

# CoronaPep: An Anti-Coronavirus Peptide Generation Tool

Aman Chandra Kaushik, Aamir Mehmood<sup>✉</sup>, Gurudeeban Selvaraj<sup>✉</sup>, Xiaofeng Dai<sup>✉</sup>,  
Yi Pan<sup>✉</sup>, and Dong-Qing Wei<sup>✉</sup>

**Abstract**—The novel coronavirus (COVID-19) infections have adopted the shape of a global pandemic now, demanding an urgent vaccine design. The current work reports contriving an anti-coronavirus peptide scanner tool to discern anti-coronavirus targets in the embodiment of peptides. The proffered CoronaPep tool features the fast fingerprinting of the anti-coronavirus target serving supreme prominence in the current bioinformatics research. The anti-coronavirus target protein sequences reported from the current outbreak are scanned against the anti-coronavirus target data-sets via CORONAPEP which provides precision-based anti-coronavirus peptides. This tool is specifically for the coronavirus data, which can predict peptides from the whole genome, or a gene or protein's list. Besides it is relatively fast, accurate, userfriendly and can generate maximum output from the limited information. The availability of tools like CORONAPEP will immeasurably perquisite researchers in the discipline of oncology and structure-based drug design.

**Index Terms**—Anti-coronavirus peptides, coronavirus, COVID-19, fingerprinting

## 1 INTRODUCTION

A range of baffling pneumonia cases was reported since mid-December 2019 in the Hubei Province (Wuhan) of China. The concerned research teams and the Chinese government were rapidly engaged to control this epidemic and thus etiological inquiries were operated. The World Health Organization (WHO) on 12th January 2020, uncertainly called it the 2019 novel coronavirus (2019-nCoV). International exigency regarding the health of the public (PHEIC) was professed on January 30, 2020, in response to the 2019-nCoV epidemic by WHO. This illness was formally called Corona Virus Disease (COVID-19) that is instigated by the 2019-nCoV. The taxonomic committee's Coronavirus study group (CSG) named this virus as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Its infection rate is quite higher as compared to that of the 2002 outbreak of Severe Acute Respiratory Syndrome (SARS) in China [1], [2]. The

SARS-CoV-2 is a type of the coronavirus, belonging to the cluster of the  $\beta$ -coronavirus. It is the 3rd recognized zoonotic coronavirus ailment after the SARS and Middle East Respiratory Syndrome (MERS). These two viruses (SARS-CoV and MERS-CoV) are also linked to clusters of the  $\beta$ -coronavirus [3]. Reports by the WHO, recognition of SARS-CoV-2 in the surrounding tastings obtained from the seafood market in Wuhan. However, it is still unconfirmed if a particular animal transmits this virus. Ji *et al.* [4] reports revealed the SARS-CoV-2 was a chimera of a bat coronavirus and a coronavirus of an unidentified source. As a result of comparative analysis, it was observed that snakes are the most probable biota source for the SARS-CoV-2 [4]. Benvenuto *et al.* [5] disclosed that the SARS-CoV-2 was merely strictly linked to the coronavirus extracts in 2015 from the Chinese chrysanthemum-headed bats, supporting the bats to the human transmission chain concept. One of the SARS-CoV-2 receptors is the angiotensin-converting enzyme 2 (ACE2) reported by Zhao *et al.* [6]. Its expression occurs on type I and II alveolar epithelial cells in healthy lungs where 83 percent of its expression occurs on the type II alveolar cells, having a higher level in men as compared to women in the alveolar cells. There is an increased expression of ACE2 initiated as a result of SARS-CoV-2 attachment, damaging the alveolar cells.

There is no vaccine yet made as of February 23, 2020, for the COVID-19. Thus, the patients currently can only be treated symptomatically. The common complications observed in the 2019 nCoV patients are acute respiratory distress syndrome (ARDS), shadowed by anemia, acute heart injuries, and subordinate infections [7]. Sufferers with obstinate hypoxemia were provided aggressive motorized ventilation. There also have been postulates that besides the antiviral impeters, inhibitors for neuraminidase and RNA synthesis inhibitors, and Chinese traditional medicine could

- Aman Chandra Kaushik and Xiaofeng Dai are with the Wuxi School of Medicine, Jiangnan University, Wuxi, Jiangsu 214122, China. E-mail: amanbioinfo@sytu.edu.cn, xiaofeng.dai@jiangnan.edu.cn.
- Aamir Mehmood and Dong-Qing Wei are with the State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China, and also with the Peng Cheng Laboratory, Shenzhen, Guangdong 518055, China. E-mail: {aamirmehmood, dqwei}@sjtu.edu.cn.
- Gurudeeban Selvaraj is with the Center of Interdisciplinary Sciences-Computational Life Sciences, College of Food Science and Engineering, Henan University of Technology, Zhengzhou 450001, China, and also with the Centre for Research in Molecular Modeling, Concordia University, Montreal, QC H4B 1R6, Canada. E-mail: gurudeeb99@haut.edu.cn.
- Yi Pan is with the Department of Computer Science, Georgia State University, Atlanta, GA 30302-5060 USA. E-mail: yipan@gsu.edu.

Manuscript received 25 Mar. 2020; revised 2 July 2020; accepted 14 Sept. 2020. Date of publication 9 Mar. 2021; date of current version 6 Aug. 2021. (Corresponding Authors: Yi Pan and Dong-Qing Wei.) Digital Object Identifier no. 10.1109/TCBB.2021.3064630

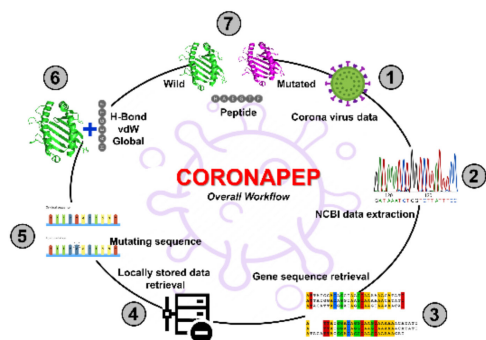


Fig. 1. This figure shows depict the overall background operations of the CORONAPEP tool during the generation of anti-coronavirus peptides. Various steps from input data requirements to the final output detail are graphically produced. The data is provided manually or in the form of the required formats which is scanned for additional information by accessing the gene information at the NCBI database and is stored locally which is later fetched for the incorporation of mutations and checking various properties to predict the final peptide that might have a better affinity with the target.

be beneficial as well for the COVID-19 therapeutic purposes [8]. Nevertheless, the efficacy of these medicines still needs to be passed through clinical tests [9], [10], [11].

Even though anti-coronavirus peptides have stupendous therapeutic significance, no tool for the anti-coronavirus peptides has been developed until now that could take clinical info e.g., VCF or SNPFF files and identify peptides from wild and mutated coronavirus protein sequences that can be used as anti-coronavirus peptides [12], [13], [14].

No efforts have been made to analyze and retrieve the important patterns from this big data that could be clinically significant for the treatment of the coronavirus. Therefore, to support the scientific portfolio for developing anti-coronavirus peptides; we have developed an anti-coronavirus peptide scanner (CORONAPEP) tool which accepts coronavirus associated clinical data including information from the above sources that is already analyzed and important evidence are retrieved along with the prediction of anti-coronavirus vaccination based on the target sequence information. We believe that CORONAPEP will be helpful for both the bioinformatics as well as for the experimental researchers working in the field of anti-coronavirus peptide-based therapeutics. This is our contribution to aid in tackling the current situation of the COVID-19 outbreak.

## 2 METHODOLOGY

The whole procedure and different techniques employed for the designing of CORONAPEP have been explained herein step-by-step. The overall workflow of the CORONAPEP is illustrated in Fig. 1.

### 2.1 Information Withdrawal for Anti-Coronavirus Targets

Several coronavirus repositories are a molecular group of precise information schemes that gather mixed archives of anti-coronavirus targets. In an attempt to procure a huge quantity of feasibly related information, several databases were taken into consideration for this purpose. The selected repositories were chosen for information hunt because of

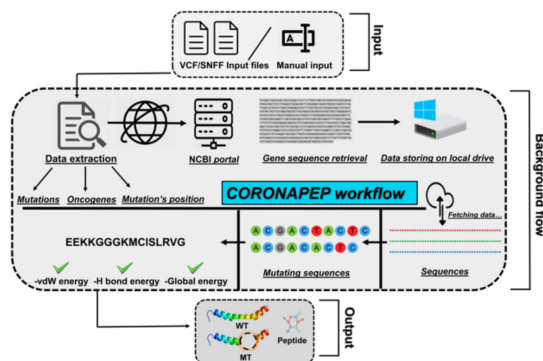


Fig. 2. Overview of the key steps taken during the peptide prediction for the Coronavirus via the CORONAPEP tool. Input files are used. Once the job is submitted, these files are used to extract information such as oncogenes, and mutations along with their exact positions. Subsequently, the CORONAPEP draws all the sequences' information one by one from the user's home directory and mutates them based on the clinical data.

their freely available clinically confirmed data samples and the magnitude of information.

### 2.2 Background Operations for the CORONAPEP Tool

The whole functional stability of a tool depends upon this model, running in the background. The accuracy, extensibility, and efficiency make an algorithm stand out among the other related developments for a similar purpose. Therefore, algorithm development is an extremely crucial step in the development of a tool.

Amidst the time interval of job submission and output display page, there transpires a series of operations in a particular style within a very less extent of only several seconds that may differ depending on the input data load, server's condition, and speed of the internet. All the background processes are briefly overviewed here.

To infiltrate the query data, input files are used. These files contain lots of information except the complete sequence of genes. Once the job is submitted, these files are used to extract information such as oncogenes, and mutations along with their exact positions. Once this information is collected, the software connects itself to the NCBI portal to retrieve the whole genes' sequences because it is absent in the input files as mentioned earlier. The sequence information is gotten and kept on a user's local directory (probably "C" on Windows or \$/home in Linux) [15], [16]. Subsequently, the CORONAPEP draws all the sequences' information one by one from the user's home directory and mutates them based on the clinical data. Till this stage, the predicted peptides are processed further for their validity testing such as their connections with the MHC-I molecule, energy of the Hydrogen bond, scores of the vdW energy global energy. Lastly, the top-ranked peptides are delivered in the results. The output data from CORONAPEP include wild type, mutant and best-ranked peptides along with their various energy values that could be used for rechecking. This whole procedure is graphically represented in Fig. 2.

### 3.1 CORONAPEP Tool Impressions

We aimed to make the proffered CORONAPEP tool very competent and reliable as well as highly user-friendly to



shadowed for a model's validation. Nevertheless, among all those techniques, a common and dependable one is to get the same or very close results for an already confirmed model. To make it simpler, we would present an example of a pharmacophore model. One of the validation techniques used for the pharmacophore method is knowing experimentally confirmed compounds that are scanned using the developed model and if it successfully delivers the accurate hits, only then it is subjected to a huge database to virtually screen drug-like compounds. The same story is repeated here in the corroboration of our proffered model where the CORONAPEP predicted peptides were checked manually and the output scoring was observed to be correct. Additionally, it's a good practice to test the model several times to evade any personal or systematic errors. Therefore, the validation of CORONAPEP is done via ten randomly selected compounds and the output observed was extremely precise. Furthermore, no error was experienced while running the program and the operations were conducted smoothly.

### 3.6 CORONAPEP Novelty and Application

Though there are a huge number of coronavirus data being produced every day and are given public access via related repositories to make these data be dissected and availed for the amelioration of manhood. Gears are required that could investigate this information and assist in the interpretation of the obtained outcomes. So far, no means are available yet that could use such clinical information and recognize possibly potent peptides. Therefore, to go parallel with the current challenges, we present the CORONAPEP that is an effective tool. It can consider anti-coronavirus targets based on their attachment affinity and offers multiple choices for the production information. The availability of options will lead to an increased usage of this software as it enlarges the work-space by leaving additional choices for the operators. Some of the major points regarding the novelty and specificity of our newly designed CORONAPEP tool are mentioned below.

- (a) This tool can only be used for the peptides prediction against the Coronavirus only because only the Coronavirus data can be used as a query.
- (b) This tool can predict peptides from the Corona whole genome as well.
- (c) Regardless of the availability, whether a gene list or protein list, the peptide can be generated.
- (d) Though there is not enough data available regarding the new Coronavirus, still this tool can generate maximum output out of the minimum input data.
- (e) If there is VCF/SNFF data in the raw information, still the peptide can be generated.
- (f) The CORONAPEP tool has a friendly user interface and requires no programming skills to operate. Thus, a user can process the data on a single click.
- (g) CORONAPEP's model is extremely fast and can process the data very quickly, causing no unnecessary lag or delay.
- (h) The peptides generated are specific and can be validated by any 3<sup>rd</sup> party technique, such as molecular docking or simulation.

## 4 DISCUSSION

The inducement for developing CORONAPEP transpired from descrying copious coronavirus data and their importance. Thus, we aspired to bestow an open-source tool that could use such data or information in the form of anti-coronavirus targets and recognize feasibly vindicated peptides. The cognizance of anti-coronavirus targets is tremendously pertinent in designing strategies to counter the coronavirus. The CORONAPEP has proven to be very helpful in this situation as it uses the information of these targets for the identification of possible peptides that have a great affinity to engage with the target. The workflow of CORONAPEP is very simple but immensely efficient as it takes the input files in the form of clinical data which is used to extract coronavirus-causing genes and also the variations in these genes. From this point, it retrieves the related protein sequence from NCBI and is altered based on the clinical information. As a result, the top-ranked peptides, are presented on the output display.

CORONAPEP's user interface is designed to be spontaneously responsive. It only requires the coronavirus data in any of the given formats as an input. The run time of the software is quite instantaneous and the output page is redirected to the user within a few seconds because of its robust coding architecture. However, the speed of the tool may vary concerning the data volume, the server's condition, and internet connectivity. The peptides predicted by CORONAPEP for these ten targets are having a very good score. It becomes evident that the tool correctly analyzes the data and predicts the potential peptides with great precision, hence confirms the validity of this tool.

So far, we believe CORONAPEP is the first attempt to let users predict peptides for the coronavirus from such clinical data in addition to supplementary useful information. This information includes wild and mutant sequences and peptides. The shortlisted protein is used for the peptide prediction and from the 2D data of the just predicted peptide and its parent protein, the global energy, vdW, and H-bond energy are also calculated. This data can be further used to affirm the validity of the prognosticated peptides by any sort of molecular docking or molecular dynamics simulation, depending on the user's choice and accessibility, though we believe it's not needed. The experimental validation of these outcomes is highly encouraged. As the peptides are originated from the clinical data, therefore *in vivo* activity is highly expected. This tool is quite advantageous in finding information about anti-coronavirus targets that are further employed for personalized medicine.

It may also be very useful in the novel target identification and proposing a potent drug for this target, which is up to 95 percent accurate. To test the workability of the CORONAPEP's Perl script, a black box technique was employed. Specifically, the prediction of peptides as a result and its ranking establishes an attractive functionality of CORONAPEP because it allows focusing the investigational assenting tests on its extremely enhanced top peptides that must consequently lead to higher potential engagements [17], [18].

## 5 CONCLUSION

The COVID-19 is an issue of international concern and threat to public health and there is an urgent need for

vaccine design. The present study offers a novel tool that can predict peptides directly from the sequences of infected patients and thus they might have a better affinity with the target and contribute to vaccine design against the COVID-19. CORONAPEP identifies peptides from the targeted coronaviruses protein which means that the predicted peptides are highly customized for the coronavirus patients.

## FUTURE DIRECTIONS

CORONAPEP provides a novel vital step towards assisting clinicians and researchers in improving their investigations via formulating new coronavirus-based targets and proposing effective peptides against such lethal conditions. In the near future, we could plan to implement drug-target interactions and small molecule-miRNA associations in the revised version of our newly designed CORONAPEP tool.

## DATA AVAILABILITY STATEMENT

The Data During and/or analyzed during the current study available from the corresponding author on request.

## ACKNOWLEDGMENTS

The Simulations in This Work Were Supported by the Center for High-Performance Computing, Shanghai Jiao Tong University. *Authors Contribution:* A.C.K., Y.P. and D.Q.W. designed the experiments. A.C.K., Y.P. and D.Q.W. did the analysis and assisted in writing the manuscript. D.Q.W., A.C.K., A.M., and Y.P. analyzed the data and wrote themanuscript. A.C.K., A.M., X.D., G.S., Y.P. and D.Q.W. read the manuscript and advised on the method development. All authors have approved the final version of the manuscript. *Fundings:* Dong-Qing Wei is supported by the grants from the Key Research Area Grant 2016YFA0501703 of the Ministry of Science and Technology of China, the National Natural Science Foundation of China (Contract no. 61832019, 61503244), the Science and Technology Commission of Shanghai Municipality (Grant: 19430750600), the Natural Science Foundation of Henan Province (162300410060), as well as SJTU JiRLMDS Joint Research Fund and Joint Research Funds for Medical and Engineering and Scientific Research at Shanghai Jiao Tong University (YG2017ZD14). These funding sources have no role in the writing of the manuscript or the decision to submit it for publication. *Conflicts of Interest:* The authors declare that they have no competing interests. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

## REFERENCES

[1] A. E. Gorbalenya *et al.*, "The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2," *Nat. Microbiol.*, vol. 5, pp. 536–544, 2020, doi: 10.1038/s41564-020-0695-z.

- [2] P. Sun, X. Lu, C. Xu, W. Sun, and B. Pan, "Understanding of COVID-19 based on current evidence," *J. Med. Virol.*, vol. 92, no. 6, pp. 548–551, 2020.
- [3] Y. Chen, Q. Liu, and D. Guo, "Emerging coronaviruses: Genome structure, replication, and pathogenesis," *J. Med. Virol.*, vol. 92, no. 4, pp. 418–423, 2020.
- [4] W. Ji, W. Wang, X. Zhao, J. Zai, and X. Li, "Cross-species transmission of the newly identified coronavirus 2019-nCoV," *J. Med. Virol.*, vol. 92, no. 4, 2020, pp. Art. no. 433.
- [5] D. Benvenuto, M. Giovannetti, A. Ciccozzi, S. Spoto, S. Angeletti, and M. Ciccozzi, "The 2019-new coronavirus epidemic: Evidence for virus evolution," *J. Med. Virol.*, vol. 92, no. 4, pp. 455–459, 2020.
- [6] Y. Zhao, Z. Zhao, Y. Wang, Y. Zhou, Y. Ma, and W. Zuo, "Single-cell RNA expression profiling of ACE2, the Receptor of SARS-CoV-2," *Am. J. Respir. Crit. Care Med.*, vol. 202, no. 5, pp. 756–759, 2020, doi: 10.1164/rccm.202001-0179LE.
- [7] C. Huang *et al.*, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *Lancet*, vol. 395, no. 10223, pp. 497–506, 2020.
- [8] H. Lu, "Drug treatment options for the 2019-new coronavirus (2019-nCoV)," *Biosci. Trends*, vol. 14, no. 1, pp. 69–71, 2020.
- [9] N. J. Edwards *et al.*, "The CPTAC data portal: A resource for cancer proteomics research," *J. Proteome Res.*, vol. 14, no. 6, pp. 2707–2713, Jun 2015, doi: 10.1021/pr501254j.
- [10] "Comprehensive molecular characterization of human colon and rectal cancer," *Nature*, vol. 487, no. 7407, pp. 330–337, 2012.
- [11] C. Wilks *et al.*, "The cancer genomics hub (CGHub): Overcoming cancer through the power of torrential data," *Database*, vol. 2014, p. bau093, 2014, doi: 10.1093/database/bau093.
- [12] S. A. Forbes *et al.*, "COSMIC: Exploring the world's knowledge of somatic mutations in human cancer," *Nucleic Acids Res.*, vol. 43, pp. D805–D811, 2015, doi: 10.1093/nar/gku1075.
- [13] B. R. Packer *et al.*, "SNP500Cancer: A public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes," (in eng), *Nucleic Acids Res.*, vol. 34, pp. D617–D621, 2006, doi: 10.1093/nar/gkj151.
- [14] M. Goldman *et al.*, "The UCSC cancer genomics browser: Update 2015," *Nucleic Acids Res.*, vol. 43, pp. D812–D817, 2015, doi: 10.1093/nar/gku1073.
- [15] A. C. Kaushik, A. Mehmood, S. Peng, Y.-J. Zhang, X. Dai, and D.-Q. Wei, "A-CaMP: A tool for anti-cancer and antimicrobial peptide generation," *J. Biomolecular Struct. Dyn.*, pp. 1–9, 2020.
- [16] A. C. Kaushik *et al.*, "CytoMegaloVirus infection database: A public omics database for systematic and comparable information of CMV," *Interdisciplinary Sci.: Comput. Life Sci.*, vol. 12, pp. 169–177, 2020.
- [17] X. Chen, N.-N. Guan, Y.-Z. Sun, J.-Q. Li, and J. Qu, "MicroRNA-small molecule association identification: From experimental results to computational models," *Brief. Bioinf.*, vol. 21, no. 1, pp. 47–61, 2020.
- [18] J. Qu, X. Chen, Y.-Z. Sun, J.-Q. Li, and Z. Ming, "Inferring potential small molecule-miRNA association based on triple layer heterogeneous network," *J. Cheminformatics*, vol. 10, no. 1, 2018, pp. Art. no. 30.

**Aman Chandra Kaushik** received the PhD degree in bioinformatics from the Gautam Buddha University, India while also researching with the Ben-Gurion University of the Negev, Israel. He is an assistant professor with the Wuxi School of Medicine, Jiangnan University, China. His research direction focuses on drug development for alzheimer, cancer and diabetes using machine learning and system biology approaches. He developed different machine learning tools Weidock, SPVec and A-CaMP have had an impact in system-based medicine. Before Jiangnan University, he was a research scholar in bioinformatics with Shanghai Jiao Tong University. He published more than 50 papers with citations of 200 times.

**Aamir Mehmood** is currently working toward the master's degree in the Department of Bioinformatics & Biostatistics, Shanghai Jiao Tong University, and is supported by the post-graduate scholarship of Shanghai Jiao Tong University. His research interests include structural biology, machine learning, DNA mismatch repair, and systems biology.

**Gurudeeban Selvaraj** received the bachelor of science degrees in biochemistry from Bharathiar University, and the master's and doctorate degrees in marine biotechnology from Annamalai University. He is working as research associate in Concordia University, Canada. He continues his research faculty position in Prof. Wei's lab, Henan University of Technology, China. His research investigates genomic and proteomic data to develop precision medicine for NSCLC using machine learning algorithms. He also involving vaccine design, statistical meta-analysis of genomic and clinical data. Prior to joining CERMM, he has completed his postdoctoral reaserach in Henan Universtiy of Technology, and Istanbul Medeniyet University. He received research grant from different funding agencies includes MITACS Global link, China Postdoctoral Science Foundation, Henan Postdoctoral Science Foundation, Henan University of Technology, The Scientific and Technological Research Council of Turkey, and University Grants Commission, India. He published more than 50 research articles with scholarly 1700 citations and H-index-21 in reputed journals including Current Medicinal Chemistry, Journal of Biomedical Informatics, Phytomedicine, The Journal of Physical Chemistry C and Current Pharmaceutical Design, Current Drug Targets and participated more than 25 different International and National conferences and workshops.

**Xiaofeng Dai is a full** professor with the Wuxi School of Medicine. Her, Research Theme: Cancer Biology and Biomarker Discovery and Biomarker discovery on breast cancer subtyping, and exploration of therapeutic strategies on triple negative breast cancers.

**Yi Pan** received the BEng and MEng degrees in computer engineering from Tsinghua University, China, in 1982 and 1984, respectively, and the PhD degree in computer science from the University of Pittsburgh, in 1991. He is currently a regents' professor and chair of computer science with Georgia State University. He has served as an associate dean and chair of Biology Department during 2013–2017 and chair of Computer Science during 2006–2013. He joined Georgia State University, in 2000, was promoted to full professor in 2004, named a distinguished University professor in 2013 and designated a regents' professor (the highest recognition given to a faculty member by the University System of Georgia) in 2015. His profile has been featured as a distinguished alumnus in both Tsinghua Alumni Newsletter and University of Pittsburgh CS Alumni Newsletter. His current research interests include parallel and cloud computing, big data, and bioinformatics. He has published more than 450 papers including over 250 SCI journal papers and 100 IEEE/ACM Transactions/journal papers. In addition, he has edited/authored 43 books. His work has been cited more than 12200 times based on Google Scholar and his current h-index is 56. He has served as an editor-in-chief or editorial board member for 20 journals including seven IEEE Transactions. Currently, he is associate editor-in-chief of *IEEE Transactions on Computational Biology and Bioinformatics*. He is the recipient of many awards including one IEEE Transactions Best Paper Award, five IEEE and other international conference or journal best paper awards, four IBM Faculty awards, two JSPS Senior Invitation Fellowships, IEEE BIBE Outstanding Achievement Award, IEEE Outstanding Leadership Award, NSF Research Opportunity Award, and AFOSR Summer Faculty Research Fellowship. He has organized numerous international conferences and delivered keynote speeches at over 60 international conferences around the world.

**Dong-Qing Wei** is a professor of bioinformatics with Shanghai Jiaotong University since 2006. Over the past three decades he has made many grand breaking contributions to the development of molecular simulation techniques and their interdisciplinary applications to systems of ever-increasing complexity. He is best known for contributions to the development of molecular simulation, more recently, AI tools and software with applications to a wide range of chemical, physical and biological systems, from electrolytes, to polar liquids, to ferroelectric liquid crystals, to combined Quantum Mechanical/Molecular Mechanical (QM/MM) systems, to membrane proteins and protein-ligand complexes applied to computer aided drug design. His most important contributions in sciences are exemplified by the discovery of ferroelectric nematic liquids crystals, the first complete ab initio MD simulation of explosion (nitromethane), and anti-aging and anti-AD drug candidate WGX-50. He published more than 300 papers with citations of 9800 times.

► **For more information on this or any other computing topic, please visit our Digital Library at [www.computer.org/csdl](http://www.computer.org/csdl).**