 out of 300 ns, after which it passes into the dissolved phase. This event is due to the fact that the shear potential, at a given time, exceeds the energy of the ligand-protein...
complex. CLC5, with a lower level of affinity for the CPH site, has less attractive ligand-protein complex dynamic stability characteristics. The stable retention period in the binding site, for this small molecule, is only 90 out of 300 ns. The next in a decreasing range of target activity is CLC2. Simulated-trajectory analysis showed that it remains in the binding site up to 193 out of 300 ns (Figure 6). It is worth noting that this significantly exceeds the retention times of the previous substances. According to biological (experimental) data, the last test compound, CLC16, is inactive. Simulated-trajectory analysis of the dynamics of the complex with HA showed a low retention time in the CPH-binding site: only 10.5 ns. For the remainder of the simulation period, the compound is in solvent. During simulation, periodic contacts between CLC16 and HA are recorded, but none of them leads to the formation of a stable, ligand-protein complex.

Thus, camphecene and its active analogues are characterized by longer retention times at the binding site. Inactive compounds, such as CLC16, may also interact with the CPH-binding site; their retention times, however, are short. Meanwhile, HA conformational rearrangements are known to occur over microsecond intervals (Lin et al., 2018). In other words, biologically inactive compounds interact (with the binding site) for short periods of time and, most likely, they do not inhibit the fusion process between viral membrane and cell membrane.

**Molecular metadynamics**

Using simulation, free energies were obtained that describe protein transition to an active conformational state at physiological and lowered pH values. We consider the active conformational state (ΔG<sub>1</sub>) to be the state of the system (HA protein and HA-ACT -‘protein-ligand’ complex) that is favorable for the onset of conformational transition from pre-fusion to post-fusion HA (Table 2). It should be noted here that reference (Benton et al., 2020) mentions a similar state.
of the HA. We compared the geometric parameters of active conformational state 1 ($\Delta G_1$) with parameters of similar structure, in state obtained by the cryo-EM reconstruction (Benton et al., 2020). RMSD is for trimeric HA form in state 0–2.507 Å, trimeric form in state 1–2.571 Å for HA monomer—2.380 (state 0), 2.412 (state 1) angstrom respectively (Details see in SM). Minimum energies were measured on the energy surface and the most energetically optimal transitions between the potential conformational states (0→1) were determined. Presumably, the starting geometry of the protein or protein-ligand complex corresponds to a certain energy minimum $\Delta G_0$ (Table 2). Then, system transition to the active conformation ($\Delta G_1$) is accompanied by overcoming the energy barrier ($\Delta G_1$) and the formation of a transition state ($\Delta G_{tr}$). Next, the system relaxes, and energy decreases with the formation of an active conformation ($\Delta G_{1}$). Thus, we can explain camphecene’s effect on protein conformational change from the standpoint of kinetics (paying attention to energy barrier height $\Delta G_{tr}$); and thermodynamics (estimating the difference between the initial ($\Delta G_0$) and final ($\Delta G_1$) system energy states). In addition, the difference in free energy values for wild type and mutant HA ($\Delta G_{WT-M}$) is interesting (Figure 7).

In the first system (Table 2—rows 1 and 2; Figure 7A—blue and green rectangles), a wild type HA, corresponding to influenza A/H1N1/PR, was considered. MD simulation confirms an energetic advantage, for protein transition into the active conformation ($\Delta G_1$), at lower pH (compared to physiological pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known logical pH) of 2 kcal/mol. This correlates well with known biological experiments and theoretical calculations, we can assume that camphecene and its analogues can bind in at least two possible hydrophobic pockets of hemagglutinin. Molecular dynamics methods show the presence of stably oriented conformational states where, in addition to the CPH-site, there is stacking in alternative, undescribed cavities. In this case, there is an assumption that CPH, when interacting with wild-type HA, involves non-specific ‘sticking’ to protein structures via hydrophobic sites. Protein structure surface analysis showed that, between the HA1 and HA2 subunits, there is a series of hydrophobic contacts that are disrupted during the transition to the post-fusion conformation.

When interacting with these areas, camphecene or its analogues stabilize the ‘hydrophobic sandwich’, thereby complicating the transition to ‘active’ protein form. It is worth noting the high density of leucine residues in the registered binding sites. This fact brings to mind the possibility of stabilizing ‘leucine zippers’, which have been previously described for proteins of this type (Cho et al., 2013; Lin et al., 2018).

However, the important V615L mutation (conferring camphecene-resistance to the low-pathogenic influenza strain) occurs at the proteolyis site. This fact allows us to consider this small hydrophobic pocket a priority. The binding site’s small size suggests that the greatest affinities should be shown by small molecules containing a rigid hydrophobic framework and a hydrophilic tail. These parameters are confirmed by experimental and molecular dynamics data. In addition, low cytotoxicity compounds are the most interesting. Thus, in addition to CPH, we highlight CLC2 and CLC4, which may be promising as potential antiviral drugs.

### Table 2. Free energy values, HA in complex with camphecene.

| Complex energy (kcal/mol) | HA-CPH Complex | pH$^*$ | $\Delta G_0$ | $\Delta G_{tr}$ | $\Delta G_1$ | $\Delta G_{WT-M}$ |
|--------------------------|----------------|-------|--------------|----------------|-------------|------------------|
| #                        | HA$^{WT}$      | 5     | -40.5        | 26.8           | -13.7       | -13.1           |
|                          | HA$^{WT}$-CPH | 7     | -39.8        | 26.7           | -13.2       | -11.0           |
|                          | HA$^{WT}$-V615L| 5     | -42.8        | 30.5           | -12.4       | -8.3            |
|                          | HA$^{V615L}$-CPH | 7     | -40.7        | 29.3           | -11.4       | -6.6            |
|                          | HA$^{V615L}$-V615L | 5     | -55.3        | 33.1           | -22.2       | -22.4           |
|                          | HA$^{V615L}$-V615L-CPH | 7     | -50.8        | 26.1           | -24.7       | -14.7           |
|                          | HA$^{V615L}$-V615L-CPH | 5     | -57.1        | 36.1           | -21.0       | -21.2           |
|                          | HA$^{V615L}$-V615L-CPH | 7     | -48.6        | 32.4           | -16.2       | -16.7           |

*$Of the medium.
Potentially, CLC4’s side chain could be extended or modified in ways that increase the molecule’s effectiveness through closer contacts with surrounding (binding site) amino acids.

Based on MD simulation data, it can be noted that CPH (when binding in the specific area at the proteolysis site) energetically stabilizes the structure of the protein as a whole, thereby preventing its transition to an active conformational state. In other words, HA in combination with CPH requires more energy for a conformational reversal and subsequent protein opening. Substitution (V615L) at the proteolysis site reduces the overall energy of the system. As a result, HA becomes a more stable structure. According to MD simulation, the system needs more energy to start the transition. When the process is already initiated, however, the transition to active conformation is facilitated.

Previously, we have shown that the camphene-resistant influenza strain, which does not differ from wild-type virus in terms of replication in cell culture, has lower fitness in lung tissue and features low pathogenicity (Zarubaev et al., 2018). Mice infected with it did not die or lose weight during influenza pneumonia. Generally, when influenza virus is cultured in vitro, excess trypsin is added to the media to permit effective replication. The protease trypsin is necessary for HA cleavage into two subunits and associated conformational flexibility.

On the other hand, increasing the thermodynamic stability of HA_{V615L} requires higher energy necessary to overcome the energy barrier for subsequent activation of the protein. Under in vivo conditions, it is possible that the supply of protease molecules is limited; a higher energy-threshold for mutant HA activation, compared to wild type, prevents conformational rearrangements, which is manifested by decreased replication and pathogenicity. Can this mechanism be associated with a decrease in the virulence of the stable, camphene-resistant influenza strain? The results presented here allow us to conclude that it is possible.

A particular mutant influenza strain is known to feature significantly reduced sensitivity to camphene (Zarubaev et al., 2018). We explain this as follows: as a result of mutation (V615L), a steric hindrance to CPH is created at the active cavity. This leads to a decrease in the compound’s affinity for the mutant strain. Despite the fact that CPH binding at the proteolysis site still occurs (Zarubaev et al., 2018), the process does not interfere with protein domain movements upon transition to the active conformation. This means that subsequent conformational change processes proceed without steric difficulty.

**Conclusion**

The theoretical work we present here allows us to assert that the location we term ‘CPH binding site’, located at the site of HA proteolysis, is the preferred site for binding of camphene (CPH) and its analogues. A hydrophobic pocket is most favorable for binding of molecules containing a rigid skeleton backbone and mobile hydrophilic substituents. We suggest that hemagglutinin’s transition from pre-fusion to post-fusion conformation proceeds through the formation of some active form (or pre-reaction state). To form a pre-reaction state, the system must also overcome the energy barrier. In agreement with known mechanisms, our MD simulation data shows this to be a pH-dependent process.

Binding of CPH at the proteolysis site increases the energy barrier and leads to the formation of an unstable, active-form of the protein-CPH complex, from the standpoint of thermodynamics. It is highly probable that the some of the steps are reversible. In other words, deactivation (or return of the protein-CPH complex from the active form to a stationary position) is more likely than subsequent conformational rearrangements. At least two factors allow us to draw such a conclusion. Firstly, when comparing free energy profiles (formation of the active protein form vs. the protein-CPH complex).
complex), the energy barrier to protein activation is much higher when CPH is present. Secondly, conformational rearrangements (pre-fusion to post-fusion) (Chernomordik & Kozlov, 2003) are difficult from the standpoint of kinetics. The presence of CPH in the binding sites also introduces thermodynamic difficulties. This explains the high antiviral activity of the compound (Zarubaev et al., 2018).

The V615L mutation in the stem region of the HA subunit leads to the formation of a system that is more stable, in terms of thermodynamics, compared to the wild-type protein. A higher activation barrier makes it difficult to switch to the active form. This is how we explain the decreased virulence of the camphecene-resistant influenza virus. Mutant HA ‘experiences’ thermodynamic and kinetic difficulties in protein activation. The substitution (V615L) leads to steric difficulties for CPH, thereby affecting its affinity. The energy profiles, for the transition from stationary to active state, of the HA(V615L) and HA(V615L−CPH) systems are similar. These data correlate well with biological results (Zarubaev et al., 2018): the CPH IC₅₀ for HA(V615L) influenza is 100-fold higher than the CPH IC₅₀ for HA wt influenza.

Thus, theoretical molecular modeling, as applied here, can indeed explain the antiviral effects of CPH and its analogues. It also suggests alternative binding options for potential HA inhibitors. In addition, we can see that seemingly insignificant changes in HA structure (from the standpoint of structural chemistry), like any other viral protein, can affect overall virus pathogenicity. In the case we analyze here, such mutation led to stabilization of surface protein-structure. This is likely behind the observed decrease in the camphecene-resistant influenza strain’s virulence and pathogenicity. It should also be noted that on the surface of the SARS-CoV-2 coronavirus, which is behind a current pandemic (2020), similar fusion proteins are present. This study may help in understanding certain mechanisms-of-action that affect the search for new, effective antiviral agents against coronavirus.

Acknowledgements
All calculations were performed on Bioinformatics Laboratory servers (Institute of Translational Medicine and Biotechnology, Sechenov University). Molecular docking and quantum chemical calculations were performed as part of a State Task (AAAA-A17-117011910028-7). Molecular dynamic simulation work was supported by I.M. Sechenov First Moscow State Medical University Strategic Development Program under the Russian Academic Excellence 5−100 Project. The authors express sincere gratitude to Edward Ramsay (Smorodintsev Research Institute of Influenza, St. Petersburg) for his invaluable help in editing and discussion of the manuscript. The authors are grateful to Ministry of Science and Higher Education, Russian Federation for the financial support (grant N 075-15-2020-777).

Disclosure statement
The authors declare no conflict of interest.

ORCID
Sophia S. Borisevich https://orcid.org/0000-0001-8481-0470
Maxim A. Gureev https://orcid.org/0000-0002-0385-922X

References
Artyushin, O. I., Sharova, E. V., Vinogradova, N. M., Genkina, G. K., Moiseeva, A. A., Klemenkova, Z. S., Orshanskaya, I. R., Shtro, A. A., Kadyrova, R. A., Zarubaev, V. V., Varoyava, O. I., Salakhutdinov, N. F., & Brel, V. K. (2017). Synthesis of camphecene derivatives using click chemistry methodology. Bioorganic & Medicinal Chemistry Letters, 27(10), 2181–2184. https://doi.org/10.1016/j.bmcl.2017.03.051
Basu, A., Antanasijevic, A., Wang, M., Li, B., Mills, D. M., Ames, J. A., Nash, P. J., Williams, J. D., Peet, N. P., Moir, D. T., Prichard, M. N., Keith, K. A., Barnard, D. L., Caffrey, M., Rong, L., & Bowlin, T. L. (2014). New small molecule entry inhibitors targeting hemagglutinin-mediated influenza A virus fusion. Journal of Virology, 88(3), 1447–1460. https://doi.org/10.1128/JVI.01225-13
Behera, A. K., Chandra, I., & Cherian, S. S. (2016). Molecular dynamics simulation of the effects of single (S221P) and double (S221P and K216E) mutations in the hemagglutinin protein of influenza A H5N1 virus: A study on host receptor specificity. Journal of Biomolecular Structure and Dynamics, 34(9), 2054–2067. https://doi.org/10.1080/07391102.2015.1106341
Benton, D. J., Gamblin, S. J., Rosenthal, P. B., & Skehel, J. J. (2020). Structural transitions in influenza haemagglutinin at membrane fusion pH. Nature, 583(7814), 150–153. https://doi.org/10.1038/s41586-020-2333-6
Benton, D. J., Nans, A., Calder, L. J., Turner, J., Neu, U., Pu Lin, Y., Ketelaars, E., Kallewaard, N. L., Corti, D., Lanzavecchia, A., Gamblin, S. J., Rosenthal, P. B., & Skehel, J. J. (2018). Influenza hemagglutinin membrane anchor. Proceedings of the National Academy of Sciences of the United States of America, 115(40), 10112–10117. https://doi.org/10.1073/pnas.1810927115
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The protein data bank. Nucleic Acids Research, 28(1), 235–242. https://doi.org/10.1093/nar/28.1.235
Bochevarov, A. D., Harder, E., Hughes, T. F., Greenwood, J. R., Braden, D. A., Philipp, D. M., Rinaldo, D., Halls, M. D., Zhang, J., & Friesner, R. A. (2013). Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. International Journal of Quantum Chemistry, 113(18), 2110–2142. https://doi.org/10.1002/qua.24481
Bochevarov, A. D., Watson, M. A., Greenwood, J. R., & Philipp, D. M. (2016). Multiconformation, density functional theory-based pKa prediction in application to large, flexible organic molecules with diverse functional groups. Journal of Chemical Theory and Computation, 12(12), 6001–6019. https://doi.org/10.1021/acs.jctc.6b00805
Bruno, V. B., Baptista, A. M., & Soares, C. M. (2012). Structural determinants for the membrane insertion of the transmembrane peptide of hemagglutinin from influenza virus. Journal of Chemical Information and Modeling, 52, 3001–3012.
Bruno, V. L., Loussa, D., Antunes, J. M., & Soares, C. M. (2015). Self-assembly molecular dynamics simulations shed light into the interaction of the influenza fusion peptide with a membrane bilayer. Journal of Chemical Information and Modeling, 55, 795–805.
Chernomordik, L. V., & Kozlov, M. M. (2003). Protein-lipid interplay in fusion and fission of biological membranes. Annual Review of Biochemistry, 72, 175–207. https://doi.org/10.1146/annurev.biochem.72.121801.161504
Cho, K. J., Lee, J.-H., Hong, K. H., Kim, S.-H., Park, Y., Lee, J. Y., Kang, S., Kim, S., Yang, J. H., Kim, E.-K., Seok, J. H., Unzai, S., Park, S. Y., Saelens, X., Kim, C.-J., Lee, J.-Y., Kang, C., Oh, H.-B., Chung, M. S., & Kim, K. H. (2013). Insight into structural diversity of influenza virus haemagglutinin. The Journal of General Virology, 94(PT 8), 1712–1722. https://doi.org/10.1099/vir.0.051136-0
Das, D. K., Govindan, R., Nikic-Spiegel, I., Krammer, F., Lemke, E. A., & Munro, J. B. (2018). Direct visualization of the conformational dynamics of single influenza hemagglutinin trimers. Cell, 174(4), 926–937. https://doi.org/10.1016/j.cell.2018.05.050
Demicheli, V., Jefferson, T., Pietrantoni, D. C., Ferroni, E., Thoming, S., Tomas, R. E., & Rivetti, A. (2018). Vaccines for preventing influenza in the elderly. Cochrane Database of Systematic Reviews, 2, 1–18.
