Quantitative Biology 2021, 9(4): 440–450
https://doi.org/10.15302/J-QB-021-0263

RESEARCH ARTICLE

Molecular docking of cyanine and squarylium dyes with SARS-CoV-2 proteases NSP3, NSP5 and NSP12

Pavel Pronkin*, Alexander Tatikolov

N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow 119991, Russia
* Correspondence: pronkinp@gmail.com

Received December 8, 2020; Revised February 2, 2021; Accepted March 5, 2021

***Background*: The outbreak and continued spread of coronavirus infection (COVID-19) sets the goal of finding new tools and methods to develop analytical procedures and tests to detect, study infection and prevent morbidity.**
***Methods*: The noncovalent binding of cyanine and squarylium dyes of different classes (60 compounds in total) with the proteases NSP3, NSP5, and NSP12 of SARS-CoV-2 was studied by the method of molecular docking.**
***Results*: The interaction energies and spatial configurations of dye molecules in complexes with NSP3, NSP5, and NSP12 have been determined.**
***Conclusion*: A number of anionic dyes showing lower values of the total energy E_{tot} could be recommended for practical research in the development of agents for the detection and inactivation of the coronavirus.**

Keywords: SARS-CoV-2; proteases; polymethine dyes; squarylium dyes; noncovalent interaction; molecular docking

Author summary: Using molecular docking modeling, the noncovalent interaction of a large number of cyanine and squarylium dyes of various classes with SARS-CoV-2 coronavirus proteases has been studied. It has been found that electrostatic ligand–protein interactions (Coulomb interactions) can play an important role in the stability of noncovalent complexes. Based on the data obtained, the selection of dyes for further practical research has been carried out with the aim of developing spectral-fluorescent probes for detection of SARS-CoV-2. In addition, it is concluded that meso-substituted thiacarbocyanines may be promising for use in photoinactivation of the coronavirus.

INTRODUCTION

The ongoing pandemic of the SARS-CoV-2 determines the need to search for new means and methods both for therapeutic purposes and for the development of analytical procedures and tests in order to detect, study, and prevent the spread of coronavirus infection.

One of the most powerful approaches currently applied in studies involving viruses include the use of fluorescent probes and labels, *e.g.*, cyanine dyes SYBR gold [1], YO, YOYO, YO-PRO and BOXTO-PRO [2]; Cy5 [3] and Cy3 [4], monomethine oxacyanines [5]. In this respect, cyanine dyes represent an attractive class of fluorescent probes and markers whose main advantages lie in the high extinction coefficients, absorption and emission in UV to NIR spectral ranges covering the optical window for biological samples, and their photophysical and photochemical properties, which depend on the molecular environment [6]. Furthermore, cyanine dyes (and related squarylium dyes) are capable of noncovalent interaction with biomolecules with fluorescence growth, which creates the prerequisites for their use as probes in biomolecular systems [7,8]. The variety of structures and properties of cyanine dyes of different classes (which contain a polymethine chain of different length, different terminal heterocycles and substituents) requires molecular modeling *in silico* in preliminary screening the compounds. Modeling using molecular docking is widely used to study the

© The Author(s) 2021. Published by Higher Education Press

interaction of drugs with the protein components of SARS-CoV-2, which provides valuable information about the binding sites of the substrate with the selected protein components and the interaction energy (or affinity) of the substrate molecule with these proteins [9–11].

In the present work, the noncovalent interaction of various classes of polymethine and squarylium dyes (about 60 compounds) with proteases NSP3, NSP5, NSP12, being nonstructural proteins of SARS-CoV-2, was modeled using the *in silico* molecular docking method. We studied polymethine dyes of various structures, differing in the length of the polymethine chain, terminal heterocycles, substituents in the polymethine chain and heterocycles. To reveal the influence of the charge of the dye molecule on this interaction, cationic, neutral, and anionic dyes were studied.

As an energetic characteristic of ligand–protein noncovalent binding, the affinity parameter (also known as an energy scoring function) is often used in docking studies [9, 12, 13]. However, in our case this parameter exhibits only slight variations as a function of the structure (or charge) of the dyes studied (see Tables 1, 2). This could be the consequence of its complex nature, which is largely based on experimental data on the dye–protein binding constants [14–17]. Since there is no such literature data on binding polymethine dyes with SARS-CoV-2 proteases (even with any proteases), the affinity values obtained using DockThor seem to be not very reliable for the system under study. As an alternative, to characterize the possible binding of the dyes to proteases, the total energy parameter (E_{tot}) is used [18, 19], which is obtained in DockThor from the MMFF94S force field and is composed of the intermolecular and intramolecular interactions according to the electrostatic and van der Waals potentials and the torsional energy of the ligand [14, 20]. We found that the E_{tot} parameter varied much greater than the affinity parameter with changing the structure of the polymethine dye; therefore, this parameter was chosen by us in the present work as a stability characteristic of the possible noncovalent binding between polymethine dyes and SARSCoV-2 proteases.

RESULTS

Molecular docking of cationic dyes

The interaction of cationic *meso*-substituted thia- and oxa-carbocyanine dyes with proteases NSP3, NSP5 and NSP12 was studied. The structural formulas of the dyes

Table 1 Results of molecular docking of anionic squarylium dyes SQ1.1–SQ1.7 with proteases NSP3, NSP5, NSP12

	Dye	Run	Aff. (kcal mol^{-1})	E_{tot} (kcal mol^{-1})	E_{VdW} (kcal mol^{-1})	E_{el} (kcal mol^{-1})
NSP3	SQ1.1	18	–7.71	19.35 ± 1.82	–10.35 ± 5.89	–35.90 ± 5.74
	SQ1.2	5	–7.79	23.20 ± 3.09	–9.35 ± 9.15	–43.81 ± 11.07
	SQ1.3	5	–8.14	–57.7 ± 2.7	0.57 ± 9.9	–53.5 ± 11.2
	SQ1.4	15	–7.04	43.56 ± 1.98	–1.30 ± 6.33	–53.09 ± 7.49
	SQ1.5	23	–7.18	44.59 ± 0.86	–6.63 ± 5.65	–46.71 ± 5.94
	SQ1.6	22	–7.25	4.80 ± 2.87	–5.73 ± 7.23	–49.20 ± 9.28
	SQ1.7	3	–7.12	–59.85 ± 1.38	–7.76 ± 7.70	–43.93 ± 8.65
NSP5	SQ1.1	24	–9.01	26.10 ± 2.59	–21.67 ± 5.84	–19.69 ± 6.38
	SQ1.2	7	–7.45	20.70 ± 2.28	–15.02 ± 2.17	–44.92 ± 4.10
	SQ1.3	12	–7.83	–68.4 ± 1.24	–12.25 ± 5.3	–54.8 ± 5.3
	SQ1.4	15	–8.41	38.06 ± 3.88	–22.83 ± 8.02	–46.24 ± 7.23
	SQ1.5	8	–8.59	38.46 ± 5.75	–29.72 ± 5.28	–41.73 ± 13.48
	SQ1.6	7	–8.632	–0.15 ± 6.69	–21.84 ± 6.86	–43.52 ± 12.02
	SQ1.7	14	–8.06	–65.97 ± 1.98	–25.65 ± 6.53	–33.01 ± 5.89
NSP12	SQ1.1	3	–7.74	–36.66 ± 3.16	–13.27 ± 3.18	–50.59 ± 5.29
	SQ1.2	24	–7.23	11.14 ± 1.19	–11.51 ± 5.34	–58.86 ± 5.96
	SQ1.3	7	–7.08	–76.9 ± 1.2	–14.02 ± 3.8	–62.5 ± 4.2
	SQ1.4	9	–6.67	23.36 ± 4.22	–15.68 ± 7.31	–62.55 ± 9.23
	SQ1.5	22	–6.70	29.99 ± 1.62	–15.99 ± 8.13	–57.97 ± 8.13
	SQ1.6	24	–7.05	–14.41 ± 1.84	–10.79 ± 5.42	–64.79 ± 5.22
	SQ1.7	22	–8.08	–77.65 ± 4.53	–12.84 ± 5.77	–58.70 ± 2.47

© The Author(s) 2021. Published by Higher Education Press

Table 2 Results of molecular docking of anionic cyanine dyes with proteases NSP3, NSP5, NSP12

Dye	X	NSP3			NSP5			NSP12		
		Run	Aff. (kcal mol^{-1})	E_{tot} (kcal mol^{-1})	Run	Aff. (kcal mol^{-1})	E_{tot} (kcal mol^{-1})	Run	Aff. (kcal mol^{-1})	E_{tot} (kcal mol^{-1})
SQ1.1	S	16	–6.55	–45.4 ± 2.1	5	–7.71	–43.8 ± 1.13	20	–7.14	–47.6 ± 2.6
SQ1.2		22	–6.21	–15.8 ± 1.51	16	–7.98	–21.6±1.96	2	–7.42	–22.4 ±3.2
SQ1.3		22	–6.95	–33.9 ± 0.51	12	–8.06	–29.3 ± 0.71	22	–7.58	–41.1 ± 3.7
SQ1.4		1	–6.56	33.4± 1.14	14	–8.01	23.0± 1.03	3	–7.34	32.1± 0.69
SQ1.5	O	15	–6.82	–62.1 ± 2.3	11	–7.40	–69.2 ± 4.2	22	–7.21	–63.2 ± 2.7
SQ1.6		21	–6.71	–54.3 ± 3.3	18	–8.18	–60.3 ± 1.64	10	–7.55	–55.9 ± 2.9
SQ1.7	$C(CH_3)_2$	15	–7.26	–35.2 ± 1.76	12	–8.19	–35.2 ±1.76	4	–6.96	–39.9 ±0.97
SQ2.1	S	22	–7.02	–19.8 ±2.8	11	–8.46	–19.23 ± 2.7	18	–7.22	–24.2 ± 1.71
SQ2.2		6	–6.97	–24.5 ± 1.96	–	–	–	–	–	–
SQ2.3		20	–8.06	17.25 ± 1.28	11	–9.50	–1.78±2.5	4	–8.50	15.5 ± 1.82
SQ2.4		19	–6.96	–5.16± 1.35	3	–9.31	–3.14±4.7	22	–7.58	–6.38± 2.4
SQ2.5	O	19	–6.76	–42.7 ±1.1	15	–7.60	–51.2 ± 3.2	24	–4.02	–41.9 ± 1.11
SQ2.6		16	–7.75	–39.1 ±1.5	21	–7.87	–47.9 ± 3.5	1	–7.58	–39.3 ± 0.98
SQ2.7		3	–6.49	58.4 ± 1.19	10	–7.99	46.48 ±1.82	10	–7.53	56.1 ± 0.87
SQ2.8		15	–6.84	–7.97±0.70	8	–9.82	–26.6 ± 2.5	20	–8.42	–7.48 ± 0.66
SQ2.9		13	–6.58	–59.2 ± 2.1	9	–9.12	–72.8 ± 2.0	21	–7.70	–59.3 ± 1.38
SQ2.10	$C(CH_3)_2$	8	–6.97	–26.7 ± 1.12	1	–8.22	–27.7 ± 3.0	23	–7.37	–32.0 ± 3.4
SQ2.11	S	11	–6.56	–22.8 ± 1.86	7	–7.41	–28.3 ± 1.45	21	–7.93	–31.3 ± 2.9
SQ2.12	O	8	–6.71	–109.0 ±0.85	19	–7.64	–124.2 ±3.6	1	–7.74	–120.1 ±0.82
SQ2.13	$C(CH_3)_2$	18	–6.96	–81.6 ± 1.4	22	–8.04	–83.3 ± 2.7	17	–7.60	–91.1 ±2.03
SQ3.1	S	21	–7.62	–35.5 ± 1.16	10	–7.24	–43.8 ± 2.4	2	–7.63	–38.2 ± 0.96
SQ3.2		8	–6.63	–39.3 ± 1.3	4	–8.08	–43.7 ± 2.9	1	–7.79	–43.8 ± 2.1
SQ3.3	O	1	–7.48	–46.2 ± 2.6	11	–9.16	–50.7 ± 1.43	19	–8.18	–42.6 ± 1.51
SQ3.4		17	–6.24	–57.2 ± 2.1	6	–8.64	–66.5 ± 2.6	6	–7.44	–59.1 ± 1.05
SQ3.5		7	–7.31	–60.8 ± 1.44	3	–7.60	–73.2 ± 1.04	22	–7.72	–66.6 ± 1.15
SQ3.6	$C(CH_3)_2$	15	–7.10	–19.01± 2.7	1	–8.21	–17.67±2.8	8	–7.31	–16.47±1.48
SQ4.1	S	23	–7.89	–4.64 ± 0.67	2	–8.68	–47.1 ± 2.4	13	–7.80	–39.1 ± 0.87
SQ4.2		9	–7.43	–25.9 ± 2.2	16	–8.76	–35.1 ± 2.7	15	–6.70	–29.4 ± 1.17
SQ4.3		24	–7.37	–18.8 ±1.09	5	–9.11	–28.7 ± 2.1	9	–7.76	–28.1 ± 3.4
SQ4.4	O	2	–6.53	–14.0 ± 1.76	7	–8.86	–14.64±0.98	4	–8.67	–18.41 ±1.47
SQ4.5		9	–7.24	–38.1 ±5.2	8	–8.12	–45.1 ± 2.2	19	–8.11	–37.6 ± 1.36
SQ4.7		4	–8.34	–39.8 ± 2.3	16	–8.31	–47.5 ± 2.0	1	–7.76	–46.2 ± 1.46
SQ4.6		15	–7.33	–46.9 ± 1.81	7	–8.03	–51.4 ± 1.38	22	–7.64	–45.5 ± 0.62
SQ4.8	$C(CH_3)_2$	18	–7.89	21.5 ± 2.5	6	–10.5	25.0 ± 2.64	9	–7.97	17.01 ±3.2

are shown in Fig. 1 (structures C1–C10), the dyes differ from each other both by substituents at position 9 of the polymethine chain (*meso* position) and by substituents in terminal heterocycles (R2 and R3 for structures C1–C10). The initial isomeric configuration of carbocyanines C1–C10 corresponds to the *cis* configuration as the most stable for *meso*-substituted carbocyanines in polar solvents [21].

All cationic carbocyanines are characterized by positive values of total energy E_{tot} when interacting with the proteases. The maximum positive E_{tot} value was obtained for dyes C6 and C10 (Supplementary Table S1). In the case of docking with NSP5 and NSP12, the highest E_{tot} values were obtained for C1 (26.71 ± 0.35 kcal mol^{-1}). The lowest E_{tot} values (but also positive) was obtained for C8 with NSP5 and NSP 12 (14.75 ± 1.50 and 16.03 ± 1.75 kcal mol^{-1} respectively). Upon docking the cationic carbocyanines with NSP3, NSP5,

© The Author(s) 2021. Published by Higher Education Press

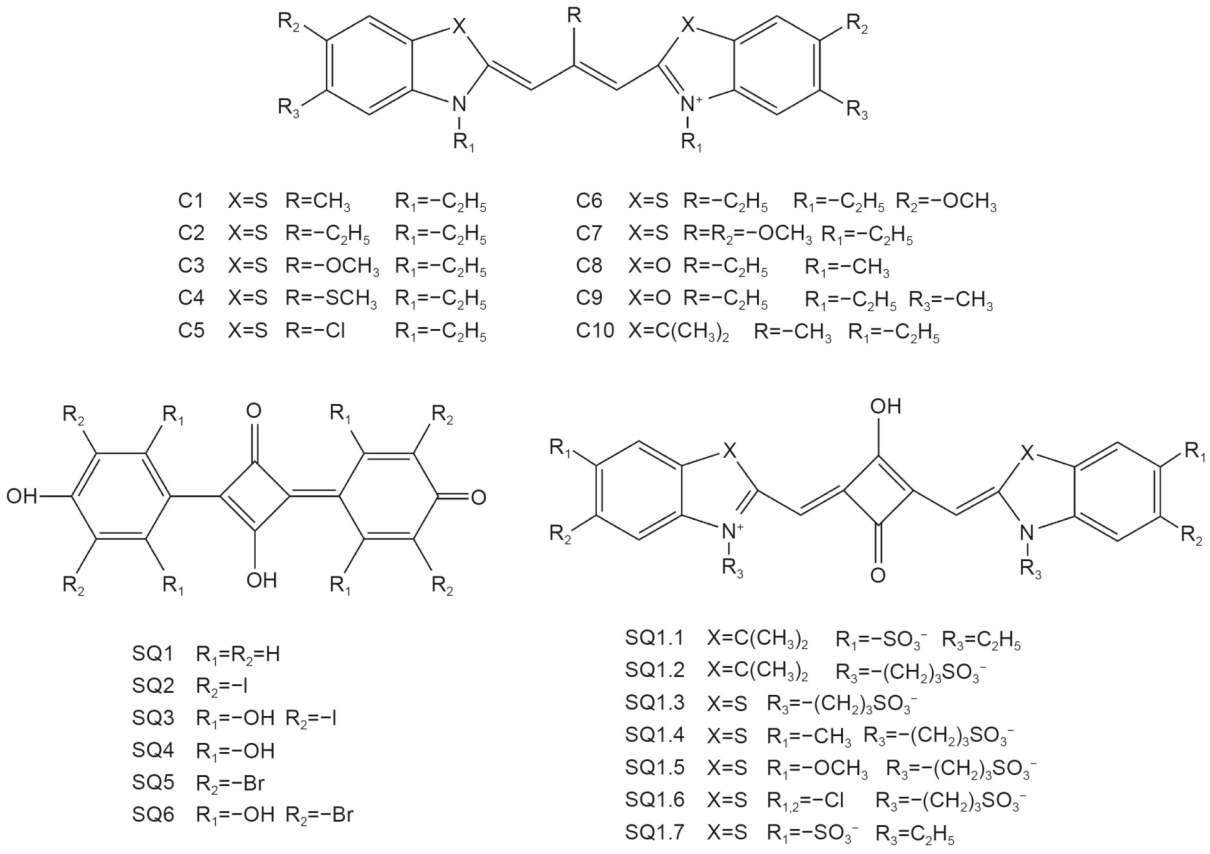


Figure 1. Structural formulas of cationic carbocyanines, as well as neutral and anionic squarylium dyes. Substituents R_n in the dye structures not mentioned in the legend are hydrogens (–H).

and NSP12, close values of the total energy of cationic carbocyanines were obtained (with the exception of C6, C8, C10), on average, 33.54 ± 5.39, 26.20 ± 5.18, and 28.24 ± 4.73 kcal mol^{-1}, respectively. The positive E_{tot} values may indicate lower stability of possible noncovalent complexes of cationic polymethine dyes with proteases NSP3, NSP5, and NSP12.

Note that for cationic thia- and oxa-carbocyanine dyes (including C1 and C10), the values of electrostatic interaction and van der Waals energies (E_{el} and E_{VdW}, respectively) were found to be negative, albeit small (Supplementary Materials).

Molecular docking with proteases NSP3, NSP5, NSP12 showed that cationic carbocyanines in complexes with these proteins are generally characterized by distorted (twisted and partially out-of-plane) *cis* configurations.

Molecular docking of squarylium dyes

The interaction with proteases NSP3, NSP5, NSP12 of six uncharged squarylium dyes was studied. The structural formulas of the dyes are shown in Fig. 1 (structures SQ1–SQ6).

Positive values of the total energy E_{tot} were obtained for docking of uncharged squarylium dyes with NSP3 (Supplementary Table S2). In particular, for dyes SQ3 and SQ4 having substituents R_1 = OH, E_{tot} = 52.66 ± 0.30 and 7.15 ± 0.81 kcal mol^{-1}, respectively, while docking of the unsubstituted dye SQ1 gave a lower (but also positive) value of E_{tot} (5.32 ± 0.66 kcal mol^{-1}). The electrostatic interaction energies E_{el} and the van der Waals energies E_{VdW} were found to be negative; their average values are –12.9 ± 2.46 and –17.98 ± 2.39 kcal mol^{-1}, respectively. The positive E_{tot} values may indicate lower stability of complexes of uncharged squarylium dyes with NSP3.

In the case of docking with NSP5 and NSP12, negative E_{tot} values were obtained for dyes SQ1 and SQ4, which may suggest the possibility of the formation of stable noncovalent complexes of dyes of this type with these proteins (Supplementary Materials). In particular, upon docking of SQ1 and SQ4 with NSP5, E_{tot} = –6.19 ± 1.36 and –1.92 ± 1.48 kcal mol^{-1}, respectively, were obtained; the docking for NSP12 gave comparable E_{tot} values (–3.51 ± 0.95 and –6.90 ± 1.33 kcal mol^{-1}, respectively). The values of E_{el} and E_{VdW} were found to be negative, while the contribution

© The Author(s) 2021. Published by Higher Education Press

of the Coulomb interaction was higher, which may indicate the predominant contribution of electrostatic interaction in the stability of the complex.

In a complex with NSP5, the molecules of uncharged squarylium dyes are located at a distance of 6–7 Å from ARG4, 19–20 Å from THR169, ~10 Å from GLY283, and 19–22 Å from TRP218 residues. For NSP12, distances of about 5–8 Å to PRO540, 12–13 Å to THR476, 11–14 Å to ALA717, and 9–10 Å to VAL86 residues were obtained.

Along with uncharged squarylium dyes, molecular docking of seven anionic squarylium dyes (structures SQ1.1–SQ1.7 in Fig. 1) with proteases NSP3, NSP5, NSP12 was performed. As a result of docking, positive values of the total energy were obtained for most of the studied anionic squarylium dyes ($E_{tot} > 0$; Table 1). The largest positive E_{tot} values were obtained for dye SQ1.5, which has OCH_3 groups in benzothiazole heterocycles (Table 1). Strongly negative E_{tot} values were obtained upon docking the squarylium dyes SQ1.1, SQ1.3, SQ1.6, and SQ1.7. Despite the different positions of the sulfonate groups, SQ1.3 and SQ1.7 have comparable negative E_{tot} values (~ –60/ –75 kcal mol^{-1}, Table 1).

The energies of electrostatic interaction and the van der Waals energies for the anionic squarylium dyes SQ1.1–SQ1.7 were found to be negative. On average, E_{el} is 1.7–4.3 times higher in absolute value than E_{VdW}, which indicates a significant contribution of electrostatic interaction forces to stabilization of dye–protein complexes.

The squarylium dyes SQ1.3 and SQ1.7 in a complex with NSP3 have twisted (close to perplanar) configurations. Docking SQ1.3 with NSP5 protease gives an almost planar structure of the dye (Fig. 2A, C), while the original isomeric configuration of the dye was retained. In the case of SQ1.7 with NSP12, the initial isomeric configuration is also generally retained, but the terminal heterocycles are somewhat twisted (by ~15–20°) relative to each other (Fig. 2B, D).

In a complex with NSP3, squarylium dye molecules are located at a distance of ~7 Å from GLU165, 7–13 Å from PRO245, and 9–13 Å from TYR205 residues. The distance from the dye molecule to THR263 is ~14 and ~20 Å for SQ1.3 and SQ1.7, respectively. In the case of NSP5, the ligand molecules are located at a distance of 8–10 Å from THR169 and GLY215, ~11–13 Å from ALA285 and ~15 Å from GLY195; the distance to GLY283 is 5 and 14 Å for SQ1.3 and SQ1.7, respectively. In a dye complex with NSP12, the distances about 5–6 Å to PRO540 and LYS718 residues, 8–10 Å to GLU87, and 12–14 Å to ASN472 were obtained.

Using UCSF Chimera [22], we performed surface binding analysis and search for hydrogen bonds between atoms of ligands and proteins, and search for interatomic

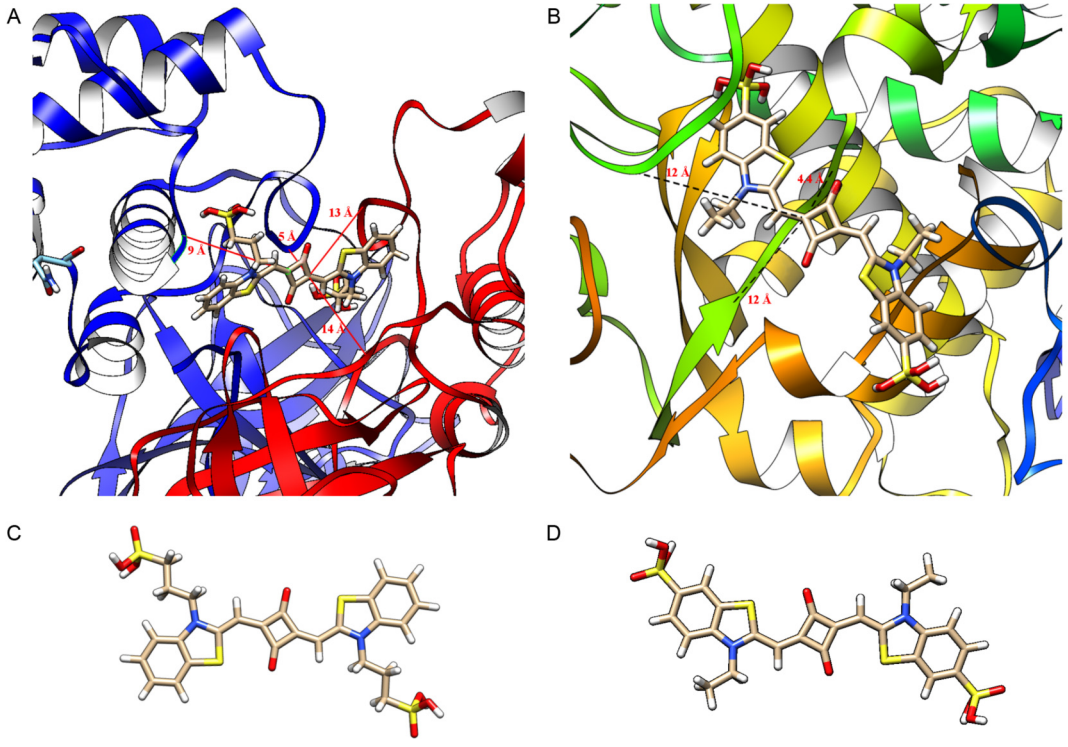


Figure 2. Result of docking of dyes SQ1.3 with NSP5 (A) and SQ1.7 with NSP12 (B), initial configurations of dyes: SQ1.3 (C), SQ1.7 (D).

© The Author(s) 2021. Published by Higher Education Press

clashes and contacts (based on van der Waals radii of molecules).

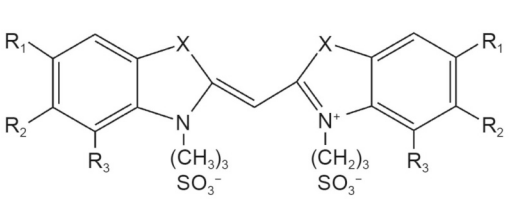
Upon the interaction of dyes SQ1.3 and SQ1.7 with NSP3, surface binding analysis did not reveal hydrogen bonding of the types with proteins. However, for both dyes, a single interatomic contact (0.62–0.65 Å) was found between one of the sulfonate groups of SQ1.3 and SQ1.7 with ASP162 and GLU201, respectively.

In the case of proteases NSP5 and NSP12, the search for interatomic collisions and contacts did not reveal direct (favorable or unfavorable) interactions. For squarylium dyes SQ1.3, SQ1.7 and protease NSP5, hydrogen bond interactions are found at distances of 2–3 Å between the terminal sulfonate groups of the dyes and ASP216 and SER284 (for SQ1.3) and GLU288 of NSP5 (for SQ1.7). Dye SQ1.3 also forms a hydrogen bond (2 Å) with LYS541 of protease NSP12, but due to the O atom of the squarylium ring (see Supplementary Fig. S1).

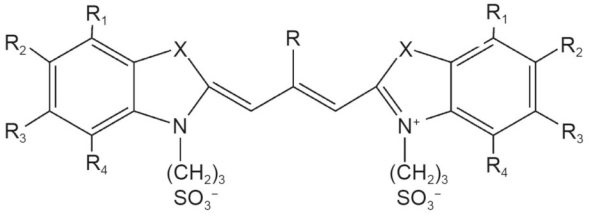
Docking of anionic cyanine dyes

Molecular docking was also performed with a series of 34 anionic cyanines with different polymethine chain lengths: 7 anionic monomethine cyanine dyes (Fig. 3, structures 1.1–1.7), 13 trimethine cyanine dyes (thia- and oxacarbocyanines, Fig. 3, structures 2.1–2.13), 6 pentamethine cyanine dyes (structures 3.1–3.6) and 8 heptamethine cyanine dyes (structures 4.1–4.8). The *meso*-substituted thia- and oxacarbocyanines corresponded to *cis* configuration.

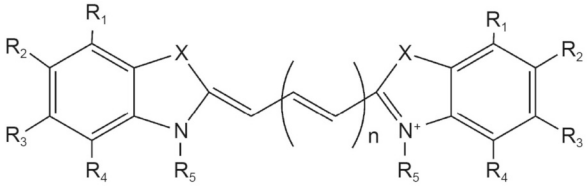
The docking has shown that most (80%) of the chosen compounds are characterized by negative values of the total energy (E_{tot} < –15 kcal mol^{-1}, see Table 2).



1.1 X=S
1.2 X=S $R_{2,3}$=−C_4H_4−
1.3 X=S R_2=−OCH_3
1.4 X=S $R_{1,2}$=−OCH_2O−
1.5 X=O
1.6 X=O R_1=−OCH_3
1.7 X=$C(CH_3)_2$



2.1 X=S R=−CH_3
2.2 X=S R=−CH_3 R_3=−Cl
2.3 X=S R=−CH_3 R_3=−C_6H_5
2.4 X=S R=−C_2H_5 R_2=−OCH_3
2.5 X=O R=−CH_3
2.6 X=O R=−CH_3 R_3=−CH_3
2.7 X=O R=−CH_3 R_3=−OCH_3
2.8 X=O R=−C_2H_5 R_3=−C_6H_5
2.9 X=O R=−CH_3 $R_{3,4}$=−C_4H_4−
2.10 X=$C(CH_3)_2$ R=−CH_3
2.11 X=S R=−CH_3 R_2=−SO_3^-
2.12 X=O R=−CH_3 R_2=−SO_3^-
2.13 X=$C(CH_3)_2$ R=−CH_3 R_2=−SO_3^-



n=2 R_5=−$(CH_2)_3SO_3^-$

3.1 X=S
3.2 X=S R_3=−OCH_3
3.3 X=O R_3=−OCH_3
3.4 X=O
3.5 X=O $R_{2,3}$=−OCH_2O−
3.6 X=$C(CH_3)_2$

n=3 R_5=−$(CH_2)_4SO_3^-$

4.1 X=S R_3=−OCH_3
4.2 X=S
4.3 X=S $R_{2,3}$=−OCH_2O−
4.4 X=O $R_{1,2}$=−C_4H_4−
4.5 X=O R_3=−OCH_3
4.6 X=O $R_{2,3}$=−OCH_2O−
4.7 X=O
4.8 X=$C(CH_3)_2$ $R_{1,2}$=−C_4H_4−

Figure 3. Structural formulas of the studied anionic cyanine dyes. Substituents R_n in the dye structures not mentioned in the legend are hydrogens (–H).

© The Author(s) 2021. Published by Higher Education Press

Significant negative E_{tot} values were obtained for anionic monomethine cyanine dyes, especially for their benzoxazole derivatives, which are significantly higher in absolute value than E_{tot} for their benzothiazole analogues. In particular, upon docking of oxa-monomethine cyanine 1.5 with NSP3, the E_{tot} value was 1.3 times higher in absolute value than that for the thia analogue 1.1. Similar results were obtained for docking the dyes with proteases NSP5 and NSP12: for oxa-dye 1.5 E_{tot} is significantly higher in absolute value than for dye 1.1 (Table 2).

Docking of trimethine cyanine dyes (carbocyanines) with proteases of SARS-CoV-2 has also shown that anionic *meso*-substituted carbocyanines could form energetically stable complexes with $E_{tot} < 0$. In particular, upon docking of oxacarbocyanines 2.5 and 2.6 with NSP3, E_{tot} was obtained in the range of –39 / –43 kcal mol^{-1}. Thiacarbocyanines show more moderate E_{tot} values (Table 2).

Molecular docking of carbocyanines with four sulfonate groups (compounds 2.11–2.13) gives significantly more negative values of the total energy (E_{tot} ~ –76.8 ± 40 kcal mol^{-1}), due to the large contribution of electrostatic interaction to stabilization of the complexes. The most negative E_{tot} values were obtained for oxacarbocyanine dyes; for example, upon docking of dye 2.12 with NSP12, E_{tot} = –120.05 ± 0.82 kcal mol^{-1}, with E_{el} = –59.3 ± 3.0 kcal mol^{-1}. Thiacarbocyanine dye 2.11 shows more moderate E_{tot} values (Table 2). Negative values of E_{tot} were obtained for docking of pentamethine cyanine (dicarbocyanine) dyes (especially for oxa-dicarbocyanines 3.3–3.5) with all proteases, which may suggest the stability of the intermolecular complexes of these dyes.

The study of the interaction with proteases of anionic heptamethine cyanine dyes (compounds 4.1–4.8) has shown that for most of these dyes noncovalent interaction could be energetically possible, leading to the formation of stable complexes. $E_{tot} < 0$ was obtained for all the dyes except 4.8 (indo-derivative). The lowest total energy values were obtained for oxa-dyes 4.5–4.7 (Table 2).

Docking of anionic cyanines with SARS-CoV-2 proteases has shown that such dyes could bind to NSP12, NSP3, NSP5 proteins in different conformations. Monomethine cyanines are characterized by twisted (nonplanar) configurations of their molecules in protease-bound states; however, for oxa-dyes 1.5, 1.6 with NSP3, NSP5, NSP12 almost planar (energetically optimal) configurations are also possible (see results for 1.5 with NSP3, Fig. 4A, C). For monomethine cyanine dyes, surface binding analysis did not reveal the presence of hydrogen bonds and interatomic clashes with NSP3, NSP5, NSP12. A planar, undistorted configuration is typical, in particular, for thiacarbocyanine 2.1 in a complex with NSP3. Docking with NSP5 gives a twisted *cis* configuration for dye 2.1, with the methyl group being almost perpendicular to the plane of one of the heterocycles, and the other being rotated by an angle of ~15°. In the case of the interaction of oxacarbocyanine 2.5 with NSP12, the calculation gives a *cis*-form twisted by ~45°. Docking of heptamethine cyanine 4.5 with NSP3 gives a twisted perplanar form; dye 4.5 with NSP5 has a crescent (bent due to the distorted bond angles of the polymethine chain, Fig. 4B, D) structure; in the case of NSP12, the structure of the dye molecule is almost planar (with the terminal heterocycles being approximately in the same plane and bonds 2–8 and 8–9 of the polymethine chain being somewhat twisted).

In a possible complex with NSP3, the molecules of the considered cyanine and squarylium dyes are located at a distance of 4–7 Å from GLU165, 7.8–11.5 Å from TYR169, 9.2–14 Å from TYR205, and 7–13 Å from LYS155 residues. Dye docking with NSP5 gives distances of 11.6–14 Å from the centers of dye molecules to ARG279 and TRP218 residues, 7.6–8.8 Å to VAL171, 11–21 Å to ALA129, and 9.4–11 Å to ASP289 (pentamethine and heptamethine cyanines are located somewhat further from ALA129, ASP289 and ARG279 than molecules of monomethine cyanines and squarylium dyes). In the case of the NSP12 protease, distances of 4.4–11.4 Å to PRO540, about 12–14 Å to GLY536 and ALA470, and 7.0–21 Å to VAL86 residues were obtained.

Surface binding analysis of thiacarbocyanine dye 2.1 with NSP3 revealed the formation of hydrogen bonds between the atoms of the sulfonate groups of the dye and residues LEU176, ASN175, GLN172 (2 Å), and a single interatomic contact (0.61 Å) with VAL200. In the case of NSP5 and NSP12, the analysis did not reveal the presence of hydrogen bonds, but revealed interatomic clashes of the dye in the case of NSP5 (~0.6 Å with ASP289). Oxacarbocyanine dye 2.13 with NSP3 forms no hydrogen bonds; the sulfonate group of the dye is in close contact with ARG164 (Supplementary Figs. S2, S3). In the case of NSP5 and 2.13, a hydrogen bond is formed between the sulfonate group of the dye and GLY282 (3 Å); the dye does not form interatomic contacts with the protein. On the contrary, 2.13 does not form hydrogen bonds with NSP12; an interatomic contact (0.67 Å) was found between the atoms of the sulfonate group of the dye and SER679.

For the thia-dye 3.2 with NSP3, an interatomic clash was detected between the atoms of the sulfonate group of the dye and GLU165 residue, but no hydrogen bonds were found. Surface binding analysis did not reveal the presence of hydrogen bonds and interatomic clashes for

© The Author(s) 2021. Published by Higher Education Press

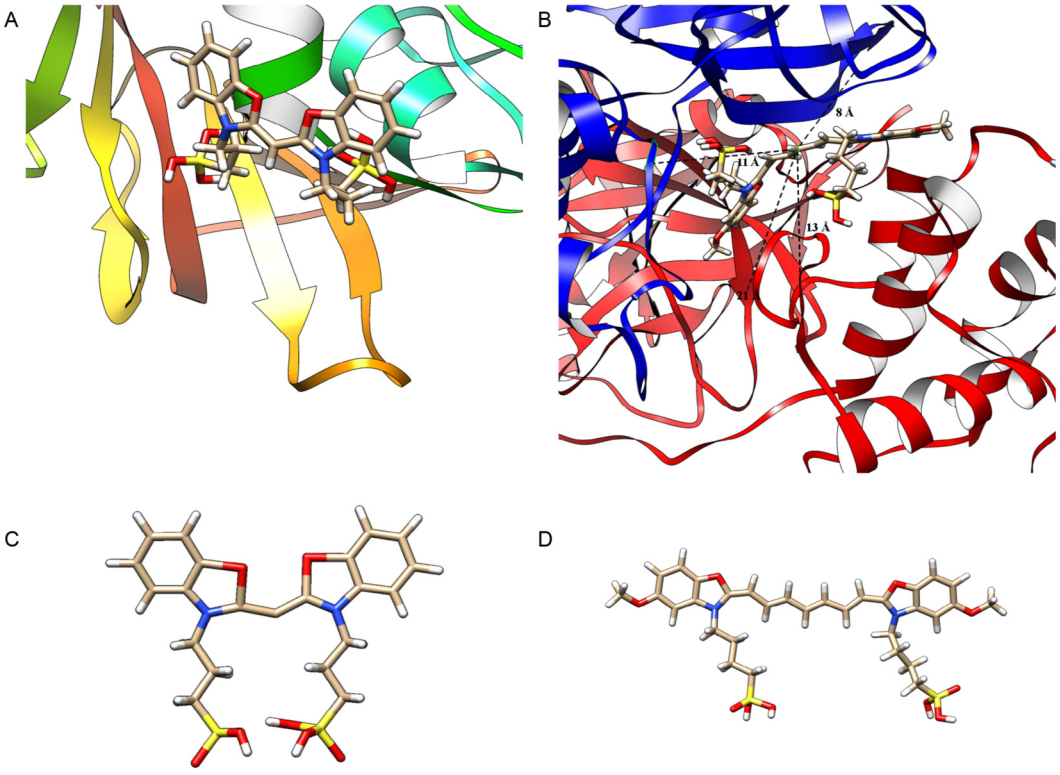


Figure 4. Results of docking of anionic dyes with SARS-CoV-2 proteases. (A) 1.5 with NSP3 and (B) 4.5 with NSP5; (C) initial configurations of dyes 1.5 and (D) 4.5.

dye 3.4 with NSP3 (Supplementary Figs. S4, S5). With NSP5, dye 3.4 forms hydrogen bonds between the sulfonate group and LEU282 (Supplementary Fig. S6). Dye 3.3 forms a hydrogen bond between the sulfonate group and GLU731 of NSP12 (Supplementary Figs. S7, S8). Surface binding analysis of the interaction of dyes 4.2 and 4.6 did not reveal hydrogen bonds and interatomic clashes with NSP3–NSP12.

DISCUSSION

The results of molecular docking obtained have shown that stability of noncovalent complexes between cyanine dyes and SARS-CoV-2 proteases is substantially affected by electrostatic ligand–protein interactions (Coulomb interactions). In particular, $E_{tot} > 0$ for cationic (C1–C9) or neutral (SQ1–SQ6) dyes, which may indicate a lower stability of complexes of such dyes with proteases. On the other hand, negative E_{tot} values were obtained for many anionic dyes, which may characterize the higher stability of the corresponding complexes. Especially low E_{tot} values were found for carbocyanines with four sulfonate groups in the molecule, which increase the negative charge of the dye. In addition, for anionic dyes the E_{el} value significantly exceeds E_{vdW} in absolute value, which also indicates the leading role of Coulomb forces in the formation of the complexes.

Structural differences in dye molecules have a significant effect on the stability of complexes with proteases NSP3, NSP5, NSP12. For squarylium dyes (SQ1.1–SQ1.7), the presence of anionic sulfonate groups in the molecule, in general, can lower the E_{tot} values, which may indicate the increased stability of complexes with proteases. In addition, the position of sulfonate groups in the molecule appears to be also capable of influencing the binding to proteases. At the same time, the introduction of other substituents ($–CH_3$, $–OCH_3$, –Cl) into the terminal nuclei of the molecule often increases E_{tot}, possibly due to steric hindrances to the formation of a stable complex. The steric factor can also explain the increase in E_{tot} on passing from thia- to indo-dyes SQ1.1 and SQ1.2.

For cyanine dyes, the presence of bulky substituents (rings) in the terminal heterocycles of their molecules can increase the E_{tot} values, which may be interpreted as destabilization of complexes (dyes 1.2, 1.4, 2.3, 2.4, 2.7, 2.8, 4.3, 4.4). This effect is probably due to greater steric hindrances upon complexation than in the case of the dyes without additional cycles.

For oxa-dyes, on the whole, more negative E_{tot} values were obtained than for the corresponding thia- and indo-

© The Author(s) 2021. Published by Higher Education Press

analogues. As an illustration, it is possible to compare the energies of dyes 1.5 – 1.1 – 1.7 (monomethine cyanines), 2.5 – 2.1 – 2.10 (anionic carbocyanines), 3.4 – 3.1 – 3.6 (dicarbocyanines), 4.4 – 4.3 – 4.8 (tricarbocyanines). Possibly, this is also due to the more compact structure of oxa-dyes than thia and indo analogues.

Hence, from the criterion of E_{tot} value, the following dyes can be selected as promising: SQ1.7, 1.5, 1.6, 2.9, 3.4, 3.5, 4.6, as well as carbocyanines with 4 sulfonate groups 2.11–2.13.

When using dyes for the spectral detection of proteins, an important role is played by the change in the spectral-fluorescent properties of the dyes upon complexation. A sharp growth of fluorescence is observed for *meso*-substituted carbocyanines (trimethine cyanines) due to a shift in the dynamic *cis–trans* equilibrium, which allows detection of very low concentrations of biomolecules when the dyes are used as fluorescent probes [7,8,23,24].

Enhancement of the process of intersystem crossing to the triplet state leads to population of triplet energy levels of carbocyanines in complexes with biomolecules [25, 26]. This determines the possibility of photochemical reactions involving excited triplet states of the dyes in these systems and can potentially lead to the formation of reactive oxygen species (photodynamic effect). Thus, cyanines may be promising for light-induced damage to the components of the virus (proteins to which the dye binds), and thus inactivation of the virus itself. Note that UV photoinactivation of MERS-CoV and SARS-CoV-2 coronaviruses in blood serum by riboflavin is currently being studied [27, 28]. From the points of view considered, *meso*-substituted thiacarbocyanines may be promising for further practical research.

MATERIALS AND METHODS

The molecular docking was performed using DockThor [29,30]. We used the target protein structures proposed on the DockThor PDB website. The proteases NSP3 (PDB code 6w9c [31], NSP5 (main protease; PDB code 6lu7 [32]), and NSP12 (PDB code 7bv2 [33]) are "wild type" (isolated in 2019 in Wuhan, China). Hydrogen atoms were added to PDB structures and hydrogen bonds were optimized (pH 7); antibodies and non-protein molecules (water, ions, cofactors and ligands) were removed (DockThor). The proposed structures of proteins NSP3, NSP5, NSP12 were used "as is"; the additional adjustment of protonation of amino acid residues was not performed.

The experiments were carried out in the blind docking mode, the approximation grating size was 40 Å, the center was at x = –30.68, y = 30.43, z = 22.38, with a discretization step of 0.42 Å. When analyzing the results, the 6 best configurations were taken into account.

To create PDB structures of ligand dyes and optimize their geometry (MMFF94 force field), the Avogadro molecular editor was used [34,35]. The UCSF Chimera package [22,36] was used for 3D visualization and analysis of docking results.

Symmetric mono-, tri-, penta-, and heptamethine cyanine dyes, both cationic and anionic (having negatively charged sulfonate groups) with various heterocycles (benzimidazolyl, benzothiazolyl, benzoxazolyl) and substituents in heterocycles, were studied. Trimethine cyanines (carbocyanines) also had various substituents in the *meso* position of the polymethine chain (CH_3, C_2H_5, CH_3O, CH_3S, Cl), since *meso*-substituted thiacarbocyanines were previously characterized as effective probes for biomacromolecules [7,8,23,24]. Also studied were squarylium dyes, including squarylium indo- and thiacyanines with anionic sulfonate groups.

The stability of possible dye–protein complexes was judged by the sign and value of the total energy of the system E_{tot} obtained as a result of docking: it was supposed that the formation of stable complexes is more probable at sufficiently low (*i.e.*, large in absolute value) negative energies.

Data and materials availability

Research data is available upon request. The raw data was generated using DockThor. Evidence supporting the findings of this study is available from the respective author P.P. upon request.

AUTHOR CONTRIBUTIONS

P.P and A.T. contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

ACKNOWLEDGEMENTS

The work was carried out under the Russian Federation State Assignment (state registration No. 001201253314; Institute of Biochemical Physicas, Russian Academy of Sciences). Molecular graphics and analyses performed with UCSF Chimera are developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

COMPLIANCE WITH ETHICS GUIDELINES

The authors Pavel Pronkin and Alexander Tatikolov declare that they have no conflict of interests.

All procedures performed in studies were in accordance with the ethical standards of the institution or practice at which the studies were

© The Author(s) 2021. Published by Higher Education Press

conducted, and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

OPEN ACCESS

This article is licensed by the CC By under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

REFERENCES

1. Mosier-Boss, P. A., Lieberman, S. H., Andrews, J. M., Rohwer, F. L., Wegley, L. E. and Breitbart, M. (2003) Use of fluorescently labeled phage in the detection and identification of bacterial species. Appl. Spectrosc., 57, 1138−1144
2. Eriksson, M., Härdelin, M., Larsson, A., Bergenholtz, J. and Akerman, B. (2007) Binding of intercalating and groove-binding cyanine dyes to bacteriophage t5. J. Phys. Chem. B, 111, 1139−1148
3. Soto, C. M., Blum, A. S., Vora, G. J., Lebedev, N., Meador, C. E., Won, A. P., Chatterji, A., Johnson, J. E. and Ratna, B. R. (2006) Fluorescent signal amplification of carbocyanine dyes using engineered viral nanoparticles. J. Am. Chem. Soc., 128, 5184−5189
4. Robertson, K. L., Soto, C. M., Archer, M. J., Odoemene, O. and Liu, J. L. (2011) Engineered T4 viral nanoparticles for cellular imaging and flow cytometry. Bioconjug. Chem., 22, 595−604
5. Vus, K., Tarabara, U., Balklava, Z., Nerukh, D., Stich, M., Laguta, A., Vodolazkaya, N., Mchedlov-Petrossyan, N. O., Farafonov, V., Kriklya, N., *et al*. (2020) Association of novel monomethine cyanine dyes with bacteriophage MS2: A fluorescence study. J. Mol. Liq., 302, 112569
6. Shindy, H. A. (2017) Fundamentals in the chemistry of cyanine dyes: A review. Dyes Pigments, 145, 505−513
7. Tatikolov, A. S. (2012) Polymethine dyes as spectral-fluorescent probes for biomacromolecules. J. Photochem. Photobiol. C., 13, 55−90
8. Tatikolov, A. S., Pronkin, P. G., Shvedova, L. A. and Panova, I. G. (2019) Meso-substituted carbocyanines as effective spectral-fluorescent and photochemical probes for structurally organized systems based on macromolecules. Russ. J. Phys. Chem. B., 13, 900−906
9. Athar, F. and Beg, M. A. (2020) Anti-HIV and Anti-HCV drugs are the putative inhibitors of RNA-dependent-RNA polymerase activity of NSP12 of the SARS-CoV-2 (COVID-19). Pharm. Pharmacol. Int. J., 8, 163−172
10. Tazikeh-Lemeski, E., Moradi, S., Raoufi, R., Shahlaei, M., Janlou, M. A. M., Zolghadri, S. (2020) Targeting SARS-COV-2 non-structural protein 16: a virtual drug repurposing study. J. Biomol. Struct. Dyn., 39, 4633−4646
11. Al-Masoudi, N. A., Elias, R. S. and Saeed, B. (2020) Molecular docking studies of some antiviral and antimalarial drugs via bindings to 3CL-protease and polymerase enzymes of the novel coronavirus (SARS-CoV-2). Biointerface Res. Appl., 10, 6444−6459
12. Guedes, I. A., de Magalhães, C. S. and Dardenne, L. E. (2014) Receptor-ligand molecular docking. Biophys. Rev., 6, 75−87
13. Guedes, I. A., Costa, L. S. C., dos Santos, K. B., Karl, A. L. M., Rocha, G. K., Teixeira, I. M., Galheigo, M. M., Medeiros, V., Krempser, E., Custódio, F. L., *et al*. (2020) Drug design and repurposing with DockThor-VS web server focusing on SARS-CoV-2 therapeutic targets and their non-synonym variants. Sci. Rep., 11, 5543
14. Guedes, I. A., Barreto, A. M. S., Marinho, D., Krempser, E., Kuenemann, M. A., Sperandio, O., Dardenne, L. E. and Miteva, M. A. (2021) New machine learning and physics-based scoring functions for drug discovery. Sci. Rep., 11, 3198
15. Li, H., Sze, K. -H., Lu, G. and Ballester, P. J. (2020) Machine-learning scoring functions for structure-based drug lead optimization. WIREs Comput. Mol. Sci., 10, e1465
16. Li, J., Fu, A. and Zhang, L. (2019) An overview of scoring functions used for protein–ligand interactions in molecular docking. Interdiscip. Sci., 11, 320−328
17. Pantsar, T. and Poso, A. (2018) Binding affinity via docking: fact and fiction. Molecules, 23, 1899−1910
18. Devgan, M. (2015) Homology modeling and molecular docking studies of DNA replication licensing factor minichromosome maintenance protein 5 (MCM5). Asian J. Pharm. Tech., 5, 17−22
19. Selvaraj, G., Kaliamurthi, S., Çakmak, Z. E. and Çakmak, T. (2017) *In silico* validation of microalgal metabolites against Diabetes mellitus. Diabetes Mellitus, 20, 301−307
20. Tutorial 1. Redocking study of HIV-1 protease in complex with amprenavir PDB code 1HPV. Laboratório Nacional de Computação Científica – LNCC/MCTI. April 2020. http://www.dockthor.lncc.br. Accessed: 12 February, 2021
21. Khimenko, V., Chibisov, A. K. and Görner, H. (1997) Effects of alkyl substituents in the polymethine chain on the photoprocesses in thiacarbocyanine dyes. J. Phys. Chem. A, 101, 7304−7310
22. Chimera, U. C. S. F. Visualization System for Research and Analysis. Version 1.13. 1. http://www.rbvi.ucsf.edu/chimera. Accessed: 12 July, 2020
23. Tatikolov, A. S., Akimkin, T. M., Kashin, A. S. and Panova, I. G. (2010) Meso-substituted polymethine dyes as efficient spectral and fluorescent probes for biomacromolecules. High Energy Chem., 44, 224−227
24. Pronkin, P. G., Shvedova, L. A. and Tatikolov, A. S. (2020) Comparative study of the interaction of some meso-substituted anionic cyanine dyes with human serum albumin. Biophys. Chem., 261, 106378
25. Pronkin, P. G., Tatikolov, A. S., Sklyarenko, V. I. and Kuz'min, V. A. (2006) Photochemical properties of meso-substituted

thiacarbocyanine dyes in solutions and in complexes with DNA. High Energy Chem., 40, 252−258

26. Pronkin, P. G., Tatikolov, A. S., Sklyarenko, V. I. and Kuz'min, V. A. (2006) Quenching of the triplet state of meso-substituted thiacarbocyanine dyes by nitroxyl radicals, iodide ions, and oxygen in solutions and in complexes with DNA. High Energy Chem., 40, 403−409
27. Keil, S. D., Bowen, R. and Marschner, S. (2016) Inactivation of middle east respiratory syndrome coronavirus (MERS-CoV) in plasma products using a riboflavin-based and ultraviolet light-based photochemical treatment. Transfusion, 56, 2948−2952
28. Keil, S. D., Ragan, I., Yonemura, S., Hartson, L., Dart, N. K. and Bowen, R. (2020) Inactivation of severe acute respiratory syndrome coronavirus 2 in plasma and platelet products using a riboflavin and ultraviolet light-based photochemical treatment. Vox Sang., 115, 495−501
29. de Magalhães, C. S., Barbosa, H. J. C. and Dardenne, L. E. (2004) Genetic and Evolutionary Computation – GECCO 2004. Lecture Notes in Computer Science. Heidelberg: Springer
30. de Magalhães, C. S., Almeida, D. M., Barbosa, H. J. C. and Dardenne, L. E. (2014) A dynamic niching genetic algorithm strategy for docking highly flexible ligands. Inf. Sci., 289, 206−224
31. Osipiuk, J., Jedrzejczak, R., Tesar, C., Endres, M., Stols, L., Babnigg, G., Kim, Y., Michalska, K., Joachimiak, A. (2020) Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. Nat. Commun., 12, 743−752
32. Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., *et al*. (2020) Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. Nature, 582, 289−293
33. Yin, W., Mao, C., Luan, X., Shen, D. -D., Shen, Q., Su, H., Wang, X., Zhou, F., Zhao, W., Gao, M., *et al*. (2020) Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science, 368, 1499−1504
34. Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E. and Hutchison, G. R. (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J. Cheminform., 4, 17
35. Avogadro: open source molecule building and visualization tool. Version 1.2. 0. http://avogadro.cc. Accessed: 12 July, 2020
36. Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C. C., Pettersen, E. F., Goddard, T. D., Meng, E. C., Sali, A. and Ferrin, T. E. (2012) UCSF Chimera, MODELLER, and IMP: an integrated modeling system. J. Struct. Biol., 179, 269−278

© The Author(s) 2021. Published by Higher Education Press