



HAL
open science

Volunteer Computing Project SiDock@home for Virtual Drug Screening Against SARS-CoV-2

Natalia Nikitina, Maxim Manzyuk, Ārtomir Podlipnik, Marko Jukić

► **To cite this version:**

Natalia Nikitina, Maxim Manzyuk, Ārtomir Podlipnik, Marko Jukić. Volunteer Computing Project SiDock@home for Virtual Drug Screening Against SARS-CoV-2. 1st International Conference on Computer Science Protecting Human Society Against Epidemics (ANTICOVID), Jun 2021, Virtual, Poland. pp.23-34, 10.1007/978-3-030-86582-5_3. hal-03746665

HAL Id: hal-03746665

<https://inria.hal.science/hal-03746665v1>

Submitted on 5 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Volunteer computing project SiDock@home for virtual drug screening against SARS-CoV-2*

Natalia Nikitina¹[0000-0002-0538-2939]✉, Maxim Manzyuk²[0000-0002-6628-0119],
Črtomir Podlipnik³[0000-0002-8429-0273], and
Marko Jukić^{4,5}[0000-0001-6083-5024]✉

¹ Institute of Applied Mathematical Research, Karelian Research Center of the Russian Academy of Sciences, 185910, Petrozavodsk, Russia
✉nikitina@krc.karelia.ru

² Internet portal BOINC.Ru, Moscow, Russia

³ Faculty of Chemistry and Chemical Technology, University of Ljubljana, 1000 Ljubljana, Slovenia

⁴ Faculty of Chemistry and Chemical Engineering, Laboratory of Physical Chemistry and Chemical Thermodynamics, University of Maribor, 2000 Maribor, Slovenia

⁵ Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, 6000 Koper, Slovenia

Abstract. In this paper, we describe a volunteer computing project SiDock@home aimed at high-throughput virtual screening of a specially developed library of small compounds against a set of targets playing important roles in the life-cycle of the virus. The originality of the screening library and the molecular docking software allows us to obtain new knowledge about chemical space in relation to SARS-CoV-2. At the same time, the existing volunteer computing community provides us with a large computational power. Having risen to a size of a modern supercomputer in several months, SiDock@home becomes an independent general drug discovery project, with its first mission targeting SARS-CoV-2.

Keywords: Distributed computing · Volunteer computing · BOINC · Desktop Grid · Virtual screening · Molecular docking · SARS-CoV-2.

1 Introduction

Since the very early stages of COVID-19 pandemic onset, scientists all over the world have employed different high-performance computing (HPC) systems to fight against SARS-CoV-2; see, e.g., [4] for a regularly updated detailed overview. There are nationwide and cross-nation research initiatives that provide scientists

* Supported by the Scholarship of the President of the Russian Federation for young scientists and graduate students (project “Game-theoretical mathematical models and algorithms for scheduling in high-performance heterogeneous computational systems”), the Slovenian Ministry of Science and Education infrastructure, project grant HPC-RIVR, and by the Slovenian Research Agency (ARRS), programme P2-0046.

with HPC resources such as projects supported by the world’s leading supercomputer Fugaku (Japan) [23], the COVID-19 High-Performance Computing Consortium (USA) [25], the Exscalate4CoV project (EU) [12] and many others.

Apart from traditional HPC systems, there are alternative capabilities such as Desktop Grids that combine non-dedicated geographically distributed computing resources (typically, desktop computers) connected to the central server by the Internet or a local access network and performing computations for the Desktop Grid in their idle time. The resources are provided by the volunteer community or by individuals and organisations related to the performed research.

Desktop Grids hold a special place among the HPC systems due to their enormous potential and, at the same time, high availability to teams of any size, even at the very early stages of research.

Today, the potential of Desktop Grids is estimated as hundreds of exaflops [1]. It is much more than the total power of all existing supercomputers. In particular, a volunteer computing project Folding@home gathered the resources of more than an exaflops in early 2020 and became the first world’s exascale system, more powerful than the top 100 supercomputers altogether [10].

Like most citizen science initiatives, BOINC projects on bio-medicine have always attracted many volunteer participants due to their socially important subjects. With the onset of a pandemic, the number of participants raised as well as the number of projects. In this paper, we try to review the BOINC projects targeting the coronavirus and present a new project named SiDock@home aimed at drug discovery and its first mission: the fight against SARS-CoV-2.

The paper has the following structure. In Section 2, we overview the current state of the BOINC middleware for organising Desktop Grids. We also describe existing BOINC-based projects targeting SARS-CoV-2. In Section 3, we describe the BOINC-based volunteer computing project named SiDock@home. In Section 4, we conclude the paper and provide plans for the project’s future.

2 BOINC-based projects targeting SARS-CoV-2

2.1 BOINC middleware

To organise and manage Desktop Grid-based distributed computations, a number of software platforms are used. The most popular platform among them is BOINC (Berkeley Open Infrastructure for Desktop Computing) [1]. Among the 157 active largest projects on volunteer computing, 89 are based on BOINC [9]; that is, BOINC can be considered a *de-facto* standard for the operation of volunteer computing projects. The platform is an actively developing Open Source software and provides rich functionality.

BOINC has a server-client architecture. The server generates a large number of tasks that are mutually independent parts of a computationally intensive problem. When a client computer is idle, it requests work from the server, receives tasks, and independently processes them. Upon finishing, it reports results back to the server. The results are then stored in the database for further usage.

Such an architecture has proven to be efficient and highly scalable for solving computationally-intensive problems of the bag-of-tasks type. The total average performance of active BOINC-based research projects is estimated as 28.5 petaflops/s [3] using an internal benchmark averaged by the last 90 days.

Due to the different benchmarks, specifics of the solved computationally-intensive problems and particular features of computing systems, it is complicated to compare the performance of BOINC-based Desktop Grids with supercomputers. Nevertheless, there exist various techniques to estimate the scale of Desktop Grids. E.g., in [22], we show that a volunteer computing project develops to the scale of a modern supercomputer in the first months of its operation (see Table 1). During periods of high workload (such as community competitions among BOINC teams), performance of a Desktop Grid increases several-fold.

Table 1: Available computational resources of the project SiDock@home during the testing phase and the main phase (in bold), as of March 23, 2021.

Average performance, teraflops/s	Peak performance, teraflops/s	Average load, active threads	Registered computers	Registered participants
23 (115)	35 (216)	1180 (5895)	934 (9193)	272 (2402)

The urgency of the SARS-CoV-2-related research has also impacted a new computational model stemming from BOINC, Science United [28]. In this model, participants specify subject areas they wish to contribute to instead of concrete projects. The administration can allocate more CPU/GPU power to priority problems, such as fighting SARS-CoV-2 [29]. Today, the combined account of Science United takes first place by daily credit [3], provided by 5 386 computers. A significant part of this capacity is performing SARS-CoV-2-related projects.

2.2 BOINC projects

For decades, the volunteer computing community has provided researchers with many computational resources ready at hand. It contributed to the development of various theoretical and empirical methods. The gained experience allowed several research groups to obtain first results at the very early stages of the fight against coronavirus. Let us overview existing public BOINC-based projects targeting SARS-CoV-2, their aims and recent results.

Rosetta@home project performs computer modelling of the end state of the folded proteins [14]. The previous experience and available volunteer resources allowed the scientific team to accurately predict the structure of the key SARS-CoV-2 spike protein (key role in pathogenesis) several weeks before its description by cryo-electron microscopy [37]. A series of previous results became a basis for *de novo* design of picomolar SARS-CoV-2 mini protein inhibitors [5].

World Community Grid, a large-scale umbrella project [36], launched the sub-project OpenPandemics [24] in May 2020 as a platform for a quick start of

research in case of any pandemic. At its current mission, OpenPandemics targets the search for potential inhibitors of various SARS-CoV-2 proteins using virtual screening of large libraries of chemical compounds.

TN-Grid project [33], in its sub-project *gene@home*, explores human gene regulatory networks for receptor proteins of SARS-CoV-2 and other viruses. To date, results include the expansion of the networks of genes associated with two non-viral diseases, identification of 22 and 36 genes to be evaluated as novel targets for already approved drugs [2].

IberCIVIS, another long-running umbrella project [13], has run the sub-project COVID-Phym aimed at the virtual screening of existing drugs for SARS-CoV-2 targets to find compounds able to block the replication mechanism of the coronavirus. The performed molecular docking allowed to deeper investigate the reasons why remdesivir and tenofovir had shown only partial evidence of improving clinical outcomes in clinical trials and observational studies [8]. The results may lay in the basis for the development of tenofovir-based drug complexes.

BOINC@TACC project hosted a large part of jobs of the virtual screening held by the University of Texas [34]. The research aimed at developing the molecular structure of a protease inhibitor that would target the coronavirus.

In Fig. 1 (a), we illustrate the performance of the active BOINC projects in teraflops/s. One observes that more than 75% of the total performance is contributed by GPU-based applications, which currently do not include SARS-CoV-2 related research (as of April 2021). The latter projects constitute only 3.6% of the total performance of BOINC.

However, as shown in Fig. 1 (b), breakdown by the number of active computers is the opposite. Note that a computer can participate in multiple BOINC projects, but detailed statistics on each computer are not available. For this reason, we do not chart intersections. Still, the data shows that SARS-CoV-2 related projects are run by between 40% and 75% of computers participating in BOINC. The data on active participants (Table 2) supports this rough estimate.

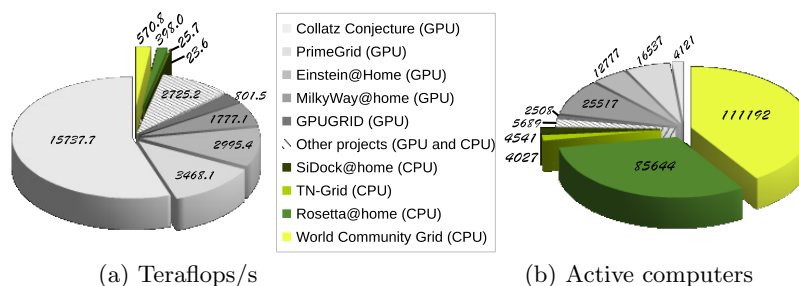


Fig. 1: Comparative performance of the active BOINC projects: Top-5 (in gray), performing SARS-CoV-2-related research (in color), and others (dashed).

Table 2: The number of active participants, as of April 5, 2021.

All BOINC projects excl. ASIC	World Community Grid	Rosetta@home	SiDock@home	TN-Grid
79 071	37 062 (47%)	30 552 (37%)	1 620 (2%)	750 (0.9%)

To summarise, existing public SARS-CoV-2-related BOINC projects base on moderately CPU-intensive applications and attract thousands of participants. Meanwhile, the BOINC community provides a large number of GPU resources that may boost a research project.

In the next section, we describe a BOINC-based research project named *SiDock@home*, created and supported by our team, aimed at drug discovery, with the first mission of fighting SARS-CoV-2.

3 SiDock@home project

3.1 Setup of a BOINC project

In March 2020, a Slovenian research group led by Dr. Črtomir Podlipnik and Dr. Marko Jukić initiated a citizen science project “*Citizen science and the fight against the coronavirus*” (COVID.SI) [11] in the field of drug design and medicinal chemistry. The project is aimed at drug discovery, and first of all, against coronavirus infection, using high-throughput virtual screening (HTVS) [16,31]. To achieve the goal, the authors performed molecular docking using a library of ten million of small molecules against multiple potential therapeutic targets.

The design of the HTVS problem allows to scale it according to the needs of researchers easily. However, even with an optimal computational process organisation, the throughput is always limited by the properties of available high-performance computing resources.

Upon the first successful results, the need for extension of the computational capacity became apparent. To complement and scale the available computational resources, we created a new SiDock@home, the BOINC-based extension of COVID.SI. The project was announced to the community in October 2020. As we have noted above, the volunteer contributions summed up to a scale of a modern supercomputer during the first five months, and SiDock@home organically grew into a sizable, independent and competent research project for general drug design. We provide the performance dynamics in more detail in [22].

The project SiDock@home [32] was created based on the BOINC middleware. The project’s server part was deployed in an Ubuntu 18.04 LTS-based machine under system configuration of 2 Xeon 6140 cores, 8 Gb RAM, 32 Gb SSD and 512 Gb HDD. At the client’s part, the project supports Windows, Linux and MacOS 64-bit operating systems, including Linux on ARM (Raspberry Pi).

3.2 High-throughput virtual screening in SiDock@home

Targets. Following the course of research of COVID.SI, we considered a set of 59 targets to screen first of all (see Table 3). 3D structural models of the targets were generated by the D-I-TASSER/C-I-TASSER pipeline [19].

We started virtual screening from 3CL^{pro}, the main viral protease, a valuable therapeutic target that possesses no mutations in the pandemically relevant variants such as B.1.1.7, B.1.351, P.1 and B.1.617. A small library of chemical compounds available to order had been already screened prior to the start of SiDock@home [17]. The study of a larger library may help to extend the structural knowledge about the target and improve the prioritisation of compounds for ordering and testing in a laboratory.

Another essential viral protease, PL^{pro}, regulating SARS-CoV-2 viral spread and innate immunity [30], was studied with two different catalytic active sites.

The next target, RdRp, is frequently used in research studies as a potential target to inhibit viral replication. Specifically, it is SARS-CoV-2 RNA-dependent RNA polymerase, an enzyme which coronavirus SARS-CoV-2 uses for the replication of its genome and the transcription of its genes.

Currently, the project is screening the library against the envelope (E) protein that is known to form a pore spanning the lipid bilayer membrane of the virus [18]. The research suggests that inhibition of the E protein viroporin could limit pathogenicity and could be of therapeutic value.

Each target corresponds to a separate computational experiment, and the next experiment's setup may depend on the results of the previous ones. If necessary, one may perform several computational experiments in parallel.

It is observed [15] that in Desktop Grids, the final part of a computational experiment may take much time to complete due to the unreliable character of the computational system. To decrease the runtime and obtain the complete picture on a considered target, we applied such a technical method as decreasing the task deadline. In BOINC, if a task does not finish upon the deadline, it is considered lost, and the server issues another replica of the task.

There is another possibility to speed up the completion of a target by announcing a competition in the BOINC community. The experience shows [22] that the performance of the project increases several-fold in such periods.

Ligands. To prepare the library of small compounds, we conglomerated multiple commercial and academic sources of 2D molecular structures, cleaned the structures, checked for errors, ionised and calculated 3D structures using the algorithms partly described in [35,26]. The resulting library contains about 1 000 000 000 compounds readily available for molecular docking software.

The relatively small size of a separate compound and the corresponding task allow graining the computational problem of HTVS as desired. In the case of a BOINC-based Desktop Grid, the common practice is to divide the problem into the tasks that run on an average desktop computer for the order of hours.

BOINC middleware allows running native applications without changing their source code using the wrapper program [38]. We used this mechanism to bring molecular docking software RxDock [27,20] to the Desktop Grid.

Table 3: Targets for the first set of computational experiments in SiDock@home.

Target ID	The protein	Organism	Source of structure	PDB Code
1-21	3CL ^{pro}	SARS-2	Snapshots from MD trajectory	
26-34	Spike Protein	SARS/ MERS/ SARS-2	Crystallographic structures	2AJF,2DD8, 3SCL, 5X58,6ACK,6LZG, 6M0J,6M17,6VW1
35-37	DHODH	Human	Crystallographic structures	4IGH,4JTU,4OQV
41-48	PL Pro	SARS/ MERS/ SARS-2	Crystallographic structures	2FE8,3MP2,4OW0, 6W9C,6WRH,6WUU, 6WX4,6WZU
49-50	FURIN	Human	Crystallographic structures	5JXH, 5MIM
51-54	Methyl Transferase	SARS-2	Crystallographic structures	6W4H,6W61, 7C2I,7C2J
55-56	E Protein	SARS/ SARS-2	NMR/ Homology model	5X29 (SARS) 5X29 Homology (SARS-2)
58-59	PL ^{pro}	SARS-2	Homology models	based on 3E9S, 5E6J, 6W9C

Following the first molecular docking results, we divided the small molecules library into packages of 2000 entries. The tasks of such size take about one-two hours to complete on an average desktop computer. However, the runtimes for different targets may vary significantly. The project’s workflow is not bound to the fixed package size and supports any other static or dynamic library division.

HTVS protocol. To reduce the computational time several-fold, we implemented a simple multi-step protocol according to [21] rather than docking the entire library in exhaustive mode. Algorithm 1 describes the workflow for a set of N ligands. Here, the values M (the maximal number of runs), R_1 (the number of runs to decide if a ligand shows a binding score better than S_1 and thus passes to the second stage), R_2 (the number of runs to decide if a ligand shows a binding score better than S_2 and thus passes to the third stage) are pre-determined using the utility `rbhtfinder` to optimise the HTVS process for a given library.

HTVS setup. BOINC provides several mechanisms to increase the effective performance of a Desktop Grid despite unreliable computational nodes.

At the moment, we have addressed two main problems that impact the efficiency of the HTVS: overdue tasks that may significantly slow down an experiment and erroneous results that cause “false-negative” outcomes. To overcome the first problem, we set a relatively short deadline of 72 hours. To overcome the second one, we set a quorum of two and implemented a result validation algorithm that checks if the log files show equal non-empty sets of ligands having been processed. The initial replication level is two instances of a task that are always sent to different users and hosts to enable cross-checking.

Algorithm 1 An HTVS protocol for N ligands and a maximum of M runs

```

1: for  $ligand = 1, 2, \dots, N$  do
2:   for  $run = 1, 2, \dots, R_1$  do ▷ First stage
3:     Run molecular docking for  $ligand$  and save the best  $score$  and  $pose$ 
4:     if  $score \geq S_1$  then go to 13
5:   end for
6:   for  $run = R_1 + 1, R_1 + 2, \dots, R_2$  do ▷ Second stage
7:     Run molecular docking for  $ligand$  and save the best  $score$  and  $pose$ 
8:     if  $score \geq S_2$  then go to 13
9:   end for
10:  for  $run = R_2 + 1, R_2 + 2, \dots, M$  do ▷ Third stage
11:    Run molecular docking for  $ligand$  and save the best  $score$  and  $pose$ 
12:  end for
13:  Output the best  $score$  and  $pose$  for  $ligand$ 
14: end for

```

Thus, a task’s input data contains a package of $N = 2000$ small molecules, the target files and the description of an HTVS protocol. Output data contains the docking results and the docking log. The latter is used to validate the results automatically and achieve a quorum.

Molecular docking software. To perform molecular docking on the first target, we used Open Source software RxDock. At the same time, we employed an independent fork CmDock [7] by the COVID.SI team, aimed at the open-science approach, utilisation of graphics processing units (GPUs) and introduction of modern docking analysis tools and improved docking algorithms. For another two targets, we used the two docking applications in parallel. Later on, we completely switched to CmDock.

Results. The described setup allowed us to perform HTVS on the complete library for four targets in five months. The results include compounds readily available to purchase as well as synthesizable compounds, and are in line with the previously started research of the group.

After initial filtering by the predicted binding energies, the results are subject to calculation of physicochemical properties and clustering analysis. In Fig. 2, we illustrate the chemical diversity of the 5141 compounds that showed predicted binding energy at most -35.0 kcal/mol in the first experiment (target 3CLpro). Here, we set the energy threshold so as to proceed on a previous study [17].

The clustering analysis performed on calculated pairwise similarities revealed 711 clusters including 519 non-singleton ones with 0.807 merge distance, according to the Kelley criterion. Therefore, the results show good chemical diversity and comply well with conventional principles of HTVS.

As of June 2021, we are post-processing the obtained results and performing HTVS for the fifth target in parallel. With hits obtained for several targets, we focus on viral proteases and are in the phase of obtaining physical samples of compounds for wet-lab testing. We are also working on SARS-CoV-2

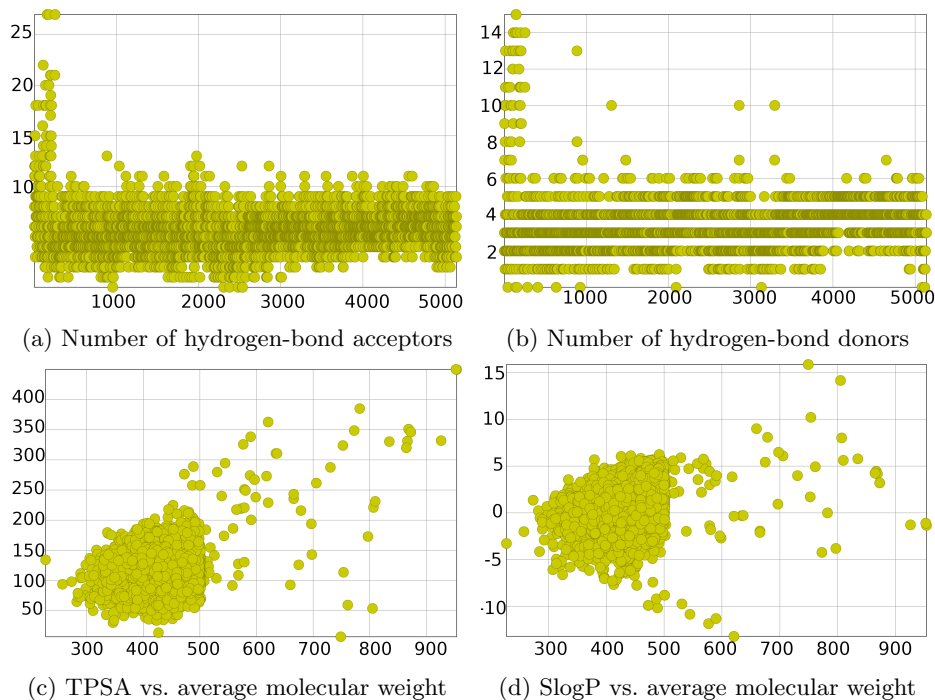


Fig. 2: Physicochemical properties of the set of selected 5141 compounds docked against target 3CLpro. Abbreviations used: TPSA – topological polar surface area; SlogP – logarithm of the enhanced atomic partition coefficient.

spike protein (Sprot) and its mutations, investigating how they are connected to pathogenicity and how to design compounds on the system.

4 Conclusion

HPC systems have been widely employed to fight against SARS-CoV-2 since the very early stages of COVID-19 pandemic onset. Desktop Grid systems serve as an efficient, highly scalable tool to complement traditional HPC systems, with an estimated potential of hundreds of exaflops. In this paper, we overviewed the existing research projects operating on BOINC-based Desktop Grids and introduced another one, SiDock@home, created and supported by our team.

The project SiDock@home is aimed at drug discovery by performing an HTVS of the specially designed small molecules library. Its first and current mission is HTVS against a set of targets related to the life cycle of SARS-CoV-2. However, SiDock@home is becoming an independent general drug discovery project. The project will eventually require large amounts of computational resources – and the BOINC community is ready to kindly provide them.

The future work is planned in several directions. Firstly, we plan to implement a GPU version of the application CmDock, which will dramatically increase the performance of the HTVS. Simultaneously, we need to increase the rate of results post-processing accordingly, most likely, in a distributed way.

Secondly, we will develop a more sophisticated HTVS protocol for optimising the computational time and hits discovery rate. Thirdly, we aim to extend the results and examine a “pan-coronavirus target”, or, more specifically, a set of different coronavirus targets to design an inhibitor that works on multiple ones.

Finally, we will continue to future potential dangers such as targets on the SADS syndrome caused by an HKU2-related coronavirus.

Acknowledgements

The first initial library (one billion of compounds) was prepared with the generous help of Microsoft that donated computational resources in the Azure cloud platform [6]. We all from COVID.SI are grateful and looking forward to future collaborations.

We wholeheartedly thank all BOINC participants for their contributions.

References

1. Anderson, D.P.: BOINC: A platform for volunteer computing. *Journal of Grid Computing* **18**(1), 99–122 (2020)
2. Blanzieri, E., Tebaldi, T., Cavecchia, V., Asnicar, F., Masera, L., Tome, G., Nigro, E., Colasurdo, E., Ciciani, M., Mazzoni, C., Pilati, S.: A computing system for discovering causal relationships among human genes to improve drug repositioning. *IEEE Transactions on Emerging Topics in Computing* pp. 1–1 (2020). <https://doi.org/10.1109/TETC.2020.3031024>
3. BOINC combined – Detailed stats — BOINCstats/BAM! <https://www.boincstats.com/stats/-5/project/detail>, [Online; accessed 30-June-2021]
4. By Editorial Team: The history of supercomputing vs. COVID-19. <https://www.hpcwire.com/2021/03/09/the-history-of-supercomputing-vs-covid-19/>, [Online; accessed 18-March-2021]
5. Cao, L., Goreshnik, I., Coventry, B., Case, J.B., Miller, L., Kozodoy, L., Chen, R.E., Carter, L., Walls, A.C., Park, Y.J., Strauch, E.M., Stewart, L., Diamond, M.S., Veesler, D., Baker, D.: De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**(6515), 426–431 (2020). <https://doi.org/10.1126/science.abd9909>, <https://science.sciencemag.org/content/370/6515/426>
6. Cloud Computing Services — Microsoft Azure. <https://azure.microsoft.com/en-us/>, [Online; accessed 30-June-2021]
7. CmDock. <https://gitlab.com/Jukic/cmdock>, [Online; accessed 30-June-2021]
8. De Salazar, P.M., Ramos, J., Cruz, V.L., Polo, R., Del Amo, J., Martínez-Salazar, J.: Tenofovir and remdesivir ensemble docking with the SARS-CoV-2 polymerase and template-nascent RNA. *Authorea Preprints* (2020)
9. Distributed Computing – Computing Platforms. <http://distributedcomputing.info/platforms.html>, [Online; accessed 30-June-2021]

10. Folding@home – fighting disease with a world wide distributed super computer. <https://foldingathome.org>, [Online; accessed 30-June-2021]
11. Home – COVID.SI. <https://covid.si/en>, [Online; accessed 30-June-2021]
12. Home [Exscalate4COV consortium]. <https://www.exscalate4cov.eu/index.html>, [Online; accessed 30-June-2021]
13. IberCIVIS. <https://boinc.ibercivis.es/>, [Online; accessed 30-June-2021]
14. Institute for Protein design, University of Washington: Rosetta’s role in fighting coronavirus. <https://www.ipd.uw.edu/2020/02/rosettas-role-in-fighting-coronavirus>, [Online; accessed 30-June-2021]
15. Ivashko, E., Nikitina, N.: Replication of “Tail” Computations in a Desktop Grid Project. In: Russian Supercomputing Days. pp. 611–621. Springer (2020)
16. Jukič, M., Janežič, D., Bren, U.: Ensemble docking coupled to linear interaction energy calculations for identification of coronavirus main protease (3CLpro) non-covalent small-molecule inhibitors. *Molecules* **25**(24), 5808 (2020)
17. Jukič, M., Škrlič, B., Tomšič, G., Pleško, S., Podlipnik, Č., Bren, U.: Prioritisation of compounds for 3CLpro inhibitor development on SARS-CoV-2 variants. *Molecules* **26**(10), 3003 (2021)
18. Mandala, V.S., McKay, M.J., Shcherbakov, A.A., Dregni, A.J., Kolocouris, A., Hong, M.: Structure and drug binding of the sars-cov-2 envelope protein transmembrane domain in lipid bilayers. *Nature structural & molecular biology* **27**(12), 1202–1208 (2020)
19. Modeling of the SARS-COV-2 Genome using I-TASSER. <https://zhanglab.ccmb.med.umich.edu/COVID-19>, [Online; accessed 30-June-2021]
20. Morley, S.D., Afshar, M.: Validation of an empirical RNA-ligand scoring function for fast flexible docking using RiboDock®. *Journal of computer-aided molecular design* **18**(3), 189–208 (2004)
21. Multi-step protocol for HTVS — RxDock 0.1.0 documentation. <https://www.rxdock.org/documentation/devel/html/user-guide/multi-step-protocol-for-htvs.html>, [Online; accessed 30-June-2021]
22. Nikitina, N., Manzyuk, M., Podlipnik, C., Jukič, M.: Performance estimation of a BOINC-based Desktop Grid for large-scale molecular docking (2021), submitted
23. Ongoing Projects — RIKEN Center for Computational Science RIKEN Website. <https://www.r-ccs.riken.jp/en/fugaku/research/covid-19/projects>, [Online; accessed 30-June-2021]
24. OpenPandemics – COVID-19 — Research — World Community Grid. <https://www.worldcommunitygrid.org/research/opn1/overview.do>, [Online; accessed 30-June-2021]
25. Projects — COVID-19 HPC Consortium. <https://covid19-hpc-consortium.org/projects>, [Online; accessed 30-June-2021]
26. Puranen, J.S., Vainio, M.J., Johnson, M.S.: Accurate conformation-dependent molecular electrostatic potentials for high-throughput in silico drug discovery. *Journal of computational chemistry* **31**(8), 1722–1732 (2010)
27. Ruiz-Carmona, S., Alvarez-Garcia, D., Foloppe, N., Garmendia-Doval, A.B., Juhos, S., Schmidtke, P., Barril, X., Hubbard, R.E., Morley, S.D.: rDock: A fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS computational biology* **10**(4), e1003571 (2014)
28. Science United. <https://scienceunited.org>, [Online; accessed 30-June-2021]
29. Science United and COVID-19. https://scienceunited.org/forum_thread.php?id=132, [Online; accessed 30-June-2021]

30. Shin, D., Mukherjee, R., Grewe, D., Bojkova, D., Baek, K., Bhattacharya, A., Schulz, L., Widera, M., Mehdipour, A.R., Tascher, G., et al.: Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature* **587**(7835), 657–662 (2020)
31. Shoichet, B.K., McGovern, S.L., Wei, B., Irwin, J.J.: Lead discovery using molecular docking. *Current opinion in chemical biology* **6**(4), 439–446 (2002)
32. SiDock@home. <https://sidock.si/sidock>, [Online; accessed 30-June-2021]
33. TN-Grid. <http://gene.disi.unitn.it/test/>, [Online; accessed 30-June-2021]
34. UTEP School of Pharmacy Developing COVID-19 Vaccine, Drug Treatments Using Supercomputing. <https://www.utep.edu/newsfeed/campus/utep-school-of-pharmacy-developing-covid-19-vaccine,-drug-treatments-using-supercomputing.html> (2021), [Online; accessed 30-June-2021]
35. Vainio, M.J., Johnson, M.S.: Generating conformer ensembles using a multiobjective genetic algorithm. *Journal of chemical information and modeling* **47**(6), 2462–2474 (2007)
36. World Community Grid – home. <https://www.worldcommunitygrid.org/>, [Online; accessed 30-June-2021]
37. Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.L., Abiona, O., Graham, B.S., McLellan, J.S.: Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**(6483), 1260–1263 (2020)
38. WrapperApp – BOINC. <https://boinc.berkeley.edu/trac/wiki/WrapperApp>, [Online; accessed 30-June-2021]