 doi:10.20944/preprints202003.0393.v1

Virtual screening and molecular dynamics simulation suggest Valproic acid Co-A could bind to SARS-CoV2 RNA depended RNA polymerase

*Anupam Patra and Neel Sarovar Bhavesh**

Transcription Regulation group, International Centre for Genetic Engineering and Biotechnology (ICGEB), Aruna Asaf Ali Marg, New Delhi 110067 INDIA.

*To whom correspondence should be addressed. E-mail: neelsb@icgeb.res.in,

Abstract

SARS-CoV2 RNA depended RNA polymerase is an essential enzyme for the survival of the virus in hosts as it helps in the replication of viral RNA. There are no human polymerases that share either sequence or structural homology with viral RNA depended RNA polymerase. These make it a good target for inhibitor discovery, as a specific inhibitor cannot cross-react with the human polymerases. We have used virtual screening, docking, binding energy calculation and simulation to show that valproic acid Co-A, a metabolite from prodrug valproic acid, forms stable interaction with nsP12 of CoV. Our results suggest valproic acid Co-A could be a potential inhibitor of nsP12 of SARS-CoV2.

Keywords: SARS-CoV2, RNA depended RNA polymerase, Valproic acid Co-A, drug repurposing

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 March 2020 doi:10.20944/preprints202003.0393.v1

Introduction

A combination of high mutation rates[1], periodic recombination[2] and the environmental adaptation dictates the evolutionary fate of the RNA viruses. From the RNA world to modern era[3], RNA viruses can live in almost all ecological niche ranging from prokaryotes to higher organism serving as host[4]. The world has witnessed several endemic or pandemic situations, which includes Polio, SARS, MERS, Dengue, Chikungunya, Zika, and the most recent coronavirus disease. WHO has characterized coronavirus outbreak as a pandemic. SARS-CoV-2 from *betacoronavirus*[5] genus comprises of a 30 kb large positive single stranded RNA, leading to severe pulmonary infection, called the **co**rona**vi**rus **d**isease (COVID-19). As of 21st March 2020 more than 280,000 confirmed cases, 89,000 recovered cases and 11,000 deaths has been reported globally. The viral outbreak begun in December 2019 in Wuhan, China and has spread to most parts of the world since then affecting all age groups. While the severity of infection by the virus depends on the host immune system, health condition and disease history of the patient pose several pharmacological challenges in controlling it.

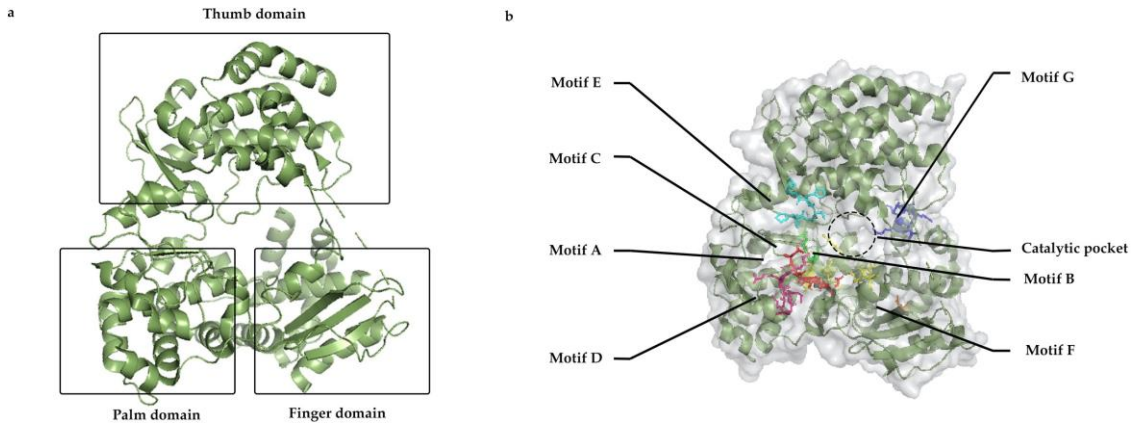

Figure 1: **a**) General structure of RdRp with three major domain; Thumb, palm and finger, **b**) distribution of motifs A to G along with inner catalytic channel.

It is known that RdRp is an essential non-structural key protein which can decipher the RNA genome to replicate and provide the viral prolongation and keep a firm hold on their natural selection[6]. Despite the versatile sequence diversity, RdRp has conserved catalytic domains. General architecture of the RdRp is like a folded right hand with catalytic palm[7] with supportive thumb and finger domain. There are at least seven motifs (A to G) around those three major domains[8]. Motif A to C represent the catalytic sites of the inner channel of the polymerase and they are responsible for the nucleotide addition cycle (NAC)[9] (Fig 1). Therefore, nsP12 the RNA dependent RNA polymerase (RdRp) of the virus could be a potent

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 March 2020 doi:10.20944/preprints202003.0393.v1

target for inhibitor design and discovery. Our *in silico* studies suggest targeting NAC on the active catalytic site of polymerase, could stop the RNA replication and because of structural difference in the viral polymerase and a specific inhibitor cannot cross-react with the human polymerase. We show here that valproic acid as a promising drug target against the virus RNA dependent RNA polymerase, nsP12 could be a promising antiviral strategy.

Material and Methods

Target selection and ligand preparation

Two targets were considered for the *in silico* study; which are RNA dependent RNA polymerases (RdRps) from Dengue (PDB id. 2J7U) and Coronavirus (SARS-CoV) (PDB id. 6NUR). The atomic structures were acquired from protein data bank RCSB (PDB)[10]. Protein preparation wizard[11] from the Schrodinger maestro platform (Schrödinger Release 2020-1: Protein Preparation Wizard; Schrödinger, LLC, New York, NY, 2020) were used to preprocess which assign bond order along with adding hydrogen atoms. The macromolecular structures were also refined with sample water orientations and optimized with the PROPKA program at pH 7.0. Further restrained minimization was performed using the OPLS3e force field[12]. Ligands or small molecules were obtained from respective databases, i.e., Drugbank[13], NCI[14], and ZINC[15]. Ligprep tool[11] (Schrödinger Release 2020-1: LigPrep, Schrödinger, LLC, New York, NY, 2020) from the maestro platform and used to process the ligand with their tautomeric combination structures along with possible state at pH 7.0 ± 2.0 under OPLS3e force field. A 20 Å receptor grid was generated around the specified residues of the catalytic sites of target macromolecules. The rest of the parameters were kept as default during Receptor Grid Generation using Glide.

Virtual Screening

The screening was set up with a combination of nearly 1.2 million small molecules in the Glide tool of maestro. Screening quality was determined by a three tier virtual screening using High throughput virtual screening (HTVS) as the first approach, followed by standard precision (SP) and extra precision (XP). From each database, the top two hundred molecules were reviewed to further extensive screening. The selection of suitable inhibitor for the macromolecule was done using docking score and conformational state penalty. The final top two hundred molecules were assorted in seven clustered groups based on their interaction patterns by interaction fingerprint tool of the maestro platform.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 March 2020 doi:10.20944/preprints202003.0393.v1

MM-GBSA score and ADME properties calculation

Ligand binding energy and ligand strain energy was calculated to validate the docking score further confirming the stability of a ligand in complex with the macromolecule. The calculation was performed using the prime tool (Schrödinger Release 2020-1: Prime, Schrödinger, LLC, New York, NY, 2020.) in the maestro platform under the OPLS3e force field and default VSGB solvation model. QikProp tool (Schrödinger Release 2020-1: QikProp, Schrödinger, LLC, New York, NY, 2020.) was used to compute the ADME and other pharmacological properties of the ligand of interest.

Molecular dynamics simulation

For simulation purposes, a system was set up to include macromolecule-ligand complex with solvent by System builder from Desmond[16] (Schrödinger Release 2020-1: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2020) application of maestro platform. TIP3P system, along with an orthorhombic box, was used to model the solvent. The water-filled box added with 150 mM of NaCl ions to neutralize the systems. Next, the Molecular dynamics tool of Desmond program was used for simulation time set up for 200 ns under normal NTP (constant number of particles (N), pressure (P), and temperature (T)) condition where the model system was relaxed before the simulation. The simulation interactive diagram tool further analyzed the simulation data in a graphical way, which provided the information of protein-ligand complex properties during the simulation time. Root Mean Square Distance (RMSD), Root Mean Square Fluctuation (RMSF), and bond interaction data were analyzed to further validate our findings in virtual docking.

Results and Discussion

Virtual screening suggests valproic acid co-A is a potential potent inhibitor

The top two hundred small molecules from three major ligand databases, *i.e.*, NCI database, Drug bank, and ZINC database upon virtual screening were selected and grouped in seven cluster shown in Table 1 and each lead molecule from the each cluster were analyzed. Interestingly 188 molecules were under clusters 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 March 2020 doi:10.20944/preprints202003.0393.v1

Table1: Seven cluster groups from virtual screening against Dengue-nsp5 represent one lead molecule and their docking score, binding energy, molecular weight, and other pharmacophore parameters are enlisted.

Cluster group	Id of lead molecule	Docking score	dG binding	Rule of 5	mw	donorHB	acceptHB	QPlogBB	QPlogS	PSA
1	DB03277	-13.147	-51.74	3	504.441	11	27	-4.658	0.602	264.510
2	DB02237	-12.126	-10.20	3	666.583	13	35	-6.588	0.931	349.932
3	DBMET01015	-10.112	-28.62	3	893.732	7	28	-9.157	-4.916	382.737
4	DBMET01907	-9.739	-37.74	3	584.960	7	28	-3.571	-4.916	218.171
5	DBMET01015	-9.238	-12.73	3	893.732	7	28	-7.684	-4.916	366.041
6	DBMET01015	-9.476	-6.68	3	893.732	7	28	-7.684	-4.916	365.368
7	DBMET02044	-9.139	-7.38	3	931.102	10	19	-5.691	-5.691	354.266

Clusters 1 and 2 were from the sugar-based core. Rests were under clusters 3, 4, 5, 6, and 7. Among them, from cluster 3, lead molecule is DBMET01015, which is valproic acid co-A showed most fitted interaction with dengue polymerase, as all catalytic motifs (A, B, and C) were involved in interaction.

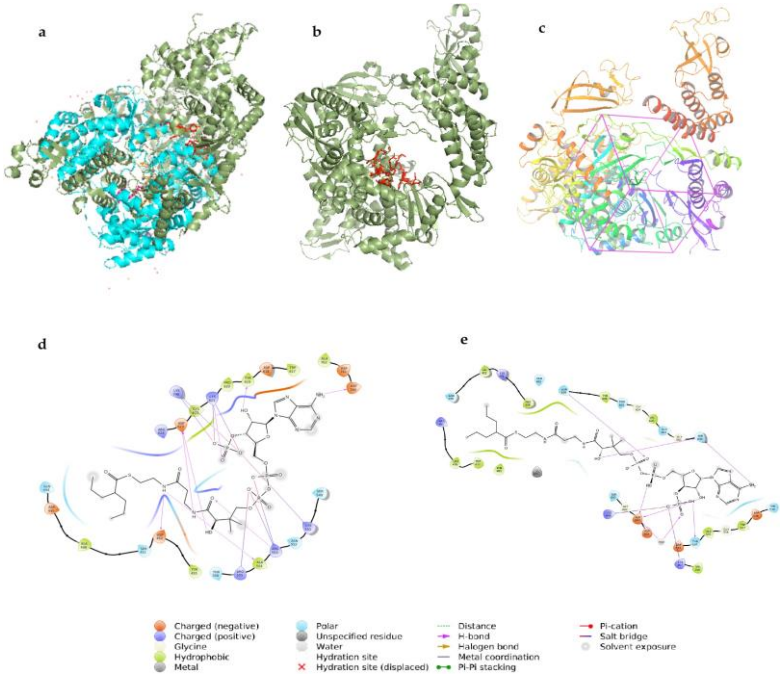

Figure 2: **a**) Super-imposed structure of Dengue-nsP5 (cyan) and CoV-nsP12 (green), **b**) Location of catalytic sites (includes motif A, B, and C) in CoV-nsp12, **c**) Grid generation for virtual docking inside the CoV-nsP12, **d**) Interacting residues in DBMET01015-Cov-nsP12, and **e**) Interacting residues in DBMET01015- Dengue-nsP5 complex.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 March 2020 doi:10.20944/preprints202003.0393.v1

The sequence and structure comparison between the two RNA dependent RNA polymerase of Dengue-nsP5 and SARS-CoV2, nsP12 showed that despite having mere 42% similarity in sequences, functional domains remain conserved having same topographic region[17] (Fig. 2a). On the basis of above finding, DBMET01015 *i.e.*, valproic acid co-A was subjected to docking with RNA dependent RNA polymerase of SARS-CoV2, nsP12 (PDB id 6NUR), which has similar catalytic architecture as dengue RdRp nsP5 (Fig. 2b). Around the catalytic sites, similar docking grid is considered for CoV-RdRp (Fig. 2c). *In silico* docking with RdRp of coronavirus shows the motif A (615-627) and motif C (754-765) involved in lead interaction with the valproic acid co-A (Fig. 2d). ASP 623, LYS 621, and TYR 619 from motif A along with ASP 760 from motif C are playing important role in interaction and stabilization of the complex. Here, valproic acid Co-A serve as a good candidate for RdRp blocker because of its interaction with the ASP residue from motif A and C are very crucial in blocking its interaction with divalent metal ions for phosphoryl transfer in the polymerase reaction[18].

Bond distances from nsP5-DBMET01015 complex, computed after interaction was found to be relevant to consider valproic acid co-A as the RdRp channel blocker involving all catalytic sites. Motif A (533-538) represents THR (534), which form H-bond (bond length 2.11 Å) with the phosphate group. Motif B (600-610) represents ASN (609) and SER (600), which form H-bond (bond length 3.09 Å and 1.78 Å respectively) with the phosphate group as well as the water molecule present inside. Motif C (662-665) represents ASP (663) as it forms the H-bond (bond length 1.58 Å and 2.03 Å) with the phosphate group as well as with the water molecule. The rest interactions belong to the finger domain of the polymerase, which complete the right-hand thumb model of RdRp. Ligand binding energy of nsP5-DBMET01015 complex is -28.62 kcal/mol, which represents the firm affinity of the ligand with stable complex structure (Fig. 2e).

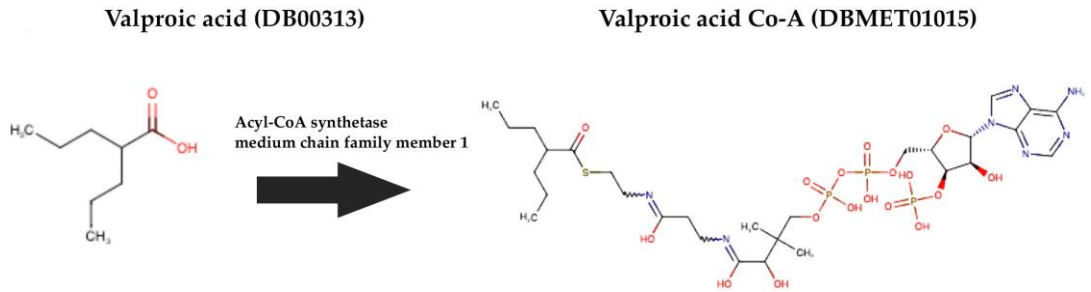

Figure 3: One step pathway to valproic acid Co-A by Acyl-CoA synthetase from prodrug valproic acid.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 March 2020 doi:10.20944/preprints202003.0393.v1

Similar contact patterns from the two virtual docking supports valproic acid co-A as a potential RdRp blocker. Valproic acid Co-A generates from prodrug valproic acid by one stem metabolic pathway (Fig. 3) by acetyl Co-A synthetase[19], which have moderate distribution on most of the human body organs. Despite of a GABA transaminase[20] blocker of neural receptor, valproic acid has extensive study history in antiviral treatment[21] including dengue infection[22].

The Molecular Mechanics/ Poisson-Boltzmann Surface Area (MM/GBSA) binding energies in both experiments are calculated i.e., Dengue-nsP5-Valproic acid Co-A with -28.62 kcal/mol and Cov-nsP12-Valproic acid Co-A with -31.70 kcal/mol. In both cases our results show strong interaction and stable contacts between the RdRp and inhibitor Valproic acid Co-A.

Molecular dynamics simulation shows valproic acid co-A forms stable complex with RdRp of SARS-CoV2

To further validate our docking result, molecular dynamics (MD) simulation system was set up for coronavirus-RdRp, nsP12 with its potent inhibitor DBMET01015 to find out the atomic positional changes during the contacts. MD simulation provides information about location of interacting amino acids and its distance patterns to form a stable inhibition complex. From the simulation data, Root mean squire deviation (RMSD) value is shown in figure 4a. Both ligand and macromolecule RMSD stays in between 2 Å, which implies the binding stability of the complex. Figures 4b and 4c show the Root Mean Square Fluctuation (RMSF) of the protein and the ligand. The docking result shows the most interacting residues of catalytic motifs *i.e.*, ASP 623, LYS 621, TYR 619, and ASP 760 and at 30% cut-off based on all contacts perspective during simulation, ASP623 and ASP760 are the key holder. Other contacts points with the Valproic acid Co-A are basically from the finger domain of inner polymerase channel. LYS 593 along with ASP 623, and ASP 760 are engaged with the two end of the ligand and increases the stability. Motif B (679-693) also takes part in the molecular contacts.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 March 2020 doi:10.20944/preprints202003.0393.v1

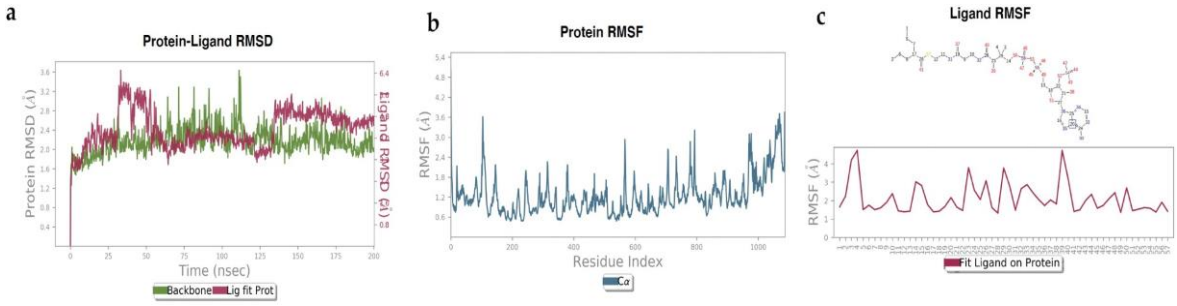

Figure 4: **a**) RMSD values during the simulation of DBMET01015-Cov-nsP12 complex, **b**) RMSF of Cov-nsP12, and c) RMSF of DBMET01015.

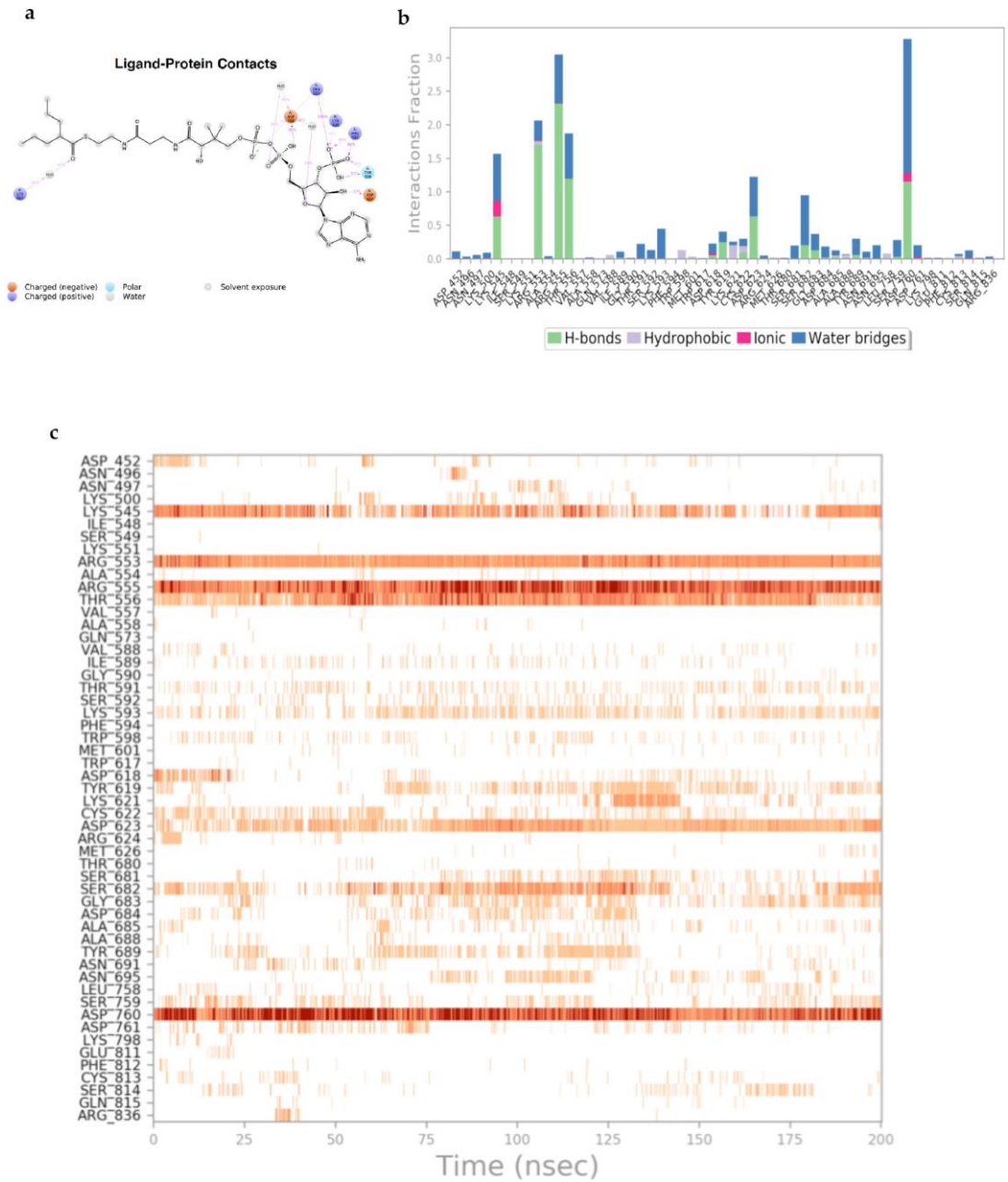

Figure 5: **a**) Protein-ligand contacts at 30% consistency of DBMET01015-Cov-nsP12 complex, **b**) Interaction patterns of DBMET01015-Cov-nsP12 complex, and **c**) Contact points in total residues during whole simulation trajectory of DBMET01015-Cov-nsP12 complex.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 March 2020 doi:10.20944/preprints202003.0393.v1

Water based interactions are also observed as LYS 593 involves with the ligand integrated with a water bridge interaction (Fig. 5a). Figure 5b shows multiple contacts between the interacting residues and the ligand, where maximum contacts are H-bond based along with water bridges. Combination of all the interactions in the allosteric site, an effective protein-ligand assembly significantly advocates a stable complex formation. All contacts throughout the simulation are shown in figure 5c. In view of all the interaction in the allosteric site, an effective protein-ligand assembly suggests a stable complex formation. Biochemical, cell culture based studies and animal studies will be required to further validate our results.

Conclusion

In view of the current outbreak and death toll scenario, novel drug design and discovery is a challenge constrained by time, but on the other hand, drug repurposing could help in the management. As valproic acid is already established drug for epilepsy management, our study with its metabolite, valproic acid Co-A could be a potential target for drug repurposing. Based on docking, binding energy calculation and simulation results our findings at this level, shows the basis of interaction and stabilization of the inhibitor with nsP12 the RNA dependent RNA polymerase (RdRp) of SARS-CoV2.

Acknowledgements

The study was funded by the ICGEB, New Delhi core research funds. Anupam Patra is supported by Senior Research fellowship from Council of Scientific and Industrial Research (CSIR), India.

Reference

1. Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. (2010) Viral mutation rates. *J. Virol.* **84**, 9733–48.
2. Becher, P. & Tautz, N. (2011) RNA recombination in pestiviruses: Cellular RNA sequences in viral genomes highlight the role of host factors for viral persistence and lethal disease. *RNA Biol.* **8**, 216–224.
3. Gilbert, W. (1986) Origin of life: The RNA world. *Nature* **319**, 618.
4. Koonin, E. V., Dolja, V. V. & Krupovic, M. (2015) Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology* **479**–**480**, 2–25.
5. Zhou, P. *et al.* (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. doi:10.1038/s41586-020-2012-7
6. Dolan, P. T., Whitfield, Z. J. & Andino, R. (2018) Mechanisms and Concepts in RNA Virus Population Dynamics and Evolution. *Annu. Rev. Virol.* **5**, 69–92.
7. Hansen, J. L., Long, A. M. & Schultz, S. C. (1997) Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure* **5**, 1109–22.
8. Jia, H. & Gong, P. (2019) A structure-function diversity survey of the rna-dependent rna polymerases from the positive-strand rna viruses. *Frontiers in Microbiology* **10**,.
9. Wu, J. & Gong, P. (2018) Visualizing the nucleotide addition cycle of viral RNA-dependent RNA polymerase. *Viruses* **10**,.
10. The Protein Data Bank. (2003) *Methods Biochem. Anal.* **44**, 181–198.
11. Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. (2013) Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* **27**, 221–234.
12. Harder, E. *et al.* (2016) OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **12**, 281–296.
13. Wishart, D. S. *et al.* (2008) DrugBank: A knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.*. doi:10.1093/nar/gkm958
14. NCI/CADD Group Chemoinformatics Tools and User Services. Available at: https://cactus.nci.nih.gov/index.html.
15. Sterling, T. & Irwin, J. J. (2015) ZINC 15 - Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **55**, 2324–2337.
16. Schrödinger Release. (2019) Desmond Molecular Dynamics System. *Schrödinger LLC*.
17. Kirchdoerfer, R. N. & Ward, A. B. (2019) Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* **10**, 1–9.
18. Yin, Y. W. & Steitz, T. A. (2004) The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* **116**, 393–404.
19. Ghodke-Puranik, Y. *et al.* (2013) Valproic acid pathway: Pharmacokinetics and pharmacodynamics. *Pharmacogenet. Genomics* **23**, 236–241.
20. Valproic Acid - an overview | ScienceDirect Topics. Available at: https://www.sciencedirect.com/topics/chemistry/valproic-acid.
21. Michaelis, M., Ha, T. A. T., Doerr, H. W. & Cinatl, J. (2008) Valproic acid interferes with antiviral treatment in human cytomegalovirus-infected endothelial cells. *Cardiovasc. Res.* **77**, 544–50.
22. Delgado, F., Cárdenas, P. & Castellanos, J. (2018) Valproic Acid Downregulates Cytokine Expression in Human Macrophages Infected with Dengue Virus. *Diseases* **6**, 59.