Chem Soc Rev

VIEWPOINT

View Article Online
View Journal

Cite this: DOI: 10.1039/d2cs00948j

Surely you are joking, Mr Docking!

F. Gentile,[a] T. I. Oprea,[b] A. Tropsha[c] and A. Cherkasov*[d]

In the wake of recent COVID-19 pandemics scientists around the world rushed to deliver numerous CADD (Computer-Aided Drug Discovery) methods and tools that could be reliably used to discover novel drug candidates against the SARS-CoV-2 virus. With that, there emerged a trend of a significant democratization of CADD that contributed to the rapid development of various COVID-19 drug candidates currently undergoing different stages of validation. On the other hand, this democratization also inadvertently led to the surge rapidly performed molecular docking studies to nominate multiple scores of novel drug candidates supported by computational arguments only. Albeit driven by best intentions, most of such studies also did not follow best practices in the field that require experience and expertise learned through multiple rigorously designed benchmarking studies and rigorous experimental validation. In this Viewpoint we reflect on recent disbalance between small number of rigorous and comprehensive studies and the proliferation of purely computational studies enabled by the ease of docking software availability. We further elaborate on the hyped oversale of CADD methods' ability to rapidly yield viable drug candidates and reiterate the critical importance of rigor and adherence to the best practices of CADD in view of recent emergence of AI and Big Data in the field.

Received 14th November 2022

DOI: 10.1039/d2cs00948j

rsc.li/chem-soc-rev

[a] Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, ON, Canada
[b] Roivant Sciences Inc, 451 D Street, Boston, MA, USA
[c] Laboratory for Molecular Modeling, Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA
[d] Vancouver Prostate Centre, Department of Urologic Sciences, University of British Columbia, Vancouver, BC, Canada. E-mail: acherkasov@prostatecentre.com

We've learned from experience that the truth will come out. Other experimenters will repeat your experiment and find out whether you were wrong or right. Nature's phenomena will agree, or they'll disagree with your theory. And, although you may gain some temporary fame and excitement, you will not gain a good reputation as a scientist if you haven't tried to be very careful in this kind of work
— *Richard Feynman, in Surely you are Joking Mr Feynman.*

F. Gentile

Francesco Gentile is an Assistant Professor in the Department of Chemistry and Biomolecular Sciences at the University of Ottawa. He obtained his PhD in Biophysics from the University of Alberta and spent three years as a postdoctoral fellow in Cherkasov's group at the University of British Columbia. Dr Gentile co-authored 25 articles, book chapters and patents. His research focuses on computational biochemistry, structure-based drug discovery, cheminformatics, and preclinical development of anticancer and antiviral drug candidates. Dr Gentile's work has been recognized by several research awards from Biophysical Society, CIHR, Michael Smith Foundation and other granting agencies.

T. I. Oprea

Tudor Oprea, who holds an MD PhD from Timisoara, Romania, is Professor and Chief, Translational Informatics Division at the UNM School of Medicine. He co-developed the lead-like concept, co-invented ChemGPS, and led the AstraZeneca strategy for compound collection enhancement for two years. At UNM, Oprea and his team contributed to the identification of selective, potent compounds for several targets. In 2002 he received the Corwin Hansch award. He is interested in translational informatics, focused on cheminformatics, drug discovery and repurposing, machine learning, and systems chemical biology. He is also Guest Professor, Center for Biological Sequence Analysis, Technical University of Denmark.

Introduction

Scientists have always strived to invent, develop, and employ approaches and technologies that would accelerate the pace of discovery. In pharmaceutical research, automated chemical synthesis, high throughput screening, and computational molecular modeling have been introduced as critical solutions to enable more effective discovery of novel medicines. Methods of computer-aided drug discovery (CADD) such as Quantitative Structure-Activity Relationships (QSAR), pharmacophore modeling, virtual screening, molecular dynamics, free energy simulations, and molecular docking have been touted at different times as powerful panaceas to the stalled drug discovery process.

In the wake of the multiple COVID-19 pandemic waves, CADD research, and molecular docking in particular, has gained popularity. The world's hopes were placed on the use of artificial intelligence (AI), molecular simulations, and super-computers to rapidly discover much needed SARS-CoV-2 therapeutics. A *ChemRxiv* preprint server "SARS-2 docking" query returns over 14 000 preprints as of today. Many if not most of these studies propose to repurpose known drugs predicted to bind to one of the COVID-19 targets, predominantly the Main viral Protease, Mpro. In most cases, predictions are not accompanied by experimental evidence. Some authors expressed confidence that massive computational studies predicting certain drug molecules to bind viral targets with high affinity can be equated with experimental evidence. A typical example of such expectation was an *Opinion* published by the *New England Journal of Medicine* in the early phases of the pandemic with an optimistic title "How to Discover Antiviral Drugs Quickly".[1]

As the time passed, however, the considerable hype associated with the promise of CADD in the beginning of the pandemic gradually morphed into frustration and criticism: For example, a recent study[2] questioned the significance of molecular docking hits followed by a popular scientific blog with similar commentary.[3] In the same vein, a recent survey of clinical studies by the then Acting FDA Commissioner, Dr Janet Woodcock, reported that there have been nearly 3000 trials of over 2500 drugs proposed as repurposing candidates against COVID-19; yet, *"the most important finding in our assessment is that the vast majority of trials of therapeutics for COVID-19 are not designed to yield actionable information; low randomization rates and underpowered outcome data render matters of safety and efficacy generally uninterpretable".*[4]

On the positive side, as we discussed in a recent review,[5] there have been rigorous computational studies that emphasized the importance of experimental validation of computational drug discovery hypotheses. Some papers incorporated experimental validation for some of the hits identified computationally. Two novel small molecule drugs, *i.e.*, Paxlovid[6] and Molnupravir[7] as well as one repurposed drug, Baricitinib, which was initially proposed based on computational studies,[8] have been approved as COVID-19 medications by the FDA. Remdesivir was also repurposed from the initial Ebola virus indication.[9] These examples suggest that continuing focused efforts by both computational and experimental researchers can result in much needed novel medications.

Herein, we reflect on molecular docking as a major computational drug discovery approach, especially in the COVID-19 context. We frame this discussion by the two aforementioned studies: first, the overly optimistic one,[1] published in the early days of the pandemic (June 2020) and the second one[2] pessimistically reflecting on the conditional end of the initial *hype* in 2022. We briefly discuss factors that led to the two extreme statements on the value of molecular docking as a drug discovery tool. We argue that these diametrically opposing views may cause readers to react by repeating the title of the famous book by Richard Feynman,[10] which we also employed in the title of this manuscript: Surely you are joking, Mr Docking! We attempt a balanced view on the *splendors and miseries* of molecular docking that transpired in the last two years of the pandemic-inspired research. As with any scientific discipline, progress can neither be accomplished *via blitzkrieg* attacks, even

A. Tropsha

Alexander Tropsha, PhD is K. H. Lee Distinguished Professor at the UNC Eshelman School of Pharmacy, UNC-Chapel Hill. Prof. Tropsha obtained his PhD in Chemical Enzymology in 1986 from Moscow State University, Russia and came to UNC-Chapel Hill in 1989 as a postdoctoral fellow. He joined the School of Pharmacy in 1991 as an Assistant Professor and became full professor in 2002. His research interests are in the areas of Computer-Assisted Drug Design, Computational Toxicology, Cheminformatics, and Biomedical Knowledge Graphs. His has authored or co-authored more than 290 peer-reviewed research papers, reviews and book chapters and co-edited two monographs.

A. Cherkasov

Artem Cherkasov is a Tier 1 Canada Research Chair in Precision Cancer Drug Design and Professor of Medicine at the University of British Columbia. Research interests include Computer-Aided Drug Discovery, Artificial Intelligence, QSAR, Cheminformatics, and development of personalized cancer therapies. Dr Cherkasov co-authored about 200 research papers, 150 patent filings and several book chapters. During his tenure at the UBC, Dr Cherkasov has out-licenced 12 technologies including 8 drug candidates to big pharma, major international venture funds and spinoff companies.

with the most modern arsenal, nor can progress evolve by equally quick yet superficial explorations of public data. We emphasize the importance of rigor in designing and executing computational experiments that should meticulously follow best practices of CADD research. We conclude by the optimistic expectation that expert applications of best practices in modern AI-driven molecular docking, backed by careful experimental validation of computational hits will likely result in the discovery of highly viable drug candidates.

The ''*Splendors*'' of molecular docking during the pandemic

In their June 2020 New England Journal of Medicine (NEJM) *Opinion*,[1] Park and Smith suggested that the dream of rapidly finding effective antiviral medicines has been finally realized. The authors posited that ultra-fast implementation of commonly used CADD techniques provides powerful tools to bring much-needed COVID-19 cures to the patients. Their own preprint was cited as evidence in support of this claim.[11] The study was widely disseminated by mainstream mass media, including CNN, who reported that simulations run by the authors on (then) the world's fastest supercomputer, Summit, identified 77 chemicals that could stop coronavirus from spreading.[12]

This claim enjoyed significant credibility having originated from one of the United States' National Laboratories, where extreme computational resources are readily available for research, and having been published in a flagship medical journal. The original study employed three well-known computational molecular modeling approaches, namely homology modeling, molecular dynamics, and molecular docking. The first approach was used to generate a structural model of the spike protein (S-protein) of SARS-CoV-2 interacting with the human ACE2 receptor; the second was used to simulate the conformational flexibility of this complex; and the third was used to identify the 77 compounds in question, from over 8000 initial molecules, including what the authors termed ''FDA-safe'' chemicals. In short, CADD methods were used to predict the ability of these compounds to perturb S-protein – ACE2 receptor binding. The computational resources that enabled this ultra-large molecular docking campaign were indeed impressive. However, the authors did not discuss well-known liabilities of the deployed CADD methods along with their results and method descriptions. The results were nevertheless immediately released, publicized, and given world-wide attention; thus, molecular docking was shining brighter than ever. But were these computational results really that impactful?

The ''*Miseries*'' of molecular docking during the pandemic

In retrospect, the much-hyped mass media reaction to a drug discovery *Opinion* that lacked any experimental validation looked surprising, given how much *in vitro*, *in vivo* and clinical experiments were already going on by June 2020. Did any of those 77 compounds identified by Park and Smith translate into viable COVID-19 drug candidates? Given the wealth of experimental antiviral campaigns that have been disclosed during this pandemic, we had the opportunity to directly assess the value of these predictions. Our assessment was facilitated by the fact that both virtual and experimental screening studies have predominantly focused on libraries of ''approved drugs'', in the effort to repurpose known drugs as COVID-19 therapies. To our knowledge, the largest of these campaigns was conducted by the NCATS (National Center for Advancing Translational Studies) team.[13] The NCATS dataset was subject to a comprehensive machine learning study, dedicated to estimating 11 different SARS-CoV-2 activities.[14] Screening the NCATS library of approved drugs[13] provided experimental data for 14 of those compounds that Parks and Smith nominated[11] as inhibitors of S-protein – ACE2 interaction. Unfortunately, none were even moderately active (Table 1), making this an illustrative case of ''digital dreams''. The phrase, coined by John Chodera in a tweet[15] describing virtual screening studies lacking experimental support, illustrates the concept of purely computational efforts not supported by experimental validation. Such work, in our opinion, should never be regarded as *drug discovery* studies.

Two years after Parks and Smith published their optimistic Opinion at the NEJM[1] and hundreds of computational studies targeting SARS-CoV-2, the gulf between expectation and experiment seems larger than ever. The growing frustration with purely computational COVID-19 CADD research has led Cerón-Carrasco to pen the ChemMedChem paper,[2] ''When Virtual Screening Yields Inactive Drugs: Dealing with False Theoretical Friends'', which was featured on the front cover of that journal.

Cerón-Carrasco discussed COVID-19 drug discovery hype, focusing on two specific docking studies[16,17] with different computational methodologies, but both processed *multibillion* compound libraries. The key idea was to retrospectively interrogate the results of these ultra-large-scale docking campaigns[16,17] using molecular mechanics/generalized Born solvation area (MMGBSA) method[18,19] as a 'gold standard' for predicting protein-ligand interaction energies. Cerón-Carrasco computed MMGBSA binding energies for 1000 potential SARS-CoV-2 Mpro inhibitors identified *via* Glide Single Precision (GlideSP)[20] scores and disclosed on the same day that the COVID-19 pandemic was declared. Since no correlation between MMGBSA and docking scores was found for the published hits, Cerón-Carrasco conclude that ''standard virtual screening (VS) simulations might not be the best strategy to increase that short list of antivirals''. This conclusion, broadly disseminated, was shortly followed by a Derek Lowe blog post, titled ''Virtual Screening for Coronavirus Protease Inhibitors: A Waste of Good Electrons? ''[3] The overall conclusion seemed to be that screening larger chemical libraries does not provide any advantage, which was in contradiction with state-of-the-art CADD observations.[21] Interestingly, Lowe had previously blogged[22] that faster virtual screenings make no difference in drug discovery either, based on the Opinion paper by Parks and Smith.[1]

Table 1 Results of direct and counter-screen assays for inhibition of S-protein – Angiotensin Converting Enzyme 2 (ACE2) interaction. These assays were conducted at the NIH NCATS.[13]

Drug name	Efficacy in AlphaLISA assay (% inhibition)[a]	Efficacy in TruHit Counter-screen Assay (% inhibition)[b]	Activity call
Acitazanolast	16	0	Inactive
Benserazide	69	41	False active
Carbazochrome sodium sulfonate	9	12	Inactive
Fidarestat	0	0	Inactive
Isoniazid pyruvate	47	0	Low quality active
Nitrofurantoin	0	0	Inactive
Pemirolast potassium	0	0	Inactive
Protirelin	0	0	Inactive
Quercetin	93	81	False active
Scutellarein	57	37	False active
Shikonin	104	99	False active
Tazobactam sodium	0	0	Inactive
Vidarabine	0	0	Inactive
Vildagliptin	0	0	Inactive

[a] This assay determines the ability of therapeutics to disrupt Spike-ACE2 interaction. [b] This counter-screen assay is used to determine if agents tested in AlphaLISA assay interfere with its readout. Compounds found active in this counter-screen assay are considered false active in the primary assay.

Virtual screening: *QUO VADIS*?

Where do these analyses leave us? Are we facing a gloomy portrait of modern computational drug discovery, and should we question its ability to help discover drugs faster and cheaper? The answer depends on our ability to comprehensively examine all available data. Computational drug discovery, enriched and empowered by the emergence of Deep Learning and AI methods, is experiencing a rapid growth[23] in both methodological aspects and, excitingly, stronger integration with experiment. As we contemplate these new approaches and increasingly powerful computers, we should pay attention to the rigor, reproducibility, and validation of *in silico* results.

For example, both the Parks and Smith[1] and Cerón-Carrasco[2] studies did not follow best practices of CADD research. Currently, the ease of access to appropriate hardware and software makes it trivial to perform virtual screening studies. However, this rapid CADD "democratization" does not eliminate the need to follow the guiding principle that all computational drug discovery efforts – including methods and applications – must be subject to rigorous "trust but verify" scrutiny. It does the scientific community no service if such non-validated work gets worldwide media attention, only to be discredited by experiment after the fact. Indeed, similar to the Park and Smith study that relied only on brute-force docking with a single approach, the Cerón-Carrasco study[2] considered MMGBSA as a sole strategy for docking pose refinement and rescoring. Given best practices in the field of molecular docking, we argue that such an assumption should be properly benchmarked using experimental results for that specific target, or at the very least, using an active-decoy set.[24] To highlight this shortcoming, we applied the exact same method,[25] implemented in Schrödinger Suite 2021-3, to a set of 62 experimentally confirmed active and 524 inactive molecules, respectively; these were identified within a Mpro virtual screening campaign using almost 40 billion molecules from the ZINC15[26] and Enamine REAL Space databases.[27] From the twenty-eight prospectively evaluated selection strategies, MMGBSA was used to re-score the inactive and active molecules obtained with single docking programs, *i.e.*, using single docking pose for each molecule. The resulting Receiver Operating Characteristic (ROC) curve to evaluate the method's capability to improve docking ranks and prioritize real hits is shown in Fig. 1. As the figure illustrates, MMGBSA shows near-random performance, producing an Area Under the Curve (AUC) of 0.52.

This outcome is not surprising: On one hand, the performance of MMGBSA to rescore structurally dissimilar compounds is far from optimal,[29] as conformational entropy changes are neglected unless explicitly estimated with quasi-harmonic approximations or normal mode analyses.[19] Such corrections were not implemented by Cerón-Carrasco.[2] On another hand, no single computational method can be considered as an error-proof means for selecting hits. Rather, consensus protocols should be deployed[21,30–32] as an integral part of the CADD Best Practices.[33] Indeed, despite significant advances in CADD methodologies, numerous studies

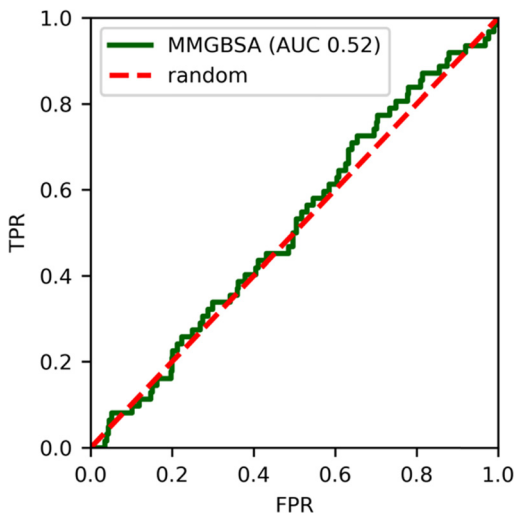

Fig. 1 ROC curve resulting from MMGBSA calculations for 62 active and 562 inactive experimentally confirmed Mpro hits; data taken from ref. 28

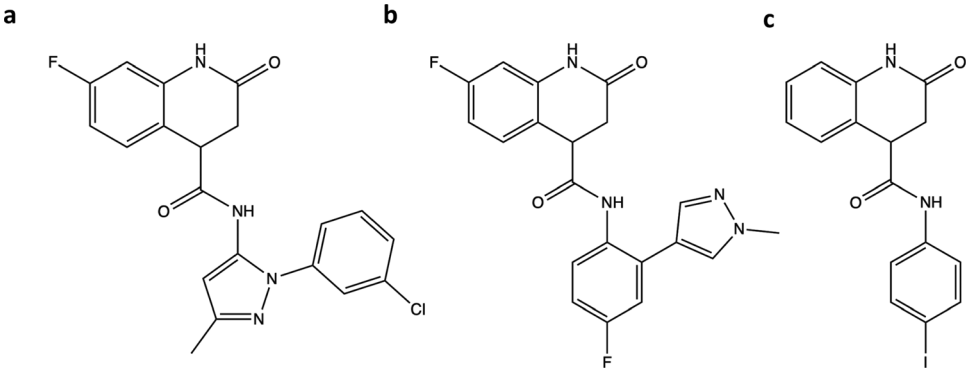

Fig. 2 Chemical structures of active dihydro-quinolinones identified by Rossetti *et al.*[34] (a) ZINC636416501 (ranked 4th in the list of 1000 potential Mpro inhibitors[16] from and measured at IC_{50} 93 μM); (b) ZINC373659060 (ranked 9th and measured at IC_{50} 58 μM); (c) Z222979552, the most potent analog identified by similarity search on the REAL Space (IC_{50} 1 μM).

unanimously emphasize the importance of expert involvement into the post-processing of molecular docking results during hit selection.[30,34,35] All recent studies that include experimental validation emphasize that 'raw' docking data represent the starting point, rather than the end-result of a CADD campaign.

To be fair, the original publication[16] criticized by Cerón-Carrasco[2] did not provide any experimental evaluation of docking predictions, given the start-of-pandemic rush to disclose the top 1000 candidate compounds. The authors made it clear that docking results represented only starting points in need of further validation and development. That list of potential SARS-CoV-2 Mpro binders (released on March 11, 2020) was subsequently evaluated by Rossetti *et al.*,[34] with two out of eight candidates selected from that top 1000 list (based on docking ranks), demonstrating modest Mpro inhibition when tested experimentally (ZINC636416501, ranked 4th in the published list and measured at IC_{50} = 93 μM and ZINC373659060, ranked 9th and measured at IC_{50}= 58 μM). Rossetti *et al.*[34] consequently used those two weak hits as templates to further explore the Enamine REAL Space database, from which a potent (1 μM), crystallographically confirmed Mpro inhibitor, Z222979552, was identified (Fig. 2). This compound displays direct antiviral activity. Furthermore, ZINC636416501 was also ranked well by the MMGBSA strategy of Cerón-Carrasco,[2] which effectively contradicts the authors' conclusion that none of the rescored molecules were active. In this context, it's relevant to mention that most experimentally validated Mpro virtual screening campaigns performed by several leading CADD groups hits with rather modest experimental activities,[36–38] which suggests that the active site of this enzyme represents a challenging target. It eventually became evident that SAR-CoV-2 Mpro assays are very dependent on experimental conditions,[37–39] and that potency evaluations require rigorous standardization. Nevertheless, these studies reinforced the idea that docking can indeed identify suitable starting points for medicinal chemistry optimization; many of those initial weak hits were consequently optimized into potent SARS-CoV-2 inhibitors with significant antiviral activity.[36–38]

Conclusions and outlook

The papers of Parks and Smith[1] and Cerón-Carrasco[2] once again highlight an important, yet frequently ignored aspect of deploying powerful supercomputers and Big Data in drug discovery. Specifically, CADD researchers must always benchmark alternative methods, employ consensus protocols, perform limited explorations around the chemical space of the initial hits, employ 'expert in the loop' iterative strategies and, experimentally validate any predictions according to the Best Practices of CADD. That is, scientists should not rush to publish unvalidated results. In the regard, it appears appropriate to re-iterate these Practices[33] previously formulated five years ago by the authors of this *Viewpoint* article, and to update them in a context of recent advances in expanding CADD to larger chemical libraries *via* physics-based and machine learning strategies:

– Study your target to gain deep knowledge of its biology and corresponding experimental assays

– Validate your screening tools and prepare your molecular databases thoroughly. Use diverse methods of virtual screening, multiple docking programs, both ligand- and structure-based approaches, and consensus scoring

– Analyze your inactive molecules and build cheminformatics models distinguishing them from your hits. Build QSAR models to rank your hits and use them to reiterate your CADD pipeline; adhere to best QSAR practices

– Analyze and visually inspect generated docking poses, and use your chemical intuition to create testable binding hypotheses

– Work closely with experimentalists during all stages of the project, learn from negative results, and fine-tune your CADD pipeline based on the wet-lab outputs

– Analyze your confirmed hits, search for analogues, explore chemical space around them, and assess the synthetic feasibility of possible derivatives

– Utilize computational tools (free energy perturbation, ADMET predictions) to guide lead optimization

In a view of several recent studies and technological advances, the original postulates could also be updated with another rule:

– The use of larger chemical databases can improve virtual screening results, when the previous rules are respected

Advances in computational hardware and software naturally lead to increased ease-of-use, *i.e.*, the ''democratization'' of CADD, which effectively enables world-wide access to highly-performant computational tools and databases. This clearly is an exciting development. The other side of this coin is that such ''democratization'' inadvertently leads to the surge on non-experts rapidly breaking into the field of CADD. While these non-experts are driven by the best of intentions, they remain unaware of the multiple challenges and complexities inherent in a field that requires not only experience in the above best practices, but also domain expertise which can only be gradually accumulated through learning and multiple rigorously designed benchmarking studies.

To conclude this *Viewpoint*, we note that CADD is witnessing an exponential growth phase. Therefore, special attention should be paid to reliability and accurate interpretation of CADD predictions, preferably matched by independent confirmation. Speaking about quantum chemistry, Ernest Davidson famously stated that ''Computers DO NOT solve the problems, PEOPLE DO''.[40] We should recall that the bridge between hypothesis and fact is called ''confirmatory experiment''. Didactive assertions that drugs can be discovered quickly by mere execution of current (molecular docking) programs *via* high-performance computers need to be subject to rigorous assessments. Inherent to the complexity of this task is the oft-forgotten fact that ''drug'' is not an intrinsic property of chemicals, since the ''drug'' attribute is human-generated, meaning it is bestowed (and sometimes removed) by regulatory agencies.[41] Indeed, compliance with Lipinski ''rule of five'' criteria[42] is often regarded as meeting ''druglike'' criteria, when in fact 40% or more of the ''non-drugs'' (from a reagents catalog) meet the same criteria.[43]

Even in times of natural disasters and emergencies, scientific research should remain rigorous. As noted by Sanjay Gupta, CNN's Chief Medical Correspondent, the story should not get ahead of the science.[44] Publishers should apply the highest level of scrutiny in evaluating the scientific merits of any publication, particularly those from a different field; and further scrutinize those where the authors make unusually bold statements that either excessively promote or unjustifiably criticize a specific approach. And researchers working in the area of computational docking and scoring should always be guided by the warning expressed by Richard Feynman and published in his famous book that we chose as an epigraph for this Viewpoint. In the absence of a comprehensive scrutiny by rigorous peer-review, unjustified or poorly justified statements will likely cause readers to respond sarcastically in a manner reminiscent of the title of Feynman's first autobiographical book:[10] Surely you are joking, Mr Docking.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

AT acknowledges support from the NIH (grant GM140154). AC acknowledges support from the CIHR (grants 170631, 172639), Innovation Fund (grant 36194) and generous donations for COVID-19 research from TELUS, Teck Resources, 625 Powell Street Foundation, Tai Hung Fai Charitable Foundation, Vancouver General Hospital Foundation. FG work is supported by the University of Ottawa startup grant. AC and FG thank Dr. Larry Goldenberg for his tireless support of research efforts on COVID-19 therapeutics development.

References

1 J. M. Parks and J. C. Smith, *N. Engl. J. Med.*, 2020, **382**, 2261–2264.
2 J. P. Cerón-Carrasco, *ChemMedChem*, 2022, **17**, e202200278.
3 D. Lowe, Virtual Screening for Coronavirus Protease Inhibitors: A Waste of Good Electrons?|Science|AAAS, **https://www.science.org/content/blog-post/virtual-screening-coronavirus-protease-inhibitors-waste-good-electrons**, (accessed 29 December 2022).
4 K. Bugin and J. Woodcock, *Nat. Rev. Drug Discovery*, 2021, **20**, 254–255.
5 E. N. Muratov, R. Amaro, C. H. Andrade, N. Brown, S. Ekins, D. Fourches, O. Isayev, D. Kozakov, J. L. Medina-Franco, K. M. Merz, T. I. Oprea, V. Poroikov, G. Schneider, M. H. Todd, A. Varnek, D. A. Winkler, A. V. Zakharov, A. Cherkasov and A. Tropsha, *Chem. Soc. Rev.*, 2021, **50**, 9121–9151.
6 D. R. Owen, C. M. N. Allerton, A. S. Anderson, L. Aschenbrenner, M. Avery, S. Berritt, B. Boras, R. D. Cardin, A. Carlo, K. J. Coffman, A. Dantonio, L. Di, H. Eng, R. A. Ferre, K. S. Gajiwala, S. A. Gibson, S. E. Greasley, B. L. Hurst, E. P. Kadar, A. S. Kalgutkar, J. C. Lee, J. Lee, W. Liu, S. W. Mason, S. Noell, J. J. Novak, R. S. Obach, K. Ogilvie, N. C. Patel, M. Pettersson, D. K. Rai, M. R. Reese, M. F. Sammons, J. G. Sathish, R. S. P. Singh, C. M. Steppan, A. E. Stewart, J. B. Tuttle, L. Updyke, P. R. Verhoest, L. Wei, Q. Yang and Y. Zhu, *Science*, 2021, **374**, 1586–1593.
7 R. A. C. Siemieniuk, J. J. Bartoszko, L. Ge, D. Zeraatkar, A. Izcovich, H. Pardo-Hernandez, B. Rochwerg, F. Lamontagne, M. A. Han, E. Kum, Q. Liu, A. Agarwal, T. Agoritsas, P. Alexander, D. K. Chu, R. Couban, A. Darzi, T. Devji, B. Fang, C. Fang, S. A. Flottorp, F. Foroutan, D. Heels-Ansdell, K. Honarmand, L. Hou, X. Hou, Q. Ibrahim, M. Loeb, M. Marcucci, S. L. McLeod, S. Motaghi, S. Murthy, R. A. Mustafa, J. D. Neary, A. Qasim, G. Rada, I. Bin Riaz, B. Sadeghirad, N. Sekercioglu, L. Sheng, C. Switzer, B. Tendal, L. Thabane, G. Tomlinson, T. Turner, P. O. Vandvik, R. W. M. Vernooij, A. Viteri-García, Y. Wang, L. Yao, Z. Ye, G. H. Guyatt and R. Brignardello-Petersen, *BMJ*, 2020, **370**, 28.
8 P. Richardson, I. Griffin, C. Tucker, D. Smith, O. Oechsle, A. Phelan and J. Stebbing, *Lancet*, 2020, **395**, e30–e31.
9 M. Wang, R. Cao, L. Zhang, X. Yang, J. Liu, M. Xu, Z. Shi, Z. Hu, W. Zhong and G. Xiao, *Cell Res.*, 2020, **30**, 269–271.

10 R. P. Feynman, *Surely You're Joking, Mr Feynman: Adventures of a Curious Character*, W. W. Norton Company, New York, 1985.
11 M. Smith and J. C. Smith, *ChemRxiv*, 2020, preprint, DOI: 10.26434/CHEMRXIV.11871402.V3.
12 The world's fastest supercomputer identified chemicals that could stop coronavirus from spreading, a crucial step toward a treatment – CNN, https://www.cnn.com/2020/03/19/us/fastest-supercomputer-coronavirus-scn-trnd/index.html, (accessed 1 July 2022).
13 NCATS, Open Science Data Portal, https://opendata.ncats.nih.gov/covid19/databrowser, (accessed 17 April 2021).
14 G. B. Kc, G. Bocci, S. Verma, M. M. Hassan, J. Holmes, J. J. Yang, S. Sirimulla and T. I. Oprea, *Nat. Mach. Intell.*, 2021, **3**(6), 527–535.
15 John Chodera (he/him) on Twitter: 'Is it really ''discovery'' of new inhibitors if there is zero experimental data? Maybe ''proposal'' of new inhibitors, but even that's a stretch. ''Digital dreams'' of new inhibitors?'/Twitter, https://twitter.com/jchodera/status/1294845832659795968, (accessed 29 December 2022).
16 A.-T. Ton, F. Gentile, M. Hsing, F. Ban and A. Cherkasov, *Mol. Inf.*, 2020, **39**, e2000028.
17 C. Gorgulla, S. S. Çınaroğlu, P. D. Fischer, K. Fackeldey, G. Wagner and H. Arthanari, *Int. J. Mol. Sci.*, 2021, **22**, 5807.
18 A. Onufriev, D. Bashford and D. A. Case, *Proteins*, 2004, **55**, 383–394.
19 S. Genheden and U. Ryde, *Expert Opin. Drug Discovery*, 2015, **10**, 449–461.
20 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739–1749.
21 J. Lyu, S. Wang, T. E. Balius, I. Singh, A. Levit, Y. S. Moroz, M. J. O'Meara, T. Che, E. Algaa, K. Tolmachova, A. A. Tolmachev, B. K. Shoichet, B. L. Roth and J. J. Irwin, *Nature*, 2019, **566**, 224–229.
22 D. Lowe, Calculating Your Way to Antivirals|Science|AAAS, https://www.science.org/content/blog-post/calculating-your-way-antivirals, (accessed 30 December 2022).
23 virtual drug screening - Search Results – PubMed, https://pubmed.ncbi.nlm.nih.gov/?term=virtualdrugscreening&timeline=expanded, (accessed 30 December 2022).
24 B. J. Bender, S. Gahbauer, A. Luttens, J. Lyu, C. M. Webb, R. M. Stein, E. A. Fink, T. E. Balius, J. Carlsson, J. J. Irwin and B. K. Shoichet, *Nat. Protoc.*, 2021, **16**, 4799–4832.
25 J. Li, R. Abel, K. Zhu, Y. Cao, S. Zhao and R. A. Friesner, *Proteins*, 2011, **79**, 2794–2812.
26 T. Sterling and J. J. Irwin, *J. Chem. Inf. Model.*, 2015, **55**, 2324–2337.
27 REAL Space – Enamine, https://enamine.net/compound-collections/real-compounds/real-space-navigator, (accessed 30 October 2022).
28 F. Gentile, M. Fernandez, F. Ban, A. T. Ton, H. Mslati, C. F. Perez, E. Leblanc, J. C. Yaacoub, J. Gleave, A. Stern, B. Wong, F. Jean, N. Strynadka and A. Cherkasov, *Chem. Sci.*, 2021, **12**, 15960–15974.
29 E. Wang, H. Sun, J. Wang, Z. Wang, H. Liu, J. Z. H. Zhang and T. Hou, *Chem. Rev.*, 2019, **119**, 9478–9508.
30 T. Tuccinardi, *Expert Opin. Drug Discovery*, 2021, **16**, 1233–1237.
31 F. Wong, A. Krishnan, E. J. Zheng, H. St€ Ark, A. L. Manson, A. M. Earl, T. Jaakkola and J. J. Collins, *Mol. Syst. Biol.*, 2022, **18**, e11081.
32 A. Fischer, M. Smieško, M. Sellner and M. A. Lill, *J. Med. Chem.*, 2021, **64**, 2489–2500.
33 F. Ban, K. Dalal, H. Li, E. LeBlanc, P. S. Rennie and A. Cherkasov, *J. Chem. Inf. Model.*, 2017, **57**, 1018–1028.
34 G. G. Rossetti, M. A. Ossorio, S. Rempel, A. Kratzel, V. S. Dionellis, S. Barriot, L. Tropia, C. Gorgulla, H. Arthanari, V. Thiel, P. Mohr, R. Gamboni and T. D. Halazonetis, *Sci. Rep.*, 2022, **12**, 1–9.
35 C. H. Zhang, E. A. Stone, M. Deshmukh, J. A. Ippolito, M. M. Ghahremanpour, J. Tirado-Rives, K. A. Spasov, S. Zhang, Y. Takeo, S. N. Kudalkar, Z. Liang, F. Isaacs, B. Lindenbach, S. J. Miller, K. S. Anderson and W. L. Jorgensen, *ACS Cent. Sci.*, 2021, **7**, 467–475.
36 E. A. Fink, C. Bardine, S. Gahbauer, I. Singh, K. White, S. Gu, X. Wan, B. Ary, I. Glenn, J. O'Connell, H. O'Donnell, P. Fajtová, J. Lyu, S. Vigneron, N. J. Young, I. S. Kondratov, A. J. O'Donoghue, Y. Moroz, J. Taunton, A. R. Renslo, J. J. Irwin, A. García-Sastre, B. K. Shoichet and C. S. Craik, *bioRxiv*, 2022, DOI: 10.1101/2022.07.05.498881.
37 Z. Li, Y. Lin, Y. Y. Huang, R. Liu, C. G. Zhan, X. Wang and H. Bin Luo, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, e2024937118.
38 M. A. M. Behnam and C. D. Klein, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, e2106095118.
39 G. Macip, P. Garcia-segura, J. Mestres-truyol, B. Saldivar-espinoza, G. Pujadas and S. Garcia-Vallvé, *Int. J. Mol. Sci.*, 2022, **23**, 259.
40 E. R. Davidson, *Reviews in Computational Chemistry*, Wiley-VCH Verlag, 2007, vol. 1, pp. 373–382.
41 O. Ursu, A. Rayan, A. Goldblum and T. I. Oprea, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2011, **1**, 760–781.
42 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 1997, **23**, 3–25.
43 O. Ursu and T. I. Oprea, *J. Chem. Inf. Model.*, 2010, **50**, 1387–1394.
44 S. Gupta, Science by press release: When the story gets ahead of the science - CNN, https://www.cnn.com/2020/06/27/health/science-by-press-release-gupta/index.html, (accessed 30 June 2020).