



OPEN

In silico study on the Hepatitis E virus RNA Helicase and its inhibition by silvestrol, rocaglamide and other flavagline compounds

Lorenzo Pedroni^{1,2}, Luca Dellaflora^{1,2}✉, Maria Olga Varrà¹, Gianni Galaverna¹ & Sergio Ghidini¹✉

Hepatitis E Virus (HEV) follows waterborne or zoonotic/foodborne transmission. Genotype 3 HEV infections are worldwide spread, especially in swine populations, representing an emerging threat for human health, both for farm workers and pork meat consumers. Unfortunately, HEV in vitro culture and analysis are still difficult, resulting in a poor understanding of its biology and hampering the implementation of counteracting strategies. Indeed, HEV encodes for only one non-structural multifunctional and multidomain protein (ORF1), which might be a good candidate for anti-HEV drugging strategies. In this context, an in silico molecular modelling approach that consisted in homology modelling to derive the 3D model target, docking study to simulate the binding event, and molecular dynamics to check complex stability over time was used. This workflow succeeded to describe ORF1 RNA Helicase domain from a molecular standpoint allowing the identification of potential inhibitory compounds among natural plant-based flavagline-related molecules such as silvestrol, rocaglamide and derivatives thereof. In the context of scouting potential anti-viral compounds and relying on the outcomes presented, further dedicated investigations on silvestrol, rocaglamide and a promising oxidized derivative have been suggested. For the sake of data reproducibility, the 3D model of HEV RNA Helicase has been made publicly available.

Hepatitis E virus (HEV) is a small positive sense single stranded RNA virus member of the *Hepiviridae* family (*Orthohepevirus* genus). It is the causative agent of the infamous Hepatitis E. As reported by the World Health Organization¹, there are 20 million HEV estimated infections per year and 3.3 million symptomatic cases of Hepatitis E worldwide. The fatality rate is usually relatively low, ranging from 0.2 to 4%, although it significantly increases for pregnant women².

The virus spreads following several routes with differences between high- and low-income countries. In the formers it is strictly related to zoonotic and foodborne transmission^{3,4} while in the latter it commonly gives waterborne outbreaks⁵. There are eight genotypes described so far (HEV-1 to -8), differing for host preferences and ways of transmission. Particularly, HEV-3, which is globally spread, and HEV-4, which is mostly limited to Asia, follow zoonotic and foodborne transmission^{6,7}. Reservoirs of HEV include deer, wild boars, cows, sheep, and goats with evidence of human infection reported to be caused by contaminated milk⁸⁻¹¹. Despite the wide range of potential hosts, the main reservoirs are pigs, and the related meat-based products are a major source of infection¹²⁻¹⁴. HEV was detected also in the berry fruit and leafy green vegetables supply chain because of irrigation with contaminated water^{15,16}.

Although this virus is widespread, in vitro culture and analysis are still difficult and consequently its molecular biology has not been fully understood yet¹⁷. One of the critical knowledge gaps concerns the HEV ORF1 gene, which encodes for a 185 kDa polyprotein with no cleavage sites reported¹⁸. ORF 1 is a multifunctional, multi-domain and non-structural polyprotein with a crucial role in the viral diffusion and replication, which may be a possible druggable target to interfere with the viral mechanisms of infection¹⁹. However, its huge dimension and multi-domain architecture make its analysis challenging via canonical molecular biology and structural

¹Department of Food and Drug, University of Parma, 43124 Parma, Italy. ²These authors contributed equally: Lorenzo Pedroni and Luca Dellaflora. ✉email: luca.dellaflora@unipr.it; sergio.ghidini@unipr.it

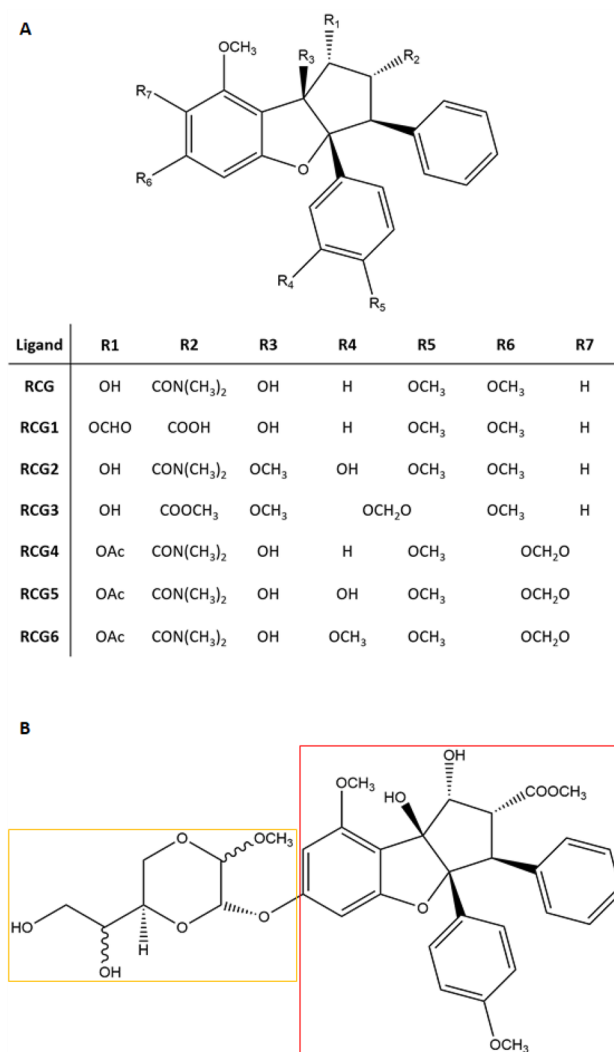


Figure 1. Chemical structure of RCG, RCG analogues and SLV. **(A)** RCG scaffold with its chemical substituents ranging from R1 to R7. **(B)** SLV chemical structure. The red square surrounds the RCG scaffold while the yellow the typical SLV dioxane portion. The latter is replaced by a hydroxyl group (–OH) in S–Ag.

approaches. In addition, the lack of available crystallographic and NMR structures hampers the identification and characterization of its druggable sites using canonical medicinal chemistry approaches. This lack of information ultimately prevents the structure-based identification of molecules targeting ORF1 domains for their possible implementation in anti-HEV strategies. In this context, *in silico* approaches can efficiently overcome the lack of structural data either to analyze proteins mechanics and druggability, or to provide a useful tool to study chemical and biological aspects of small molecules, including antiviral compounds^{20–23}. For this reason, an *in silico* procedure has been developed and applied to target HEV ORF1. Based on previous studies showing the inhibitory activity of the natural plant secondary metabolite silvestrol (SLV; Fig. 1) against HEV-3^{24–26}, our study provided a reliable model to: I) investigate the underpinning mechanisms and viral target; and II) to estimate the activity of SLV analogues for further dedicated investigations. To do so, the HEV RNA Helicase domain was modelled and refined via a homology modelling approach based on an innovative, hybrid structure- and sequence-based big-data analysis targeting the whole set of crystallographic data available so far in the Protein Data Bank (nearly 190.000 structures; last database access 28th February 2022). Then, the interaction of the model with the already known RNA Helicase inhibitor rocaglamide (RCG) and a set of 9 natural-related compounds²⁷ was calculated through docking and molecular dynamics (MD) simulations.

Results and discussion

Building and refining the HEV RNA Helicase model. The HEV RNA Helicase domain of the HEV ORF1 polyprotein was meant to be modelled via homology modelling (HM), as reported in the "Homology modelling" section, due to the absence of a crystallographic structure in the Protein Data Bank. This method allows the successful modelling of proteins whose 3D structure is missing and it is particularly useful when crystallographic investigations are challenging, as in the case of polyproteins²⁸.

Specifically, the HEV ORF1 domain's coding sequence of HEV RNA Helicase was chosen based on the results reported by Karpe and Lole²⁹ who succeeded to recombinantly express the putative HEV1 ORF1 RNA Helicase region (from amino acid 960 to 1204) proving its activity. The focus on HEV3 (UniProt AC Q6J8G2) was due to its prevalent foodborne/zoonotic transmission and its worldwide spread³⁰. The localization of the RNA Helicase domain on the HEV3 ORF1 was achieved based on the alignment with the HEV1 RNA Helicase sequence proved as active by Karpe and Lole. Particularly, the target sequence (Tar-Seq) HEV3 ORF1 region ranging from amino acid 975–1219 shared 90.6% identity and 95.5% similarity with the HEV1 RNA Helicase domain (Fig. S3, Supplementary materials).

Concerning the template selection for HM, it is typically obtained identifying homologous sequences with available 3D structure via the BLAST UniProt web-interface, which is a gold benchmark standard in this kind of studies³¹. However, using the Tar-Seq as input for a canonical BLAST search gave no results. This was likely due to the nature of the BLAST heuristic algorithm that are likely to fail aligning polyprotein domains over their entire sequence. Indeed, the BLAST UniProt web-interface searches target sequences within the whole primary protein sequences associated with 3D data and not limiting the search to the actual sequence resolved in the crystal structure. However, in most of the medium/large proteins, such as polyproteins, the crystals related to the UniProt ACs partially cover the protein primary structure causing BLAST to fail in identifying homologs with known structures. To overcome this intrinsic weakness, a successful workflow was setup (see in "Sequence and template selection" section for further details). The modification in the search space used in this approach made it robust, reusable, and successful to find a useful template to build a model for a polyprotein domain. In more detail, all the entries having a 3D structure belonging to the Prokaryotic reign with E.C 3.6.4.13 (RNA Helicase activity, according to Brenda classification³²) were downloaded from PDB. Then, 3D structure files were converted to FASTA files via an in-house script setting up a database which included the "crystallome" of the prokaryotic RNA Helicase available at the time of analysis. This database was finally searched for Tar-Seq homologous to identify a proper template for HM.

The best aligning protein among the 1940 PDB entries annotated as prokaryotic RNA Helicases at the time of analysis was chosen as template (see in the "Homology modelling" section for further details). It was the PDB entry having code 3WRY³³, which showed identity and similarity percentage to HEV RNA Helicase of 24.5% and 39.2%, respectively, and with alignment score and e-value of 63.5 bits and 1e-12, respectively. Of note, the identity percentage was high enough with respect to the alignment length to ensure a reliable modeling procedure in agreement with previous evidence³⁴. Such protein belongs to the *Tomato mosaic virus* (Tmv), a positive sense single stranded RNA virus belonging to the same HEV class. Before using it as template to model the HEV RNA Helicase domain, the last 160 N-terminal residues, which were organized in a self-standing sub-domain, were removed being not covered by our Tar-Seq. The model and its refinement process, reported in the "Model generation and refinement" section, allowed to obtain a reliable and stable model having 90% of the residues within the mostly favored regions, 10% in allowed regions and with no residues in generously or disallowed regions according to the Ramachandran plot (Fig. S4, Supplementary Materials). Furthermore, the model was also checked on the ProSA-web Server obtaining a Z-score of 6.14 which is within the range of scores typically found for native proteins of comparable size³⁵ further confirming the model reliability. The last check was the comparison of the model with the one produced by the blind deep learning-based web resource trRosetta³⁶: their structural alignment resulted in an RMSD lower than 2.6 Å, confirming the reliability of the entire procedure described above (Fig. S5, Supplementary materials).

The poly-purine RNA added via docking to the modelled protein to complete the HEV RNA Helicase complex was derived from the PDB structure 6JIM³⁷, an RNA Helicase of the *Chikungunya virus*, which is a species belonging to the *Alsuviricetes* class like HEV and Tmv^{38,39}.

Docking ligands and analyzing complex stability. For the sake of identifying compounds with potential anti-viral properties, the model was targeted with natural bioactive compounds belonging to the cyclopenta[*b*]benzofuran/flavagline class to identify promising candidates to test in further dedicated investigations. This class of compounds has been described to have a broad spectrum of activity including insecticidal, antifungal, anti-inflammatory and anticancer activities to cite but a few⁴⁰. However, the mechanisms of action underpinning those activities still need to be clarified, though RCG and other flavagline analogues proved to target prohibitins and to exert inhibitory activity towards RNA Helicases²⁷. SLV (Fig. 1B) was also included in the study as it was an already known inhibitor of HEV replication, although the underpinning mechanisms and molecular targets still need clarifications^{24–26}. Moreover, silvestrol aglycone (S–Ag;CID 24178739, Fig. 1B) was analysed due to the chemical similarity with both RCG and SLV, along with two virtual decoys to assess the procedural performances (ZINC ID ZINC8584442 and ZINC8387186; further details are reported in the "Docking" section).

The position to dock the ligands within the ligand binding site was defined based on the PDB structure 5ZC9 showing RCG engaged in a well-defined base–base stacking interaction (Fig. S2, Supplementary Materials).

As reported in Table 1, RCG showed the highest docking score followed by all its analogues. All of them clearly showed a stacking interaction except for RCG6 and the two decoys. Considering that the higher the score, the stronger the interaction, according to previous evidence^{41,42}, this result suggested the preferred interaction, and possibly the higher inhibitory activity, of RCG compared to the other analogues. On the other hand, SLV obtained a higher docking score than virtual decoys but sensibly lower than both RCG and S–Ag. Of note, the inactive decoys served to validate the model. In particular, the lack of activity was imputable to their incapability to keep a stable stacking interaction resulting in an enzyme inhibition. Therefore, the fact that they both showed no stacking interactions corroborated the model reliability.

Ligand	PLP score
RCG	256.075
RCG1	224.929
RCG2	246.042
RCG3	224.077
RCG4	231.621
RCG5	228.042
RCG6	169.175
SLV	184.600
S-Ag	236.039
ZINC8584442	178.539
ZINC8387186	150.527

Table 1. Docking PLP score obtained running GOLD. The highest the score the more likely the docking pose is optimal.

Once obtained the docking pose for each ligand, 25 ns long MD simulations were run to investigate the evolution of the built complexes. Based both on the 5ZC9 crystal structure and other evidence from the literature claiming the RCG stabilization of the protein-RNA complex⁴³, the following geometrical rules were associated to ligands theoretically able to interact with and inhibit HEV RNA Helicase. Specifically, they should: I) keep the stacking interaction with the RNA stably; and II) preserve the stability of the RNA-protein complex avoiding RNA detachment.

RCG and SLV were both engaged in a base-base stacking interaction, also stabilizing the protein-RNA complex with no appreciable RNA detachment (Fig. 2).

RCG1 did not show a proper stacking interaction and promoted the RNA detachment, pointing to its limited theoretical inhibitory activity against the viral RNA Helicase. Interestingly, this result is in line with the Pan et al. study reporting more than 500-fold activity decrease with respect to RCG against the human eIF4A1 ATP-dependent RNA helicase²⁷. RCG2 caused a self-collapse of the RNA, promoting its detachment from the protein, and no stacking interactions were observed. This also suggested a low inhibitory potential, in line with the inactivity reported by Pan et al.²⁷ against human RNA Helicase eIF4A1. RCG3, a less cytotoxic RCG analogue²⁷, showed a single base interaction, rather than a base-base stacking interaction, and promoted the RNA detachment. Therefore, it was not deemed an efficient inhibitor. RCG4 and RCG6 both favoured the RNA detachment, with the latter reported as less active than the former, although RCG4 was found in an RCG-like stacking interaction with the RNA. They were not considered able to appreciably inhibit the HEV RNA Helicase activity.

S-Ag did not interact at all with the RNA. Shortly after the beginning of the MD simulation it slipped out the RNA chain and kept a likely unspecific surface interaction with a near protein portion (Fig. 2). This was probably due to the presence of the hydrophilic hydroxyl group in position R6 (Fig. 1), which prevented a proper a stacking interaction, which requires hydrophobic interactions, and to the absence of the SLV dioxane portion. Indeed, this part participates in the SLV stacking interaction confirming the Cencic and co-workers' hypothesis claiming its crucial role for the SLV activity⁴⁴. Concerning the two decoys, they both promoted RNA detaching, particularly ZINC8584442. This, along the lack of base-base stacking, was expected for inactive compounds and confirmed the procedural performances.

Lastly, a novel RCG analogue, i.e. RCG5, was rationally designed and evaluated based on the calculated pose of other RCG analogues included in this study. Particularly, starting from docking poses it was noticed there was a hydrogen bond between the hydroxyl group in position R4 of RCG2 and an RNA base (Fig. 3). Moreover, looking at the progression of RCG2 MD simulation, the starting H-bond was not kept over time but a new one between the hydroxyl group in R4 and the RNA backbone was formed. This interaction was thought as favorably stabilizing the complex and theoretically able to enhance the activity of the compound. Therefore, in RCG5, the R4 of RCG4 (i.e., -H) was replaced with a hydroxyl group. RCG5 showed a stable stacking interaction while keeping the RNA-protein complex stable, steadily forming an H-bond with a phosphate group of the RNA backbone (Fig. 3). This newly designed compound needs further investigation and experimental evaluation to validate its HEV RNA Helicase inhibitory activity.

Conclusion

Keeping in mind the intrinsic issues in running experimental trials on HEV related to its biology and to the shortage of a robust cell-culture system supporting its life cycle⁴⁵, *in silico* approaches may ensure a useful framework of analysis to advance HEV understanding and design counteracting strategies. In this context, the present work provided a useful prioritization of compounds, most of which are not commercially available, supporting a knowledge-based and informed selection either for their synthesis or purification in future dedicated works. Overall, this study: (i) presented mechanistic insights on HEV RNA Helicases and its SLV/RCG-dependent inhibition; (ii) expanded the current understanding of the structure-activity relationship for SLV and RCG-related compounds; (iii) provided a blueprint for further analysis targeting HEV RNA Helicases. Moreover, this study described for the first time the inhibition of HEV RNA Helicase as a plausible mechanism of action of SLV, which was already described reducing HEV capability to penetrate cells. This ancillary mechanism, which deserves

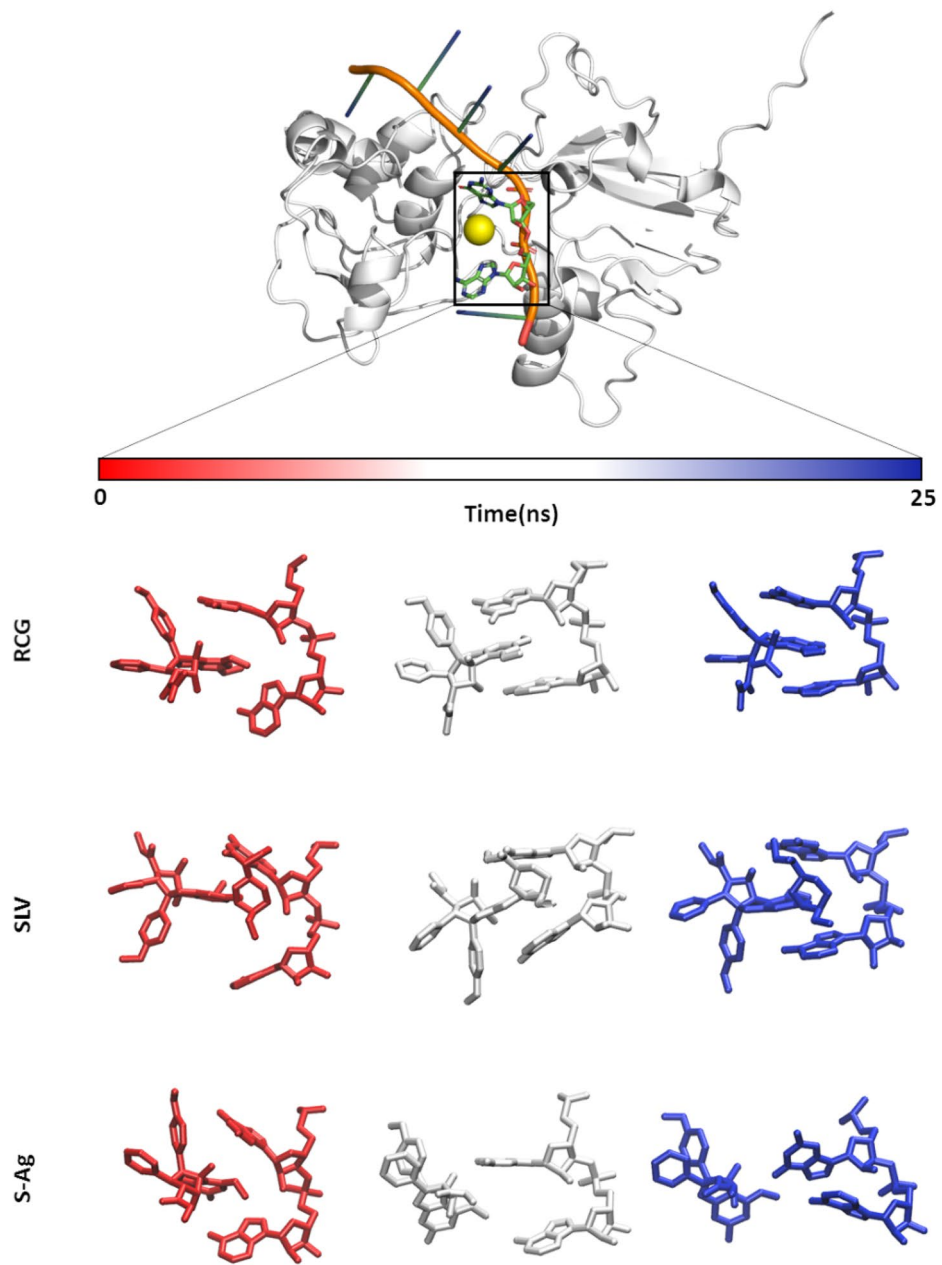


Figure 2. Protein-RNA-ligand complex and evolution over time of RCG, SLV and S-Ag with respect to the RNA bases. The protein is represented as a white cartoon, the RNA as an orange/green/blue cartoon while the two bases involved in the interaction with the ligand as green sticks. The yellow sphere represents the ligand docking-site. Under the time bar, starting from the top and moving to the bottom we can alternatively see the time-step sticks representation of both the RNA bases and RCG, SLV and S-Ag trajectories.

further investigation with high priority, was found plausible also for other RCG/SLV-like compounds and may complement the already reported activity of this HEV impairing agents on host RNA Helicases. Furthermore, there are evidence reporting SLV as well-tolerated in animals⁴⁰ making it an interesting candidate as a feed additive. The described methodology, starting from the template selection moving to the actual modelling and testing, succeeded to build a reliable model able to qualitatively discriminate several ligands. In addition to this, the pipeline is highly versatile, flexible and it can be replicated on other HEV ORF1 domains, such as the RNA Dependent RNA Polymerase which has already been proved as a suitable target to inhibit virus replication⁴¹.

As a general remark, the high identity percentage among the HEV RNA Helicases of genotypes 1 to 4 (all above 88%; see Fig. S6, Supporting material), which are those available in the manually curated section of UniProt database (last accessed 3rd August 2022), might suggest a broad inter-genotype activity for RCG, SLV and RCG6.

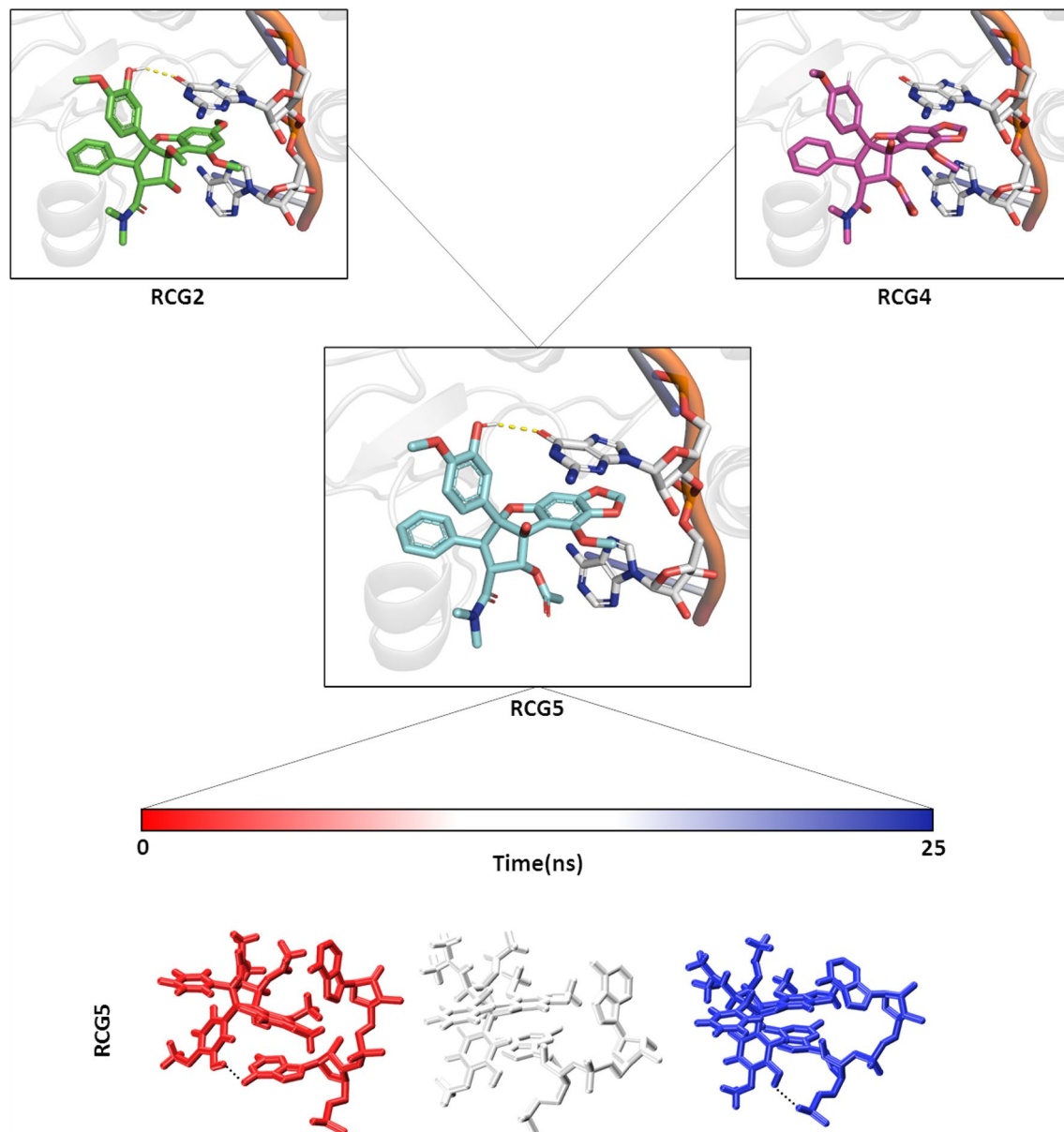


Figure 3. RCG2 is represented as green sticks, RCG4 as magenta sticks, RCG5 as pale-blue sticks. RNA bases involved in the interaction are represented as white sticks while the RNA backbone as orange cartoon. The yellow dashed lines represent the Hydrogen bond occurring between the R4 of both RCG5 and RCG2, and the RNA base. The time-step representation of the RCG5/RNA-bases trajectories is reported under the time bar. The black dashed lines represent both the starting Hydrogen bond and the final one.

For the sake of data reproducibility and to support further studies, the 3D model of HEV RNA Helicases used in this work has been made publicly available (https://github.com/FC-MMLAB-UniPr/HEV3_Helicase_Model).

Materials and methods

Homology modelling. *Sequence and template selection.* The primary protein sequence used to model the 3D structure of HEV RNA Helicase was stored in the UniProt database (release 2021_04) with the Accession Code (AC) Q6J8G2⁴⁶. As no crystallographic structures of HEV RNA Helicase were available in the Protein DataBank (PDB) at the time of analysis (last database access January 2022) a HEV RNA Helicase model was obtained via HM. To do so, a hybrid structure- and sequence-based strategy was developed as the search of publicly available PDB structures via the webserver BLAST interface (basic local alignment search tool, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>)⁴³, which is a goldstandard in this kind of study, did not gave useful structures for HM. Specifically, a set of non-eukaryotic proteins annotated as RNA Helicases (E.C 3.6.4.13, according to Brenda classification³²) with available 3D structure in the Protein Data Bank was collected (228 entries at the time of analysis). To retrieve the 3D structure of this set of proteins, their UniProt ACs were downloaded and mapped towards the Protein Data Bank using the “Retrieve/ID Mapping” tool available on UniProt. Starting

from their PDB IDs their 3D structures were iteratively downloaded (1940 at the time of analysis) in the .pdb format using a PyMol script developed in-house (available upon request). Each chain belonging to a .pdb file was converted to a fasta formatted file using an ad hoc script developed in-house (available upon request) and subsequently concatenated in a unique *multi.fasta* file. This *multi.fasta* file was converted into a local database of sequences via the *makeblastdb* command set with default parameters (ncbi-blast+ suite; version 2.11.0+) ⁴⁷ selecting the tabular output format. The best hit in terms of alignment score, e-value and identity percentage was considered to develop the homology model of HEV RNA Helicase.

Model generation and refinement. Modeller version 10.0 interfaced to Chimera (version 1.15) ⁴⁸ was used to generate the model of HEV RNA Helicase. The chosen template was the crystallographic structure with PDB code 3WRY (only the residues aligning the RNA Helicase domain were considered). Fifty models were generated, setting the inclusion of non-water heteroatoms, using a thorough optimization and choosing the best scored model according to zDOPE for subsequent analysis.

After computing a tentative model and checking its Ramachandran Plot with PROCHECK v3.5 ⁴⁹, regions with improper dihedrals (namely, residues 30–33; 101–103, 111–113 and 97–99) were stepwise refined with the Modeller loop-refinement tool (version 10.0 interfaced to Chimera ⁵⁰ version 1.15) using the DOPE modelling protocol, generating 5 models and carrying forth to the analysis the best model according to the zDOPE scoring.

Before using the model, a final assessment was performed by re-building its Ramachandran Plot and checking its Z-score on the ProSa-web Server ³⁵ to verify the proper topology.

Building the model-RNA complex. The HEV RNA Helicase-RNA complex was obtained docking the RNA sequence to the previously built RNA Helicase model via ClusPro 2.0 ³⁸ set with default parameters. The input RNA structure was chosen based on the following protocol. First, all the structures of RNA Helicase containing RNA were downloaded from Protein Data Bank and aligned to the model previously obtained. The RNA of the structure with the most similar organization (i.e., with the lowest RMSD value calculated in PyMol using the super command) was chosen (PDB ID 6JIM). In the last step, the RNA of 6JIM and the HEV RNA Helicase model were uploaded on ClusPro 2.0 ³⁸. Out of the obtained complexes the analysis of one of the mostly hydrophobic-favoured poses, showing the RNA molecules arranged in the same area as 6JIM structure, was carried forth. The sequence was then edited to a poly-purine fragment in agreement with previous studies reporting its suitability to interact with SLV-related compounds ⁵¹.

Docking. Docking studies were performed with GOLD (Genetic Optimization for Ligand Docking; version 2021.3) ⁵² to provide a plausible binding architecture for a set of flavagline compounds. The 3D structure of RCG (Fig. 1A), SLV (Fig. 1B) and S-Ag (Fig. 1B) were downloaded from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>; CID 331783, CID 11787114 and CID 24178739, respectively) ⁵³ in the .sdf format. The other RCG analogues (Fig. 1A) were generated editing the RCG structure using the PyMol Builder tool (version 2.3.0) and further optimized using Chimera (version 1.15) ⁵⁰ with the Minimize Structure tool (5000 steepest descent steps and 100 conjugate gradient steps).

The RNA Helicase model was used as input structure and the space to arrange ligands was set based on the architecture of binding of the RCG in the human eIF4A1 ATP-dependent RNA helicase having PDB code 5CZ9. The structure was visually aligned to the HEV RNA Helicase model, and the binding site was defined in a 10 Å-radius sphere around the centroid of the inter-bases space occupied by RCG in 5CZ9 structure (Fig. S1, Supplementary Materials). RCG was docked first generating 100 poses with no positioning constraints, setting the ligand fully flexible and allowing polar protein hydrogens free to rotate. The best scored pose according to PLP Scoring function (256 units; the higher the score, the more probable and favoured the ligand interaction) showed a comparable binding architecture to the crystallographic binding pose of RCG. Such pose was then used as a position restraint setting the similarity option with shape overlap (weight constraint 200). This helped docking SLV and other analogues facilitating their proper arrangement into the binding site.

Virtual decoys were also generated to test procedure performances via the DUD-E database Generate tool (<http://dude.docking.org/>) ⁵⁴. The 50 decoys generated were ranked according to chemical similarities to SLV using LiSiCA algorithm ⁵⁵ and the two extremes (the most similar and the most dissimilar compound; ZINC ID ZINC8584442 and ZINC8387186, respectively) (Fig. S2 Supplementary materials) were docked for the sake of procedure validation (see in the “[Docking ligands and analyzing complex stability](#)” section for further details).

Molecular dynamics. Molecular dynamics (MD) simulations were performed to investigate the overall geometrical stability of HEV RNA Helicase-ligands complexes over time. The adopted software was GROMACS (version 2019.4) ⁵⁶ with CHARMM27 all-atom force field parameters support ⁵⁷. All the ligands have been processed and parameterized with CHARMM27 all-atom force field using the SwissParam tool (<http://www.swissparam.ch>) ⁵⁸. Input structures were solvated with SPC/E waters in a rhombic dodecahedron periodic boundary condition, and counter ions (Na⁺ or Cl⁻) were added to neutralize the system. Prior to perform molecular dynamic simulations, the systems were energetically minimized to avoid steric clashes and to correct improper geometries using the steepest descent algorithm with a maximum of 50,000 steps. Afterwards, all the systems underwent isothermal (300 K, coupling time 0.1 ps) and isobaric (1 bar, coupling time 2 ps) 100 ps simulations before running a 25 ns MD simulation.

Multiple sequence alignment. A multiple sequence alignment has been performed to infer the activity of RCG, SLV and RCG6 over the HEV RNA Helicases of different HEV genotypes. The analysis focused on genotypes 1 to 4 as they were those publicly available on the manually curated and reviewed section of UniProt

database (last accessed 3rd August 2022). We performed the analysis using Clustal Omega Web Server (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) with default parameters⁵⁹.

Data availability

Data are available upon a formal request to the corresponding author. The HEV RNA Helicase model used in this study is available at https://github.com/FC-MMLAB-UniPr/HEV3_Helicase_Model.

Received: 6 June 2022; Accepted: 5 September 2022

Published online: 15 September 2022

References

- World Health Organization. *Hepatitis E*, <<https://www.who.int/news-room/fact-sheets/detail/hepatitis-e>> (2021).
- Webb, G. W. & Dalton, H. R. Hepatitis E: An underestimated emerging threat. *Ther. Adv. Infect. Dis.* <https://doi.org/10.1177/2049936119837162> (2019).
- Meng, X. J. Zoonotic and Foodborne Transmission of Hepatitis E Virus. *Semin. Liver Dis.* **33**, 41–49. <https://doi.org/10.1055/s-0033-1338113> (2013).
- Yugo, D. M. & Meng, X. J. Hepatitis E virus: Foodborne, waterborne and zoonotic transmission. *Int. J. Environ. Res. Public Health* **10**, 4507–4533. <https://doi.org/10.3390/ijerph10104507> (2013).
- Carratala, A. & Joost, S. Population density and water balance influence the global occurrence of hepatitis E epidemics. *Sci. Rep.* <https://doi.org/10.1038/s41598-019-46475-3> (2019).
- Dalton, H. R., Kamar, N. & Izopet, J. Hepatitis E in developed countries: current status and future perspectives. *Future Microbiol.* **9**, 1361–1372. <https://doi.org/10.2217/fmb.14.89> (2014).
- Okamoto, H. Genetic variability and evolution of hepatitis E virus. *Virus Res.* **127**, 216–228. <https://doi.org/10.1016/j.virusres.2007.02.002> (2007).
- Dziedzinska, R., Krzyzankova, M., Bena, M. & Vasickova, P. Evidence of Hepatitis E virus in goat and sheep milk. *Viruses-Basel* <https://doi.org/10.3390/v12121429> (2020).
- Huang, F. *et al.* Excretion of infectious hepatitis E virus into milk in cows imposes high risks of zoonosis. *Hepatology* **64**, 350–359. <https://doi.org/10.1002/hep.28668> (2016).
- Obaidat, M. M. & Roess, A. A. Individual animal and herd level seroprevalence and risk factors of Hepatitis E in ruminants in Jordan. *Infect. Genet. Evol.* <https://doi.org/10.1016/j.meegid.2020.104276> (2020).
- Takahashi, K., Kitajima, N., Abe, N. & Mishiro, S. Complete or near-complete nucleotide sequences of hepatitis E virus genome recovered from a wild boar, a deer, and four patients who ate the deer. *Virology* **330**, 501–505. <https://doi.org/10.1016/j.virol.2004.10.006> (2004).
- Boxman, I. L. A. *et al.* Porcine blood used as ingredient in meat productions may serve as a vehicle for hepatitis E virus transmission. *Int. J. Food Microbiol.* **257**, 225–231. <https://doi.org/10.1016/j.ijfoodmicro.2017.06.029> (2017).
- Lopez-Lopez, P. *et al.* Risk factors associated with hepatitis E virus in pigs from different production systems. *Vet. Microbiol.* **224**, 88–92. <https://doi.org/10.1016/j.vetmic.2018.08.020> (2018).
- Pallerla, S. R. *et al.* Hepatitis E virus genome detection in commercial pork livers and pork meat products in Germany. *J. Viral Hepatitis* **28**, 196–204. <https://doi.org/10.1111/jvh.13396> (2021).
- Kokkinos, P. *et al.* Virological quality of irrigation water in leafy green vegetables and berry fruits production chains. *Food Environ. Virol.* **9**, 72–78. <https://doi.org/10.1007/s12560-016-9264-2> (2017).
- Maunula, L. *et al.* Tracing enteric viruses in the European berry fruit supply chain. *Int. J. Food Microbiol.* **167**, 177–185. <https://doi.org/10.1016/j.ijfoodmicro.2013.09.003> (2013).
- Wang, B. & Meng, X. J. Structural and molecular biology of hepatitis E virus. *Comput. Struct. Biotechnol. J.* **19**, 1907–1916. <https://doi.org/10.1016/j.csbj.2021.03.038> (2021).
- Ropp, S. L., Tam, A. W., Beames, B., Purdy, M. & Frey, T. K. Expression of the hepatitis E virus ORF1. *Adv. Virol.* **145**, 1321–1337. <https://doi.org/10.1007/s007050070093> (2000).
- Purdy, M. A. *et al.* ICTV virus taxonomy profile: And ITHepviridae&IT. *J. Gen. Virol.* **98**, 2645–2646. <https://doi.org/10.1099/jgv.0.000940> (2017).
- Seyedi, S. S. *et al.* Computational approach towards exploring potential anti-chikungunya activity of selected flavonoids. *Sci. Rep.* **6**, 24027. <https://doi.org/10.1038/srep24027> (2016).
- Patel, C. N. *et al.* Identification of antiviral phytochemicals as a potential SARS-CoV-2 main protease (Mpro) inhibitor using docking and molecular dynamics simulations. *Sci. Rep.* **11**, 20295. <https://doi.org/10.1038/s41598-021-99165-4> (2021).
- Dwivedi, V. D. *et al.* Mechanistic insights into the Japanese encephalitis virus RNA dependent RNA polymerase protein inhibition by bioflavonoids from *Azadirachta indica*. *Sci. Rep.* **11**, 18125. <https://doi.org/10.1038/s41598-021-96917-0> (2021).
- Alencar, W. L. M. *et al.* Interactions of Co, Cu, and non-metal phthalocyanines with external structures of SARS-CoV-2 using docking and molecular dynamics. *Sci. Rep.* **12**, 3316. <https://doi.org/10.1038/s41598-022-07396-w> (2022).
- Glitscher, M. *et al.* Inhibition of Hepatitis E virus spread by the natural compound silvestrol. *Viruses-Basel* <https://doi.org/10.3390/v10060301> (2018).
- Parvez, M. K. & Subbarao, N. Molecular analysis and modeling of hepatitis E virus helicase and identification of novel inhibitors by virtual screening. *Biomed. Res. Int.* <https://doi.org/10.1155/2018/5753804> (2018).
- Todt, D. *et al.* The natural compound silvestrol inhibits hepatitis E virus (HEV) replication in vitro and in vivo. *Antiviral Res.* **157**, 151–158. <https://doi.org/10.1016/j.antiviral.2018.07.010> (2018).
- Pan, L., Woodard, J. L., Lucas, D. M., Fuchs, J. R. & Kinghorn, A. D. Rocaglamide, silvestrol and structurally related bioactive compounds from *Aglaia* species. *Nat. Prod. Rep.* **31**, 924–939. <https://doi.org/10.1039/c4np00006d> (2014).
- Withana-Gamage, T. S., Hegedus, D. D., Qiu, X. & Wanasundara, J. P. D. In silico homology modeling to predict functional properties of cruciferin. *J. Agric. Food Chem.* **59**, 12925–12938. <https://doi.org/10.1021/jf201979a> (2011).
- Karpe, Y. A. & Lole, K. S. NTPase and 5' to 3' RNA duplex-unwinding activities of the hepatitis E virus helicase domain. *J. Virol.* **84**, 3595–3602. <https://doi.org/10.1128/jvi.02130-09> (2010).
- Li, P. F. *et al.* The global epidemiology of hepatitis E virus infection: A systematic review and meta-analysis. *Liver Int.* **40**, 1516–1528. <https://doi.org/10.1111/liv.14468> (2020).
- Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303. <https://doi.org/10.1093/nar/gky427> (2018).
- Jeske, L., Placzek, S., Schomburg, I., Chang, A. & Schomburg, D. BRENDA in 2019: a European ELIXIR core data resource. *Nucleic Acids Res.* **47**, D542–D549. <https://doi.org/10.1093/nar/gky1048> (2019).
- Ishibashi, K. *et al.* Structural basis for the recognition-evasion arms race between Tomato mosaic virus and the resistance gene Tm-1. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3486–E3495. <https://doi.org/10.1073/pnas.1407888111> (2014).
- Rost, B. Twilight zone of protein sequence alignments. *Protein Eng.* **12**, 85–94. <https://doi.org/10.1093/protein/12.2.85> (1999).

35. Wiederstein, M. & Sippl, M. J. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **35**, W407–W410. <https://doi.org/10.1093/nar/gkm290> (2007).
36. Du, Z. Y. *et al.* The trRosetta server for fast and accurate protein structure prediction. *Nat. Protoc.* **16**, 5634–5651. <https://doi.org/10.1038/s41596-021-00628-9> (2021).
37. Law, Y. S. *et al.* Structural insights into RNA recognition by the Chikungunya virus nsP2 helicase. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9558–9567. <https://doi.org/10.1073/pnas.1900656116> (2019).
38. Desta, I. T., Porter, K. A., Xia, B., Kozakov, D. & Vajda, S. Performance and Its limits in rigid body protein-protein docking. *Structure* **28**, 1071. <https://doi.org/10.1016/j.str.2020.06.006> (2020).
39. Lwasaki, S. *et al.* The translation inhibitor rocaglamide targets a bimolecular cavity between eIF4A and polypurine RNA. *Mol. Cell* **73**, 738. <https://doi.org/10.1016/j.molcel.2018.11.026> (2019).
40. Greger, H. Comparative phytochemistry of flavaglines (= rocaglamides), a group of highly bioactive flavolignans from Aglaia species (Meliaceae). *Phytochem. Rev.* **21**, 725–764. <https://doi.org/10.1007/s11101-021-09761-5> (2022).
41. Fliszar-Nyul, E. *et al.* Interaction of mycotoxin alternariol with serum albumin. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms20092352> (2019).
42. Del Favero, G. *et al.* A target fishing study to spot possible biological targets of fusaric acid: Inhibition of protein kinase-A and insights on the underpinning mechanisms. *Food Chem. Toxicol.* <https://doi.org/10.1016/j.fct.2021.112663> (2022).
43. Chu, J. *et al.* Rocaglates Induce Gain-of-Function Alterations to eIF4A and eIF4F. *Cell Rep.* **30**, 2481. <https://doi.org/10.1016/j.celrep.2020.02.002> (2020).
44. Cencic, R. *et al.* Antitumor Activity and Mechanism of Action of the Cyclopenta b benzofuran Silvestrol. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0005223> (2009).
45. Cao, D. J. & Meng, X. J. Molecular biology and replication of hepatitis E virus. *Emerg. Microb. Infect.* <https://doi.org/10.1038/emi.2012.7> (2012).
46. Bateman, A. *et al.* UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515. <https://doi.org/10.1093/nar/gky1049> (2019).
47. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410. <https://doi.org/10.1006/jmbi.1990.9999> (1990).
48. Webb, B. & Sali, A. in *Protein Structure Prediction, 3rd Edition* Vol. 1137 *Methods in Molecular Biology* (ed. Kihara, D.) 1–15 (2014).
49. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. Procheck - a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291. <https://doi.org/10.1107/s0021889892009944> (1993).
50. Pettersen, E. F. *et al.* UCSF chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612. <https://doi.org/10.1002/jcc.20084> (2004).
51. Muller, C. *et al.* Comparison of broad-spectrum antiviral activities of the synthetic rocaglate CR-31-B (-) and the eIF4A-inhibitor Silvestrol. *Antiviral Res.* <https://doi.org/10.1016/j.antiviral.2020.104706> (2020).
52. Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267**, 727–748. <https://doi.org/10.1006/jmbi.1996.0897> (1997).
53. Kim, S. *et al.* PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res.* **49**, D1388–D1395. <https://doi.org/10.1093/nar/gkaa971> (2021).
54. Mysinger, M. M., Carchia, M., Irwin, J. J. & Shoichet, B. K. Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *J. Med. Chem.* **55**, 6582–6594. <https://doi.org/10.1021/jm300687e> (2012).
55. Legnik, S. *et al.* LiSiCA: A software for ligand-based virtual screening and its application for the discovery of butyrylcholinesterase inhibitors. *J. Chem. Inf. Model.* **55**, 1521–1528. <https://doi.org/10.1021/acs.jcim.5b00136> (2015).
56. Pall, S., Abraham, M. J., Kutzner, C., Hess, B. & Lindahl, E. in *2nd International Conference on Exascale Applications and Software (EASC)*. 3–27 (2015).
57. Best, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273. <https://doi.org/10.1021/ct300400x> (2012).
58. Zoete, V., Cuendet, M. A., Grosdidier, A. & Michielin, O. SwissParam: A fast force field generation tool for small organic molecules. *J. Comput. Chem.* **32**, 2359–2368. <https://doi.org/10.1002/jcc.21816> (2011).
59. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* <https://doi.org/10.1038/msb.2011.75> (2011).

Acknowledgements

This research benefits from the HPC (High Performance Computing) facility of the University of Parma, Italy. The authors would also like to acknowledge the Emilia-Romagna Region for founding the PhD fellowship and project within the context of “BIG DATA per una regione europea più ecologica, digitale e resiliente” (Deliberazione della G.R. n. 752 del 24/05/2021; Rif. PA 2021-15857/RER).

Author contributions

Conceptualization: L.P., L.D.; methodology: L.P., L.D.; formal analysis: L.P., L.D.; investigation: L.P., L.D.; data curation: L.D., L.P.; writing—original draft preparation: L.P., L.D., G.G., S.G., M.O.V.; funding acquisition: L.D., S.G.; project administration: L.D., S.G.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-19818-w>.

Correspondence and requests for materials should be addressed to L.D. or S.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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

© The Author(s) 2022