bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424283; this version posted December 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The SARS-CoV-2 S1 spike protein mutation N501Y alters the protein interactions with both hACE2 and human derived antibody: A Free energy of perturbation study

Filip Fratev[1, 2]

[1] Micar Innovation (Micar21) Ltd., Persenk 34B, 1407, Sofia, Bulgaria;

[2] Department of Pharmaceutical Sciences, School of Pharmacy, The University of Texas at El Paso, 1101 N Campbell St, El Paso, TX 79968, USA;

*Corresponding author e-mail: fratev@micar21.com

Abstract

The N501Y mutation in Covid-19 arise many question but a small amount of data are currently available. An urgent understanding of N501Y mechanism of action at molecular level is highly required. Here, we present the preliminary results of our Free energy perturbation (FEP) and Molecular dynamics (MD) calculations for the interaction of the spike S1 receptor binding domain (RBD) with both the ACE2 receptor and an antibody, STE90-C11, derived from COVID-19 patients. The results shown that the S1 RBD-ACE2 interaction was increased whereas those with the STE90-C11 antibody significantly decreased (over about 160 times). This may explain the observed in UK more spread of the virus but also emerge an important question about the possible human immune response and already available vaccines. Indeed, the latter may still act well but our data indicate some possible reduction of their effect. Further studies of N501Y mutation are need.

Letter to the editor

An increased discussion about the N501Y COVID-19 mutation continues arising many question but small amount of data are currently available [1]. This mutation was mainly observed in UK during the last weeks, led to new restrictions and many countries closed their borders for the travelers coming from the island. A little is known about the N501Y but its position and well established interaction with the human ACE2 protein, which is responsible for the virus cell entry, deserve a special attention. Moreover, it has been shown that N501Y significantly increases virus adaption in a mouse model [2]. In addition, an enhancement of the virus transmission in humans with about 70% was reported [1]. Thus, a data about the N501Y is urgently need.

Although the computational demands, the Free energy of binding (FEP) approach is one of the most successful and precise in silico techniques for accurate prediction of both the ligand selectivity [3-4], protein-protein interactions [5-6] and protein stability [7-8]. It outperforms significantly the traditional molecular dynamics based methodologies, such as for example MM/GBSA and empirical solutions like

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424283; this version posted December 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

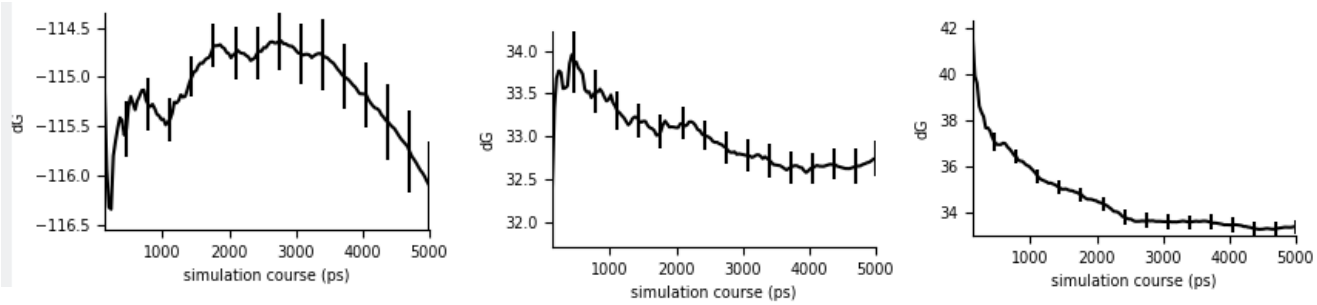

Figure 1. Observed convergence of the calculated Bennett free energies for the complex legs for complexes of the mutations: ***(A)*** N501D-ACE2, ***(B)*** N501Y-STE90-C11 and ***(C)*** N501Y-ACE2.

FoldX and etc. It often precisely predicts the free energy differences between the mutations with a RMSE of about only1.2 kcal/mol [7-8]. It has been also shown that the better sampling approaches lead to much better results and most importantly for more than 90% correct predictions of the effect; i.e. whether the effect will be positive or negative after certain ligand or protein substitutions [9].

To calculate the differences in the free energy of binding for each complex in this study we employed the Desmond/Schrodinger's FEP/REST approach described in details previously [3-8]. Initially, the default sampling protocol was applied with the number of lambda (λ) windows either 12 or 24, in dependence of the mutation charge. A pre-equilibration and 5 ns-long replica exchange solute tempering (REST) simulations in a muVT ensemble was further conducted. Only the mutated atoms was included in the REST "hot region" region. OPLS3e force field was used for the all simulations [10]. A set of N501 mutations for which experimental data is available was selected to validate the calculations. The experimental structure of S1 RBD-hACE2 (pdb: 6M0J) was used as a starting point. After the solvation with SPC waters the complex consisted of over 102 000 atoms. For the study of S1 RBD interactions with the neutralizing antibody STE90-C11 selected from COVID-19 patients [11], which was well tolerated to the known mutations, we use the very recently published X-ray structure with a PDB access number of 7B3O.

Table 1. FEP results for the selected mutations. See the text for details. Note that in the second update of this manuscript the all data from both 20 ns extended default FEP and our sampling protocol will be added. *Data from 20ns FEP

Mutation	ΔΔG Bennett (5 ns)	ΔΔG Bennett (20 ns)*	ΔΔG Error	Energy convergence	CC convergence
N501Y-STE90-C11	3.78	N/A	0.29	Good	N/A
N501Y-ACE2	-0.5	-1.75	0.21	Fair	Fair
N501D-ACE2	6.25	N/A	0.46	Bad	Bad
N501T-ACE2	-1.69	N/A	0.24	Fair	Fair
N501F	0.78	0.31	0.25	Bad	Bad

Table 1 presents the results from FEP calculations. As one can see the energy convergences in some cases were not so good for both the Bennett and the cycle closure (CC) approaches (see **Figures 1A**). In fact, the error of the calculated $\Delta\Delta G_{CC}$ predictions suffered from much larger standard deviations thus we discarded these values and they were not considered at this point of the study. We used only the Bennett values to make our conclusions herein. The convergence is indeed a very important issue during the FEP calculations. Hence, we are performing now a new set of FEP calculations using both the default protocol but extended to 20 ns-long REST sampling and our own developed sampling protocol [9]. In

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424283; this version posted December 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the particular case, before the REST procedure we use an equilibration of each λ for 50 ns and then run 20 ns-long REST simulations. The results from these calculations will be published in the second version

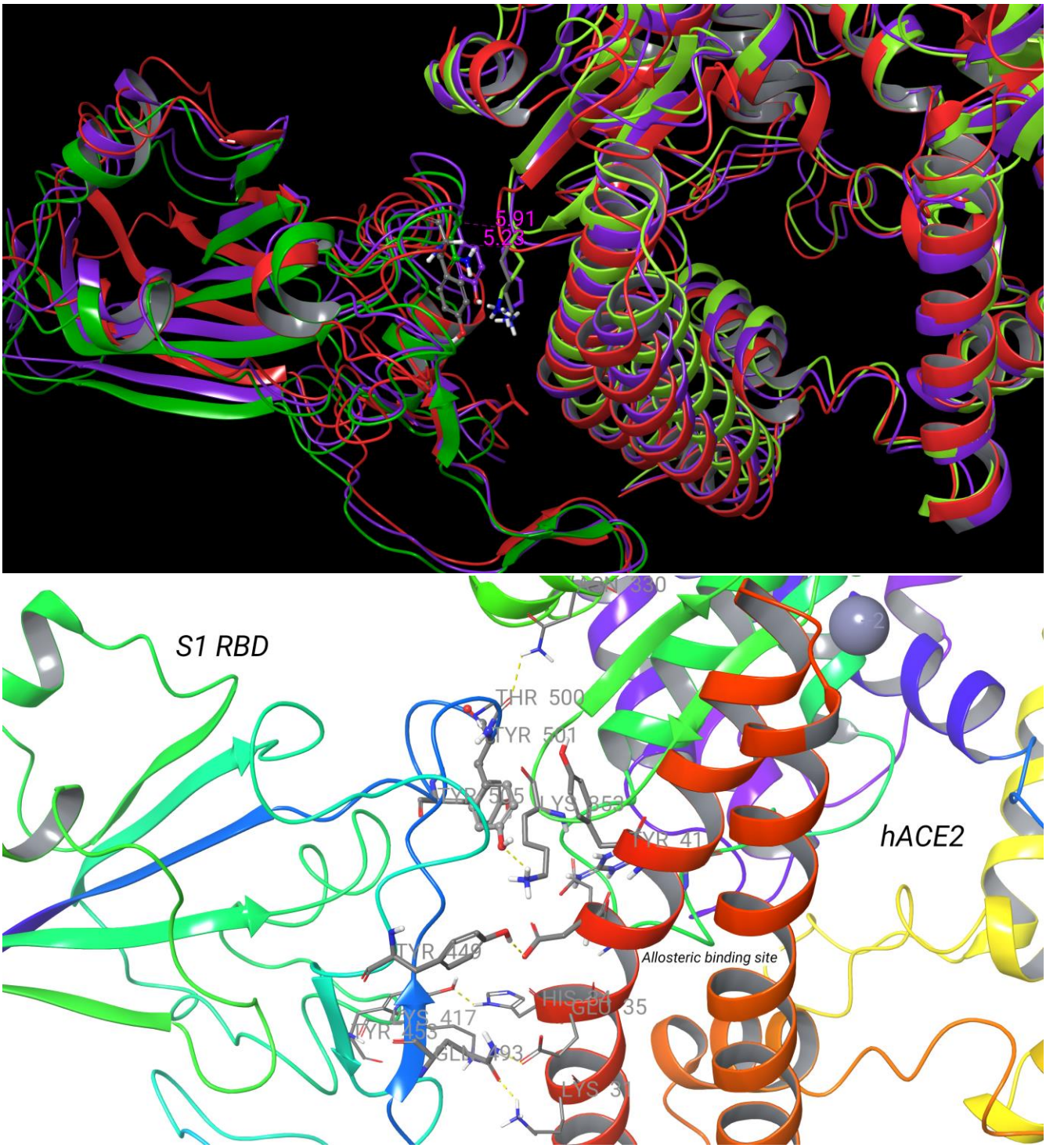

Figure 2. The identified changes in the interactions between S1 RDB and hACE2. ***(A)*** An alignment of the obtained structures after FEP simulation of WT (5ns; in green color), N501Y mutant (5ns; in blue color) and 500ns-long classical MD of N501Y mutant (in red color). The distances between Cα atoms of Tyr501 and Lys355 are also shown ***(B)*** A close look of identified by FEP simulation interactions of N501Y mutant. With yellow dot lines are shown the H-bonds.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424283; this version posted December 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of this manuscript. The better sampling protocols could make the free energies much more precise and improve the CC ΔΔG values, however, it is not likely to change qualitatively the results present here; i.e. whether one mutation can results either in an increase or decrease in the interactions. This conclusion is also based on our experience with such type of simulations with both the ligand-protein complexes and protein mutations [9, 12-13].

The main results of the FEP study were:

1) We observed significant decrease of the binding between S1 RBD and STE90-C11 antibody by ΔΔG of **3.78 kcal/mol**. This is a significant value and the observed convergence was good (**Table 1** and **Figure 1B**). The binding energy of the antibody can be roughly estimated based on the published value of IC_{50}=0.56nM in a plaque-based live SARS-CoV-2 neutralization assay [11]. This is equivalent to an dG value of about -12.7 kcal/mol ($\Delta G=RTln(IC_{50})$). Thus, our calculations predict that the N501Y mutant will produce a decrease of the binding to ΔG = -8.92 kcal/mol or 293nM. This is about 161 times lower than the wild type.

2) We detected an increase of the binding between S1 RBD and human ACE2 by ΔΔG value of only -0.5 kcal/mol after 5 ns of FEP calculations. However, considering the free energy trend of the complex leg (see **Figure 1C**; it decreased by about 1 kcal/mol only for the last 2 ns of the FEP simulations) and 500 ns-long classical MD simulation we expected this value to be much more significant after extension of the REST sampling to 20 ns. During the writing of the final version of this manuscript we obtained ΔΔG value from this extended 20 ns sampling for N501Y mutation equal to **-1.75 kcal/mol** which confirmed our hypothesis. In support to our data are also the *in vivo* studies of N501Y on mice [2].

These results can be considered as a trustful because reproduced well the experimental data [14, see Fig 4A]. For instance, for the N501D mutation we calculated a ΔΔG value of 6.25 kcal/mol meaning that it greatly reduces the S1 RBD binding to ACE2, in accordance to the experimentally observed change of over 100 times. In contrary, the N501T mutation transformed the S1 RBD to a better binder (ΔΔG = -1.69 kcal/mol), in an excellent agreement with the experimental data, but more precise values will be available from the better sampling protocols which are much more time consuming. Additional sets of calculations for other mutations, such as for example the transformation of N501D to N501Y (ΔΔG = -4.32 kcal/mol), were also performed in order the CC values to be obtained. The latter residue substitution provides an additional support that the N501Y mutation increases significantly the binding compared to the N501D one.

Further, we identified the conformational changes in the N501Y mutated S1 RBD-ACE2 complex. In particular, we compared the most representative structures after 5 ns of FEP simulation for the wild type and 5 ns sampling of the N501Y mutation, respectively, and also the average structure obtained by 500ns-long classical MD simulation (**Figure 2A**). As one can see, the S1 RBD rotates about 20 degree as shown by the 500 ns-long MD simulation (in red color on **Figure 2A**). This rotation is also well visible even only after 5 ns-long FEP/REST MD run compared to the same simulation of the wild type. In a result, the RBD can approach deeper to the center of the binding site with ACE2. This is an additional argument

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424283; this version posted December 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that one can expect lower FEP value for the N501Y mutant after applying of our better sampling protocol with an equilibration of 50 ns for each λ before the REST procedure.
The position of the surface residues was also changed. The Tyr501 makes a stable H-bond with the crucial for the ACE2 binding residue Lys355 but the longer 500 ns-long MD simulation showed that this bond is not so pronounced and mainly the hydrophobic and π-π stacking of Tyr501 increase the binding strength (**Figure 2B**). The Leu455 was in a much closer position to ACE2 interacting with the surface helix residues. A conformational change of other residues were also detected as such for example in Thr500, Tyr505, Tyr449, Tyr453, Gln493 and other. These simulations explain also why the G502 and L455 mutations are so sensitive to the ACE2 interactions.

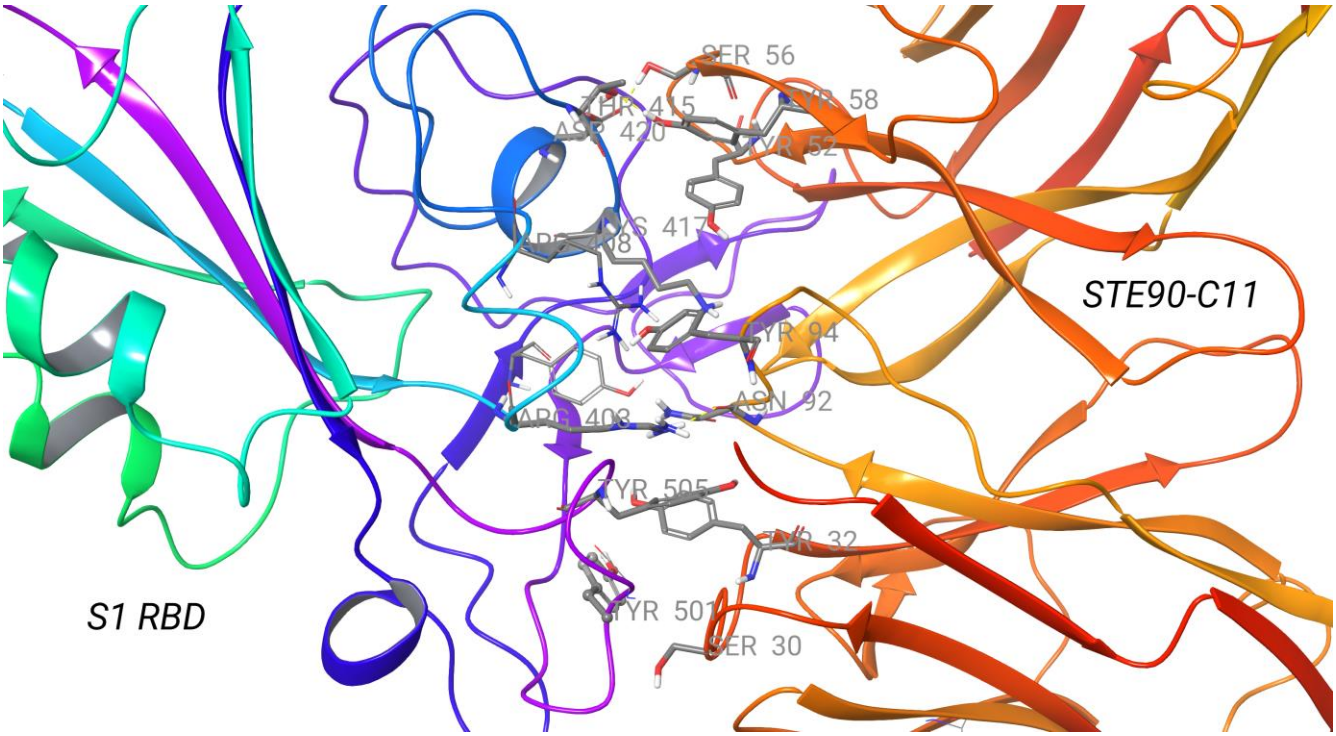

Figure 3. The identified changes in the interactions between S1 RBD and STE90-C11 human derived antibody due to the N501Y mutation. With yellow dot lines are shown the H-bonds.

Recently we identified a set of ACE2 allosteric modulators which bind to ACE2 but did not produce any significant reduction in the virus replication (unpublished results). We targeted the binding site located inside ACE2 which is close to the virus's S1 RBD (see **Figure 2B**). Thus, it seems that the virus can act in a different way, as such for example Neuropilin-1 entry mediation [15] or other processes are also possible. To our best knowledge there are no other similar ACE2 binders developed up to the moment which are able to reduce the virus replication, not only the interactions with S1 RBD. However, it is reasonable to expect that the action of such type of inhibitors will be not affected greatly by the virus's RBD mutations.

As we shown by FEP calculations the reduction of the S1-RBD binding to STE90-C11 was well pronounced. Thus, we also studied the structural changes due to the N501Y mutation based on most represented structure from the FEP MD ensemble. Two equivalent antibodies can bind to the virus spike

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424283; this version posted December 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

S1-RBD. Thus, the N501Y mutation can affect the binding by both via direct interactions with only one of them and also to produce a change in the interactions between the individual STE90-C11 units.
One of the obvious alters detected was the disruption of the formed by Gln498 H-bond with Ser30. This is also valid for Thr500 – Ser30 hydrogen bond and in general the hydrophobic interactions in this part of the S1 RBD binding surface (see **Figure 3**). Indeed, the Asn501-Ser30 and Gly502-Gly28 H-bonds were also removed. The Tyr501 did not provided any significant interactions with the antibody. The Tyr58 located in the second chain of the antibody dramatically changed its conformation leaving the central point of the binding to the antibody without any stable hydrophobic stabilization and also disrupting the hydrogen bond network formed by Ser56 of the antibody. These changes were introduced because of the Tyr501 stabilization role on the conformation of Arg403 and Arg408. Tyr58-Thr415 and Ser56-Asp420 H-bonds were also altered by the conformation of Tyr58. All of these conformational changes were not observed during the wild type simulation. These data should be further confirmed by long term MD simulations which are underway in our lab and more details would be revealed.

Acknowledgments

Thanks due to the Suman Sirimulla for the helpful discussion and Nikolay Tsvetanov from Micar21 and Apolo LLC for the financial support. We thank also to Dimitar Dimitrov for the management of several activities.

References

[1] K. Kupferschmidt. Mutant coronavirus in the United Kingdom sets off alarms, but its importance remains unclear. Science, December 20, **2020**; https://www.sciencemag.org/news/2020/12/mutant-coronavirus-united-kingdom-sets-alarms-its-importance-remains-unclear
[2] Gu H, Chen Q, Yang G, et al. (2020) Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. *Science* **2020**, 369, 1603–1607.
[3] Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M.K.; Greenwood, J.; et al. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field *J. Am. Chem. Soc.*, **2015**, 137 (7), 2695–2703.
[4] Abel, R.; Wang, L.; Harder, E.D.; Berne, B.J.; Friesner, R.A., "Advancing Drug Discovery through Enhanced Free Energy Calculations," *Acc. Chem. Res.*, **2017** 50 (7), 1625-1632.
[5] A. J. Clark, C. Negron, K. Hauser, M. Sun, L. Wang, R. Abel, R. A. Friesner. Relative Binding Affinity Prediction of Charge-Changing Sequence Mutations with FEP in Protein–Protein Interfaces, *J. Mol. Biol.*, **2019**, 431, 7, 1481-1493.
[6] Clark AJ, Gindin T, Zhang B, Wang L, Abel R, Murret CS, Xu F, Bao A, Lu NJ, Zhou T, Kwong PD, Shapiro L, Honig B, Friesner RA. Free Energy Perturbation Calculation of Relative Binding Free Energy between Broadly Neutralizing Antibodies and the gp120 Glycoprotein of HIV-1. *J Mol Biol.* **2017**; 429(7):930-947.
[7] Ford MC, Babaoglu K. Examining the Feasibility of Using Free Energy Perturbation (FEP+) in Predicting Protein Stability. *J. Chem. Inf. Model*. **2017** 57(6):1276-1285.
[8] Duan J, Lupyan D, Wang L. Improving the Accuracy of Protein Thermostability Predictions for Single Point Mutations. *Biophys J*. **2020**119(1):115-127.

bioRxiv preprint doi: https://doi.org/10.1101/2020.12.23.424283; this version posted December 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[9] Fratev F, Sirimulla S. An Improved Free Energy Perturbation FEP+ Sampling Protocol for Flexible Ligand-Binding Domains. *Sci. Rep*. **2019**; 9(1):16829.

[10] Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J.Y.; Wang, L.; Lupyan, D.; Dahlgren, M.K.; Knight, J.L.; Kaus, J.W.; Cerutti, D.S.; Krilov, G.; Jorgensen, W.L.; Abel, R.; Friesner, R.A., OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput*, **2016**, 2 (1), 281–296

[11] F. Bertoglio, V. Fühner, M. Ruschig, P. A. Heine, A SARS-CoV-2 neutralizing antibody selected from COVID-19 patients by phage display is binding to the ACE2-RBD interface and is tolerant to known RBD mutations. *BioRxiv*, **2020**, December 3, doi: https://doi.org/10.1101/2020.12.03.409318

[12] Fratev F, Miranda-Arango M, Lopez AB, Padilla E, Sirimulla S. Discovery of GlyT2 Inhibitors Using Structure-Based Pharmacophore Screening and Selectivity Studies by FEP+ Calculations. *ACS Med. Chem. Lett*. **2019**; 10(6):904-910.

[13] Fratev F, Steinbrecher T, Jónsdóttir SÓ. Prediction of Accurate Binding Modes Using Combination of Classical and Accelerated Molecular Dynamics and Free-Energy Perturbation Calculations: An Application to Toxicity Studies. *ACS Omega*. **2018**; 3(4):4357-4371.

[14] Starr, T. N., Greaney, A. J., Hilton, S. K., Crawford, K., Navarro, M. J., Bowen, J. E., Tortorici, M. A., Walls, A. C., Veesler, D., & Bloom, J. D. (2020). Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* **2020**, 182, 5, 3, 1295-1310.e20

[15] Cantuti-Castelvetri L, Ojha R, Pedro LD, Djannatian M, Franz J, Kuivanen S, van der Meer F, Kallio K, Kaya T, Anastasina M, Smura T, Levanov L, Szirovicza L, Tobi A, Kallio-Kokko H, Österlund P, Joensuu M, Meunier FA, Butcher SJ, Winkler MS, Mollenhauer B, Helenius A, Gokce O, Teesalu T, Hepojoki J, Vapalahti O, Stadelmann C, Balistreri G, Simons M. Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. Science. **2020**; 370(6518):856-860.