JAP&T TPA-Timişoara

Available online at http://journal-of-agroalimentary.ro

Journal of Agroalimentary Processes and Technologies 2021, *27* (2), 236-242

Journal of Agroalimentary Processes and Technologies

QSAR and molecular docking approaches on a wide range of natural and synthetic antivirals against MERS-CoV and SARS-CoV-2

Daniel I. Hădărugă

Department of Applied Chemistry, Organic and Natural Compounds Engineering, Polytechnic University of Timişoara, 300001-Timişoara, Carol Telbisz 6, România

Abstract

Various natural and synthetic compounds were evaluated in the last decades for their inhibitory activity against coronaviruses. They belong to many chemical classes such as terpenoids, alkaloids, peptides, phenothiazines, or steroids. Twenty one compounds from these classes were evaluated as antivirals (50 % effective concentration for virus binding inhibition, EC_{50}) against Middle East respiratory syndrome coronavirus (MERS-CoV). The goal of the study was to evaluate the main structural characteristics of these antivirals that influence the anti-MERS-CoV activity through the quantitative structure-activity relationships (QSARs) and to use them for prediction of the inhibitory activity against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by molecular docking experiments. Valuable QSARs were obtained with the logarithm of the octanol/water partition coefficient, log*P*, for a parabolic equation with r = 0.905 for the non-peptide related set of anti-MERS-CoV compounds, acting by inhibiting the virus binding to the human cell. On the other hand, better results were obtained for the second set, mainly of peptide-related anti-MERS-CoV, which inhibit virus replication. The most important parameters were the centered Broto-Moreau autocorrelation - lag 0 / weighted by I-state (*ATSC0s*, with r = 0.965), the largest absolute eigenvalue of Burden modified matrix - n 4 / weighted by relative I-state (*SpMax4_Bhs*, r = 0.965), and count of E-states for (strong) hydrogen bond donors (*nHBd*, r = 0.937). The most active antivirals were subjected to molecular docking into the MERS-CoV and SARS-CoV-2 receptor sites corresponding to S1 protein:dihydrofolic acid and protease:2-chlorocyclohexylmethyl-2-hydroxymethyl-6-nitrocyclohexanol complexes, respectively. In both cases, the best interaction energies were obtained for bufalin, a steroid derivative with two *H*-donor and four *H*-acceptor groups, suggesting that the inhibition mechanism can be similar and can be further studied as starting structural architecture for new compounds with possible effects against SARS-CoV-2.

Keywords: MERS-CoV, SARS-CoV-2, Covid-19, QSAR, molecular docking, antiviral drug

1. Introduction

Coronaviruses are RNA viruses that especially infect the respiratory tract. In the last decades, Middle East respiratory syndrome coronavirus (MERS-CoV) and severe acute respiratory syndrome coronaviruses types 1 and 2 (SARS-CoV-1 and SARS-CoV-2) generates epidemics and pandemics, including the present one [1-5]. There are many theories related to the origin of such viruses that involve African bats and camels, but other mammals, i.e. sheep and goats can be infected. However, all these viruses have high death rates [1,5].

There are many studies on the coronavirus infection and replication mechanisms [3-6]. They involves spike proteins (and other proteins) with various subdomains that contain the binding receptor sites for attaching and fusing to the human cells. Angiotensin converting enzyme 2 (ACE2) is involved in the mechanism of cell entry. These receptor sites also bind the corresponding antiviral drugs, which inhibit the binding to human cells [3]. On the other hand, virus replications involve

* Corresponding author: daniel.hadaruga@upt.ro

enzymes for replication. Most common are proteases such as 3C-like protease (3CLpro), or RNA polymerases [4-6].

MERS-CoV virus has a binding domain on spike protein named S1-NTD, which is a receptor for various antiviral drugs, but also for the reference compound dihydrofolic acid [7-11]. On the other hand, protease inhibitors bind to the specific sites of MERS-CoV or SARS-CoV-2 proteases such as 3CLpro [5,6,11]. A novel inhibitor that bind to this enzyme is 2-chlorocyclohexylmethyl-2-hydroxymethyl-6-nitrocyclohexanol, the 3CLpro:antiviral drug complex being characterized by X-ray diffractometry [12,13].

In this study, the main structural characteristics of twenty one antiviral drugs from various chemical classes were evaluated through their influence as anti-MERS-CoV activity. The quantitative structure-activity relationships (QSARs) for anti-MERS-CoV drugs were used to further evaluate the activity of the most active compounds against SARS-CoV-2 by molecular docking experiments.

2. Materials and Method

2.1. Antiviral drugs selection

Twenty one antiviral drugs from various chemical classes having inhibitory activity against MERS-CoV were selected from literature [11,12]. They were split into two classes. Set A consists of antiviral drugs that act by inhibiting the virus binding to the human cell and Set B consists of peptide-related anti-MERS-CoV drugs, which inhibit virus replication. The antiviral activity was the 50 % effective concentration (mol/L) for virus binding inhibition, EC_{50}, expressed as $pEC_{50} = -\log EC_{50}$ (Tables 1 and 2).

2.2. Molecular modeling, conformational analysis, structural descriptors and QSARs

Molecular modeling, conformational analysis and some of structural descriptors determination for the antiviral molecules were performed with the appropriate programs from the HyperChem 7.52 package [14]. First, structures were built and modeled using *MM+* molecular mechanics program, while conformational analysis was performed using the same molecular mechanics approach by *Conformational Search* program. In all cases, Polak Ribiere algorithm, RMS gradient of 0.05 kcal/mol, flexible bonds and rings, as well as a limit of 4 kcal/mol above best for the retained conformations were considered. The most stable conformations of compounds were used for calculating structural descriptors. *QSAR Properties* from HyperChem and *PaDEL Descriptors 2.21* [15] were used for calculating more than 1500 descriptors from various classes (constitutional, 2D and 3D descriptors such as topological, electronic, steric and hydrophobic descriptors). The selection of the best monoparametric QSARs were achieved through *Correlation matrices* from Statistica 7.1 package (StatSoft, Inc., Tulsa, OK, USA) and the QSAR equations for both antiviral sets were obtained through *Multiple Linear Regression* program form the same package. The quality of QSAR models was evaluated through the standard errors for coefficients, correlation coefficient, *r*, standard deviation for the equation, *s*, and the Fischer value. Validation of QSAR models were obtained by cross-validation, using odd-even leave-half-out (LHO) method, by means of cross-validation correlation coefficient, q_{cv}.

2.3. Molecular docking of the best antiviral drugs into the MERS-CoV and SARS-CoV-2 receptor sites

Two receptor sites corresponding to the binding of MERS-CoV through spike protein (S1-NTD) and SARS-CoV-2 replication through the 3CLpro protease were used for complexation of the most active antivirals resulted from QSAR analysis. The starting conformations of these biomacromolecules were obtained from Protein Data Bank (PDB, https://www.rcsb.org/). The reference compounds from the PDB complexes (dihydrofolic acid for MERS-CoV and 2-chlorocyclohexylmethyl-2-hydroxymethyl-6-nitrocyclohexanol for SARS-CoV-2) were replaced by the most active antivirals, maintaining the *H*-bonding group and hydrophobic moiety (van der Waals interactions) similarities. The new protein/enzyme:antiviral drug complexes were optimized using the same *MM+* program, using the same modeling parameters.

The receptor:antiviral drug interaction energy was obtained as the difference between the sum of the energies of "empty" biomacromolecule and antiviral drug (the most stable conformation) and the complex energy (kcal/mol). The energy values depend on the modeling technique used. As a consequence, the interaction energy was only used for comparing various complexes with the same protein/enzyme, in order to have a qualitative evaluation of the antiviral drug binding level and further to the antiviral activity.

Table 1. Antiviral drugs having inhibitory activity against MERS-CoV binding to the human cell (Set A). Inhibitory activity was expressed as pEC_{50}, with EC_{50} in mol/L [11,12]

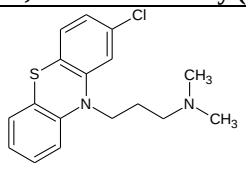
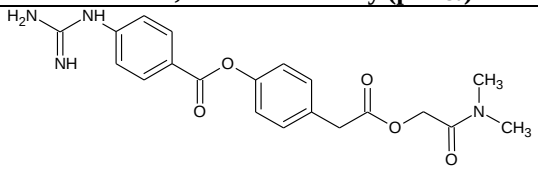
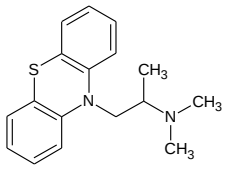
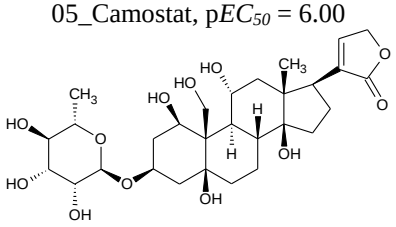
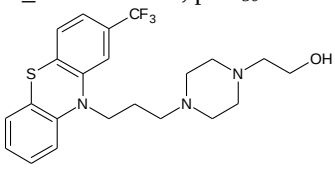
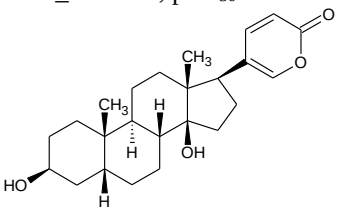
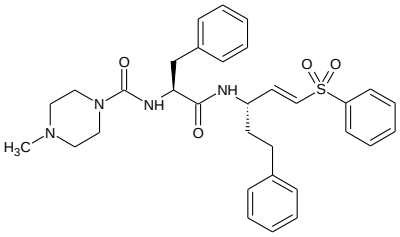
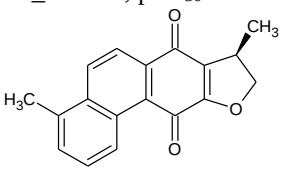
Structure, code and activity (p*EC*$_{50}$)	**Structure, code and activity (p*EC*$_{50}$)**
	
01_Chlorpromazine, pEC_{50} = 4.31	05_Camostat, pEC_{50} = 6.00
	
02_Promethazine, pEC_{50} = 4.78	06_Ouabain, pEC_{50} = 7.30
	
03_Fluphenazine, pEC_{50} = 4.82	07_Bufalin, pEC_{50} = 7.90
	
04_K11777, pEC_{50} = 7.34	08_Dihydrotanshinone, pEC_{50} = 5.57

Table 2. Antiviral drugs having inhibitory activity against MERS-CoV virus replication (Set B). Inhibitory activity was expressed as pEC_{50}, with EC_{50} in mol/L [11,12]

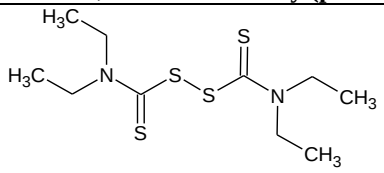
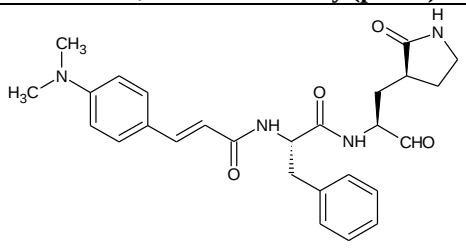
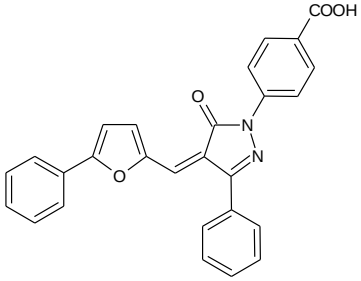
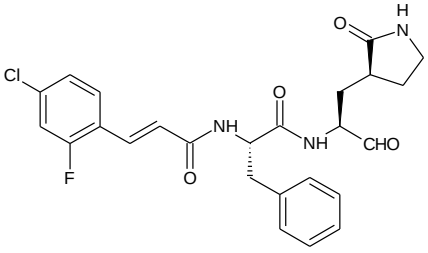
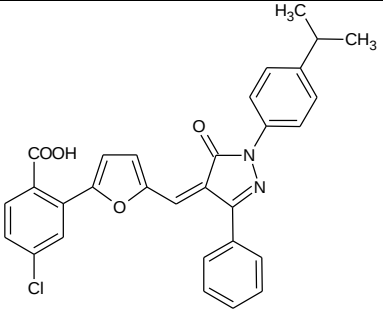
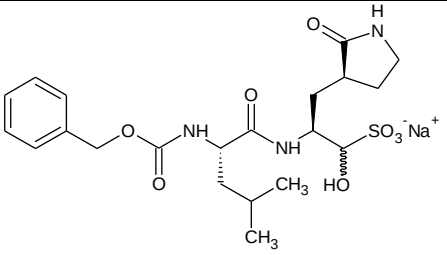
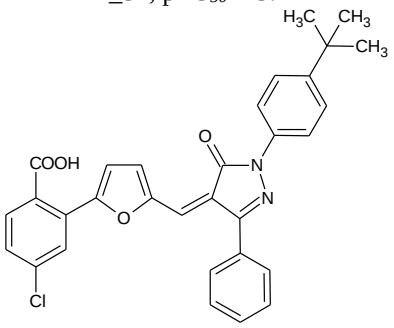
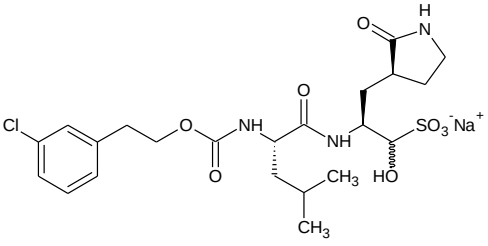
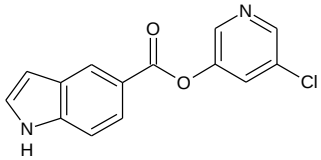
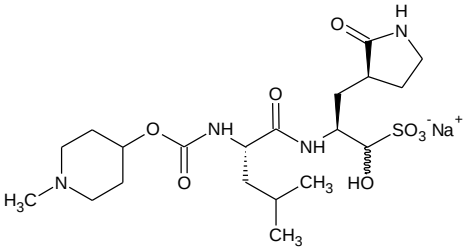
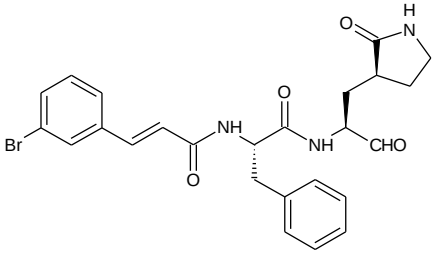
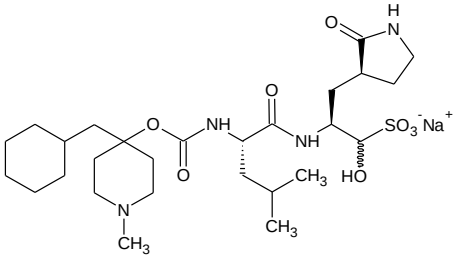
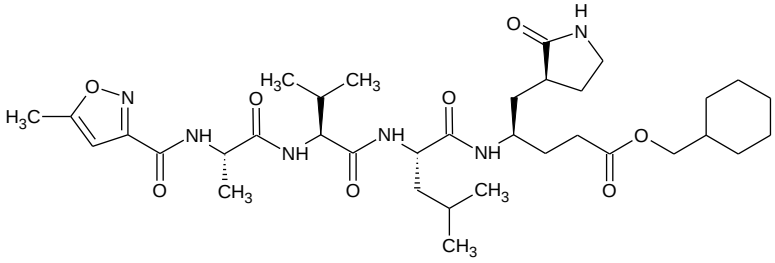
Structure, code and activity (p*EC*$_{50}$)	**Structure, code and activity (p*EC*$_{50}$)**
	
09_Disulfiram, pEC_{50} = 4.64	15_6c, pEC_{50} = 5.92
	
10_3k, pEC_{50} = 5.24	16_6d, pEC_{50} = 6.22

Table 2. (continued)

Structure, code and activity (pEC_{50})	Structure, code and activity (pEC_{50})
	
11_3h, pEC_{50} = 5.14	17_GC376, pEC_{50} = 6.05
	
12_3i, pEC_{50} = 5.13	18_GC813, pEC_{50} = 6.30
	
13_CE5, pEC_{50} = 4.90	19_10a, pEC_{50} = 6.30
	
14_6b, pEC_{50} = 5.85	20_10c, pEC_{50} = 6.10
	
21_N3, pEC_{50} = 6.55	

3. Results and Discussion

3.1. QSARs for anti-MERS-CoV

Valuable QSARs were obtained for both antiviral sets, especially with structural parameters related to hydrophobicity and constitution. However, many parameters are intercorrelated such as in the case of topological descriptors.

The best results for hydrophobicity descriptors were obtained in the case of *AlogP* parameter (Ghose-Crippen *LogKow*) for the set A of antivirals in a parabolic model (Eq. 1). This model suggest a hydrophobicity of $AlogP_{optimum}$ = -2.66, close to compounds 04_K11777, 06_Ouabain and 07_Bufalin, all with more hydrophilic structures. It means that more hydrophilic antivirals are more active for inhibiting the MERS-CoV binding to the human cells. It is possible that the accession to the virus receptor site from the extracellular side (aqueous) is facilitated for more hydrophilic drugs. Better results were obtained for the 3D topological descriptor *TDB2m* (3D topological distance based autocorrelation - lag 2 / weighted by mass), in a monoparametric QSAR model, with a correlation coefficient of r = -0.937 and a cross-validation correlation coefficient of q_{cv} = 0.910 (Eq. 2). However, *ALogP* and *TDB2m* parameters are intercorrelated ($r_{intercorr.}$ = 0.550) and lower *TDB2m* values indicate anti-MERS-CoV drugs with higher activity ($TDB2m_{optimum}$ = 188).

$$pEC_{50} = 6.68(\pm0.34) - 0.85(\pm0.19)\cdot AlogP - 0.16(\pm0.06)\cdot(AlogP)^2 \quad \text{(Eq. 1)}$$

n=8, r=0.905, q_{cv}=0.269, s=0.47, F=7.9

$$pEC_{50} = 17.93(\pm1.83) - 0.059(\pm0.009)\cdot TDB2m \quad \text{(Eq. 2)}$$

n=8, r=-0.937, q_{cv}=0.910, s=0.52, F=43.0

Hydrophobic parameter also correlate with the anti-MERS-CoV activity for virus replication inhibitors from the set B. Thus, an inverse linear correlation with *logP* having a correlation coefficient of r = -0.774 and with molecular surface, *S*, of r = 0.794 (positive correlation) have been obtained, while the intercorrelation coefficient was low ($r_{intercorr.}$ = -0.44). Better QSARs were obtained with topological descriptors, *ATSC0s* (centered Broto-Moreau autocorrelation - lag 0 / weighted by I-state, Eq. 3), *SpMax4_Bhs* (largest absolute eigenvalue of Burden modified matrix - n 4 / weighted by relative I-state, Eq. 4) and *nHBd* (count of E-states for (strong) hydrogen bond donors, Eq. 5) with high r values of 0.956, 0.965 and 0.937, respectively. These descriptors also have high cross-validation correlation coefficients, q_{cv}, of 0.935-0.963, but with intercorrelation coefficients of 0.96-0.97. The optimum values for these descriptors are the lowest *ATSC0s* values of 202-210, *SpMax4_Bhs* of 5.06-5.22 and *nHBd* of 5 that correspond to compounds 13_CE5, 19_10a and 20_10c, having two *N*-containing heterocyclic moieties and multiple *H*-bonding groups (both *O*- and *N*-based).

$$pEC_{50} = 4.54(\pm0.12) + 0.0087(\pm0.0008)\cdot ATSC0s \quad \text{(Eq. 3)}$$

n=13, r=0.956, q_{cv}=0.956, s=0.19, F=117.5

$$pEC_{50} = 2.07(\pm0.30) + 0.83(\pm0.07)\cdot SpMax4\_Bhs \quad \text{(Eq. 4)}$$

n=13, r=0.965, q_{cv}=0.963, s=0.17, F=150.1

$$pEC_{50} = 4.83(\pm0.12) + 0.30(\pm0.03)\cdot nHBd \quad \text{(Eq. 5)}$$

n=13, r=0.937, q_{cv}=0.935, s=0.23, F=79.0

3.2. Molecular docking of antiviral drugs into the receptor sites of MERS-CoV and SARS-CoV-2

The antiviral drug binding at the receptor site (spike protein or protease) involves the orientation of the most stable conformation structures in the cavity of the receptor site. Structural features such as *H*-bonding and hydrophobic groups or moieties were oriented according to the reference molecules (dihydrofolic acid for MERS-CoV and 2-chloro-cyclohexylmethyl-2-hydroxymethyl-6-nitrocyclo-hexanol for SARS-CoV-2). Three antiviral drugs having high, medium and low activities have been selected for molecular docking and comparison of the protein/enzyme-drug interaction energies. Thus, 07_Bufalin with pEC_{50} = 7.90, 21_N3 with medium activity of 6.55 and the less active 01_Chlorpromazine, with pEC_{50} of only 4.31 were

considered. All interaction energies were favorable, but the highest one was obtained for the most active compound, 07_Bufalin (Figures 1 and 2). However, the medium active antiviral 21_N3 has the lowest interaction energy in both cases. This docking experiments support further researches on the bufalin-related compounds for obtaining novel anti-coronavirus drugs.

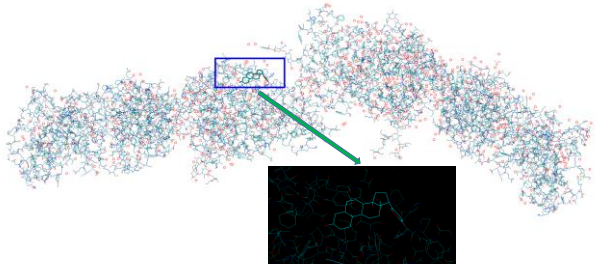


Figure 1. The optimized complex of 07_Bufalin and MERS-CoV spike protein (S1-NTD domain) at the dihydrofolic acid site (3D structure of the protein was based on the XRD from Protein Data Bank, https://www.rcsb.org/ [9]). The detail of the interaction site is presented in the bottom-right corner

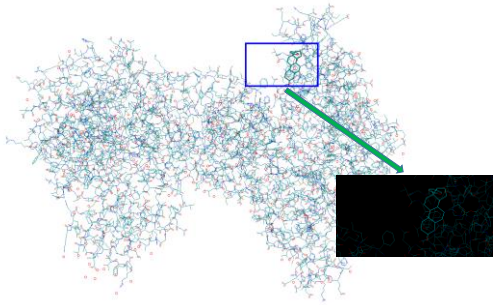


Figure 2. The optimized complex of 07_Bufalin and SARS-CoV-2 at the 3CL protease specific antiviral site (3D structure of the enzyme was based on the XRD from Protein Data Bank, https://www.rcsb.org/ [13]). The detail of the interaction site is presented in the bottom-right corner

4. Conclusion

In this study, valuable and statistically significant QSAR models for the inhibition of the MERS-CoV binding to the human cells were obtained. Both hydrophobicity through the parabolic model involving *logP* descriptor and the number of *H*-donor atoms are important for the anti-MERS-CoV activity. This aspects indicate the possibility of stabilization of the ligand-receptor complex by several such bonds, as well as the transport to the protein/enzyme surface.

The interaction energies between the antiviral drugs and specific protein/enzyme of the MERS-CoV and SARS-CoV-2 coronaviruses were high for the compound bufalin. This chemical architecture consist of a steroid basic structure with two *H*-donor and four *H*-acceptor groups, indicating possible similarities between the inhibition mechanisms of these two antiviral actions. Moreover, this structure can be further studied as starting structural architecture for new compounds with possible effects against SARS-CoV-2 and other coronaviruses.

Acknowledgments. The author wants to thank S. Funar-Timofei and M. Mracec ("Coriolan Drăgulescu" Institute of Chemistry, Romanian Academy) for the help with Statistica 7.1 and HyperChem packages. The author also wants to thank D.I. Codrea for the contribution to molecular modeling of antivirals.

Compliance with Ethics Requirements. The author declares that he respect the journal's ethics requirements. The author declares that he has no conflict of interest.

References

[1] Al Sulayyima, H.J.; Khorshid, S.M.; Al Moummar, S.H., Demographic, clinical, and outcomes of confirmed cases of Middle East Respiratory Syndrome coronavirus (MERS-CoV) in Najran, Kingdom of Saudi Arabia (KSA); A retrospective record based study, *Journal of Infection and Public Health* **2020**, *13*(9), 1342-1346, https://doi.org/10.1016/j.jiph.2020.04.007

[2] Du, L.; Tai, W.; Zhou, Y.; Jiang, S., Vaccines for the prevention against the threat of MERS-CoV, *Expert Review of Vaccines* **2016**, *15*(9), 1123-1134, https://doi.org/10.1586/14760584.2016.1167603

[3] Xia, S.; Liu,, Q.; Wang, Q.; Sun, Z.; Su, S.; Du, L.; Ying, T.; Lu, L.; Jiang, S., Middle East respiratory syndrome coronavirus (MERS-CoV) entry inhibitors targeting spike protein, Virus Research 194 (2014) 200–210, http://dx.doi.org/10.1016/j.virusres.2014.10.007

[4] Ganta, N.M.; Jadav, S.S.; Kumar, A.; Dan, N.; Mohanty, N.P., The updates on Middle East Respiratory Syndrome coronavirus (MERS-CoV) epidemiology, pathogenesis, viral genome and currently available drugs, *Journal of Pharmaceutical Chemistry* **2016**, *3*(2), 10-18

[5] Hu, B.; Guo, H.; Zhou, P.; Shi, Z.L., Characteristics of SARS- CoV-2 and COVID-19, *Nature Reviews. Microbiology* **2021**, *19*, 141-154, https://doi.org/10.1038/s41579-020-00459-7

[6] Jackson, C.B.; Farzan, M.; Chen, B.; Choe, H., Mechanisms of SARS-CoV-2 entry into cells, *Nature Reviews. Molecular Cell Biology* **2022**, *23*, 3-20, https://doi.org/10.1038/s41580-021-00418-x

[7] Pallesen, J.; Wang, N.; Corbett, K.S.; Wrapp, D.; Kirchdoerfer, R.N.; Turner, H.L.; Cottrell, C.A.; Becker, M.M.; Wang, L.; Shi, W.; Kong, W.-P.; Andres, E.L.; Kettenbach, A.N.; Denison, M.R.; Chappell, J.D.; Graham, B.S.; Ward, A.B.; McLellan, J.S., Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen, *PNAS* **2017**, *114*(35), E7348-E7357, https://doi.org/10.1073/pnas.1707304114

[8] Park, Y.-J.; Walls, A.C.; Wang, Z.; Sauer, M.M.; Li, W.; Tortorici, M.A.; Bosch, B.-J.; DiMaio, F.; Veesler, D., Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors, *Nature Structural & Molecular Biology* **2019**, *26*, 1151-1157, https://doi.org/10.1038/s41594-019-0334-7

[9] Wang, N.; McLellan, J.S., Crystal Structure of MERS-CoV S1-NTD bound with G2 Fab, *PDB Protein Data Bank* **2019**, https://doi.org/10.2210/pdb6PXH/pdb

[10] Wang, N.; Rosen, O.; Wang, L.; Turner, H.L.; Stevens, L.J.; Corbett, K.S.; Bowman, C.A.; Pallesen, J.; Shi, W.; Zhang, Y.; Leung, K.; Kirchdoerfer, R.N.; Becker, M.M.; Denison, M.R.; Chappell, J.D.; Ward, A.B.; Graham, B.S.; McLellan, J.S., Structural Definition of a Neutralization-Sensitive Epitope on the MERS-CoV S1-NTD, *Cell Rep* **2019**, *28*, 3395-3405.e6, https://doi.org/10.1016/j.celrep.2019.08.052

[11] Liang, R.; Wang, L.; Zhang, N.; Deng, X.; Su, M.; Su, Y.; Hu, L.; He, C.; Ying, T.; Jiang, S.; Yu, F., Development of Small-Molecule MERS-CoV Inhibitors, *Viruses* **2018**, *10*, 721, https://doi.org/10.3390/v10120721

[12] Chang, K.-O.; Kim, Y.; Groutas, W.C.; Hua, D.; Saif, L.J., Broad-spectrum antivirals against 3C or 3C-like proteases of picornavirus-like supercluster: picornaviruses, caliciviruses and coronaviruses, **US Patent 9474759 B2**, 25 October **2016**

[13] Su, H.X., Zhao, W.F., Li, M.J., Xie, H., Xu, Y.C., SARS-CoV-2 3CL protease (3CL pro) in complex with a novel inhibitor, *PDB Protein Data Bank* **2020**, https://doi.org/10.2210/pdb6M2N/pdb

[14] HyperChem, version 7.52, **2005**, Hypercube, Inc., Gainesville, FL, USA

[15] Yap, C.W., PaDEL-Descriptor: An open source software to calculate molecular descriptors and fingerprints, *Journal of Computational Chemistry* **2011**, *32*(7), 1466-1474