Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Artificial Neural Network Based Study Predicting GS-441524 As Potential Inhibitor of SARS COV-2 Activator Protein Furin: A Polypharmacology Approach

Dhanalakshmi Menamadathil
Research and Development Centre, Bharathiar University, Marudhamalai Rd, Coimbatore, TamilNadu, 641046, India.

Kajari Das
Department of Biotechnology, College of Basic Science and Humanities, Odisha University of Agriculture and Technology.Bhubaneswar- 3, Odisha

Medha Pandya
The KPES Science College, M.K Bhavnagar University

Sejal Shah
Department of Microbiology, School of Science, RK. University, Rajkot, Gujarat, India

Ayushman Gadnayak
Centre for Genomics & Biomedical Informatics, IMS and SUM Hospital, Siksha “O” Anusandhan (Deemed to be University), Bhubaneswar, Odisha 751003, India

Sushma Dave
Department of Applied Sciences,JIET,Jodhpur,Rajasthan,India

Jayashankar Das (dasjayashankar@gmail.com)
Centre for Genomics & Biomedical Informatics, IMS and SUM Hospital, Siksha “O” Anusandhan (Deemed to be University), Bhubaneswar, Odisha 751003, India

---

Research Article

Keywords: Furin, Covid-19, Drug repurposing, ANN, SOM, LB-VS,SB-VS, Poly pharmacology, GS-441524, Antiviral

Posted Date: May 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-513443/v1

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

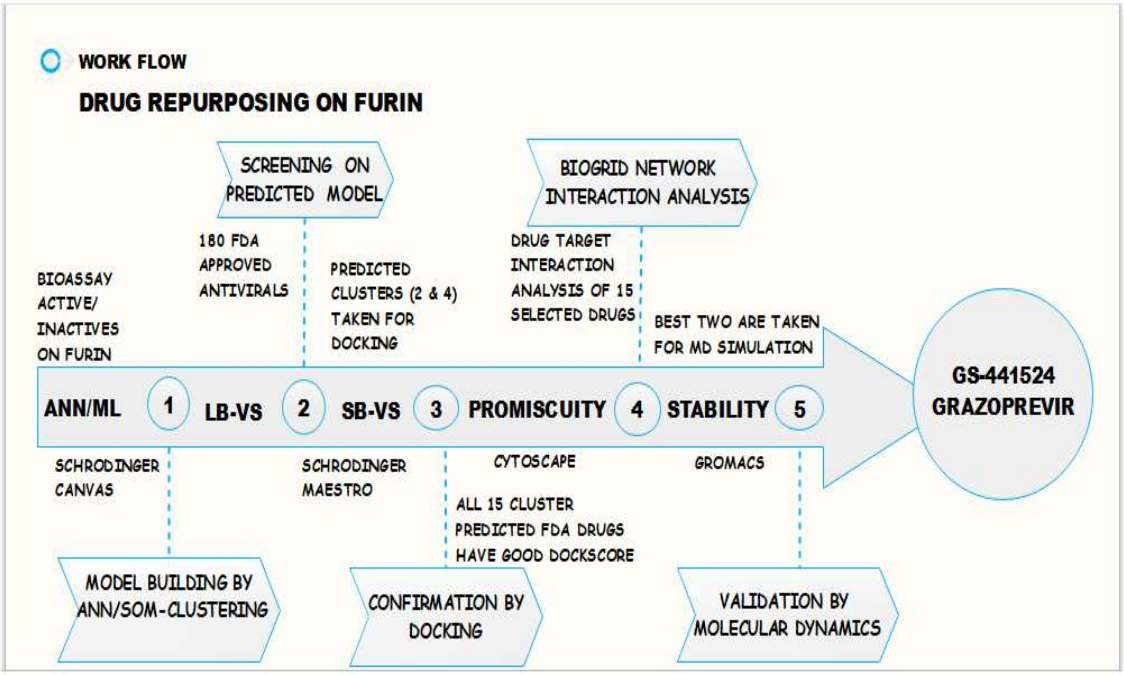
Artificial neural network based study predicting GS-441524 as potential inhibitor of SARS COV-2 activator protein Furin: A polypharmacology approach

Dhanalakshmi. M[1],Kajari Das[2], Medha Pandya[3], Sejal Shah[4], Ayushman Gadnayak[5], Sushma Dave[6*], Jayashankar Das[5*]

1. Research and Development Centre, Bharathiar University, Marudhamalai Rd, Coimbatore, Tamil Nadu, 641046, India.
2. Department of Biotechnology, College of Basic Science and Humanities, Odisha University of Agriculture and Technology. Bhubaneswar-3, Odisha.
3. The KPES Science College, M.K Bhavnagar University, Bhavnagar, Gujarat, India.
4. Bio Research and Characterization Centre, School of Science, RK University, Rajkot, Gujarat-360020
5. Centre for Genomics & Biomedical Informatics, IMS and SUM Hospital, Siksha "O" Anusandhan (Deemed to be University), Bhubaneswar, Odisha 751003, India.
6. Department of Applied Sciences, JIET Jodhpur, Rajasthan, India.

Corresponding Authors: Sushma Dave sushmadave26@gmail.com,**Jayashankar Das** dasjayashankar@gmail.com

Graphical Abstract



Abstract

Furin, a pro-protein convertase, plays a significant role of biological scissor in bacterial, viral, and even mammalian substrates which in turn decides the fate of many viral and bacterial infections along with the numerous ailments caused by cancer, diabetes, inflammations, and neurological disorders. In the wake of the current pandemic caused by the virus SARS COV-2, furin has become the center of attraction for researchers. In the present work, we have searched for novel inhibitors against this interesting human target from FDA-approved antivirals. To enhance the selection of new inhibitors we employed Kohonen's-artificial neural network-based self-organizing maps for ligand based virtual screening. Promising results were obtained which can help in drug repurposing and network pharmacology studies addressing the errors due to promiscuity/polypharmacology. We found 15 existing FDA antivirals having the potential to inhibit furin. Among these, six compounds have targets on other important human proteins (LDLR, FCGR1A, PCK1, TLR7, DNA and PNP) also. These 15 drugs inhibiting furin could be studied in patients having many viral infections including SARS COV-2, which is known to have many interacting motifs like NSPs, ORFs, and spike protein. We also propose two promising candidate FDA drugs GS-441524 and Grazoprevir (MK-5172) to repurpose as inhibitors of furin. The best results were observed with GS-441524.

Key Words: Furin, Covid-19, Drug repurposing, ANN, SOM, LB-VS,SB-VS, Poly pharmacology, GS-441524, Antiviral.

Introduction

New drugs or treatments approved by USFDA have remained constant though there is a huge development in the pharma industries since the last two decades[1]. Repurposing of the drug can shorten the time for designing and is cost-effective with randomized clinical trials[2].It can be mediated by using different computational platforms with novel testable hypotheses for systematic drug repositioning[3],[4]. Due to the unavailability of 3D-based protein structures, traditional structure-based methods are the limitation. Furthermore, drug resistance is a big challenge due to the single virus protein target as viral genome evolution is observed rapidly[5]. Our group recently reported for the first time in India in 2021 about two lineages B.1.112 and B.1.99 and also revealed 1,143 unique single nucleotide variants along with 73 novel variants using COVIDSeq approach for SARS-CoV-2[6]. Most antiviral therapies do not completely eradicate the viral infection. Instead, they limit the immune recovery, and their long-term use results in resistance and toxicity; there is always a requirement of new antiviral classes with new targets. Specific antiviral agents should essentially target interrupting the viral life cycle or immune clearance. There are two categories of targets required for the completion of virus replication and infection; viral molecules and host molecules. In our recent work on molecular docking-based screening, direct antiviral activity by remarkable binding of vitamins as inhibitors to various important proteins of SARS-Cov-2 are suggested[7]. Moreover, this work revealed vitamin B12 might be employed

as an efficient nutraceutical for almost all the nCoV-2 drug targets including the host protease furin which is a human host molecule and one of the promising targets for therapeutic intervention. It cleaves more than 150 mammalian, viral and bacterial proteins including the disastrous virus Coronaviruses (CoVs). Researchers in their recent works have foundfurin inhibitors as a new avenue for the treatment of COVID-19[8]. The mechanism for furin-mediated protein processing for health and disease is well-known. Furin is a member of the pro-protein convertase family, PCSK 3 representing furin. The viral and bacterial infection is caused by the production of various toxins by proteolytic cleavage such as Anthrax toxin, Diphtheria toxin, Pseudomonas exotoxin A, Shiga toxin,etc. Furin activation of viral proteins is essential for their entry to the host cell. Furin has basic amino acid motifs as a cleavage site "PACE" paired with basic amino acid cleaving enzymes. It is located at chromosome 15 q26.1 and expressed by the FUR (FES upstream region) [9]. The hot spot cleavage site of furin is described as R-X-K/R-R↓[10]. Bioinformatic analyses have revealed more than 100 furin cleavage sites in humans such as growth factors, cytokines, collagens, adhesion molecules, MMP (metalloproteinases), receptors, etc.[11].Many recent reports have mentioned peptidic and nonpeptidic inhibitors for therapeutic inhibition of furin resulting in its reduced proteolytic activity. Numerous pharmaceutical companies have come up with solutions for the current global pandemic caused by SARS-CoV-2 with different vaccines and existing medicines. However, vaccination is going on in spite of which many cases are being observed and no effective medications are there due to the appearance of double mutant and triple mutant strains. Hence the need of the hour is a better prognosis of the Covid patients.

Cheminformatics incorporating machine learning are breakthrough techniques imperative in the pharmaceutical field today as we live in an era of infectious outbreaks, rising lifestyle diseases and the age-old time-consuming procedures in drug discovery[12]. Artificial intelligence (AI) and machine learning (ML) tools are applied for furin targeting while screening of molecules including higher-dimensional similarity/differences and networking observation has been performed to establish a relationship (interactome) with selected drug compounds and their respective important host factors along with furin.

Kohonen's Self-organizing maps (SOM) are extensively used today to establish the relationship of structure to biological activity and drug discovery in many pharmaceutical research. This methodology is widely applied in drug repurposing and scaffold hopping[13]. Self-Organizing Maps (SOM) are mainly an artificial neural network (ANN) that implements virtual screening for assembling thematic molecular libraries. The concept of the SOM technique, case studies, and potential future applications are presented in the report presented by Schneider et al[13]. Here employs the ANN to identify the nonlinear connections of arbitrary

dimensions into processable real numbers of one or two-dimensional maps at less computational costs. This mapping technique permits high-dimensional input space to be transformed into lower-dimensional output space[14]. SOM is considered to be unsupervised training in which networks learn from their individual arrangement of training data without supervision. Studies with supervised SOM methods were proven more accurate than standard linear QSAR methods[15]. ANN is not a single algorithm; instead, it comprises a framework of machine learning algorithms to process complex data. Kohonen's self-organizing map is employed here to screen and predict the active FDA antivirals on Furin basically to transform a pattern of arbitrary dimension into a 2D map adaptively.

Managing large amounts of heterogeneous data in a short period is becoming possible through AI technology. AI/ML, systems biology, including computer-aided drug repurposing are speedily becoming an integral part of the global workflow, to prioritize novel drug targets, new chemical entities, and to evaluate off-label or drug repositioning proposals[16]. All these above-mentioned techniques are utilized in our present work where we focused to present an integrative antiviral drug repurposing methodology for the first time with the combination of systems pharmacology-based network medicine platform that quantifies the interplay between the virus-host interaction and drug targets in the human PPI network. ANN employed a polypharmacology based computational approach has been reported in the present paper with multiple antiviral drugs for furin which uncovers the therapeutic interventions.

Results

Building of model for antiviral drugs via SOM clustering

The Furin actives and in-actives are manually curated from the PubChem database (supplementary table S1and S2)for model building for the ligand based virtual screening(LB-VS). ANN-based self-organizing maps were generated using unsupervised machine learning clustering algorithms in the canvas Schrodinger suite and using the above mentioned model the activity of the FDA approved/investigational antiviral drugs were screened (supplementary table S3). The descriptors generated in canvas were applied for the ANN modeling inputs and the SOM was generated by the property-based clustering. However, SOM is considered as a supervised training method in this study and we employed the actives as input data without labeling while submitting the job, most of the time the output-data analysis became more straightforward with known labels[17]. The number of molecules from each inadequate dataset is presented in Table 1.

Cluster no	Clusters	Actives	Inactive	FDA	Considered for analysis	Reason of consideration of molecules
	Number of molecules					
1	69	__	__	69	Discarded	No Actives
2	33	20	__	13	Selected	FDA + Active & No Inactives
3	29	29	__	__	Discarded	No FDA Drugs
4	27	25	__	2	Selected	FDA+Active& No Inactives
5	23a	19	__	4	Selected	FDA+Active& No Inactives
6	23b	14	__	9	Selected	FDA+Active& No Inactives
7	21	21	__	__	Discarded	No FDA Drugs
8	20a	19	__	1	Selected	FDA+Active& No Inactives
9	20b	__	__	20	Discarded	No Actives
10	20c	19	__	1	Selected	FDA+Active& No Inactives

Table 1: Different molecular clusters of antivirals generated by SOM considered for analysis.

The SOM results have been presented in distinct rows and columns in specific colour codes that help us to scrutinize the high dimensional data of complex datasets into lower dimensional space. In this study, the properties of Furin actives and in-actives accompanying with FDA drugs had undergone a mapping to place each molecule into individual clusters (10*10 cells) possessing varied populations, a maximum of showing in red color (having 69 molecules) and most clusters having least members, null values shown in blue color (see fig 1a). Figure 2b illustrates the Euclidean distances in the hyperspace between the molecules. If the difference between Euclidean distances of data points were high, then that ensemble would have been in red, and if the distance was low, it would have been in blue. The variations among the population and the distance were interpreted by various colors between blue and red. Our results revealed that bioassay actives are in blue. The slightly variant molecules depicted in green for the distance analysis (not much distant) and they are resembling in property space as well. The greener section owned higher numbers of FDA drugs.

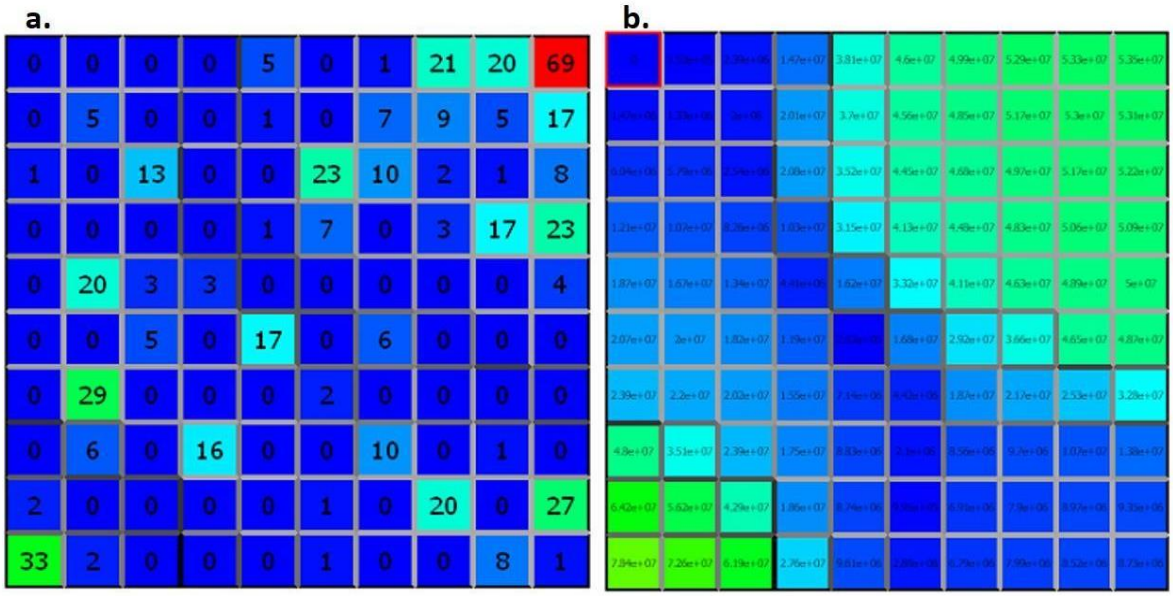


Figure 1:(a) Cell Population Analysis of antiviral FDA drugs and bioassay actives of Furin (b) Euclidean distance between the antiviral FDA drugs and bioassay actives of Furin in the property based SOM. (SOM Colour Schema: If the difference is high, then that ensemble is represented in red, and if the distance is low, it is represented in blue. The various colours between blue and red are used to represent the variations among the properties, population, and the distance)

Each cluster was manually examined to confirm the most desirable molecules comprising only actives along with FDA drugs and restrict in-actives molecules. The clusters in the combination of all three active /inactive/FDA were discarded. Out of 15 clusters, the clusters holding more than 10 molecules were selected for further analysis. Cluster number 2 holding 33 molecules has the maximum number of FDA antivirals and actives together (20 actives + 13 FDA). The cluster number 4 carrying 27 molecules (25 actives + 2 FDA) which is the second-largest membered cluster yet possessing just two FDA drugs (GS-441524 and Grazoprevir). The cluster 4 possessing active ligand$C_{28}H_{37}N_{13}O_2$(PubChem CID: 44593169) which was in interaction with furin from protein structure database (PDB Id 5MIM) acknowledged as the reference active molecule in this study[18]. Cluster number 5, 6 and 8 also own the active and FDA drugs screened out and granted for additional analysis.

The physicochemical descriptors are tabulated in table 2. The higher values of MW, AlogP, RB, Ring count, A higher number of N atoms, and a lower number of O atoms, all can make a compound promiscuous[19]. The FDA drug Isatorbine and Grazoprevir have a sulfur atom and that is why exhibiting high value for the topological distance between N-S and O-S. The molecule Grazoprevir is bigger compared to the reference ligand even though showing similar structure and properties has higher values. GS-441524 is smaller and all physicochemical properties are promising compared to the reference ligand. The list of both clusters 2 and 4 with their structures is listed in supplementary table S4 and S5.

Name	AlogP	HBA	HBD	RB	Ring Count	Atom Count	Bond Count	Bonds in Ring System	Distances Between N..S*	Distances between O..S*
Porfimer Sodium	13.5137	7	6	18	10	87	96	56	0	0
Isatoribine	-0.9417	7	5	2	3	21	23	15	13	27
GppCp	-3.4787	14	4	8	3	32	34	15	0	0
Valomaciclovir	-0.7498	6	3	9	2	25	26	10	0	0
Valaciclovir	-0.64	6	2	8	2	23	24	10	0	0
Lobucavir	-1.558	5	4	3	3	19	21	14	0	0
Didanosine	-0.7108	5	2	2	3	17	19	15	0	0
Entecavir	-1.3654	5	4	2	3	20	22	15	0	0
Valganciclovir	-1.1507	7	3	9	2	25	26	10	0	0
GS-441524	-1.4269	7	4	1	3	21	23	15	0	0
Penciclovir	-1.4945	5	4	5	2	18	19	10	0	0
Valomaciclovir stearate	7.1392	7	2	27	2	44	45	10	0	0
Aciclovir	-1.3847	5	3	4	2	16	17	10	0	0
Gancyclovir	-1.8954	6	4	5	2	18	19	10	0	0
Grazoprevir	4.1725	11	3	8	7	54	60	39	45	68
$C_{28}H_{37}N_{13}O_2$ (ACTIVE)	0.8933	6	1	10	4	43	46	24	0	0

Table 2*:* Pharmacokinetic Properties of FDA drugs from cluster 2 and 4(* Sum of topological distances)

Screening for FDA antivirals against furin using molecules elected from SOM clustering via molecular docking

Molecular docking analysis is a vital tool for structure based virtual screening (SB-VS) to predict the promising structural conformations amongst ligand and active sites of an enzyme/receptor. The SOM examination unveiled five clusters 2, 4, 5, 6, and 8 are notable for additional knowledge hence all active ligands along with FDA drugs were docked against the furin. The docking result has shown the conformation and the compound with the least

docking score was taken as the most reliable docked antiviral. The best 15 FDA antiviral drugs fall in two clusters 2 and 4 with an immeasurable docking score (Fig.2.c) with furin. The figure 2a shows that all 60 (33+27) molecules of cluster 2 and 4 have docking scores close to each other and in the model and it is evident that most of their energy falls in the -40 to -90 range in figure 2b.

The best docked FDA drug was an outlier while comparing with furin active molecules with reference to Pharmacokinetic and topological descriptors as well. (Table 2) These drugs and their former targets were further considered for the network analysis.

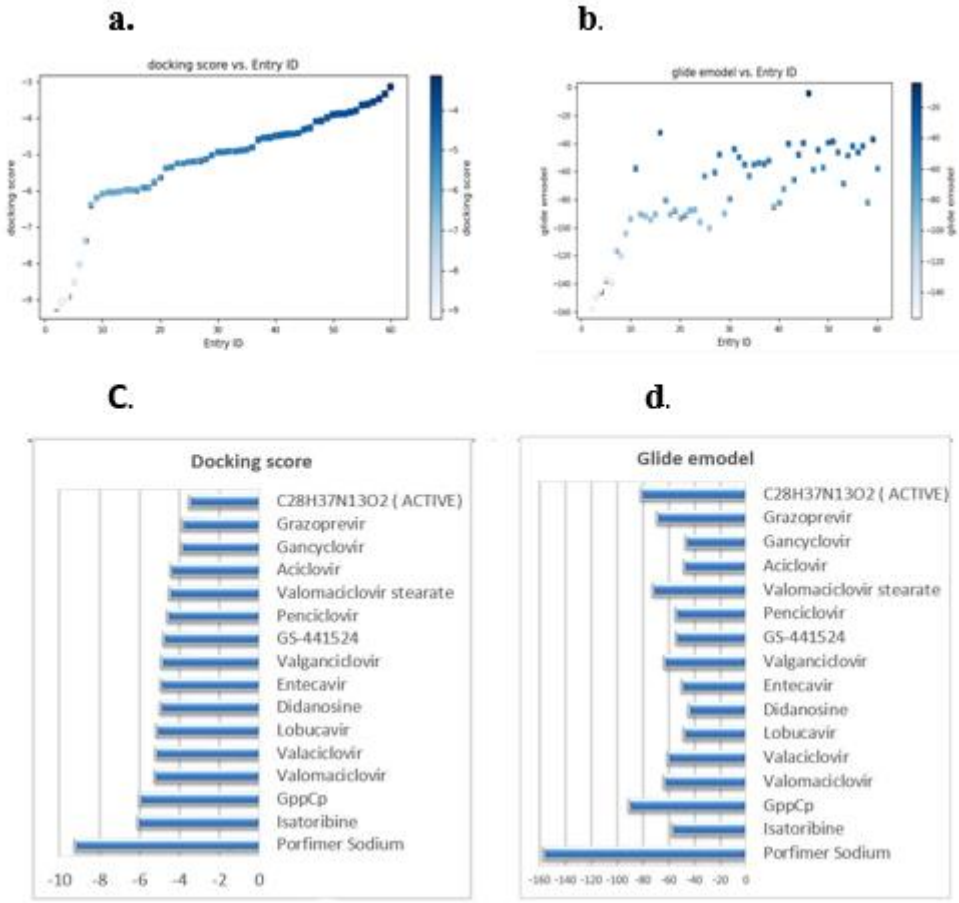


Figure 2: (a) Molecular docking score of all molecules from cluster 2 and cluster 4 (b) Glide e-model of all molecules from cluster 2 and cluster 4 (c) Molecular docking score of FDA drugs from cluster 2 and cluster 4 (d) Glide e-model of FDA drugs from cluster 2 and cluster 4

Unraveling furin inhibitor by network based polypharmacology study

The pharmacokinetic descriptors are high for most of the drugs listed in table 2 and can have promiscuity[19]. It is necessary to analyze all possible polypharmacology before we practice FDA-approved drug candidates on to any other targets. For this purpose, we utilized the network-based interaction study of the drugs selected to repurpose. This analysis aims to check the interaction network of present targets before we move further, whether it is in host or pathogen and how costly it becomes while repurposing it on furinConcerning FDA drugs with multiple targets to be active against furin, its target proteins should be within or near the

corresponding subnetwork in the host-pathogen interactome. The structure based virtual screening outcomes uncovered 15 promising antiviral FDA drugs for repurposing against the furin even though one was an outlier from the reference ligand values.The molecular targets of FDA antivirals are arranged in table 3.

Drug bank_ID	**Name of drug**	**Drug targets**	**Organism**	**Inference**
DB00707	Porfimer Sodium	LDLR, FCGR1A	HUMAN	Two targets and too many interactions with human
DB04860	Isatoribine	TLR	HUMAN	Only two interaction but with an important target in human
DB03725	5'-Guanylyl methylene bis phosphonate	PCK1	HUMAN	One target but many interactions in human
DB00900	Didanosine	PNP,RNaseH	HUMAN,HIV	One target and many interactions in human
DB00442	Entecavir	DNA	HUMAN	No interactions but costly target in human
DB01610	Valganciclovir	DNA	HUMAN	
DB15686	GS-441524	Replicase polyprotein 1ab, RNA-directed RNA polymerase L	SARS-CoV, Zaire ebolavirus	Max interactions with SARS-COV2,and common targets with furin, No interactions with Zaire ebolavirus
DB00577	Valaciclovir	TK, DNA polymerase (UL30)	HHV1	Only 3 interactions, all on HHV,can be considered for further studies but need to analyse more data
DB00299	Penciclovir	TK, UL 30	HHV	
DB00787	Aciclovir	TK, UL 30	HHV1,HHV3	
DB01004	Ganciclovir	TK, UL 30	HHV	
DB11575	Grazoprevir	NS3/4A protein	Hepatitis C Virus	No interactions to analyze but can be considered for further studies but more data required
DB06575	Valomaciclovir	No target but in use for HHV infection		No targets known and so is the interactions but can be taken for further studies but more data required
DB12531	Lobucavir	No target but in use for HIV infection		
DB15651	Valomaciclovir stearate	No target but in use for HHV infection		

Table 3: Top 15 antiviral drugs and their targets scrutinized for polypharmacology study

The drug-target network exploration represents (figure 3) that nodes (red with label) at the center are human targets and white with label are their interactions. The plotted networks manifest six molecules in the interaction network shown in red as Porfimer Sodium (LDLR, FCGR1A), Isatoribine (TLR7), GppCp (PCK1), Didanosine (PNP), ValganciclovirandEntecavir. A drug with multiple targets can be efficient for a disease or may cause off-target side effects. Hereabouts, the targets are associated with multiple significant biological pathways (Fig.3) therefore FDA drugs associated with these targets are eliminated in the further analysis. The proteins that act as drug target for one disease may also be appropriate drug target for possible antiviral infection due to common protein-protein interactions (PPI) and functional pathways explained by the human interactome[20]. From network analysis, we found that human protein furin is directly associated with viral drug targets in the PPIs. The most abundant group related tofurin is SARS CoV-2 related interactome.We identified that furin is exclusively found in the GS-441524 drug's subnetwork. Conversely, Valaciclovir, Penciclovir, Aciclovir, and Ganciclovir presently acting on Human Herpesvirus (HHV) protein UL30 (DNA polymerase), and no interaction is observed with human proteins. Grazoprevir is associated with the NS3/4A protein of the Hepatitis C virus but no interactions to analyse. The network-based drug-disease vicinity sheds light on the connection between drug targets and molecular factors in disease pathobiology modules within the PPIs and can serve as a useful tool for efficient screening of potentially new suggestions for FDA-approved drugs. In brief, the up-to-date interactome examination disclosed that GS-441524 can be a repurposed drug on furin.

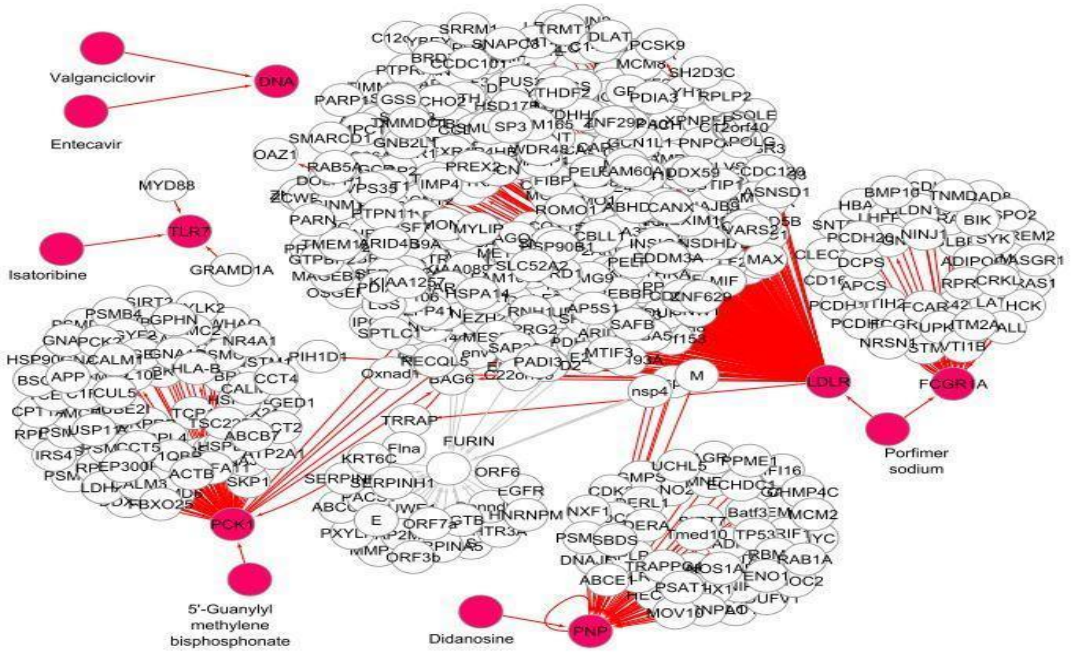


Figure 3:Drug target network analysis for drugs acting on human targets, the red circles are drugs red lines are edges connecting their targets, white with labels are their interacting proteins .

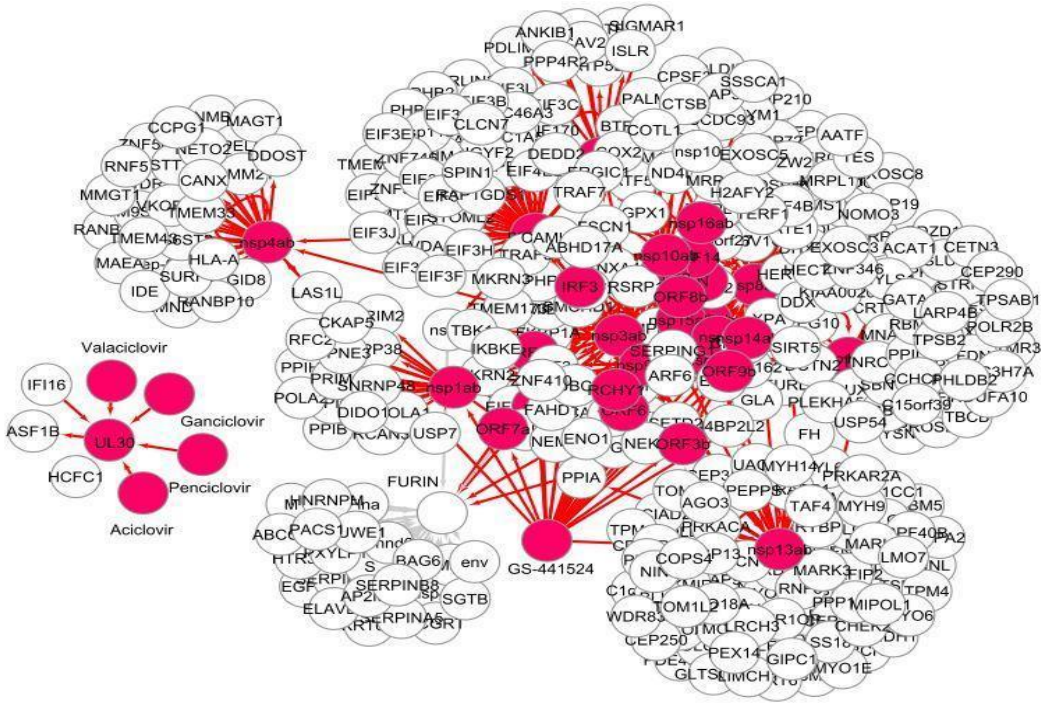


Figure 4*:*Drug target network analysis for viral host interactome: furin with drugs acting on viral targets, GS-441524 having maximum interactions with SARS-COV2 related proteins, the red circles are drugs red lines are edges connecting their targets and it reaches to furin through common targets

Insight into structural perturbation and interaction profile of Furin bound with GS-441524, Grazoprevir and reference ligand

The molecular dynamics simulation approach has been systematically employed to inspect features of conformational changes leading to modifications in protein-ligand interactions, molecular dynamics, and protein folding[21]. In this study, according to SOM clustering and network-based polypharmacology analysis, Grazoprevir and GS-441524 can be repurposed on human furin without many complications as per the available pieces of information today. Taking into consideration the parameters such as root means square deviation (RMSD), root means square fluctuation (RMSF), the radius of gyration (Rg), and solvent-accessible surface area (SASA), MD simulation was performed for 50 ns to observe the change (if any) in the conformation of furin with drugs, as compared to that of the active reference ligand ($C_{28}H_{37}N_{13}O_2$).

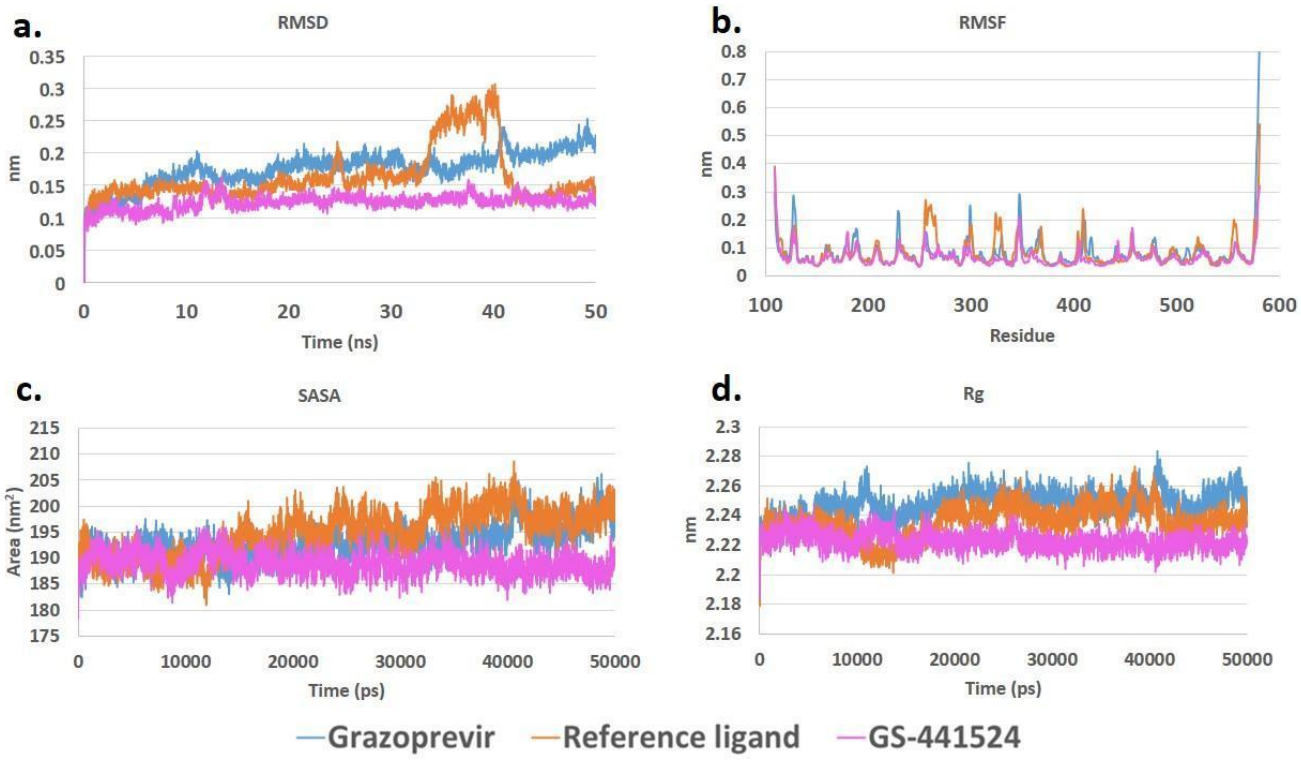


Figure 5*:* Molecular dynamics simulation analysis of protein-ligand complex of furin-Grazoprevir, furin-reference ligand, and furin-GS-441524: (a) Root means square deviation (RMSD), (b) Root means square fluctuation (RMSF), (c) solvent-accessible surface area (SASA),(d) Radius of gyration (Rg).

The RMSD value of the backbone atoms of the ligand GS-441524-furin complexes was observed about 0.15A° which is found to be as good as the RMSD of Grazoprevir-furin (0.2A°). However, in comparison to the reference furin inhibitor complex, different types of deviation were observed throughout the simulation. Overall, RMSD ranges of ~0.1–0.15, ~0.1–0.25 nm, and ~0.1–0.15 perceived for the GS-441524, Grazoprevir, and the reference complexes respectively (Fig.5a.). There is retreating of RMSD in the ligand orientation near~30 ns and ~40 ns in reference ligand complex. The GS-441524 complex scale of fluctuation and the dissimilarity between the average RMSD values acclaimed that simulation produced stable trajectories. Therefore, these outcomes established that FDA drug GS-441524 is more enduring complex with furin in correlation to the other two molecules. The backbone RMS fluctuations were also calculated as they can postulate the clue about backbone mobility. The RMSF study correlates the functional analysis (Fig 5b.), the highest fluctuation ~0.3 nm observed between 245-265 residues in the reference complex, furthermore 120-130 and 340-350 residues in Grazoprevir. In the furin-GS-441524 complex, the highest fluctuations ~0.2 nm and ~0.15 recorded near 345-370 and 120-130 residues respectively. These described major reductions in scores of particular residues propose their direct contribution in enhancing the overall stability of FDA drug bound furin complex in divergence to reference complex. In figure 6d, the radius of gyration (Rg) graph illustrates that the circulation of atoms of a protein around its axis in all the complexes is stable. The Rg of furin- GS-441524 complex is ~2.2 to ~2.22 nm which qualifies one to assess the compactness changes of a complex. Similarly, all

complexes exhibited a virtuous solvent accessible area graph (Fig.5c.). Overall analysis spectacles furin in complex with GS-441524 was observed to maintain equilibrium throughout the simulation. Afterward MD simulation, the protein-ligand complex is proximal to the in-vitro environment, hence further wet-lab experiments may qualify the GS-441524 as furin inhibitor.

Discussion

In the present article, we focused on our particular interest in drug repurposing, since already approved drugs can go through shortened clinical trials. As our knowledge is exponentially expanding with constantly emerging "omics" (e.g., transcriptomic, proteomic, metabolomic), data on drug/disease-associated mechanisms could easily be incorporated and analyzed. Our group working in this area has reported through immunoinformatics approach for immunogenic protein from the proteome of Leptospira and a codon based analytical study for susceptible hosts of Nipah Virus [22],[23]. This also helps in the analysis of quantitative structure-activity relationships, pharmacophores, homology models, and other molecular modeling approaches as well as network analysis of biological functions, machine learning, and any other analysis tools that include using a computer.

In this study, each molecule is considered as a data point, no labeling was done during submission of data for analysis, but generated descriptors propose the input into the neuron vectors which in turn make the input nodes. The data processing can be categorized into three broad interconnected networks in ANN, the input layer, the hidden layer, and the output layer. The connections in the network were assigned with weighted values which help the learning process and create the nodes in the hidden layers and finally the optimized signal after all the iterations are passed on to the output layer. According to the input given through descriptors of each data point, the molecules, within the different neurons in the network, there will be competitive and cooperative learning iterations by the end of which the winning neuron among the neuron vectors was determined and will be the closest to the molecule or the data point. After the determination of the winning neuron, update it with the weight of the neighboring neurons to place each one in possible closest positions, this process is continued until the best optimization is achieved and the clustered molecules have fallen into the cells of the map[24](Fig 1a). Each cluster was manually analyzed to confirm the best ones containing actives together with FDA drugs and have no in-actives. Out of 180 FDA antivirals, 15 predicted by cluster 2 and 4 demonstrated good results in molecular docking studies. However, it is necessary to analyze polypharmacology before we take a possible drug onto any other targets and not that binding affinity and matching properties alone. In recent years researchers are becoming dependent on a computational approach for their study on prediction of targets, off-targets, and poly-pharmacology of hit compounds. Thus the network pharmacology

comprises components functioning in stepwise order "multi drugs to multi-targets and multi-diseases[25].

There are many important host proteins like FCGRIA, LDLR, PNP, PCK1, and TLR 7 that have well-established interactions with the selected drugs. All of them are having more or less connections with various diseases (bacterial, viral, tumouretc) positively or negatively. Purine-nucleoside phosphorylase (PNP) is a very important enzyme that catalyzes nucleotide synthesis in the salvage pathway. Inherited deficiency of PNP causes T cell deficiency and abnormal B cell function in some patients as adenosine deaminase activity is affected. The enzyme deficiency was observed in HIV patients[26]. FCGRIA is a leucocyte membrane receptor protein capable of binding the Fc domain of IgG with high affinity and signaling for a variety of inflammatory and proinflammatory responses. In human immunodeficiency virus (HIV) infection, it has a role in sepsis and rheumatoid arthritis. FCGRIA has been proposed as an attractive target for immunotherapy by various workers [27]. Phosphoenolpyruvate carboxykinase (PEPCK also known as PCK, EC number 4.1.1.32) a cytoplasmic protein catalyzes gluconeogenesis implicating its important role in cancer biology and virology. Anti-tumor treatment based on gluconeogenesis has become an advanced concept nowadays. According to a recent report, gluconeogenic pathway inhibition of PCK1 enhanced HBV replication[28]. Several repeated amino acid sequences are present in the ligand-binding domain for LDL in LDLR (low-density lipoprotein receptor). Lipoprotein receptor-mediated HCV entry into the host cell is also well documented.

We target furin as part of the treatment of infectious diseases by making sure that the other targets plotted here will not get disturbed. But if we take the drugs presently acting on the 5 targets of human (FCGRIA, LDLR, PNP, PCK1, and TLR 7) and repurpose them on furin, their existing network of interactions will be disrupted; All the four drugs which are known to interact with these targets namely Porfimer Sodium, Isatorbine, 5'-Guanylyl methylene bisphosphonate, and Didanosine, are not to be carried further to repurpose without detailed studies. We strongly recommend collecting detailed data regarding the effect on furin from the patients who take these medicines and the repercussions due to this while infected by viral diseases like SARS- CoV-2. The drugs Entecavir and Valganciclovir also need proper studies before attempting to repurpose as their known current target is human DNA. Although no interactions are available, no inference can be made from this study at present. We collected information from drug bank details about all the drugs under discussion the two best-docked molecules Porfimer Sodium and Isatoribine are presently used, former for esophageal cancer treatment and the latter a compound that elevates levels of interferon-alpha.

After the poly-pharmacology analysis followed by molecular dynamics simulation, we found GS-441524 and Grazoprevir to be the most promising molecules to repurpose against the furin. GS-441524 is the nucleoside, a metabolite of the phosphoramidate prodrug

Remdesivir. GS-441524 was initially developed by Gilead Sciences, Inc for the treatment of filovirus infections. It is a 1′-cyano substituted C-nucleoside derived from 4-aza-7,9-dideazaadenosine. Remdesivir (GS-5734), a prodrug of GS-441524 used for the treatment of Ebola infections[29]. Current investigations demonstrate that this molecule has been designed to target the RNA-dependent RNA polymerase (RdRp) of several viruses, including the Ebola virus, MERS-CoV, SARS-CoV, and potentially SARS-CoV-2. The latest study has further confirmed that the direct use of GS-441524 can potently inhibit the replication of SARS-CoV-2 in a mouse model[30]. In present study, another prominent antiviral molecule Grazoprevir screened out as a furin inhibitor. Grazoprevir is a hepatitis C protease inhibitor given in combination with Elbasvir (under the brand name Zepatier) for the treatment of chronic hepatitis C virus (HCV) infection in adults. Up-to-date examinations unveiled that it has striking anti- COVID activity for the essential proteins involved in viral entry into host cells ACE2 and TMPRSS2[31]. The HCV drugs Simeprevir, Vaniprevir, Paritaprevir, and Grazoprevir in combination with Remdesivir gained the effectiveness of Remdesivir alone by 10-fold[32]. Nguyenla et al (2020), published that Remdesivir has a synergistic effect with Velpatasvir and Elbasvir the HCV drugs[33]. Additionally in combination with HCV partner drugs Sofosbuvir and Grazoprevir promote efficacy, synergistically with Remdesivir in a human lung epithelial cell line infected with a clinical isolate of SARS-CoV-2. Herewith we strongly recommend that Remdesivir analog GS-441524 alone or in combination with Grazoprevir may have a possible inhibitory effect on furin. Nonetheless further wet-lab experiments may shed light on this investigation.

A number of very recent researches are evidencing the importance of GS-441524 as a potent antiviral drug due to its strong binding properties with NSP3 macrodomain and NSP12-NSP8-NSP7 of SARS Cov 2[34],[35]. It has also been proven that this molecule has very less retention value inside host body as most of the residue is eliminated with urine suspecting no side effect[36]. Amazingly out of the majority of Remdesivir in clinical trial experiments, recovered in urine was found to be in form of GS-441524; whereas only 10% recovered as remdesivir[37],[38]. Moreover, no report has shown any host protein interaction except some metabolising enzymes like CPY450. Our work has achieved the interesting interaction of GS-441524 with furin, though indirectly. Hence results of our work could be a very powerful support to the scientific community to take a final decision towards repurposing GS-441524 as SARS Cov 2 antiviral drug.

Materials and Methods

Datasets:
Human furin protein had 29 bioassays containing 221066 tested compounds in PubChem Repository on 07-07-2020 (**https://pubchem.ncbi.nlm.nih.gov/protein/EAX02111**). All furin actives and inactives are taken for analysis except inactive of AID 463115 (~2.2 lakh) to

minimize the bias due to imbalance of diversity of actives and inactives[39] . The 386 actives and 93 in-actives in structure format are downloaded from the PubChem repository which was tested on human Furin-Paired basic amino acid cleaving enzyme. Descriptor generation and subsequent model building is done using these compounds.

Among the 21 PDB structures available, 5MIM along with the bound ligand (C28H37N13O2, CID - 44593169,IUPAC Name:2-[4-[(1R,2R,4S,5S)-2-[bis (azanyl) methylideneamino]-5-[4-[bis(azanyl)methylideneamino]phenoxy]-4-[[4-[bis(azanyl)methylideneamino]phenyl]amino]cyclohexyl]oxyphenyl]guanidine) which is considered as reference molecule in this study) is downloaded from the protein data bank as it has a single chain and zero mutations with a good X-ray Structure Validation Report (**https://www.rcsb.org/structure/5MIM**). Antiviral drugs which come under approved/investigational category in drugbank (**https://go.drugbank.com/**) are selected as screening dataset and their SDFs (total 180) are downloaded from PubChem repository.

Model building by SOM clustering:

The SDF files of furin actives and in-actives downloaded from PubChem bioassays were used to build a model using self-organizing maps {(the unsupervised machine learning clustering algorithm in Canvas of the Schrodinger suite (Schrodinger, Inc., LLC, New York, USA)} and to screen the activity of the FDA approved/investigational antiviral drugs. The physicochemical, topological, ligfilter and the qikprop descriptors generated fed as input into the neural network which in turn had undergone artificial neural network-based competitive learning to develop the model cluster. The SOM is generated by the property-based clustering using all available descriptors of categories mentioned above generated by the software.

Ligand docking:

The protein was pre-processed in the protein preparation wizard of maestro9.3 (Schrodinger, Inc., LLC, New York, USA) the states were generated at a pH 7.0 ± 2.0. The metals and the water molecules (Beyond 5A^{o} from the hit groups) were removed and minimized in OPLS-3e force field. The ligand is also prepared using the ligprep wizard of maestro taking all stable conformations up to 32. The grid for docking is prepared by taking the centroid of the PDB inbuilt inhibitor in Glide. Ligand docking is done with the prepared protein on the generated grid with the prepared ligand (the ligands from best clusters). The glide score along with glide emodel are tabulated and analyzed for binding affinity.

Network interaction analysis

The interactome was built for (i) FDA drug and human target network and (ii) FDA drug and viral target network in connection with human target furin. All the target proteins examined in this study of antiviral FDA drugs (supplementary table S6 and S7)which are considered active on furin and obtained from the drug bank. The information of targets for each drug is collected

from UniProt id and the interactions downloaded from bio grid interactions. The network is built and visualized in Cytoscape 3.8.2 (https://cytoscape.org/).

Molecular Dynamics simulation studies

Molecular dynamic simulations on furin-reference ligand, furin-GS-441524 and furin-Grazoprevir complexes were carried out using the GROMACS 5.1 software and the regular CHARMM36 force field. The ligand topology was optimised in a high-throughput way using CGenFF. The CHARMM General Force Field (CGenFF) tool combines atom typing and analogy assignment of parameters and charges. Atom typing is done by a testable configurable decision tree[40]. The role of a protein or enzyme is inextricably related to its internal and external dynamics. Since there are no experimental options, the only way to test protein dynamics is by quantitative analytical techniques. A customised GROMOS966 was used to run a 50-ns MD simulation to investigate protein-ligand binding and unbinding in terms of energy and geometry shifts, as well as the mechanism and functional behaviour of the furin protein after binding to both ligands subsequently. Furthermore, the energy of systems was reduced by using the steepest descent approaches. To keep the systems running for 100 ps, the canonical ensemble of NPT and NVT was used. For 50 ns, molecular dynamics simulation was done. The protein-ligand complexes main chain backbone atoms' root-mean-square divergence (RMSD), root-mean-square fluctuation (RMSF) the radius of gyration (Rg), and solvent-accessible surface area (SASA) values were calculated.

Acknowledgement

The authors are grateful to the support given by the ministry of education and innovation cell for software facilities as part of drug discovery Hackathon 2020, Govt of India.All authors are grateful to the “Covid Omics Research Consortium” (CORC) initiated by DrJayashankar Das, especially all members of CORC repurposing wing for instilling enthusiasm igniting the passion in these tough times.

Authors’ contributions

DM, MP and JD conceptualized the project and methodology. DM performed the data collection, data analysis, validation data curation and prepared the draft of the manuscript. MP, SS, SD and KD supported the analysis, writing of the manuscript and scientific discussions. AG supported during the validation, and preparation of the final manuscript. All authors contributed to the manuscript writing. All authors read and approved the final manuscript.

Reference

1. Greene, J. A. &Loscalzo, J. Putting the Patient Back Together — Social Medicine, Network Medicine, and the Limits of Reductionism. *N. Engl. J. Med.* 377, 2493–2499 (2017).
2. Santos, R. et al. Europe PMC Funders Group Europe PMC Funders Author Manuscripts A comprehensive map of molecular drug targets. 16, 19–34 (2019).
3. Cheng, F., Hong, H., Yang, S. & Wei, Y. Individualized network-based drug repositioning infrastructure for precision oncology in the panomics era. *Brief. Bioinform.* 18, 682–697 (2017).
4. Cheng, F. In silico oncology drug repositioning and polypharmacology. *Methods Mol. Biol.* 1878, 243–261 (2019).
5. Zumla, A., Chan, J. F. W., Azhar, E. I., Hui, D. S. C. & Yuen, K. Y. Coronaviruses-drug discovery and therapeutic options. *Nat. Rev. Drug Discov*. 15, 327–347 (2016).
6. Bhoyar, R. C. *et al.* High throughput detection and genetic epidemiology of SARS-CoV-2 using COVIDSeq next-generation sequencing. *PLoS One***16**, 1–16 (2021).
7. Pandya, M., Shah, S., Dhanalakshmi, M., Juneja, T. & Patel, A. Unravelling Vitamins as Wonder Molecules for Covid-19 Management via Structure-based Virtual Screening. Preprint athttps://doi.org/10.21203/rs.3.rs-144177/v1) 13 January (2021).
8. Wu, C. *et al.*Furin: A Potential Therapeutic Target for COVID-19. *iScience***23**, 101642 (2020).
9. Uhlén, M. et al. Tissue-based map of the human proteome. *Science* (80-.). 347, (2015).
10. Tian, S. A 20 Residues Motif Delineates the Furin Cleavage Site and its Physical Properties May Influence Viral Fusion. *Biochem. Insights 2*, BCI.S2049 (2009).
11. Tian, S., Huang, Q., Fang, Y. & Wu, J. FurinDB : A Database of 20-Residue Furin Cleavage Site Motifs , Substrates and Their Associated Drugs. 1060–1065 (2011) doi:10.3390/ijms12021060.
12. Selzer, P. &Ertl, P. Applications of self-organizing neural networks in virtual screening and diversity selection. *J. Chem. Inf. Model.* 46, 2319–2323 (2006).
13. Schneider, P., Tanrikulu, Y. & Schneider, G. Self-Organizing Maps in Drug Discovery: Compound Library Design, Scaffold-Hopping, Repurposing. *Curr. Med. Chem.* 16, 258–266 (2008).
14. Hristozov, Dimitar P., Tudor I. Oprea, and Johann Gasteiger. "Virtual screening applications: a study of ligand-based methods and different structure representations in four different scenarios." *Journal of computer-aided molecular design* 21, no. 10 (2007): 617-640.
15. Xiao, Y. De *et al.* Supervised self-organizing maps in drug discovery. 1. Robust behavior with overdetermined data sets. *J. Chem. Inf. Model.* 45, 1749–1758 (2005).
16. Levin, J. M. et al. Artificial intelligence, drug repurposing and peer review. Nat. Biotechnol. 38, 1127–1131 (2020).
17. Jinuraj, K. R. et al. Feature optimization in high dimensional chemical space: Statistical and data mining solutions. *BMC Res. Notes* 11, 1–7 (2018).
18. Dahms, Sven O., Guan-Sheng Jiao, and Manuel E. Than. "Structural studies revealed active site distortions of human furin by a small molecule inhibitor." *ACS chemical biology* 12, no. 5 (2017): 1211-1216.
19. Meanwell, N. A. Improving drug candidates by design: A focus on physicochemical properties as a means of improving compound disposition and safety. *Chem. Res. Toxicol.* 24, 1420–1456 (2011).
20. Cheng, F. *et al.* Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.***9**, 1–12 (2018).
21. Pandya, M., Jani, S., Dave, V., &Rawal, R. Nanoinformatics: An Emerging Trend in Cancer Therapeutics. Nanobiotechnology: Concepts and Applications in Health, Agriculture, and Environment, 135. (2020).
22. Lata, K. S. *et al.* Exploring Leptospiral proteomes to identify potential candidates for vaccine design against Leptospirosis using an immunoinformatics approach. *Sci. Rep.***8**, 1–15 (2018).

23. Khandia, R. *et al.* Analysis of nipah virus codon usage and adaptation to hosts. *Front. Microbiol.***10**, 1–18 (2019).
24. G. Schneider, Paul Wrede b .,Artificial neural networks for computer-based molecular design *Progress in Biophysics & Molecular Biology 70*, 175±222 (1998)
25. Anighoro, A., Bajorath, J. &Rastelli, G. Polypharmacology : Challenges and Opportunities in Drug Discovery Department of Life Science Informatics , B-IT , LIMES Program Unit Chemical Biology and Medicinal. *J. Med. Chem*. 57, 7874–7887 (2014).
26. Murray, L., Loftin, C. & Gwyneth, C. Elevated Activity. 65, 1318–1323 (1985).
27. Thepen, T., et al. "Fcγ receptor 1 (CD64), a target beyond cancer." *Current pharmaceutical design* 15.23 (2009): 2712-2718.
28. Tang, X. et al. Increased expression of phosphoenolpyruvate carboxykinase cytoplasmic isoform by hepatitis B virus X protein affects hepatitis B virus replication. J. Med. Virol. 91, 258–264 (2019).
29. Warren, Travis K., et al. "Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys." *Nature* 531.7594 (2016): 381-385.
30. Li, Y. *et al.*Remdesivir Metabolite GS-441524 Effectively Inhibits SARS-CoV-2 Infection in Mouse Models. *J. Med. Chem.* (2021) doi:10.1021/acs.jmedchem.0c01929.
31. Behera, Santosh Kumar, et al. "Computational drug repurposing study elucidating simultaneous inhibition of entry and replication of novel corona virus by Grazoprevir." *Scientific reports* 11.1 (2021): 1-11.
32. Bafna, Khushboo, et al. "Hepatitis C Virus Drugs Simeprevir and Grazoprevir Synergize with Remdesivir to Suppress SARS-CoV-2 Replication in Cell Culture." *bioRxiv* (2020).
33. Nguyenla, Xammy, et al. "Discovery of SARS-CoV-2 antiviral synergy between remdesivir and approved drugs in human lung cells." *bioRxiv* (2020).
34. Ni, X. *et al.* Structural Insights into Plasticity and Discovery of Remdesivir Metabolite GS-441524 Binding in SARS-CoV-2 Macrodomain. *ACS Med. Chem. Lett.* (2021) doi:10.1021/acsmedchemlett.0c00684.
35. Jung, L. S., Gund, T. M. & Narayan, M. Comparison of Binding Site of Remdesivir and Its Metabolites with NSP12-NSP7-NSP8, and NSP3 of SARS CoV-2 Virus and Alternative Potential Drugs for COVID-19 Treatment. *Protein J.***39**, 619–630 (2020).
36. Humeniuk, R. *et al.* Pharmacokinetic, Pharmacodynamic, and Drug-Interaction Profile of Remdesivir, a SARS-CoV-2 Replication Inhibitor. *Clin. Pharmacokinet.* (2021) doi:10.1007/s40262-021-00984-5.
37. Yang, K. What Do We Know About Remdesivir Drug Interactions? *Clin. Transl. Sci.***13**, 842–844 (2020).
38. Li, D. et al. Reviews on Biological Activity, Clinical Trial and Synthesis Progress of Small Molecules for the Treatment of COVID-19. Topics in Current Chemistry vol. 379 (*Springer International Publishing*, 2021).
39. Gad, A. et al. Virtual screening and repositioning of inconclusive molecules of beta-lactamase Bioassays—A data mining approach. *Comput. Biol. Chem*. 70, 65–88 (2017).
40. Vanommeslaeghe, K. and MacKerell Jr, A.D., “Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing”. *Journal of chemical information and modeling,* 52(12), pp.3144-3154.(2012)

Figures

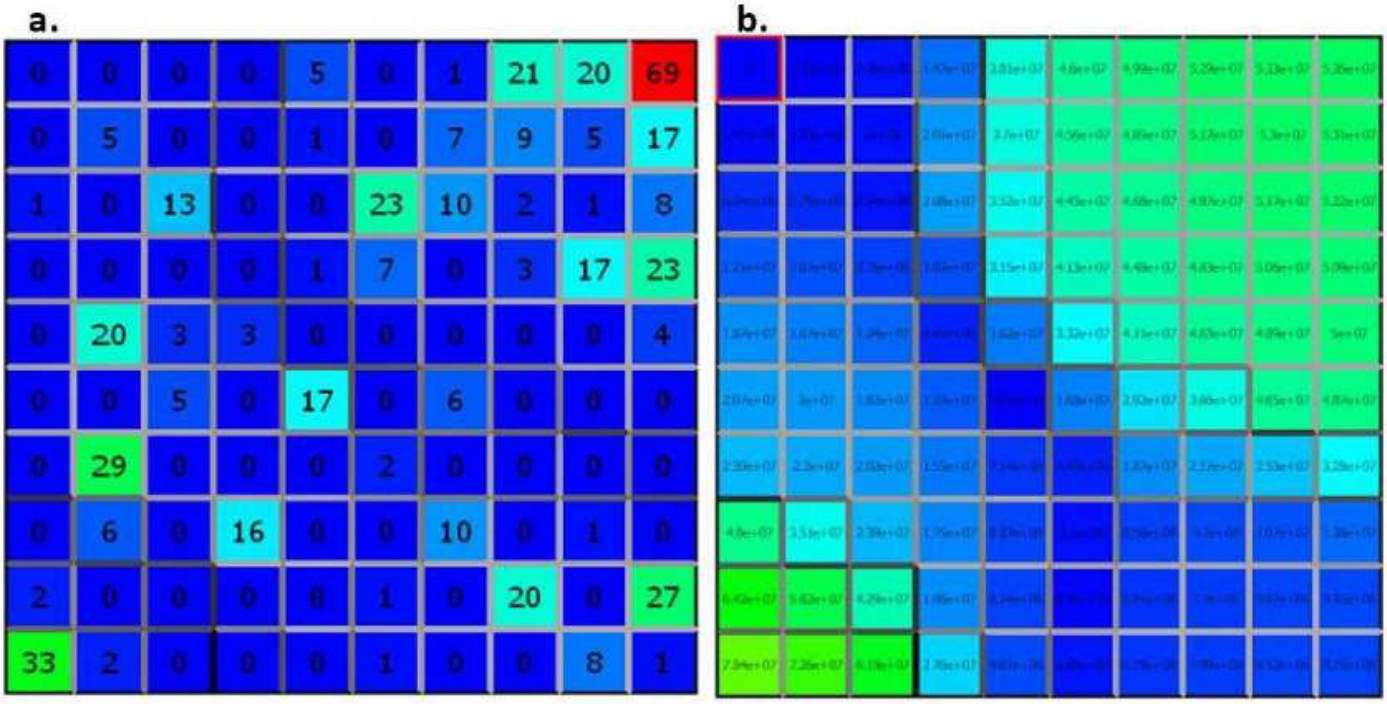


Figure 1

(a) Cell Population Analysis of antiviral FDA drugs and bioassay actives of Furin (b) Euclidean distance between the antiviral FDA drugs and bioassay actives of Furin in the property based SOM. (SOM Colour Schema: If the difference is high, then that ensemble is represented in red, and if the distance is low, it is represented in blue. The various colours between blue and red are used to represent the variations among the properties, population, and the distance)

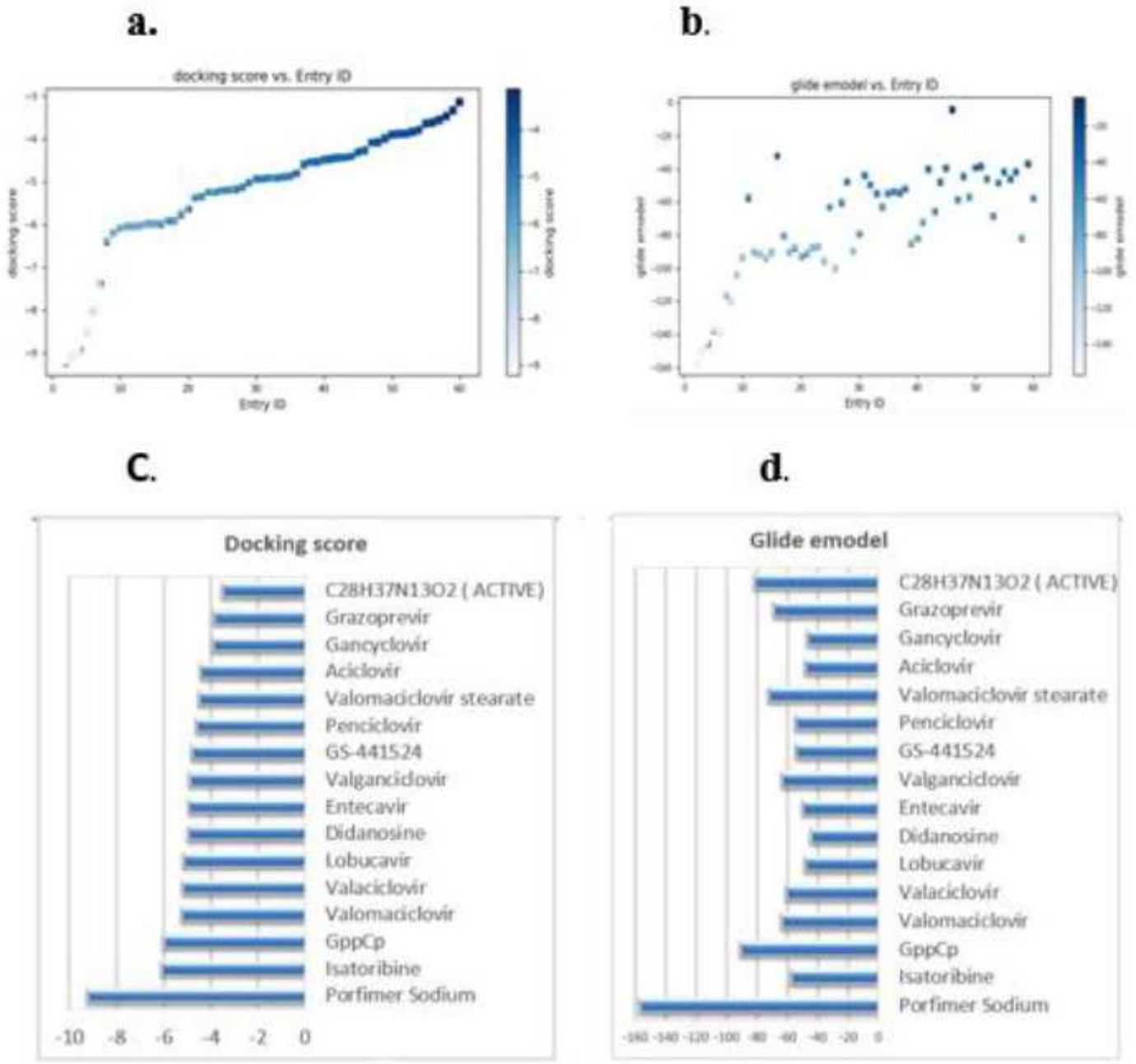


Figure 2

(a) Molecular docking score of all molecules from cluster 2 and cluster 4 (b) Glide e-model of all molecules from cluster 2 and cluster 4 (c) Molecular docking score of FDA drugs from cluster 2 and cluster 4 (d) Glide e-model of FDA drugs from cluster 2 and cluster 4

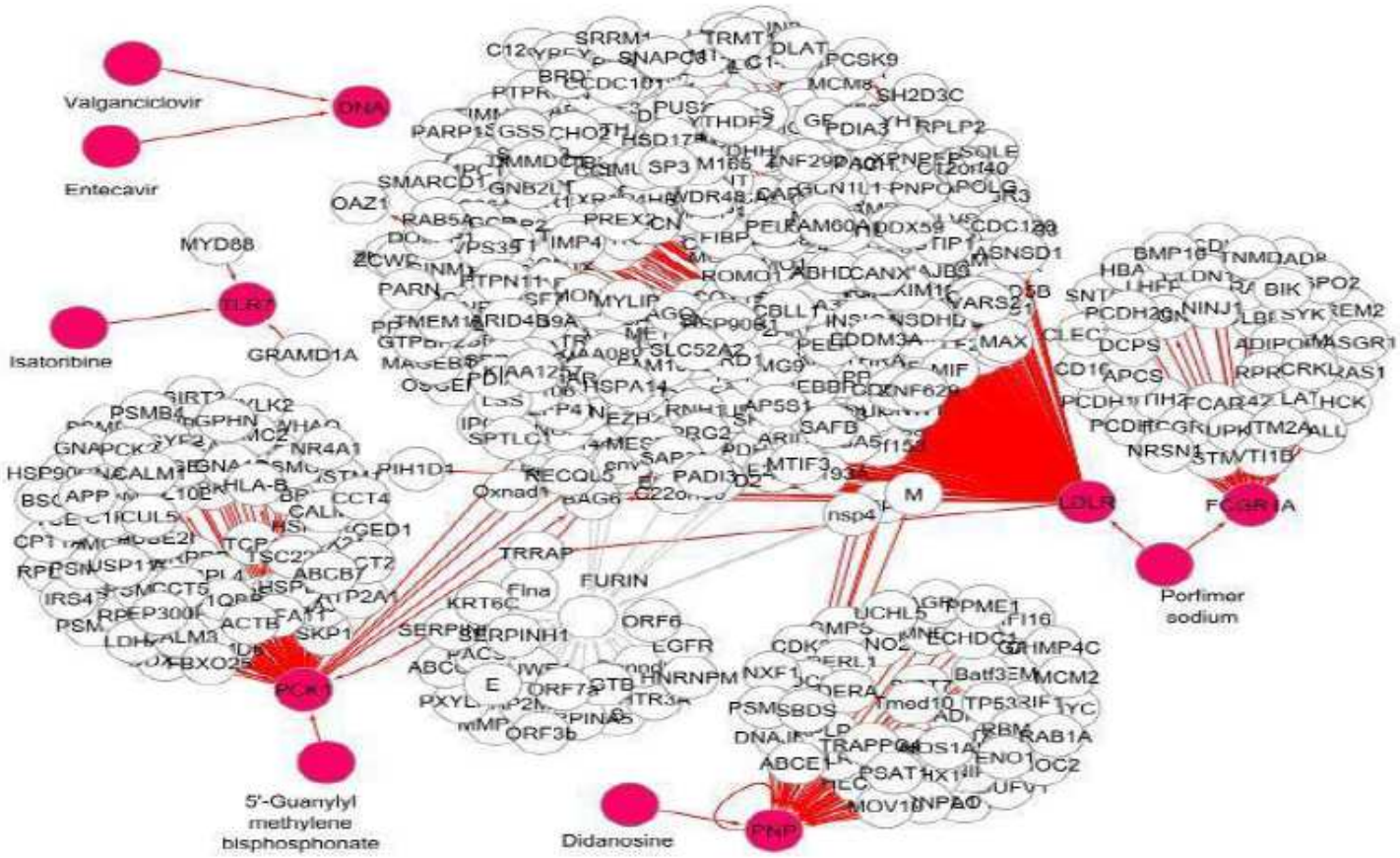


Figure 3

Drug target network analysis for drugs acting on human targets, the red circles are drugs red lines are edges connecting their targets, white with labels are their interacting proteins .

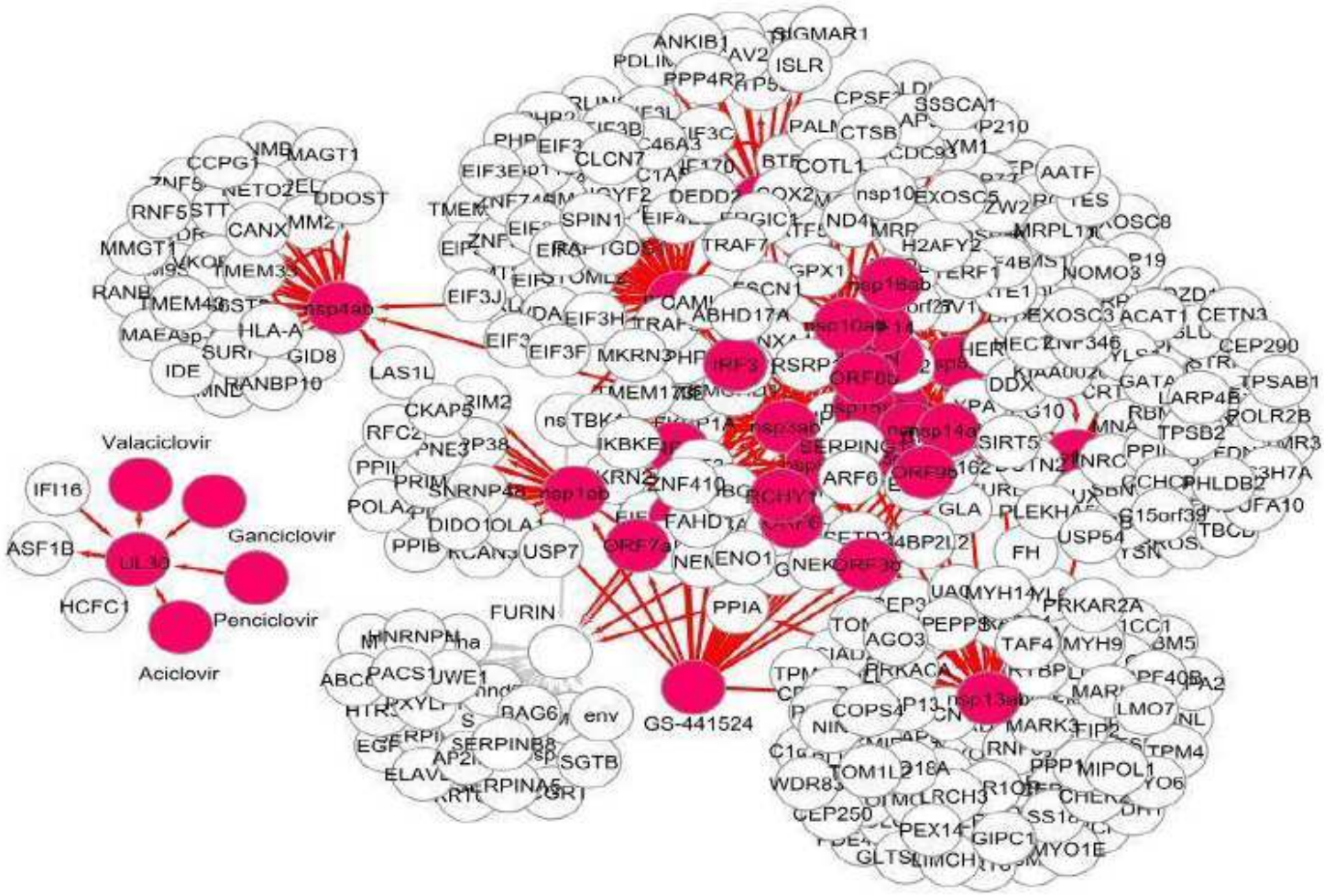


Figure 4

Drug target network analysis for viral host interactome: furin with drugs acting on viral targets, GS-441524 having maximum interactions with SARS-COV2 related proteins, the red circles are drugs red lines are edges connecting their targets and it reaches to furin through common targets

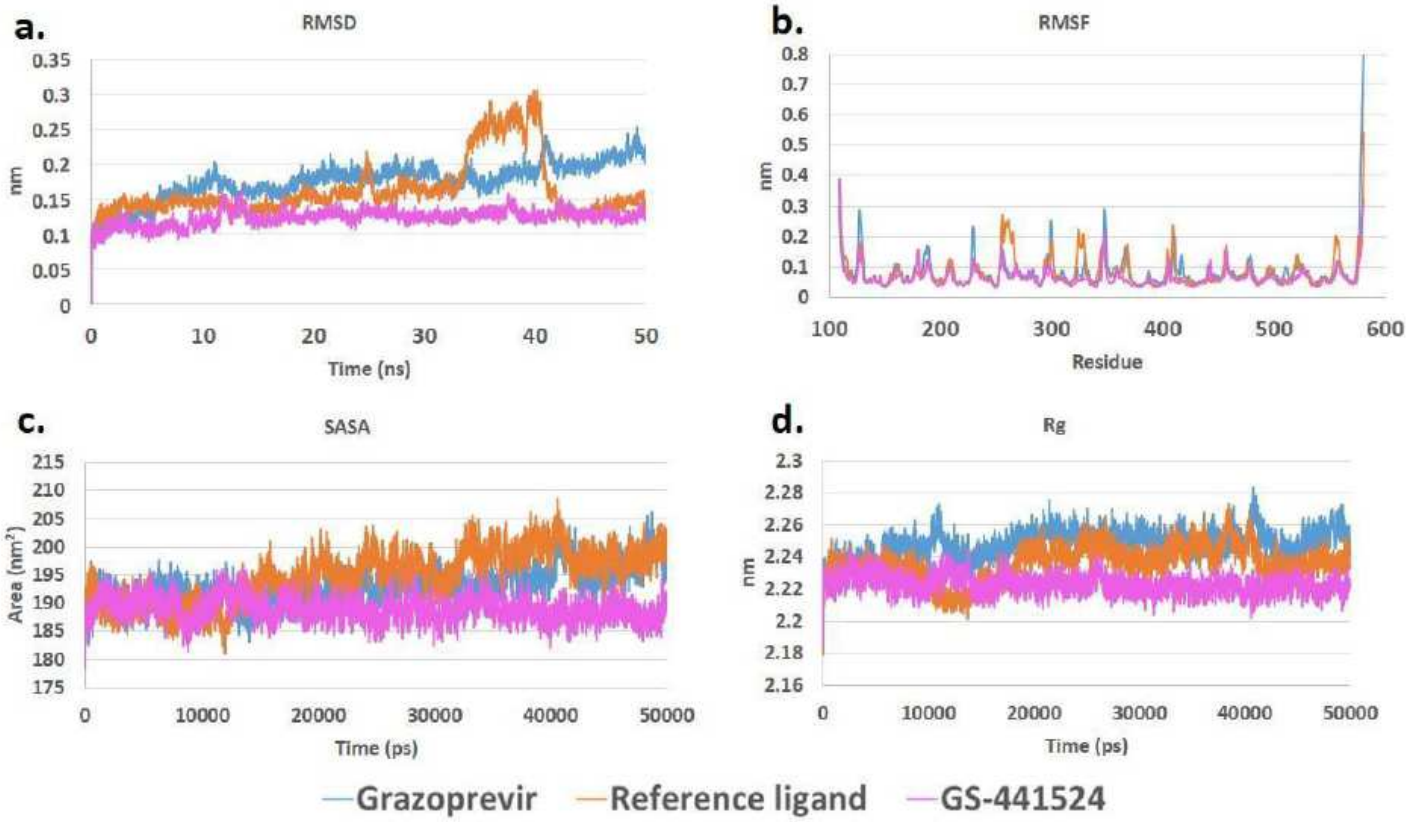


Figure 5

Molecular dynamics simulation analysis of protein-ligand complex of furin-Grazoprevir, furin-reference ligand, and furin-GS-441524: (a) Root means square deviation (RMSD), (b) Root means square fluctuation (RMSF), (c) solvent-accessible surface area (SASA),(d) Radius of gyration (Rg).

Supplementary Files

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

- SupplementaryTableS1S2S3.xlsx
- SupplementaryTableS4S5.pdf
- SupplementaryTableS6S7.xlsx
- GraphicalAbstract.png