Quantum Deep Learning Functional Similarities on Remdesivir, Drug Synergies to Treat COVID-19 in Practice.

Ioannis Grigoriadis (✉ jgrigoriadis@biogenea.gr)

Biogenea Pharmaceuticals Ltd

Research article

Keywords: Deep Learning, Quantum mechanics, Colchicine, Remdesivir, Ursolic acid, COVID-19

Posted Date: December 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-127218/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Novel SARS coronavirus 2 (SARS-CoV-2) of the family Coronaviridae starting in China and spreading around the world is an enveloped, positive-sense, single-stranded RNA of the genus betacoronavirus encoding the SARS-COV-2 (2019-NCOV, Coronavirus Disease 2019). Remdesivir drug, or GS-5734 lead compound, first described in 2016 as a potential anti-viral agent for Ebola disease and has also being researched as a potential therapeutic agent against the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the coronavirus that causes coronavirus disease 2019 (COVID-19). Computer-aided drug design (CADD), Structure and Ligand based Drug Repositioning strategies based on parallel docking methodologies have been widely used for both modern drug development and drug repurposing to find effective treatments against this disease. Quantum mechanics, molecular mechanics, molecular dynamics (MD), and combinations have shown superior performance to other drug design approaches providing an unprecedented opportunity in the rational drug development fields and for the developing of innovative drug repositioning methods. We tested 18 phytochemical small molecule libraries and predicted their synergies in COVID-19 (2019- NCOV), to devise therapeutic strategies, repurpose existing ones in order to counteract highly pathogenic SARS-CoV-2 infection and the BRD4- conserved residue associated COVID-19 pathology. We anticipate that our quantum deep learning similarity approaches can be used for the development of anticoronaviral drug combinations in large scale HTS screenings, and to maximize the safety and efficacy of the Remdesivir, Colchicine and Ursolic acid drugs already known to induce synergy with potential therapeutic value or drug repositioning to COVID-19 patients.

Introduction

Ancient Indian medicinal scriptures including Athurveda, Phyllanthus emblica, Rig-Veda, Phyllanthus emblica, Vitis vinifera, and (1) Charka Sanhita have been widely used since ancient times are and are well known for boosting the immune system. They are a remarkable natural source of (1,2) high value alkaloids, phenols, flavonoids, coumarines, chalcones, lignans, alkanes, polyketides, alkenes, simple aromatics, Emblicanin B, alkynes, Vitilagin, peptides, Chebulanin, terpenes, Punigluconin and steroids for the treatment of various human ailments and for the increase of the human immunity (1,2) and longevity. In the current era of computer-aided drug discovery, enormous medicinal drug likeness properties of ancient plants allows the pharmacists and researchers to exclusively use them for the In-Silico discovery of drug-like natural molecules. Ursolic acid (3- β -3-hydroxy-urs-12-ene-28-oic-acid) and oleanolic acid (3- β -hydroxyolean-12-en-28-oic acid) are pentacyclic triterpenoid compounds and have been isolated from diverse plants, such as Chamaedorea tepejilote [1], Ledum groenlandicum [2], Lantana hispida [3], and Uncaria rhynchophylla [4], with a widespread occurrence throughout the plant kingdom (1,2). Their antibacterial [1,2,4], antiparasitic [2,3,4], antiviral [1,3,4], antitumor, antioxidant, anti-inflammatory [4], and hepatoprotective therapeutic properties and biological activities have been reported and in-vitro evaluated in human macrophage-models at inhibitory concentrations. Bromodomain protein 4 (BRD4) is a member of the extraterminal (BET) family proteins that binds to acetylated histones through elongation factor b (P-TEFb). The bromodomains act as a reader in cell growth and a writer in cell cycle progression of

histone acetylation that plays important roles in different biological processes including memory formation, replication, mitochondrial oxidative phosphorylation, transcription, DNA repair and DNA damage response (5). Molecular structure can be determined in heterodox interpretations by solving the time-independent (11) Schrödinger equation: QM methods, (12) vertex prizes and edge costs including ab initio (13) Density Filed Theories (DFT) and semi-empirical in place of the quantum processor and (14) energy among other observables, (15) under simulated (16) error as well as (17) to reposition drugs (18) about bonding may represent the (19) similarities and dissimilarities (20) between drugs and (21) repurposed viral (22) proteins respectively. (23) However, the Schrödinger equation (24) cannot actually be solved for any (25) but a one- data-driven (26) electron system methods (the hydrogen atom), (27) and approximations (28) need to be made. (29) In this research article (30) we present a drug- repositioning strategy and a Quantum (31,32,33,34,35,36,37) Deep Learning SARS-CoV-2 main protease, Mpro, DPP4 ADABP, ADCP2, CD26, DPPIV, TP103, FURIN, FUR, PACE, PCSK3, FURIN-ADAMTS1-ROR-GAMMA-SRRM2-ROR γ - NF- κ B/RelA-STAT3A-ROR γ - NF- κ B/RelA-STAT3B, Nsp15, TCEB2 $\rightarrow$ ASB8 $\rightarrow$ TCEB1, ACE2, ACEH angiotensin I converting enzyme 2, Remdesivir, Colchicine and Ursolic acid drugs network-based (13,14,15,18-38) prioritization method based on a (22,24,26,28-39) heterogeneous network (30,31,32,33,36-40) integrating similarity to (30,31,32,37,39,41) combine drugs and herbals that can be used against devastating diseases such as COVID-19. (12-42) This deep learning drug combination technology to predict new (17-43) therapeutic indications for drugs and novel treatments for diseases (18-44) has the potential to infer novel combined (2,3,5-17,18,22-45) treatments for COVID-19 diseases (19,21,25,28,30-46) as an evidence-based (19,32,33,34,37,45,46,47,48) drug repurposing AI-COVID-19 tool.

Materials And Methods

Sequences retrieval and alignment

All structural alignments were performed using the Dali server (<http://ekhidna2.biocenter.helsinki.fi/dali/>) (Holm, 2020). nCOVSNVs with ID NC_045512 were identified and cross-matched with BLASTp for the COVID-19 virus (SARS-CoV-2) receptor

binding domain (RBD) and subdomain-1 (319th to 591st aa) of the the spike glycoprotein [bat coronavirus RaTG13] with about 74% sequence identity in the single receptor-binding domain for the up configuration and the S1 protein partial [SARS coronavirus GD322] with 100% sequence identity. Total of 1652 SARS-CoV-2 S protein complete sequence alignments of COVID-19 RBD subdomain-1 (319th to 591st) amino acid available at the (2,3,7-39) NCBI Virus portal were retrieved. We then sought to rank the COVID-19 virus RBD subdomain-1 (319th to 591st) amino acid sequences based on the predicted probability that the RNA folds into the MEA structure with 73.96% sequence identity and not other structures (8,9,12-30) according to the patient's corresponding sample information, dates of sample collections, and geographical locations among others which have shown strong clinical evidence of the presence of SARS-CoV by immunohistochemistry, high-resolution oligonucleotide array, electron microscopy, and real-time reverse transcription-qRT-PCR.

Preparation of the protein structures

For this purpose, (1,2,4-6,23) we initially selected the three-dimensional structures of the non-structural proteins of the Nsp3, Nsp5 (PLpro domain), (1-5,6,7-48) Nsp12 (RdRp), the structural Spike proteins, the Nsp15 (endoribonuclease), and the nucleocapsid protein (N protein) between the SARS-CoV-2 Mpro structure and the SARS-CoV Mpro, PDB entry 6LU7 as the best-characterized drug targets among coronaviruses. (2,5,6,7,9-27) For the N protein, we clustered 31 conformations with (11,13,14-29,32) Glu174 present in an opened residue conformation out of a total of 40 ionization states (16,19,21-46,47,48) present in the globular cluster of five protease helices, which is involved through a salt-bridge interaction in modulating the homodimerization of the NMR-derived structure of the (PDB code 6YI3 49) Mpro, mainly between the binding cavities of the Arg4 of one protomer and the Glu290 domains of other to select a short linear peptide subset representative (22,24,26-42,43,46,48) within the protease flexibility. (17,18,21-43,48) The aliphatic carbon atoms from the the closed conformation of the Glu174/Glu166 side chain were prepared as a part of the phosphate binding site which was lead in the binding site to steric clashes with potential inhibitors by removing all water molecules using PyMOL software.

Collecting gold-standard pairwise drug combinations

In this study, we focused on pairwise drug combinations on clinically and FDA-approved investigational drugs by assembling the publicly available clinical data sources from the updated human interactome and the multiple commonly used small molecule databases. (23,24) Each approved drug in combinations was required to have the experimentally screening system (25,26) and validated COVID-19 protein target information based on GeneCards and on high-quality PPIs: each EC_{50} , IC_{50} , K_i , or $K_d \leq 10 \mu M$. (27,28) Compound name, gene expression data, metabolic associations, evolutionary analysis, commercial name or generic name of each drug and protein-drug datasets were included and (29,30) standardized by MeSH and UMLS vocabularies (22,28-47) and further transferred to DrugBank ID (30-32) from the DrugBank database (v4.3)40. (30-38, 40) Duplicated drug pairs were removed. (40, 42, 44) In total, 681 unique pairwise drug combinations by literature-derived low-throughput experiments (46,48) connecting 362 drugs were retained. We also compiled clinically reported adverse (30,32,37-47,48) drug–drug interactions (DDIs) data from the (38,39,40,42-48) the Therapeutic Target Database (TTD, v4.3.02)41, the PharmGKB database (December 30, 2015) and the FDA approved DrugBank databases using reported binding affinity data such as the dissociation constant (K_d), inhibition constant/potency (K_i), median inhibitory concentration (IC_{50}) $\leq 10 \mu M$ or median effective concentration (EC_{50}). Drug–Protein target interactions, (42-48) Chemical similarity analysis and (32,35,36-48) Collecting adverse drug–drug interactions of (54,58) FDA approved drug pairs were acquired from the publicly available DrugBank database (v4.3).

Gene Ontology (GO) similarity analysis and Clinical similarity of drug pair analysis

The Gene Ontology (GO) annotation for all the drug-gene targets of the SARS- CoV-2 main protease, Mpro, DPP4 ADABP, ADCP2, CD26, DPPIV, TP103, FURIN FUR, PACE, PCSK3,FURIN-ADAMTS1-ROR-GAMMA-SRRM2-ROR γ - NF- κ B/RelA-STAT3A- ROR γ - NF- κ B/RelA-STAT3B, Nsp15, TCEB2-->ASB8->>TCEB1 coding genes was downloaded from the website: <http://www.geneontology.org/>. We used three types of the

experimentally validated or (4,5,7-22,48) literature-derived evidences (32,33,36-48) biological processes (BP), (31,32-48) molecular function (MF), (37,39,42-45) and cellular component (CC), (37-48) excluding annotations inferred (36-47,48) computationally. The semantic comparison of GO annotations offers quantitative (32,33-45-48) ways to compute similarities between (22,27,20-32,41-45) genes and drug-gene (22,23,25,27-48) combination therapy products. The overall GO similarity of the drug target-SARS-CoV-2 main protease, Mpro, DPP4 ADABP, ADCP2, CD26, DPPIV, TP103, FURIN FUR, PACE, PCSK3, FURIN-ADAMTS1-ROR-GAMMA-SRRM2-RORY- NF-κB/RelA-STAT3A-RORY- NF-κB/RelA-STAT3B, Nsp15, TCEB2->ASB8->>TCEB1 coding genes binding to the Remdesivir, Colchicine and Ursolic acid drugs was determined by Eq. $(x' = f(x, t, \eta), x(t_0) = x_0(\eta) \quad D = \{(t_k, y \sim k)\}_{k=1}^{nt} \quad y \sim ik = y_i(t_k) + \epsilon_{ik}, \epsilon_{ik} \sim N(0, \sigma_i^2 (y \sim ik - y_i(t_k))^2 \sigma_i^2, p(\theta|D) = p(D|\theta)p(\theta)p(D) \quad p(D|\theta) \frac{1}{\beta} p(\theta) \quad \sigma_{10} \sigma_{20}$, averaging all pairs of protein target-SARS-CoV-2 main protease, Mpro, DPP4, ADABP, ADCP2, CD26, DPPIV, TP103, FURIN FUR, PACE, PCSK3, FURIN-ADAMTS1-ROR-GAMMA-SRRM2-RORY- NF-κB/RelA-STAT3A-RORY- FURIN-LRP1-MMP17-MMP2-MMP25-NF-κB/RelA-STAT3B, Nsp15, TCEB2->ASB8->>TCEB1 coding genes and b with $\in$ and $\in \langle GO \rangle$ $= 1 \text{ pairs} \sum \{p(D|\theta) = \prod_{i=1}^n y \prod_{k=1}^{nt} \frac{1}{\sigma_i^2 \pi} \exp\}$, $GO(p(\theta|D) = p(D|\theta)p(\theta)p(D) \quad p(D|\theta) \frac{1}{\beta} p(\theta) \quad \sigma_{10} \sigma_{20})$ (6).

Preparation of the datasets with known e-Drug3D drugs

The dataset containing the (5,7,9-48) FDA-approved drugs and active metabolites was constructed from the e-Drug3D commercial small molecule and fragments of 1305 different FDA-approved drugs that currently contains 1519 annotated 3D structures of molecular weight <2000 that are represented by a single conformation dataset, (10,12,14-48) a dataset updated annually and freely available for the scientific community at <https://chemoinfo.ipmc.cnrs.fr/MOLDB/index.php> (13,15-48). In this In-Silico effort we prepared mol2 files of the essential dataset of the chemical structures from the two supplementary e-Drug3D collections which were carefully constructed and have been shown similar inhibition in infectious models and resulting from (i) the generation of multiple conformers of high-quality and curated structures of FDA-approved drugs with molecular weight less than 2000 for the: $\{[5-(1\lambda^4,3\lambda^4,7\lambda^4\text{-purin-6-ylsulfanyl})-1\text{-methyl-4H-}1\lambda^5\text{-imidazol-4-ylidene}]\text{(oxo)imino}\}$ - λ^1 -oxidanylidene-heptahydrogen ring systems and (ii) the calculation of the most probable ionic and tautomeric states at pH 7.4. Smaller drugs (backbone with less than 20 heavy atoms) were also prepared in the .pdb, .pdbqt and .mol2 format files (26,27-42,46-48) present in the e-Drug3D dataset. (28,30-48) The ensemble of 3D molecule conformations (31,32-45,48) of the larger drugs from e-Drug3D were provided on a separated drug library dataset.

Drugs under clinical trials (COVID-19 repurposing dataset)

Our approach was focused on identifying a cluster of (1-48) similar chemotypes followed by parallel docking grid generation using the advanced to the next MM-PBSA-WSAS, KNIME-HTS-HVS filter, and chemical structure preparation wizard as provided by the BiogenetoligandoroITM cluster of algorithms that had the potential to target the (2-48) FURIN-ADAMTS1-ROR-GAMMA-SARS-COV-2 conserved domains and fit the geometric constraints without any restraints or constraints when filling in the open valence of the SARS-COV-2-ACE2-RORY-BRD4-FURIN (1-30,41-48) binding pocket residues. These were

then (6-48) converted into substructure searches of one copy of SARS-CoV-2 main protease (7-45) which were used to mine commercially available (8-42) compounds and covalently bonded SARS-COV-2 inhibitors using the (9-32) eMolecules database. The dataset of the selected hit drugs followed the force field parameters as applied to the partial atomic charges of the selected ligands which were derived using the RESP and collected from published articles (3,5,7,13,48) and approved drugs listed on the (2-39) DrugBank database in the "Clinical Trial Summary by Drug" section to fit the HF/6-31G* electrostatic potentials generated using the Gaussian 16 software package. (4,6-18) We intended to generate these from two different branches to model the viral protein using the isotropic position scaling algorithm. (7,9-46,48) One branch of the selected hit elements was referred to as the "Remdesivir Literature Substructures" branch, (8,11-47) which was based on the Remdesivir, Colchicine and Ursolic acid substructures using the Antechamber module as extracted (15,46) from published bromodomain inhibitors.

Structure-Based Pharmacophore, Docking, Machine Learning (QSAR) Methods.

Molecular docking, quantum mechanical LigandorITM-inspired physarum-prize-collecting Neural Matrix Factorizations (12-13) and drug repositioning scoring and constrained analysis using the SHAKE algorithm were implemented (13-14) to a collection of the (14-15) ZINC databases. Protein-molecule complexes, (12-21). followed by structural relaxation (13-22) for the cocrystal ligand of SARS-CoV-2 main protease, N3 were generated through (17-18) flexible-ligand:rigid-receptor (19) The molecular (4) docking in this local energy minimization of the equations of motion was conducted at a time step of 1 fs to optimize (15) protein-molecule (22-24) interactions for the relaxation phase and 2 fs capping (24-27) for the N- and C-terminal of each (26) fragment. i-GEMDOCK (26-27) software deployed at 298 K, 1 bar for 10 ns for the equilibrium and sampling phases through cycles in amino-acids 15-23+ within 4 Å to calculate the full electrostatic energy of any (27-28) docked molecule. In addition, following our (37-38) static analysis, we conducted some (38-39) preliminary molecular dynamics studies (40-41) on a potential "latch" for the Down state protomer. (38) Explicit solvent molecular dynamics (MD) (39) in both the minimization and MD simulation stages of novel coronavirus spike protein (42-43) were performed (44-45) using the NAMD2 program (45). We used the CHARMM-Gui (46) with the CHARMM36m force field (47-48) along with TIP3P water molecules to explicitly (30-39,44) solvate the proteins of a unit cell in a macroscopic lattice and add any (32-43) missing residues of repeating images from the experimental structure atom files selected from the sampled snapshots. (45) Simulations were carried out (35-45) maintaining the number of simulated particles, (27-48) pressure and temperature (the NPT ensemble) (11-48) constant with the Langevin piston method (23-39) specifically used to maintain a constant pressure of 1 atm. (12-46) We employed periodic (18-40) docking boundary conditions to enrich repurposing drug candidates (20-43) using a more accurate and more efficient method, the BiogenetoligandorITMQMM-PBSA-WSAS pipeline for a water box simulation (43) volume as well as the particle nanomesh (24-45) Ewald (PME) method with a 20 Å cutoff distances (24-46) between the simulated SARS-COV-2 protein (16-47) and water box edge since the most chemical ligands and herbals comprising antiviral and antibacterial activities are large and consist of many rotatable bonds. (15) The integration time entropy step was 2 femtoseconds using a method coined WSAS (weighted solvent

accessible surface area) with our protein simulations (17-48) as conducted under physiological conditions (37 C, pH of 7.4, (69,70) physiological ionic strength) based on the MM-PBSA-WSAS binding free energies. To emphasize the (8,9,12-47) computational efficiency of the method (9,10-46) we would like to mention that the average computing time (12,14,15-48) for one solution was <10 s on a single GPU processor by using solvent assessable surface area, 0.054. (15,16,17-48) A significant variability in the solutions can be observed. (15,17,19-48). The docking calculation (20,21,23-48) for estimating the nonpolar solvation energy was performed on the prepared dataset of 195 SARS-COV-2 receptors and FDA ligands by using the BiogenetoligandoroITM Cluster of docking softwares and workflows based on default parameters.

Results

In this drug combination In-Silico project, a comparative parallel docking analysis of the S protein has been performed and the selected ligands that had docking scores better than -60.5 kcal/mol with one reference strain were selected as the hit candidates for each CoV type and thereby taking into account only the most prevalent residue harbored at the conserved given positions. To overcome this, we employed the conservation analysis using the total number of protein sequences treated for each CoV type, (Supp.Material), (Figures1a,1b,2,3a,3b,3c,3d,3e,4a,4b,4c,4d,4e,4f,4g,5a,5b,5c,5d), (Tables1a,1b,2,3a,3b) using the reliable BiogenetoligandoroITMQMM-PBSA-WSAS binding free energy HTS-HVS strategy. The Remdesivir, Colchicine and Ursolic acid drugs bind into the SD1 stands the most conserved druggable domain among all four Beta-CoVs analyzed. The Remdesivir small molecule when combined to the Ursolic acid and Colchicine chemical structures binds inside the amino acid 287 | LEU | A | 9000 | 6Q5 | A | 3.99 | 3849 | 200 | -19.472, 23.825, -15.927 | -15.959, 22.345, -17.092 | + | 323 | HIS | A | 9000 | 6Q5 | A | 3.56 | 3848 | 516 | -25.010, 21.669, -15.561 | -24.824, 21.905, -12.016 | (Supp.Material), (Figures1a,1b,2,3a,3b,3c,3d,3e,4a,4b,4c,4d,4e,4f,4g,5a,5b,5c,5d), (Tables1a,1b,2,3a,3b). Each separate analysis returned positive results (Tables1, 3a, 3b, 3c)(Figure2, and Figures3a, 3b, 3c, 3d, 3e, 4a, 4b), indicating the Recomborovir compounds we selected might directly inhibit 2019-nCoV. The Remdesivir, Colchicine and Ursolic acid small molecules targeted the PLpro (M2, M3, M7, M9, M10, M11 and M13) dindng domains in the amino acids of the | 164 | HIS | A | 5 | PJE | C | 2.16 | 3.07 | 153.73 || 2408 | N3 | 1266 | O2 | with the docking energy values of the -12.282, 14.994, 67.123 | -15.161,15.336, 68.144 | in the region between the thumb and palm of the amino acid of the | 144 | SER | A | 803 | DMS | A | | 3.66 | 3.99 | 102.68 | | 1114 | O3 | 4736 | O2 | with the docking energies of the 35.403, -33.742, -8.029 | 37.550, -32.180, -11.001 |, which might interfere with substrate entering this enzyme's active |166 | GLU | A | 803 | DMS | A | 5.01 | | Sulfonium | 4735 | 36.185, -32.686, - 7.387 | 36.922, -36.568, -4.314 | sites, located at the bottom of the two domains (29). This molecule combination named Recomborovir reported to inhibit the 3CLpro (M1, M2, M3, M4, M5, M7, M8, M10, M11, M12 and M13) mainly entered the region between domains 2 and 3, and this | 295 | ASP | A | 813 | DMS | A | 5.38 | | Sulfonium | 4739 | 35.134, -46.698, -27.502 | 30.989, -48.448, -24.558 | region is important for 3CLpro to form a dimmer (30). Recomborovir was reported to inhibit viral entry, accordingly it bound the fusion cone of spike protein; this cone structure is important for viral membrane fusion (31) (Supp.Material),

(Figures1a,1b,2,3a,3b,3c,3d,3e,4a,4b,4c,4d,4e,4f,4g,5a,5b,5c,5d), (Tables1a,1b,2,3a,3b,3c). The Ursolic acid when combined with the Remdesivir and Colchicine small molecules generated Hydrogen bonds, Salt bridges and Metal complexes of the POP (composite ligand, containing Diphosphate, dihydrogen) ION binding sites into the 7bv2 protein targets The Ursolic acid small molecule when combined with the Remdesivir chemical structure generated ****Salt Bridges**** into the amino acid of the 836 | ARG | A | 18 | U | P | 4.57 || Phosphate | with the docking energies of the 8240,8240,8243,8228,8241,8242 | 84.767, 82.285, 113.039 | 88.794, 80.678, 111.585 849 | LYS | A | 17 | U | P | 4.64 || 8220,8220,8205,8222,8223,8221 | 81.688, 84.935, 117.166 | 84.063, 86.886, 120.648 |. This drug combination of the Remdesivir, Colchicine and Ursolic acid generated ****pi-Cation Interactions**** inside the binding cavities of the amino acid of the 555 | ARG | A | 101 | F86 | P | 3.61 | 1.61 | | Aromatic | with the docking energies of the 8618,8608,8610,8621,8625 | 90.830, 92.165, 105.289 | 93.435, 92.714, 107.732 | and ****Metal Complexes****10 | G | P | 1005 | MG | A | 8606 | Mg | 8283 | O | 1 | 2.33 | | 0.00 | NA | 1 | with the docking values of the 92.442, 84.162, 100.658 | 90.927, 83.719, 102.373 |.The Recombovir drug combination targets the backbone of the Leu278- Gly279- Ile280 loop which is coordinating the P+2 (Asp) side chain, while P+1 (Ser) side chain is coordinated by the Ile280 carbonyl. This interaction is further strengthened by the backbone interaction between the amide hydrogen/carbonyl of the Recomborovir compounds at Gly284 and the P-3 (Ile) carbonyl/amide hydrogen. Figures1a,1b,2,3a,3b,3c,3d,3e,4a,4b,4c,4d,4e,4f,4g,5a,5b,5c,5d), (Tables1a,1b,2,3a,3b,3c).

Discussion

In this research report, we found that the Cluster of the Recombovir-(Drug Combination) which were identified during screening of a compound diversity set performed by the BiogenetoligandorolTM cluster of algorithms on the intersection track (Lys711 and Arg355/SARS-CoV2 PLpro and Lys711 and Arg355. The chemical structures of the Remdesivir, Colchicine and Ursolic acid) targeted into the Lys711 and Arg355 residues and inside the residues of the Phe19, Trp23, and Leu26, which are located in an alpha-helical region of the SARS- CoV2 PLpro N terminus that binds to the N-terminal Lys711 and Arg355 hydrophobic pocket (17). The druggable scaffold of this drug combination of the Remdesivir, Colchicine and Ursolic acid small molecules target into the binding domain of these three critical SARS-CoV2 PLpro residues; the combination of the hit compounds therefore competes with endogenous SARS-CoV2 PLpro for binding to Lys711 and Arg355. We created a new track to display the contacts of this drug combination with each of those: Lys711 and Arg355, and SARS-CoV2 PLpro. Interestingly, this drug combination consisted of the Remdesivir, Colchicine and Ursolic acid chemical structures targets the Lys711 and Arg355 homo-dimerization site and intersects within the Lys711 and Arg355-Recombovir-(Drug Combination) binding sites, suggesting that they may also interfere within the binding pockets of the Lys711 and Arg355 homodimerizations. The key residues of the seven hot spot residues that contribute to S spike glycoprotein trimerization can be potentially In-Silico down-regulated by the docking combination of the therapeutic agents of the Remdesivir, Colchicine and Ursolic acid drugs in order to significantly block the quaternary structure assembly of the SARS-CoV-2 replication protein. Furthermore, the combination of the drugs of the Remdesivir, Colchicine and Ursolic acid targets the BRD4, JQ-1,

ISG15, IFN- β , IL-1 β , I-BET 151 and OTX-015 exhibited viral inhibition with a short hairpin RNA (shRNA; shBRD4) which is involved in the recognition of molecular patterns and mediate severe inflammatory responses, may also interact within the binding domains of the S protein via 10 active residues located in the S1 subunit. I also suggested that the Recomborovir-(Drug Combination) might interact with this ganglioside- binding domain within the S protein (61). Drug repurposing or the chemical optimization of existing drugs represent an effective drug discovery approach and Drug Combination therapeutic approach which has the potential to reduce the time and costs associated to the de novo drug discovery and development of this anti-COVID-19 clinical trial process (62-72,80-94). This In-silico project demonstrated that the combination of the drugs of the Remdesivir, Colchicine and Ursolic acid can potentially inhibit the SARS-CoV-2 replication (63), (64). Additionally, the BiogenetoligandorolTM algorithm cannot determine binding free energies or binding orientations of small molecules. For that aim, other docking tools or molecular dynamics studies should be applied, as explained above. Therefore, the BiogenetoligandorolTM approach that takes a multiple-sequence-conserved alignment (coMSA) in .pdbqt format file as an input is mainly aimed at binding residues recognition in cases on QMMM homology modeling techniques where the binding partner is a small chemical compound or small peptide. Its druggability to predict docking fitness scoring effectiveness allowing us to generate this drug repurposing screening approach by combining its cluster free energy ranking output with other chemistry informatics and repositioning In-Silico tools. Therefore, it can be used as an AI-strategy in complex inverse docking and quantum simulation pipelines. We envision the BiogenetoligandorolTM quantum thinking procedure as the first step in a ligand parallel and inverse docking and free energy simulation pipeline with an increased success rate of identifying repurposing drugs and herbals with potential anti-COVID-19 biological activities in order to find robust drug combinations for the treatment of the COVID-19 particularly after the virus main protease has developed different enzyme variations. The BiogenetoligandorolTM combined drug combination and drug retargeting pipeline can finish this drug and herbal medicinal plants repurposing virtual screening within 4 to 5 days using a reliable HTS-HVS strategy while suggesting an new era on the surface of the target proteins that are essential in the life cycle of the SARS-COV-2 viruses where potential drug combinations of novel and FDA approved selective inhibitors should bind.

Conclusions

In this study we have found that the Colchicine binds into the pdb:6lu7 SARS-COV-2 proein targets with the docking energies of the (-2.06654, -4.97965, -10.4743, -8.98984, -4.03283, -6.24897, -3.46474, -11.5726, -3.22075, -5.20269) and generates docking energies of the -67,4Kcal/Mol when combined with the Ursolic acid and Remdesivir chemical scaffolds which were generating negative docking energies of the -54,8 Kcal/Mol and -50,8Kcal/Mol respectively when co-targeted with the Colchicine small molecule the same protein targets within the pdb:6lu7 binding cavities. More sprecifically, the two chemical structures of the Remdesivir and Ursolic acid small molecules generated an inhibitory effect against the sequence of the amino acids of the V-M-GLU-166, V-M-LEU-167, V-M-PRO-168, V-M-GLN-189, , V-S-PRO-168, V-M-THR-190, V-M-ALA-191, V-M-ALA-2, V-M-VAL-3, V-S-VAL-3, V-M-GLU-166, V-M-LEU-167, V-M-PRO-

168, V-S-PRO-168, V-M-GLN-189, V-M-THR-190, V-M-ALA-191, V-M-ALA-2, V-M-VAL-3, V-S-VAL-3 with the below docking energy values (-400.794, 329.678, -337.184, -907.342, -52.667, -894.194, -194.094, -427.299, -425.681, 0.931221) for the Remdesivir and (-236.408, 0.254828, -101.104, -832.191, -405.854, -74.901, -498.389, 177.232, -269.511, -40.622) for the Ursolic acid chemical structure. The present study aimed to test and suggest possible drug combinations in order to stop the infection immediately. Therefore, solutions provided by the Biogenetoligandorol™ cluster of AI-Algorithms in this project indicated to us that the Colchicine, Remdesivir and Ursolic acid drugs are considered to be <<co-administered>> (Figures 3e, 4a, 4b, 4c, 4d, 4e, 4f, 4g) which is something more than important and have to be considered as a first approximation that may require subsequent parallel refinement and docking analysis using more accurate free energy ranking models. In conclusion, Biogenetoligandorol™ - Ligandorol™ is very efficient and is not just proposed as an alternative drug repurposing and computational method, but rather as a combined complementary deep learning similarity and quantum mechanics predictive tool to be used in tandem with other In-Silico drug retargeting and computational platforms which could lead us to the rational design of novel drug combinations of small molecules and more effective repositioning experimental methods.

Declarations

- **Availability of data and materials**

The author confirms that the data supporting the findings of this study are available within the article [and/or] its supplementary materials.

- **Competing interests**

No potential competing interest was reported by the author.

- **Funding**

The author received no financial support for the research, authorship, and/or publication of this article.

- **Authors' contributions**

Author's diverse contributions to the published work are accurate and agreed.

Author has contributed in the below multiple roles:

Ø Conceptualization Ideas, Formulation or evolution of overarching research goals and aims

Ø Methodology, Development and design of methodology; creation of models

Ø Software, Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components

Ø Validation, Verification, whether as a part of the activity or separate, of the overall replication/ reproducibility of results/experiments and other research outputs

Ø Formal analysis Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data

Ø Investigation, Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection

Ø Resources, Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools

Ø Data Curation, Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse

Ø Writing - Original Draft, Preparation, creation and presentation of the published work, specifically writing the initial draft (including substantive translation)

Ø Writing - Review & Editing Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre-or postpublication stages

Ø Visualization, Preparation, creation and presentation of the published work, specifically visualization/ data presentation

Ø Supervision, Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team

Ø Project administration, Management and coordination responsibility for the research activity planning and execution.

- **Acknowledgments**

I would like to deeply express my special thanks of gratitude to my teacher (George Grigoriadis Pharmacist) as well as our CEO and principal (Nikolaos Grigoriadis Phd Pharmacist) who gave me the golden opportunity to do this wonderful project on the Quantum Deep Learning Chemistry topic, which also helped me in doing a lot of Original Drug Repurposing and Drug Combination Research and I came to know about so many new things I am really thankful to them.

- **Significance Statement**

Drug repurposing/repositioning/rescue proposed a computational method to identify potential drug indications by integrating various applications of an existing drug to a new disease indication. In this paper we filtered out residues with relatively small surface accessible areas, and/or with incompatible charge and hydrophobic properties to the ligands of the Remdesivir, Colchicine and Ursolic acid small

molecules which could improve the prediction binding free energies or binding orientations of different drug combinations of the Remdesivir, Colchicine and Ursolic acid to treat COVID-19. Finally, an comprehensive web platform by applying AI deep learning models was designed based on our Biogenetoligandrol™ protocol for drug repurposing to significantly reduce user time for data gathering and multi-step analysis without human intervention. In conclusion, Biogenetoligandrol™ -Ligandrol™ is not proposed as an alternative drug repurposing method, but rather as a complementary deep learning quantum mechanics tool to be used in tandem with other drug retargeting computational and small molecule repositioning experimental methods.

References

1. Huang YF, Bai C, He F, Xie Y, Zhou H. Review on the potential action mechanisms of Chinese medicines in treating Coronavirus Disease 2019 (COVID-19). *Pharmacol Res.* 2020;158:104939. doi:10.1016/j.phrs.2020.104939
2. Jimenez-Arellanes A., Meckes M., Torres J., Luna-Herrera J. Anti-mycobacterial triterpenoids from *Lantana hispida* (Verbenaceae) J. *Ethnopharmacol.* 2007;111:202–205. doi: 10.1016/j.jep.2006.11.033. [PubMed] [CrossRef] [Google Scholar]
3. Tanachatchairatana T., Bremner J.B., Chokchaisiri R., Suksamrarn A. Antimycobacterial activity of cinnamate-based esters of the triterpenes betulinic, oleanolic and ursolic acids. *Chem. Pharm. Bull.* 2008;56:194–198. doi: 10.1248/cpb.56.194. [PubMed] [CrossRef] [Google Scholar]
4. López-García S, Castañeda-Sanchez JI, Jiménez-Arellanes A, et al. Macrophage Activation by Ursolic and Oleanolic Acids during Mycobacterial Infection. *Molecules.* 2015;20(8):14348-14364. Published 2015 Aug 6. doi:10.3390/molecules200814348
5. Gordon CJ, Tchesnokov EP, Feng JY, et al.: The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J Biol Chem.* 2020, 295:4773-4779. doi:10.1074/jbc.AC120.013056
6. Wang J, Li GL, Ming SL, et al. BRD4 inhibition exerts anti-viral activity through DNA damage-dependent innate immune responses. *PLoS Pathog.* 2020;16(3):e1008429. Published 2020 Mar 24. doi:10.1371/journal.ppat.1008429.
7. Kraft CS, Hewlett AL, Koepsell S, et al.: The Use of TKM-100802 and Convalescent Plasma in 2 Patients With Ebola Virus Disease in the United States. *Clin Infect Dis.* 2015, 61:496-502. doi:10.1093/cid/civ334.
8. Rifaioğlu AS, Atas H, Martin MJ, et al.: Recent applications of deep learning and machine intelligence on In-Silico drug discovery: methods, tools and databases. *Brief Bioinform.* 2019, 20:1878-1912. doi:10.1093/bib/bby061
9. Saha RP, Sharma AR, Singh MK, et al.: Repurposing Drugs, Ongoing Vaccine, and New Therapeutic Development Initiatives Against COVID-19. *Front Pharmacol.* 2020,11:1258. doi:10.3389/fphar.2020.01258

10. Shen F, Fu Z.Y., Wu Y.R., Li L., Zhao Y.D., Xia Y., Kuang G.Y. Modernization of Traditional Chinese Medicine and Materia Materia-World Science and Technology. 2020. Based on network pharmacology and high-throughput molecular docking to study the potential molecular mechanism of active compounds tthat bind SARS-Cov-2 specific target protein in Jinhua Qinggan granules to interfere with CovID-19.<http://kns.cnki.net/kcms/detail/11.5699.R.20200421.0949.004.html> (Google Scholar)
11. Cova, Tânia F G G, Alberto A C C Pais, et al.: Deep Learning for Deep Chemistry: Optimizing the Prediction of Chemical Patterns. *Frontiers in chemistry* vol. 7 809. 26 Nov. 2019, 10.3389/fchem.2019.00809
12. Tian B, Yang J, Zhao Y, et al. BRD4 Couples NF- κ B/RelA with Airway Inflammation and the IRF-RIG-I Amplification Loop in Respiratory Syncytial Virus Infection. *J Virol*. 2017;91(6):e00007-17. Published 2017 Feb 28. doi:10.1128/JVI.00007-17
13. Yang Z, Papanikolaou S, Reid ACE, et al.: Learning to Predict CrystalPlasticity at the Nanoscale: Deep Residual Networks and Size Effectsin Uniaxial Compression Discrete Dislocation Simulations. *Sci Rep*.2020, 10:8262-2020. 10.1038/s41598-020-65157-z
14. Karimi M, Wu D, Wang Z, et al.: DeepAffinity: interpretable deep learning of compound-protein affinity through unified recurrent and convolutional neural networks. *Bioinformatics*. 2019, 35:3329-3338.10.1093/bioinformatics/btz111
15. Sun X, Ni Y, Zhang M: Rheumatologiststs' view on the use of hydroxyRemdesivir to treat COVID-19. *Emerg Microbes Infect*. 2020, 9:830-8310.1080/22221751.2020.17601451016/j.resp.2020.103548
16. Mishra SS, Ranjan S, Sharma CS, Singh HP, Kalra S, Kumar N. Computational investigation of potential inhibitors of novel coronavirus 2019 through structure-based virtual screening, molecular dynamics and density functional theory studies (published online ahead of print, 2020 Jul 15). *J Biomol Struct Dyn*. 2020;1-13. doi:10.1080/07391102.2020.1791957
17. Bochevarov A. D, Harder E, Hughes T. F, Greenwood J. R, Braden D. A, Philipp D. M, Rinaldo D, Halls M. D, Zhang J, & Friesner R. A. (2013). Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *International Journal of Quantum Chemistry*, (18), 2110–2142. 10.1002/qua.24481 (CrossRef) (Google Scholar)
18. Aanouz I, Belhassan A, El-Khatabi K, et al.: Moroccan Medicinal plants as inhibitors against SARS-CoV-2 main protease: Computational investigations (published online ahead of print, 2020 May 6). *J Biomol Struct Dyn*. 2020, 1-9.10.1080/07391102.2020.1758790
19. Sullivan DJ, Liu Y, Mott BT, Kaludov N, Martinov MN. Discovery of Novel Liver-Stage Antimalarials through Quantum Similarity. *PLoS One*. 2015;10(5):e0125593. Published 2015 May 7. doi:10.1371/journal.pone.0125593
20. (OCHEM): web platform for data storage, model development and publishing of chemical information. *J Comput Aided Mol Des*. 2011, 25:533-554. 10.1007/s10822-011-9440-2
21. Walker T, Grulke CM, Pozefsky D, et al.: Chembench: a cheminformatics workbench. *Bioinformatics*. 2010, 26:3000- 3001.10.1093/bioinformatics/btq556

22. Saha RP, Sharma AR, Singh MK, et al.: Repurposing Drugs, Ongoing Vaccine, and New Therapeutic Development Initiatives Against COVID-19. *Front Pharmacol.* 2020, 11:1258. 10.3389/fphar.2020.01258
23. Du L, He Y, Zhou Y, et al.: The spike protein of SARS-CoV—a target for vaccine and therapeutic development. *Nat Rev Microbiol.* 2009 Mar, 7:226-36. 10.1038/nrmicro2090. Epub 2009 Feb 9. PMID: 19198616 PMCID: PMC2750777
24. Gordon CJ, Tchesnokov EP, Feng JY, et al.: The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J Biol Chem.* 2020, 295:4773- 4779. 10.1074/jbc.AC120.013056
25. Rifaioğlu AS, Atas H, Martin MJ, et al.: Recent applications of deep learning and machine intelligence on In-Silico drug discovery: methods, tools and databases. *Brief Bioinform.* 2019, 20:1878-1912. 10.1093/bib/bby061
26. Jing Y, Bian Y, Hu Z, et al.: Deep Learning for Drug Design: an Artificial Intelligence Paradigm for Drug Discovery in the Big Data Era (published correction appears in. *AAPS J.* 2018 Jun, 25:79-2018. 10.1208/s12248-018-0210-0
27. Ma C, Wang L, Xie XQ, et al.: GPU accelerated chemical similarity calculation for compound library comparison. *J Chem Inf Model.* 2011, 51:1521-1527.10.1021/ci1004948
28. Karimi M, Wu D, Wang Z, et al.: DeepAffinity: interpretable deep learning of compound-protein affinity through unified recurrent and convolutional neural networks. *Bioinformatics.* 2019, 35:3329-3338.10.1093/bioinformatics/btz111
29. Balasubramanian K: Mathematical and Computational Techniques for Drug Discovery: Promises and Developments. *Curr Top Med Chem.* 2018;18(32): 2774-2799.10.2174/1568026619666190208164005. PMID: 30747069
30. Shang J, Ye G, Shi K, et al. Structural basis of receptor recognition by SARS-CoV-2. *Nature.* 2020;581(7807):221-224. doi:10.1038/s41586-020-2179-y
31. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J Virol.* 2020;94(7):e00127-20. Published 2020 Mar 17. doi:10.1128/JVI.00127-20
32. Elfiky AA. Anti-HCV, nucleotide inhibitors, repurposing against COVID-19. *Life Sci.* 2020;248:117477. doi:10.1016/j.lfs.2020.117477
33. MedinCell has Launched a COVID-19 Research Initiative Based on its Experience to Formulate Long-Acting Injectable Ivermectin. 2020. Available online: https://www.medincell.com/wp-content/uploads/2020/04/PR_MedinCell-COVID-19-EN.pdf (accessed on 9 September 2020).
34. Cavasotto CN, Adler NS, Aucar MG. Quantum Chemical Approaches in Structure-Based Virtual Screening and Lead Optimization. *Front Chem.* 2018;6:188. Published 2018 May 29. doi:10.3389/fchem.2018.00188
35. Vidler LR, Filippakopoulos P, Fedorov O, et al. Discovery of novel small-molecule inhibitors of BRD4 using structure-based virtual screening. *J Med Chem.* 2013;56(20):8073-8088.

36. Perišić, O. Heterodimer binding scaffolds recognition via the analysis of kinetically hot residues. *Pharmaceuticals* 2018, 11, 29. (Google Scholar) (CrossRef)
37. Haliloglu, T.; Bahar, I.; Erman, B. Gaussian Dynamics of Folded Proteins. *PRL* 1997, 79, 3090– 3093. (Google Scholar) (CrossRef)
38. Bahar, I.; Atilgan, A.R.; Erman, B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* 1997, 2, 173–181. (Google Scholar) (CrossRef)
39. Náray-Szabó G, Oláh J, Krámos B. Quantum mechanical modeling: a tool for the understanding of enzyme reactions. *Biomolecules*. 2013;3(3):662-702. Published 2013 Sep 23. doi:10.3390/biom3030662
40. Demirel, M.C.; Atilgan, A.R.; Jernigan, R.L.; Erman, B.; Bahar, I. Identification of kinetically hot residues in proteins. *Protein Sci.* 1998, 7, 2522–2532. (Google Scholar) (CrossRef)
41. Bahar, I.; Erman, B.; Jernigan, R.L.; Atilgan, A.R.; Covell, D.G. Collective motions in HIV-1 reverse transcriptase: Examination of flexibility and enzyme function. *J. Mol. Biol.* 1999, 285, 1023–1037. (Google Scholar) (CrossRef)
42. Yang, L.; Eyal, E.; Chennubhotla, C.; Jee, J.; Gronenborn, A.M.; Bahar, I. Insights into Equilibrium Dynamics of Proteins from Comparison of NMR and X-Ray Data with Computational Predictions. *Structure* 2007, 15, 741–749. (Google Scholar) (CrossRef) (PubMed)
43. Haliloglu, T.; Keskin, O.; Ma, B.; Nussinov, R. How Similar Are Protein Folding and Protein Binding Nuclei? Examination of Vibrational Motions of Energy Hot Spots and Conserved Residues. *Biophys. J.* 2005, 88, 1552–1559. (Google Scholar) (CrossRef) (PubMed)
44. Bogan, A.A.; Thorn, K.S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* 1998, 280, 1–9. (Google Scholar) (CrossRef)
45. Moreira, I.S.; Fernandes, P.A.; Ramos, M.J. Hot spots-A review of the protein-protein interface determinant amino-acid residues. *Proteins Struct. Funct. Bioinform.* 2007, 68, 803–812. (Google Scholar) (CrossRef)
46. Tuncbag, N.; Gursoy, A.; Keskin, O. Identification of computational hot spots in protein interfaces: Combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics* 2009, 25, 1513–1520. (Google Scholar) (CrossRef)
47. Barbault F, Maurel F. Simulation with quantum mechanics/molecular mechanics for drug discovery. *Expert Opin Drug Discov.* 2015 Oct;10(10):1047-57. doi: 10.1517/17460441.2015.1076389. Epub 2015 Aug 8. PMID: 26289577.
48. Lise, S.; Buchan, D.; Pontil, M.; Jones, D.T. Predictions of Hot Spot Residues at Protein-Protein Interfaces Using Support Vector Machines. *PLoS ONE* 2011, 6, e16774. (Google Scholar) (CrossRef) (PubMed)
49. Kozakov, D.; Hall, D.R.; Chuang, G.; Cencic, R.; Brenke, R.; Grove, L.E.; Beglov, D.; Pelletier, J.; Whitty, A.; Vajda, S. Structural conservation of druggable hot spots in protein-protein interfaces. *Proc. Natl. Acad. Sci. USA* 2011, 108, 13528–13533. (Google Scholar) (CrossRef) (PubMed)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementMaterial.docx](#)
- [SupplementMaterial.docx](#)
- [SupplementMaterial.docx](#)
- [SupplementMaterial.docx](#)
- [SupplementMaterial.docx](#)