Targeting SARS-CoV-2 Nsp12/Nsp8 interaction interface with approved and investigational drugs: an in silico structure-based approach

Ozal Mutlu\textsuperscript{a}, Osman Mutluhan Ugurel\textsuperscript{b,c}, Emrah Sarici\textsuperscript{d}, Oguz Ata\textsuperscript{e}, Tugba Gul Incib\textsuperscript{b}, Erennur Ugurel\textsuperscript{b}, Sinem Kocer\textsuperscript{f} and Dilek Turgut-Balik\textsuperscript{b}\textsuperscript{g}

\textsuperscript{a}Faculty of Arts and Sciences, Department of Biology, Marmara University, Istanbul, Turkey; \textsuperscript{b}Faculty of Chemical and Metallurgical Engineering, Department of Bioengineering, Yildiz Technical University, Istanbul, Turkey; \textsuperscript{c}School of Engineering and Natural Science, Department of Basic Science, Altinbas University, Istanbul, Turkey; \textsuperscript{d}Vocational School of Health Services, Medical Laboratory Techniques, Artvin Coruh University, Artvin, Turkey; \textsuperscript{e}School of Engineering and Natural Science, Department of Software Engineering, Altinbas University, Istanbul, Turkey; \textsuperscript{f}Faculty of Pharmacy, Department of Pharmaceutical Biotechnology, Istanbul Yeni Yuzuyil University, Istanbul, Turkey

Communicated by Ramaswamy H. Sarma

\textbf{CONTACT}
Dilek Turgut-Balik \textsuperscript{b}  \texttt{dilekbalk@gmail.com}  \textsuperscript{b} Faculty of Chemical and Metallurgical Engineering, Department of Bioengineering, Yildiz Technical University, Davutpasa Campus, Esenler, Istanbul 34210, Turkey

\textcopyright{} 2020 Informa UK Limited, trading as Taylor & Francis Group

\section{Introduction}

Novel coronavirus disease (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in approximately 19,000,000 confirmed cases and 700,000 deaths worldwide as of 5 August 2020 (World Health Organization, 2020a). It was announced by the WHO as a new pandemic on 11 March 2020, with rapidly increasing incidence of cases and deaths from all over the world (World Health Organization, 2020b). There is no specific treatment or drug developed against COVID-19 caused by SARS-CoV-2 currently (Wang et al., 2020). While scientists from all over the world are working to develop various treatment options, currently one of the fastest methods for identification of potent drugs for COVID-19 is drug repositioning of FDA-approved drugs, by targeting enzymes essential for the life cycle of SARS-CoV-2 (Zhou et al., 2020). Drug repositioning decreases cost, saves time and also decreases the risk of failure (Pushpakom et al., 2019). The efficacy of some FDA-approved drugs against coronaviruses like severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV) and SARS-CoV-2, have been tested via in silico studies (De Wilde et al., 2014; Pillaiyar et al., 2020; Wu, Liu, et al., 2020). Recently, repositioning of drugs like Remdesivir and methylcobalamin has been suggested as potential inhibitory drugs against SARS-CoV-2 (Chakrabortya, 2020; Narayanan & Nair, 2020; Zhou et al., 2020). However, new drug studies by repositioning still need to be performed until an effective drug against SARS-CoV-2 can be found.

Viral genome replication has one of the most important roles for SARS-CoV-2 life cycle as it polymerizes the RNA, and makes proofreading and final capping activities by using various viral and host proteins (Romano et al., 2020). Human coronaviruses including SARS, MERS and SARS-CoV-2 have 16 non-structural proteins (Nsps) encoded by open reading frames 1a and 1b (ORF 1a/1b) (Wu, Peng, et al., 2020), and conserved Nsps play a part in several coronavirus replication and transcription processes (Ahn et al., 2012; Chan et al.,...
Nsp12 is one of the conserved Nsp among coronaviruses (Gorbalenya et al., 2020). It has ‘capped right hand’ structure with the subdomains including a finger domain, a palm domain, a thumb domain and a N-terminal extension domain with the total of 932 amino acids in length (Kirchdoerfer & Ward, 2019; Posthuma et al., 2017; Ruan et al., 2020; Sexton et al., 2016). The N-terminal extension domain has a nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain and an interface domain (Kirchdoerfer & Ward, 2019). Viral growth is associated with NiRAN domain for SARS-CoV (Romano et al., 2020). Studies suggest that role of NiRAN domain for SARS-CoV-2 Nsp12 exhibits the structural property of kinase-like folds (Kirchdoerfer & Ward, 2019; Romano et al., 2020). Nsp12 is also part of the C-terminal domain of coronavirus replication/transcription complex, this is an RNA-dependent RNA polymerase (RdRb) domain which shows weak primer-dependent RNA synthesis activity without co-factors (te Velthuis et al., 2010). Nsp12 is a vital enzyme that functions as an RdRp (Perlman & Netland, 2009; Wu, Liu, et al., 2020), in addition, Nsp8 is also found to be associated with RdRp (te Velthuis et al., 2012). The role of Nsp8 in RdRp is defined as a primase for RNA synthesis (Perlman & Netland, 2009; te Velthuis et al., 2012). Structural analysis of Nsp8 revealed that it complexes with Nsp7 during their primer-dependent RdRp activity, in the study a Nsp7–Nsp8 complex of SARS-CoV protein structure is identified in 3D with 2.4 Å resolution (Subissi et al., 2014; te Velthuis et al., 2012; Zhai et al., 2005). Moreover, Nsp7 and Nsp8 are identified as co-factors of SARS-CoV Nsp12 that stimulate polymerase activity of Nsp12 by forming a complex, which is vital for SARS-CoV replication (Kirchdoerfer & Ward, 2019; Subissi et al., 2014). Nsp12 binds Nsp7 and Nsp8 and includes a kinase-like fold for SARS-CoV (Kirchdoerfer & Ward, 2019). The Nsp12–Nsp7–Nsp8 complex and its stoichiometry are confirmed for SARS-CoV along with some of their surface interactions (Kirchdoerfer & Ward, 2019). There are 96.35%, 98.8% and 97.5% similarities among Nsp12, Nsp7 and Nsp8 of SARS-CoV and SARS-CoV-2, respectively (Ruan et al., 2020). The Nsp12–Nsp7–Nsp8 complex is targeted for SARS-CoV inhibition studies and also targeted for SARS-CoV-2 (Ruan et al., 2020; Yin et al., 2020). The structure of Nsp12–Nsp7–Nsp8 along with the structure of an RdRp activity inhibitor Remdesivir is also studied for SARS-CoV-2 (Yin et al., 2020).

In this study, FDA-approved and investigational drugs were screened with wild-type and mutant SARS-CoV-2 Nsp12–Nsp8 interfaces, prioritizing high-affinity drugs, then complexes were approved by molecular dynamics (MD) simulations for their stabilities.

2. Materials and methods

2.1. Sequence analysis

SARS-CoV-2 reference genome sequence (NC_045512.2) (Wu, Peng, et al., 2020) was obtained from GenBank for this research (Benson et al., 2003), and all other SARS-CoV-2 isolated genome sequences with high coverage were downloaded from GISAID EpiCoV (https://www.epicov.org/) for comparative analysis.

In this study, the whole genome sequences of 30,366 SARS-CoV-2 isolates were aligned by Python language-based software ODOTool, which was developed as a strategy-based multiple sequence alignment tool and was previously tested for variation detection among SARS-CoV-2 genomes (Uğurel et al., 2020). The genotype differences between isolated genomes of SARS-CoV-2 were defined visually using the JalView v2.10.5 alignment viewer (Waterhouse et al., 2009).

Mutations that caused amino acid exchanges in essential genes for the life cycle of SARS-CoV-2 and present in more than 50% of the isolate genome sequences were identified. Reverse analysis has shown that two of the most common mutations are on the Nsp12 region and near the active site of RNA polymerase. Thus, this enzyme was designated as the target protein in this study. Sequences of SARS-CoV-2 reference Nsp12 amino acid sequence were collected from GenBank database and the nucleotide sequence mutant type was translated to protein sequence by ExPASY Translate tool (Gasteiger et al., 2003) using standard codon table (Peabody, 1989). Nsp12 wild-type protein sequences (Wuhan-Hu-1: YP_009725307.1) and the translated sequence of first mutant isolate (EPI_ISL_418218) were aligned using the European Bioinformatics Institute (EMBL-EBI) MUSCLE multiple sequence alignment tool (http://www.ebi.ac.uk/Tools/msa/muscle) (Edgar, 2004) to conduct multiple protein sequence analysis.

2.2. Virtual high-throughput screening of approved and investigational drugs

Structures of FDA-approved and investigational drugs were retrieved from the ZINC DB (Sterling & Irwin, 2015) and Pubchem/ClinicalTrials sources. Drug structures were prepared by LigPrep (LigPrep, Schrödinger, LLC, New York, NY, 2020) at neutral pH. SARS-CoV-2 RNA-dependent RNA polymerase (Nsp12) Cryo-EM structure with 2.90 Å resolution was obtained from the RCSB PDB (https://www.rcsb.org) with accession number of 6M71. Adding hydrogens, filling missing side chains, protonation at neutral pH and 0.5 Å restraint minimization was conducted by Protein Preparation Wizard (Protein Preparation Wizard, Schrödinger, LLC, New York, NY, 2020) from Maestro (Maestro, Schrödinger, LLC, New York, NY, 2020). Mutation (P323L) of the wild-type Nsp12 protein was created in Maestro program and mutant structure was prepared by following the same procedure as wild-type. The docking site, characterized as the Nsp12–Nsp8 interaction site, was set up using the Receptor Grid Generation Panel in Maestro. Virtual screening and molecular docking of drugs were carried out in the interaction interface by following Glide ‘Virtual Screening Workflow’ (Glide, Schrödinger, LLC, New York, NY, 2020) (Friesner et al., 2004; Halgren et al., 2004). Prime was used for MM-GBSA binding energy calculation (Jacobson et al., 2004) (Prime, Schrödinger, LLC, New York, NY, 2020) of the best poses selected based on XP GScores. Druggability of the Nsp12–Nsp8 interface was assessed by DogSiteScorer Server (Volkamer et al., 2012) from ProteinsPlus web portal (https://proteins.plus/). Nsp12/
Nsp8 complex interface residue interactions were assessed by the PDBePISA server (https://www.ebi.ac.uk/pdbe/pisa/pistart.html) (Krissinel & Henrick, 2007). ADME properties of RX-3117 were identified by using SwissADME (http://www.swissadme.ch/index.php) (Daina et al., 2017).

### 2.3. Molecular dynamics simulations

The charge parameters for all of the bonds, angles, dihedrals and atom types in the system of RX-3117 and Nebivolol were assigned using AM1-BCC atomic charge (Jakalian et al., 2002) and general AMBER force field (GAFF) parameters (Wang et al., 2004) via Antechamber (Wang et al., 2006). For the starting simulation, the topology and coordinate files of complexes were prepared using AMBER14 tLeap. A truncated octahedron box 5 Å in size was created in a periodic box of TIP3P water (Mark & Nilsson, 2001). The systems were neutralized by adding Cl⁻ and Na⁺ counter ions using a Coulombic potential on a grid. Energy minimization was performed to reduce bad interactions and clashes in the complexes before starting the simulation. The energy minimization was started with 10,000 cycles of steepest descent (Fletcher & Powell, 1963) and switch on conjugate gradient algorithms for 90,000 steps (Møller, 1993). Steepest descent algorithm quickly eliminates severe clashes and bad interactions. This is followed by conjugate gradient algorithms, this is more effective and slower at reaching an energy minimum. The AMBER ff14SB (Maier et al., 2015) force field was used for all simulations. The minimized systems were heated slowly from 10 K to 300 K with constant volume and pressure using Langevin thermostat (Davidchack et al., 2009) with a collision frequency of $\gamma = 10.0 \text{ ps}^{-1}$ for during 200 ps with 1 femtosecond steps. The complexes of RX-3117/Nsp12 wild-type, Nebivolol/Nsp12 wild-type, RX-3117/Nsp12 mutant, Nebivolol/Nsp12 mutant were simulated for 100 ns with a time step of 2 fs and hydrogen atoms constrained using the SHAKE algorithm (Ryckaert et al., 1977) was applied to restrain all atoms during dynamics run. The temperature during production runs was controlled using Berendsen weak-coupling algorithm (Berendsen et al., 1984). The coordinates of the system were written every 10 ps during 100 ns MD simulation. The simulation results were evaluated using the root mean square deviation (RMSD) and $\beta$-factor plot using CPPTRAJ (Roe & Cheatham, 2013) from MD simulation trajectory over 100 ns for complexes. Schrödinger Maestro (Maestro, Schrödinger, LLC, New York, NY, 2020) was used in 2D ligand-protein interaction analysis and PyMOL Molecular Graphics System (The PyMol molecular graphics system, version 2.0. Schrödinger, LLC, New York, NY, 2017) was used in 3D superimposition representations.

### 3. Results and discussion

#### 3.1. Sequence analysis

About 30,366 genome sequences isolated from all around the world were downloaded from either NCBI or GISAID on 20 May 2020 and deposited in a database as reported in one of our studies (Uğurel et al., 2020). These data were further analysed towards the repositioning of FDA-approved and investigational drugs against SARS-CoV-2.

The SARS-CoV-2 genome consists of 11 gene regions in total, named: ORF1ab, S, ORF3a, E, M, ORF6, ORF7a, ORF7b, ORF8, N and ORF10 in reference genome (NC_045512.2) (Wu, Peng, et al., 2020). Although there are many studies concerning variations across whole SARS–CoV-2 genome, few of them showed the structural impact on proteins (Gao et al., 2020; Kirchdoerfer & Ward, 2019; Shannon et al., 2020). In our study, we focused on the Nsp12 variations that alter the codon sequence and cause amino acid exchanges on the protein sequence. Three variations were found on the gene sequence encoding Nsp12: C14408T (first reported by Pachetti et al., 2020), C14805T and C15324T. These were present in all 30,366 isolates at frequencies of 70.42%, 10.01% and 2.4%, respectively (Uğurel et al., 2020). Only C14408T variations caused amino acid exchange (P323L) on Nsp12 of SARS-CoV-2, shown in Figure 1. Studies on SARS-CoV have shown that Nsp12 is a vital protein for replication and pathogenesis and also a potential target for antiviral development studies (Perlman & Netland, 2009; Wu, Liu, et al., 2020; Yin et al., 2020). Therefore, in this study, the effects of this amino acid exchange on the Nsp12 was further evaluated by structural studies to determine whether Nsp12 with and without this specific mutation in SARS-CoV-2 could possibly be targeted by the repositioning of the FDA-approved and investigational drugs to possibly cure COVID-19 infections.

#### 3.2. Virtual high-throughput screening of approved and investigational drugs

A drug library containing both FDA-approved and investigational drugs currently in clinical trials was screened for potential interaction within the Nsp12–Nsp8 interface region. Mutation of P323L on Nsp12 was found near to the interaction interface where Nsp8 protein binds and residues between 100 and 124 of Nsp8 forms an $\alpha$-helix-loop-$\alpha$-helix structure in the interaction site (Figure 2). Furthermore, based on PDBePISA interface results, P323 forms H bond with Asn118 of Nsp8 structure. Previous studies on SARS-CoV Nsp12/Nsp8 complexes showed that mutations of D99A, P116A, P183A, R190A in Nsp8 caused a decrease in primer extension and de novo polymerase activities (Subissi et al., 2014). The decrease in enzymatic activity was explained by the loss of Nsp12–Nsp8 interaction due to mutations in these critical and essential interface residues (Subissi et al., 2014). For these reasons, studies should be addressed in order to reveal the impact of P323L mutation on Nsp12–drug interactions and Nsp12/Nsp8 complex formation. Therefore, the Nsp12–Nsp8 interaction cleft where the P323L mutation was found is a significant region for docking of drugs which aim to disrupt the Nsp12–Nsp8 interaction. According to the DogSiteScorer server results, both wild-type and mutant Nsp12 structures had one of the highest drug (0.83) and simple (0.44 and 0.42 for the wild-type and mutant, respectively) scores for the Nsp8 interaction pocket when compared with the other site on the Nsp12 enzyme. This also emphasizes
the significance and the vulnerability of the Nsp12–Nsp8 interaction interface which forms a groove and allows small chemicals to bind.

Based on virtual high-throughput screening of FDA-approved and investigational drugs on wild-type and mutant Nsp12 protein, drugs with XP GScores lower than 7.0 kcal/mol was evaluated as possible inhibitors (Tables 1 and 2).

All drugs occupied the interface pocket where Nsp8 protein binds. The pocket itself was surrounded mainly by hydrophobic residues including Valine, Leucine, Phenylalanine, Alanine, Tyrosine and Proline. Nebivolol, RX-3117 (fluorocyclopentenyl cytosine), Fenoterol and Glucocorticoid interacted with both wild-type and mutant Nsp12 structures. The common residues interacting with the most drugs were found to be Thr324, Phe326, Pro328 and Val330, which also interacted with the Nsp8 protein (Figure 3). Only Fenoterol was shown to form H bond with the Pro323 residue in wild-type Nsp12 structure. Also, Fenoterol made H bonds with Leu270, Phe326 and Val330 in both structures. Moreover, Nebivolol made H bonds with Thr324, Phe326 and Pro328 of both mutant and wild-type proteins. However, H bonding interaction of Nebivolol with Val330 has been seen in only mutant Nsp12 structure. RX-3117 made a π–π stacking interaction with Tyr272 of wild-type Nsp12. Additionally, RX-3117 interacted with Leu270, Thr324, Pro328, Val330, and Leu270, Phe326, Thr324, Val330 of wild-type and mutant Nsp12 structures, respectively. Docking position of the Glucocorticoids was nearly same in both structures and made H bond with the same residues including Thr324 and Pro328. When comparing these four drugs, Nebivolol had the lowest MM-GBSA dG binding energy and therefore was the most leading drug in the pocket based on the docking results. Nebivolol is a β1-adrenoreceptor antagonist and carries highly selective properties, it is generally used for the treatment of hypertension patients (Giles et al., 2017; Marketou et al., 2017; Moen & Wagstaff, 2006). To the best of our knowledge, any antiviral effects of Nebivolol have not been reported yet. On the other hand, RX-3117 (fluorocyclopentenyl cytosine) is a fluorocyclopentenyl cytosine derivative which is in phase 2 research for advanced bladder cancer and pancreatic cancer. RX-3117 was designed from cytosine analogues that were shown to be active against both DNA and RNA viruses including HIV, West Nile virus, HSV-1 and cytomegalovirus (Choi et al., 2012; Marquez et al., 1988; Song et al., 2001). ADME properties of the investigational compound RX-3117 were analysed by SwissADME (http://www.swissadme.ch/index.php) (Daina et al., 2017). ADME results showed that molecular weight of RX-3117 is 257.22 g/mol, topological polar surface area is 121.60 Å², number of rotatable bonds is 2, number of hydrogen bond donors is 4, number of hydrogen bond acceptors is 6, Log Po/w (predicted octanol/water partition coefficient) is 1.29, LogS, (predicted aqueous solubility) is 0.01, Gastrointestinal absorption is high, Log Kp (skin permeation) is 9.74 cm/s, Lipinski violations is zero. These results were found to appropriate for Lipinski’s rule of five and Veber’s rule (Lipinski, 2004; Veber et al., 2002), therefore the compound could be identified a drug-like compound.

Fenoterol is a 4-hydroxyphenyl derivative of metaproterenol and therefore the drug is chemically similar to orciprenaline. The compound is a β2-adrenoceptor antagonist used for the treatment of asthma, bronchiolitis and other respiratory diseases (Heel et al., 1978; Svedmyr, 1985). Although fenoterol might cause some adverse effects based on drug doses (Wong et al., 1990), the drug was applied for asthma treatment (Heel et al., 1978; Svedmyr, 1985). Glucocorticoids belong to the group of steroid hormones and are clinically used based on anti-inflammatory, immunosuppressive and
anti-allergic properties (Jiang et al., 2015). Glucocorticoids also decrease inflammation of the airway and are therefore used for the treatment of respiratory diseases like asthma (Thomas et al., 2014; Wood & Barnes, 1995). However, this hormone increases the infection rate of respiratory viruses and an in vitro study showed that additional therapy would be necessary for this treatment (Thomas et al., 2014).

Figure 2. Cartoon view of the Nsp12–Nsp8 interaction interface in wild-type (a and b) and mutant structures (d and e). Surface view of the wild-type (c) and mutant (f) Nsp12 with Nsp8 structure. Leu323 and Pro323 are shown as red stick structures.

Other than the drugs which were found to be interacting with both wild-type and mutant structures, tetrahydourididine, dobutamine and saquinavir were found to interact only with the wild-type Nsp12 residues including Thr324, Phe326, Pro328, Val330 and Ser664 and had the best XP GScores among others. Tetrahydourididine is a competitive cytidine deaminase inhibitor and was analysed for treatment of sickle
cell disease and some types of cancers due to the compound’s effect in gene regulation (Desimone et al., 1985; Funamizu et al., 2012; Molokie et al., 2017). In addition, it was observed that tetrahydrodouridine had an enhanced antiviral effect on Herpes Simplex Virus (North & Mathews, 1981). Saquinavir is an antiretroviral protease inhibitor that is an aspartic acid derivative and a FDA-approved drug used in the treatment of HIV (National Center for Biotechnology Information, 2020a). Saquinavir has been proposed for the inhibition of SARS-CoV-2 Nsp14 (Liu et al., 2020), Nsp12–Nsp7–Nsp8 (Ruan et al., 2020) and S protein, RdRp (nsp12 and nsp8), exonuclease (nsp14), 2′-O methyltransferase (nsp16), helicase (nsp13) and Nsp10 (Xu et al., 2020). Dobutamine is a synthetic catecholamine and FDA-approved direct-acting inotropic agent (National Center for Biotechnology Information, 2020b). The antiviral properties of dobutamine were screened in silico for the SARS-CoV-2 main protease (Farag et al., 2018) and for the nucleocapsid (N) protein of porcine epidemic diarrhoea virus which belongs to Coronaviridae family (Deejai et al., 2017). In mutant Nsp12, Canagliflozin had the best XP GScore when compared with other drugs in mutant Nsp12–Nsp8 interaction interfaces and interacted with Val330 in this region. Canagliflozin was approved by the FDA in 2013 and is used in the treatment of type II diabetes. This drug is a C-glycosyl derivative compound and acts as a sodium-glucose transport protein subtype 2 (SGLT2) inhibitor (Kaushal et al., 2014). Muscle loss is common in patients with type II diabetes, therefore

| Drug name         | Glide emodel | XP GSscore | XP HBond | XP LipophilicEvdW | XP Electro | MM-GBSA dG bind |
|-------------------|--------------|------------|----------|-------------------|------------|-----------------|
| Wild-type Nsp12   | 37.732       | -8.41      | -4.02    | -1.461            | -1.128     | -38.08          |
| Dobutamine        | -55.552      | -8.213     | -2.33    | -4.263            | -1.233     | -56.45          |
| Saquinavir        | -87.088      | -7.887     | -1.906   | -5.192            | -1.058     | -60.94          |
| Nebivolol         | -57.774      | -7.699     | -1.625   | -4.719            | -1.057     | -55.64          |
| RX-3117           | -47.046      | -7.386     | -2.446   | -2.149            | -0.877     | -33.16          |
| TG100-115         | -55.208      | -7.307     | -1.373   | -3.175            | -0.62      | -35.72          |
| NOP-1A            | -59.667      | -7.264     | -1.33    | -3.709            | -0.997     | -30.27          |
| Fenoterol         | -53.694      | -7.101     | -2.109   | -3.281            | -1.385     | -41.11          |
| Glucocorticoid    | -61.155      | -7.094     | -2.105   | -4.191            | -1.059     | -53.11          |
| Thiabarine        | -46.843      | -7.088     | -2.36    | -1.837            | -0.969     | -38.86          |

| Drug name         | PubChem/ZINC ID | Function                                           | Current state                        |
|-------------------|-----------------|----------------------------------------------------|-------------------------------------|
| Epicatechin       | 182232          | Natural phenolic antioxidant.                     | Investigational, Phase I-II         |
| Nebivolol         | 71149993        | Pre-diabetic                                       | Approved (FDA)                      |
| Ipragliflozin     | 10453870        | ß-1 adrenergic receptor antagonist.                | Investigational, Approved in Japan and Russia |
| Fenoterol         | 3343            | ß-2 adrenergic                                    | Investigational, Approved in Japan and Russia |
| Glucocorticoid (SCHEMBL291203) | 11502129 | Anti-inflammatory, immunosuppressive and anti-allergic activities | Investigational, Phase I-IV |
| Tetrahydrouridine | 9543506         | Cytidine deaminase inhibitor.                      | Investigational                      |
| TG100-115         | 10427712        | PI 3-kinase inhibitor                              | Investigational, Phase I-II         |
| Secondary metabolite of Morus bark | 5,42,22,687 | N/A                                                | Investigational                      |
| Thiabarine        | 168566          | Antitumor activity on solid tumors                 | Investigational                      |
| NOP-1A            | 52951810        | Imaging for substance-related disorder             | Investigational, Phase I            |
| Canagliflozin     | 24812758        | Inhibitor of sodium-glucose transporter 2 (SGLT2) with anti-hyperglycemic activity | Investigational, Phase I-II         |
| RX-3117 (fluorocyclopentenyl cytosine) | 11242315 | Anti-cancer activity                               | Investigational                      |
| Saquinavir        | ZINC000029416466 | HIV protease inhibitor                             | Approved                            |
| Dobutamine        | ZINC000000057278 | Adrenergic and dopaminergic agent Inotropic agent  | Approved                            |
Canagliflozin has been found to have a positive effect on the cardiovascular system in clinical trials in the following years. Canagliflozin has been started to be used for preventing muscle loss with FDA approval in 2018 (Jakher et al., 2019), to best of our knowledge there are no publications on the antiviral effects of Canagliflozin in the literature yet. In addition to Canagliflozin, Epicatechin (a natural phenol), Ipragliflozin (sodium-glucose cotransporter-2 inhibitor) and a secondary metabolite of Morus bark (PubChem ID 54222687) were also found to interact only with the mutant Nsp12–Nsp8 interaction cleft and interacted pocket residues including Phe326, Pro328 and Val330.

3.3. Molecular dynamics simulations

Structural dynamics of protein structures and mobility analysis is a crucial step in molecular dynamic methods and has an important role in determining structure-function relationships in proteins. For these analyses, the RMSD and root mean square fluctuations (RMSF) of the structures were calculated during the MD simulation (Martínez, 2015). Two ligands which have the lowest binding energy in both proteins were selected from the library and docked to wild and mutant (P323L) forms of Nsp12 protein. The complexes (RX-3117/Nsp12 wild-type, Nebivolol/Nsp12 wild-type, RX-3117/Nsp12 mutant, Nebivolol/Nsp12 mutant) were simulated for 100 ns using MD methods in order to analyse the structural mobility of these complexes and observe any differences between both structures. For this purpose, the simulation trajectories were processed, and the RMSD and atomic fluctuation (β-factors) of complexes were plotted (Figure 4). The average deviation of RX-3117/Nsp12 wild-type and Nebivolol/Nsp12 wild-type complexes during 100 ns were calculated as 2.44 Å and 2.28 Å, respectively. Likewise, the average deviation of RX-3117/Nsp12 mutant and Nebivolol/Nsp12 mutant complexes during 100 ns was calculated as 2.29 Å and 2.51 Å, respectively.
Complexes without the Nebivolol/Nsp12 mutant complex structurally stabilized after about 20 ns. The deviation of Nebivolol/Nsp12 mutant complex increased dramatically after 20 ns but reached equilibrium after 50 ns. Also, the deviation of RX-3117/Nsp12 mutant complex increased after 90 ns and began to equilibrate again. The calculated average deviation of this complex is the lowest compared to other complexes as 2.29 Å, and the 3D structure is analysed to determine why this occurs after the simulation at 90 ns (Figure 4(a)). This graph demonstrates whether the compounds that are docked to proteins remain in the binding site and their binding positions. In general, it can be concluded that all complexes are stable for 100 ns, given the initial coordinates. The β-factor plots refer to the average displacement of the Cα
backbone carbon atom of each residue according to the initial structure of a particular atom. Generally, \( \beta \)-factors calculated from X-ray experiments are also called temperature factors (Bormot et al., 2011). These values obtained in this study were calculated from the RMSF values from molecular dynamic simulations (Figure 4(b)). The highest fluctuation in the region between residues 268–275 was observed in the Nebivolol/Nsp12 wild-type complex. The ligand Nebivolol is a larger compound than the other ligands and therefore caused greater change in the cavity. Since the region between residues 268–275 did not form a secondary structure, Nebivolol could be easily displaced. It was observed that local fluctuations were higher in both the wild-type and mutant protein to which Nebivolol was bound (Figure 4(b)). RX-3117 was docked to the hydrophobic cavity of the Nsp12 wild-type protein using molecular docking methods (Figure 5(a)). The ligand formed two hydrogen bonds with Val330 as well as a hydrogen bond with both Thr-324 and Pro-328 after 100 ns simulation, and the first position of RX-3117 remained stable (Figure 5(b)). However, the ligand lost

Figure 6. Two-dimensional representation of protein–ligand interactions. (a) RX-3117/Nsp12 mutant complex in the initial position, (b) RX-3117/Nsp12 mutant complex after 100 ns MD simulation, (c) Nebivolol/Nsp12 mutant complex in the initial position, (d) Nebivolol/Nsp12 mutant complex after 100 ns MD simulation.
the hydrogen bond with Leu270 and moved closer to the hydrophobic pocket. Nebivolol was also placed in the hydrophobic cavity using molecular docking methods (Figure 5(c)). After 100 ns simulation, it formed hydrogen bonds with Thr324 and also expanded the hydrophobic cavity by pushing the region between the residues 268–275 outwards (Figure 5(d)). Therefore, the ligand had lost hydrogen bonds with Pro328 and Phe326. It was revealed that both ligands were stable, forming hydrogen bonds in the hydrophobic cavity in Nsp12 wild-type protein after simulation of 100 ns MD simulation.

The mutation in the Nsp12 protein converted Proline-323 to Leucine, and this mutation was located inside the hydrophobic cavity in the ligand-binding region. RX-3117 was docked to the hydrophobic cavity using molecular docking methods, it did not interact with the Leu323 and was entirely located in the cavity (Figure 6(a)). After 100 ns MD simulation, the ligand formed hydrogen bonds with Thr324 (a polar residue) but Val330 and Leu323 did not interact with the ligand due to its location inside the cavity (Figure 6(b)). Besides, the hydrogen bond with Leu270 and Phe326 was not preserved after 100 ns. Likewise, Nebivolol was docked to the same cavity by molecular docking methods (Figure 6(c)). Nebivolol formed a hydrogen bond with Thr324 and Pro328 and remained stable in the binding position (Figure 6(d)).

The effect of the mutation on Nsp12 protein 3D structure and ligand binding sites was investigated in terms of dynamic behaviour of the mutant protein with respect to the wild-type protein. Proline, which is located at position 323 in the wild-type Nsp12 protein, has a different cyclic structure compared to other amino acids. The side chain is bound to the protein backbone twice and therefore there is no alpha hydrogen that can form a hydrogen bond in the structure. Therefore, secondary structures in most proteins do not involve proline, and proline limits alpha-helix structures (Lehninger, 2004). Leucine in the mutant Nsp12 protein has an aliphatic side chain and is a hydrophobic residue. They
are generally located within alpha-helix secondary structures in the hydrophobic cavities (Betts & Russell, 2003).

The positions of the ligands after 100 ns MD simulation are given in comparison with the wild-type and mutant Nsp12 protein. RX-3117 maintained its position in both proteins, but the secondary structure of mutated region was changed. For the reasons described above, since there was a Leucine instead of Proline in the hydrophobic pocket in the mutant Nsp12 protein, the boundary of the α-helix secondary structure had disappeared and the structure had become flexible through the prediction of local structures. Proteins, 79(3), 839–852. https://doi.org/10.1002/prot.22922

Chakraborty, H. J. (2020). Drug Repurposing against SARS-CoV-2 RDRP – A computational quest against CoVid-19. Nature Research, 1–19. https://doi.org/10.21203/rs.3.rs-22079/v1

Chan, J. W. F., Lau, S. K. P., To, K. K. W., Cheng, V. C. C., Woo, P. C. Y., & Yuen, K. Y. (2013). Middle East respiratory syndrome coronavirus: Another zoonotic betacoronavirus causing SARS-like disease. Clinical Microbiology Reviews, 28(2), 465–522. https://doi.org/10.1128/CMR.00102-14

Choi, W. J., Chung, H.-J., Chandra, G., Alexander, V., Zhao, L. X., Lee, H. W., Nayak, A., Majik, M. S., Kim, H. O., Kim, J.-H., Lee, Y. B., Ahn, C. H., Lee, S. K., & Jeong, L. S. (2012). Fluorocyclopentenyl-cytosine with broad spectrum and potent antitumor activity. Journal of Medicinal Chemistry, 55(9), 4521–4525. https://doi.org/10.1021/jm3004009

Daina, A., Michielen, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports, 7, 42717. https://doi.org/10.1038/srep42717

Davidchack, R. L., Handel, R., & Tret’yakov, M. V. (2009). Larginve thermo-stat for rigid body dynamics. Journal of Computational Physics, 130(23), 1–15. https://doi.org/10.1063/1.3149788

De Wilde, A. H., Joehmans, D., Posthuma, C. C., Zevenhoven-Dobbe, J. C., Van Nieuwkoop, S., Bestebroer, T. M., Van Den Hoogen, B. G., Nteyts, J., & Snijder, E. J. (2014). Screening of an FDA-approved compound library identifies four small-molecule inhibitors of Middle East respiratory syndrome coronavirus replication in cell culture. Antimicrobial Agents and Chemistry, 58(8), 4875–4884. https://doi.org/10.1128/AAC.03011-14

Deejai, N., Roshorm, Y. M., & Kubera, A. (2017). Antiviral compounds against nucleocapsid protein of porcine epidemic diarrhea virus. Animal Biotechnology, 28(2), 120–130. https://doi.org/10.1080/10495398.2016.1232268

Desimone, J., Heller, P., Molokie, R. E., Hall, L., & Zwiers, D. (1985). Tetrahydrouridine, cytidine analogues, and hemoglobin F. American Journal of Hematology, 28(3), 283–288. https://doi.org/10.1002/ajh.2830180310

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Farag, A. B., Wang, P., Ahmed, M. S., & Sadek, H. A. (2018). Corresponding authors. In J. Collado-Vides & R. Hofstädter (Eds.), Gene regulation and metabolism. MIT Press. https://doi.org/10.7551/mitpress/3215.003.0017

Fletcher, R., & Powell, M. J. D. (1963). A rapidly convergent descent method for minimization. The Computer Journal, 6(2), 163–168. https://doi.org/10.1093/comjnl/6.2.163

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., & Shenkin, P. S. (2004). Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. Journal of Medicinal Chemistry, 47(17), 1739–1749. https://doi.org/10.1021/jm0306430

Funamizu, N., Lacy, C. R., Fujita, K., Furukawa, K., Misawa, T., Yanaka, K., & Manome, Y. (2012). Tetrahydrouridine inhibits cell proliferation through cell cycle regulation regardless of cytidine deaminase expression levels. PLoS One, 7(5), e37424. https://doi.org/10.1371/journal.pone.0037424

References

Ahn, D. G., Choi, J. K., Taylor, D. R., & Oh, J. W. (2012). Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. Archives of Virology, 157(11), 2095–2104. https://doi.org/10.1007/s00705-012-1404-x

Benson, D. A., Karsh-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2003). GenBank. Nucleic Acids Research, 31(1), 23–27. https://doi.org/10.1093/nar/gkg057

Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. The Journal of Chemical Physics, 81(8), 3684–3690. https://doi.org/10.1063/1.448118

Betts, M. J., & Russell, R. B. (2003). Amino acid properties and consequences of substitutions. In M. R. Barnes & I. C. Gray (Eds.), Bioinformatics for geneticists (pp. 289–316). Wiley. https://doi.org/10.1002/0470867302.ch14

Bornot, A., Etchebest, C., & De Brevern, A. G. (2011). Predicting protein flexibility through the prediction of local structures. Proteins, 79(3), 839–852. https://doi.org/10.1002/prot.22922

Acknowledgements

Numerical calculations including MD simulations were carried out at TUBITAK ULAKBIM (Turkey), High Performance and Grid Computing Center (TRUBA).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Ozal Mutlu https://orcid.org/0000-0003-4551-5780
Osman Mustułan Uğurel https://orcid.org/0000-0002-5365-0950
Emrah Sarýer https://orcid.org/0000-0003-1721-0314
Oğuz Ata https://orcid.org/0000-0003-4511-7694
Tugba Gul Inci https://orcid.org/0000-0002-2801-8021
Erennur Ugurel https://orcid.org/0000-0002-5504-660X
Sinem Kocer https://orcid.org/0000-0003-0517-7422
Dilek Turgut-Balık https://orcid.org/0000-0003-2417-4104

O. MUTLU ET AL.
