

The Chemical Mechanism of Pestilences or Coronavirus Disease 2019 (COVID-19)

Dongdong Zhang¹, Lin Fang¹, Li Wang¹, Zhirui Pan¹, Zhongyuan Lai^{1,2}, Mengqu Wu¹, Kun Tang¹, Ludan Lei¹, Dahong Qian^{1,3}, Zhende Huang^{1,3}, Xudong Wang^{1*}, Haibo Chen⁴, Kuo-Chen Chou^{5*}

¹Deepblue Academy of Sciences, Deepblue Technology (Shanghai) Co., Ltd., 369 Weining Rd, Shanghai 200050, China; ²Department of Physics, Fudan University, Shanghai 200433, China; ³School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; ⁴Deepblue Technology (Shanghai) Co., Ltd., 369 Weining Rd, Shanghai 200050, China; ⁵Gordon Life Science Institute, Boston, MA 02418, USA

Correspondence to: Dongdong Zhang, zhangdongdong@deepblueai.com; Lin Fang, fangl@deepblueai.com; Li Wang, wangli194@163.com; Zhirui Pan, panzhr@deepblueai.com; Zhongyuan Lai, abrikosoff@yahoo.com; Mengqu Wu, Mengquwu@gmail.com; Kun Tang, tangk@deepblueai.com; Ludan Lei, 2205498580@qq.com; Dahong Qian, dahong.qian@sjtu.edu.cn; Zhende Huang, Davidhuang10203@sina.com; Xudong Wang, wangxd@deepblueai.com; Haibo Chen, chenhb@deepblueai.com; Kuo-Chen Chou; kcchou@gordonlifescience.org or kcchou38@gmail.com

Keywords: Coronavirus Disease, COVID-19, Vaccine, 5-Steps Rule, PseAAC

Received: October 12, 2020

Accepted: November 17, 2020

Published: November 20, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

ABSTRACT

In this paper, the chemical mechanism of the coronavirus disease 2019 (COVID-19) has been explored and clearly revealed.

1. INTRODUCTION

Around the end of 2019, a new viral species caused large-scale transmissions and infections, discovered in Wuhan (WHO Emergencies Preparedness, Response, 2020) and subsequently around the world (WHO COVID-19 Disease Dashboard, 2020). Symptoms caused include coughing, shortness of breath, and fever. Around 1% to 5% (Worldometer, 2020) of confirmed infections have resulted in deaths, mainly due to severe respiratory failure (CDC, 2020). Genealogical tree studies of the new virus strains have later revealed them to be phylogenetically intimate relatives of the Severe Acute Respiratory Syndrome Coronavirus, namely (SARS-CoV), first identified in 2003 [1]. This new virus has been named SARS-CoV-2 by the International Committee on Taxonomy of Viruses (ICTV) (Gorbalenya *et al.*, 2020) on February 11th, 2020 [2-29].

2. MATERIALS AND METHOD

For the materials used in this study, see [Figure 1](#) and [Figure 2](#).

SARS-CoV	939	QALNTLVKQLSSNFGAISSVLDILSRDKVEAEVQIDRLITGRQLSLOTQYVVTQQLIRAA
Bat-CoV	953	QALNTLVKQLSSNFGAISSVLDILSRDKVEAEVQIDRLITGRQLSLOTQYVVTQQLIRAA
SARS-CoV-2	957	QALNTLVKQLSSNFGAISSVLDILSRDKVEAEVQIDRLITGRQLSLOTQYVVTQQLIRAA
pangolin-CoV	951	QALNTLVKQLSSNFGAISSVLDILSRDKVEAEVQIDRLITGRQLSLOTQYVVTQQLIRAA
SARS-CoV	999	EIRASANLAATKMSECVLQSKRVDFCGKGYHLMSPFQAAPHGVVFLHVTVYVPSQERNF
Bat-CoV	1013	EIRASANLAATKMSECVLQSKRVDFCGKGYHLMSPFQAAPHGVVFLHVTVYVPSQERNF
SARS-CoV-2	1017	EIRASANLAATKMSECVLQSKRVDFCGKGYHLMSPFQAAPHGVVFLHVTVYVPSQERNF
pangolin-CoV	1011	EIRASANLAATKMSECVLQSKRVDFCGKGYHLMSPFQAAPHGVVFLHVTVYVPSQERNF
SARS-CoV	1059	TAPAICHGKAEFPREGVVFVNGTSHWFITQRNFYEPQIITTDNTFVSGNCDVVIGIVNNT
Bat-CoV	1073	TAPAICHGKAEFPREGVVFVNGTSHWFITQRNFYEPQIITTDNTFVSGNCDVVIGIVNNT
SARS-CoV-2	1077	TAPAICHGKAEFPREGVVFVNGTSHWFITQRNFYEPQIITTDNTFVSGNCDVVIGIVNNT
pangolin-CoV	1071	TAPAICHGKAEFPREGVVFVNGTSHWFITQRNFYEPQIITTDNTFVSGNCDVVIGIVNNT
SARS-CoV	1119	VYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNES
Bat-CoV	1133	VYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNES
SARS-CoV-2	1137	VYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNES
pangolin-CoV	1131	VYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNES
SARS-CoV	170	IDVSEKSGNFKNLREFVFKNIDGFLYVYKGYQPIDVVRDLPSPGNTLKPFIKLPFLGINIT
Bat-CoV	177	MDLEGKQGNFKNLREFVFKNIDGFKIYKSHTPINLVRDLPQGFSALEPLVDLPIGINIT
SARS-CoV-2	177	MDLEGKQGNFKNLREFVFKNIDGFKIYKSHTPINLVRDLPQGFSALEPLVDLPIGINIT
pangolin-CoV	175	MDLEGKQGNFKNLREFVFKNIDGFKIYKSHTPINLVRDLPQGFSALEPLVDLPIGINIT
SARS-CoV	1179	LIDLQELGKYEQYIKWPWYIWLGFIAGLIAIMVTIMLCCMTSCCCLKGCSCGSCCKF
Bat-CoV	1193	LIDLQELGKYEQYIKWPWYIWLGFIAGLIAIMVTIMLCCMTSCCCLKGCSCGSCCKF
SARS-CoV-2	1197	LIDLQELGKYEQYIKWPWYIWLGFIAGLIAIMVTIMLCCMTSCCCLKGCSCGSCCKF
pangolin-CoV	1191	PIDLQELGKYEQYIKWPWYIWLGFIAGLIAIMVTIMLCCMTSCCCLKGCSCGSCCKF
SARS-CoV	284	AELKCSVKSFEIDKGIYQTSNFRVPSGDVVRFPNITNLCPPFGEVFNATKFPFSVYAWERK
Bat-CoV	297	SEIKCTLKSFIVBKGIYQTSNFRVPTDTSIVRFPNITNLCPPFGEVFNATKFPFSVYAWERK
SARS-CoV-2	297	SEIKCTLKSFIVBKGIYQTSNFRVPTDTSIVRFPNITNLCPPFGEVFNATKFPFSVYAWERK
pangolin-CoV	295	SEIKCTLKSLTIVBKGIYQTSNFRVPTDTSIVRFPNITNLCPPFGEVFNATKFPFSVYAWERK
SARS-CoV	344	KISNCVADYSVLYNSTSFSTFKCYGVSATKLNLDLCFINNVYADSFVWGGDEVROIAPGQTG
Bat-CoV	357	RISNCVADYSVLYNSTSFSTFKCYGVSATKLNLDLCFINNVYADSFVWGGDEVROIAPGQTG
SARS-CoV-2	357	RISNCVADYSVLYNSTSFSTFKCYGVSATKLNLDLCFINNVYADSFVWGGDEVROIAPGQTG
pangolin-CoV	355	RISNCVADYSVLYNSTSFSTFKCYGVSATKLNLDLCFINNVYADSFVWGGDEVROIAPGQTG
SARS-CoV	404	VIADYNYKLPDDFPGCVIAWNSKHIDAKEGNPNYIYRLEFRKANLKPFFERDISTEIYOAG
Bat-CoV	417	KIADYNYKLPDDFPGCVIAWNSKHIDAKEGNPNYIYRLEFRKANLKPFFERDISTEIYOAG
SARS-CoV-2	417	KIADYNYKLPDDFPGCVIAWNSKHIDAKEGNPNYIYRLEFRKANLKPFFERDISTEIYOAG
pangolin-CoV	415	VIADYNYKLPDDFPGCVIAWNSKHIDAKEGNPNYIYRLEFRKANLKPFFERDISTEIYOAG
SARS-CoV	464	GKPCPTPLPANCYVPLNDYGFYFTTGGYQYQPRVVVLSFELLNAPATVCGPKLSTDLIKN
Bat-CoV	477	SKPCNGQTEINCYPYLYRYGFYFTDGVGHQPRVVVLSFELLNAPATVCGPKLSTNLVKN
SARS-CoV-2	477	SKPCNGQTEINCYPYLYRYGFYFTDGVGHQPRVVVLSFELLNAPATVCGPKLSTNLVKN
pangolin-CoV	475	STPCNGQVGLNCYVPLERYGFHPTTIGVNYQPRVVVLSFELLNAPATVCGPKLSTTLVKD
SARS-CoV	523	QCVNFNENGLTGTGVLTPSSKRFLPFQFGRDVSDFDTSVRDPKTSSEILDITPCSFGGVS
Bat-CoV	537	KCVNFNENGLTGTGVLTPSSKRFLPFQFGRDIADTTDAVRDPQTEILDITPCSFGGVS
SARS-CoV-2	537	KCVNFNENGLTGTGVLTPSSKRFLPFQFGRDIADTTDAVRDPQTEILDITPCSFGGVS
pangolin-CoV	535	KCVNFNENGLTGTGVLTPSSKRFLPFQFGRDISDTTDAVRDPQTEILDITPCSFGGVS
SARS-CoV	583	VITPGTNASSEVAVLYQDVNCTEVPVAIHADQLTPAWRYSAGANVVFQTRAGCLVGAEHV
Bat-CoV	597	VITPGTNASNOVAVLYQDVNCTEVPVAIHADQLTPAWRYSAGANVVFQTRAGCLVGAEHV
SARS-CoV-2	597	VITPGTNASNOVAVLYQDVNCTEVPVAIHADQLTPAWRYSAGANVVFQTRAGCLVGAEHV
pangolin-CoV	595	VITPGTNASNOVAVLYQDVNCTEVPMAIHAEQLTPAWRYSAGANVVFQTRAGCLVGAEHV
SARS-CoV	643	DTSYECDIPIGAGICASYHTVSL...RSTSQKSIIVAYTMSLGAENSVAYSNNIAIPT
Bat-CoV	657	NNSYECDIPIGAGICASYQTQNS...RSVASQSIIVAYTMSLGAENSVAYSNNIAIPT
SARS-CoV-2	657	NNSYECDIPIGAGICASYQTQNS...RSVASQSIIVAYTMSLGAENSVAYSNNIAIPT
pangolin-CoV	655	NNSYECDIPIGAGICASYHSMSSL...RSVNSQSIIVAYTMSLGAENSVAYSNNIAIPT
SARS-CoV	699	NFSISITTEVMPVSMKTSVDCNMYICGDSIECANLLLOYGFSFCTQLNRALSGLIAEQDR
Bat-CoV	713	NFTISVTTTEILPVSMKTSVDCNMYICGDSIECANLLLOYGFSFCTQLNRALTGLIAEQDK
SARS-CoV-2	717	NFTISVTTTEILPVSMKTSVDCNMYICGDSIECANLLLOYGFSFCTQLNRALTGLIAEQDK
pangolin-CoV	711	NFTISVTTTEILPVSMKTSVDCNMYICGDSIECANLLLOYGFSFCTQLNRALTGLIAEQDK
SARS-CoV	759	NTQEVFAQVKQMYKTPILKDFGGFNFSQILPDPKPKSKRSFIEDLLFNKVTADAGFIKQ
Bat-CoV	773	NTQEVFAQVKQMYKTPILKDFGGFNFSQILPDPKPKSKRSFIEDLLFNKVTADAGFIKQ
SARS-CoV-2	773	NTQEVFAQVKQMYKTPILKDFGGFNFSQILPDPKPKSKRSFIEDLLFNKVTADAGFIKQ
pangolin-CoV	771	NTQEVFAQVKQMYKTPILKDFGGFNFSQILPDPKPKSKRSFIEDLLFNKVTADAGFIKQ

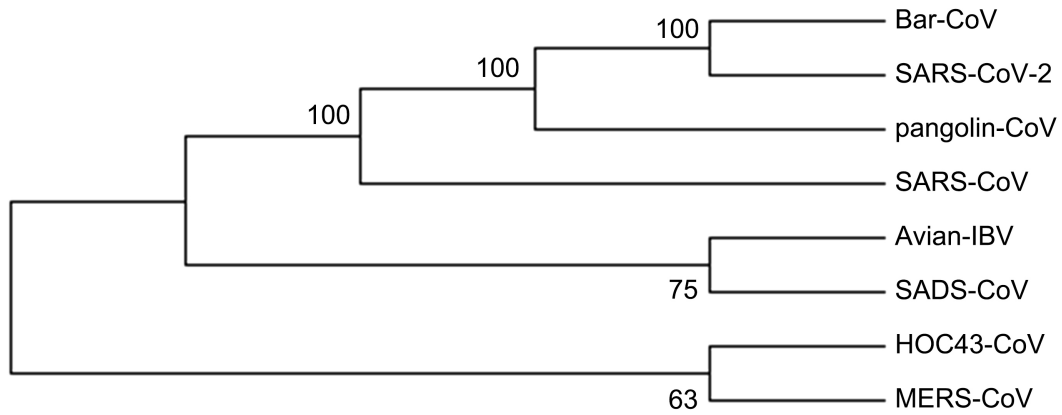


Figure 1. The alignment of the amino acid sequences of the S proteins from 4 coronaviruses and the phylogenetic tree of the genes of the receptor binding proteins from 8 different single strand RNA viruses with the MEGA software. Top: The alignment of the amino acid sequences of the S proteins from 4 coronaviruses, namely Bat Coronavirus RaTG13, Pangolin CoV (GX), SARS-CoV (NC_004718) and SARS-CoV-2 (NC_045512). Receptor Binding Domain (RBD) is labeled with the color boxes and the black arrows indicate the beginning and the end of RBD.

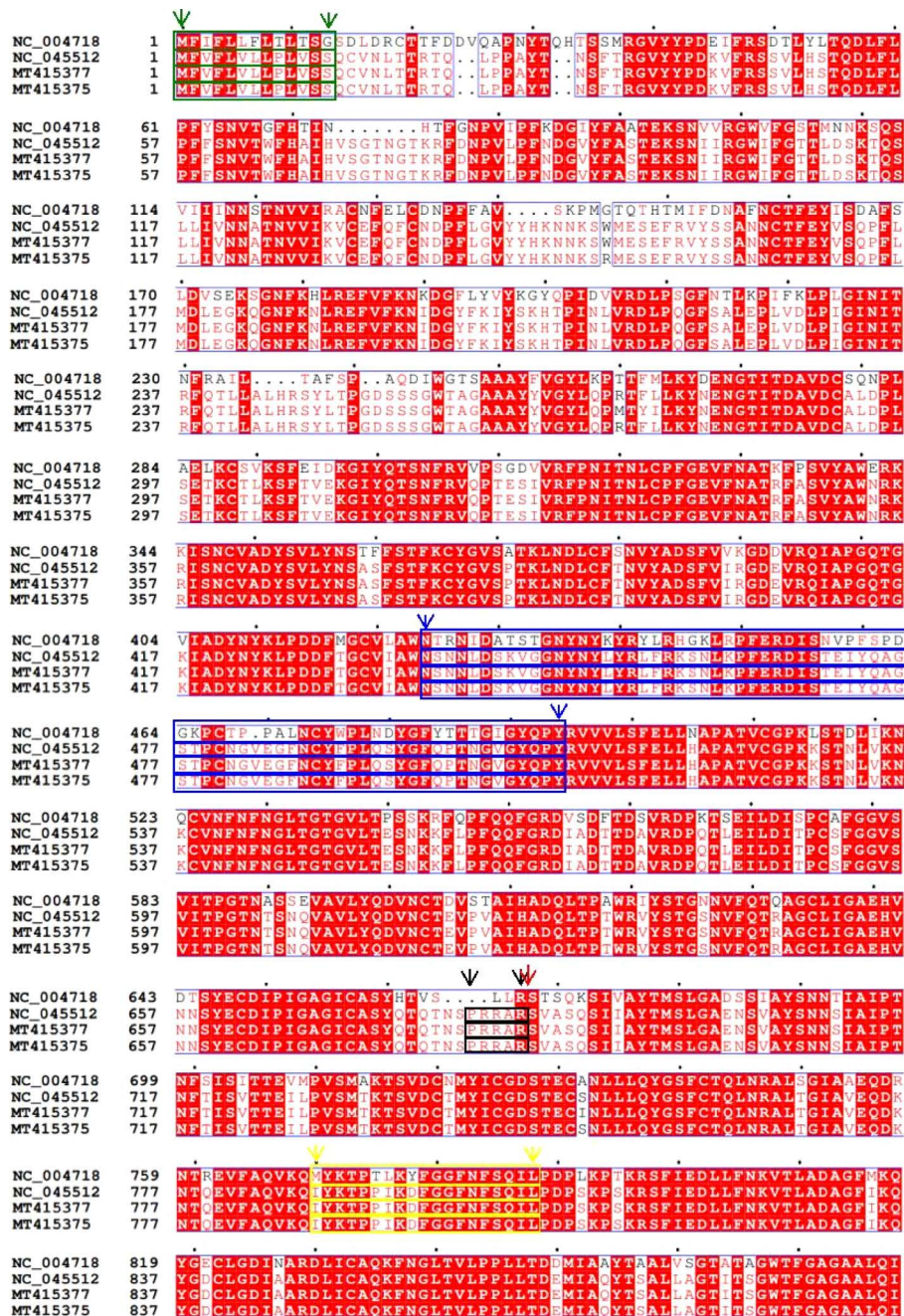


Figure 2. The alignment of the spike (S) protein of 3 strains of SARS-CoV-2 and 1 strain of SARS-CoV (NC_004718.3), SARS-CoV-2 (NC_045512), MT415375 and MT415377. Important functional regions, such as the Signal Peptide (SP), Receptor Binding Motif (RBM), Fusion Peptide (FP) and the inserted amino acid (amino acids PRRA)/host protease recognition site and cleavage site (amino acids RRAR)/cleavage site (between amino acids 685R and 686S) have been marked on the CA pictures. SP: green Box (the green arrows showing the beginning and end of the SP); RBM: blue box (the blue arrows showing the beginning and end of the RBM); PRRAR: black box (the black arrows showing the beginning and end of the inserted/host protease recognition amino acid sites; the red arrow showing the host protease cleavage site between amino acids 685R and 686S) (T. Meng, *et al.*); FP: yellow box.

The “Chou’s 5-step rules” or “Chou’s 5-steps rule” has also been applied because it has been widely and increasingly used by many scientists to study biological sequence analyses [2-56].

To develop a really useful predictor for a biological system, one needs to go through the following five steps: 1) select or construct a valid benchmark dataset to train and test the predictor; 2) represent the samples with an effective formulation that can truly reflect their intrinsic correlation with the target to be predicted; 3) introduce or develop a powerful algorithm to conduct the prediction; 4) properly perform cross-validation tests to objectively evaluate the anticipated prediction accuracy; 5) establish a user-friendly web-server for the predictor that is accessible to the public. Papers presented for developing a new sequence-analyzing method or statistical predictor by observing the guidelines of Chou’s 5-step rules have the following notable merits: 1) crystal clear in logic development, 2) completely transparent in operation, 3) easily to repeat the reported results by other investigators, 4) with high potential in stimulating other sequence-analyzing methods, and 5) very convenient to be used by the majority of experimental scientists.

Accordingly, the webserver for the current paper has been currently underway; once done, it will be announced through the addendum of this paper.

3. CONCLUSION

The chemical principle proposed in this paper is extremely useful in fighting against the pandemic COVID-19, particularly in developing very powerful vaccine against this kind of horrible disease.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest regarding the publication of this paper.

REFERENCES

1. Hyndman, D., Bauman, D.R., Heredia, V.V. and Penning, T.M. (2003) The Aldo-Keto Reductase Superfamily Homepage. *Chemico-Biological Interactions*, **143-144**, 621-631. [https://doi.org/10.1016/S0009-2797\(02\)00193-X](https://doi.org/10.1016/S0009-2797(02)00193-X)
2. Akbar, S., Rahman, A.U., Hayat, M., *et al.* (2020) cACP: Classifying Anticancer Peptides Using Discriminative Intelligent Model via Chou’s 5-Step Rules and General Pseudo Components. *Chemometrics and Intelligent Laboratory (CHEMOLAB)*, **196**, Article ID: 103912. <https://doi.org/10.1016/j.chemolab.2019.103912>
3. Charoenkwan, P., Schaduangrat, N., Nantasenammat, C., Piacham, T. and Shoombuatong, W. (2020) Correction: Shoombuatong, W., *et al.* iQSP: A Sequence-Based Tool for the Prediction and Analysis of Quorum Sensing Peptides via Chou’s 5-Steps Rule and Informative Physicochemical Properties. *International Journal of Molecular Sciences*, **21**, 75. <https://doi.org/10.3390/ijms21072629>
4. Yang, L., Lv, Y., Wang, S., Zhang, Q., Pan, Y., Su, D., Lu, Q. and Zuo, Y. (2020) Identifying FL11 Subtype by Characterizing Tumor Immune Microenvironment in Prostate Adenocarcinoma via Chou’s 5-Steps Rule. *Genomics*, **112**, 1500-1515. <https://doi.org/10.1016/j.ygeno.2019.08.021>
5. Kabir, M., Ahmad, S., Iqbal, M. and Hayat, M. (2020) iNR-2L: A Two-Level Sequence-Based Predictor Developed via Chou’s 5-Steps Rule and General PseAAC for Identifying Nuclear Receptors and Their Families. *Genomics*, **112**, 276-285. <https://doi.org/10.1016/j.ygeno.2019.02.006>
6. Nosrati, M., Mohabatkar, H. and Behbahani, M. (2020) Introducing of an Integrated Artificial Neural Network and Chou’s Pseudo Amino Acid Composition Approach for Computational Epitope-Mapping of Crimean-Congo Haemorrhagic Fever Virus Antigens. *International Immunopharmacology*, **78**, Article ID: 106020. <https://doi.org/10.1016/j.intimp.2019.106020>
7. Khan, Y.D., Amin, N., Hussain, W., Rasool, N., Khan, S.A. and Chou, K.C. (2020) iProtease-PseAAC(2L): A Two-Layer Predictor for Identifying Proteases and Their Types Using Chou’s 5-Step-Rule and General PseAAC.

Analytical Biochemistry, **588**, Article ID: 113477. <https://doi.org/10.1016/j.ab.2019.113477>

8. Charoenkwan, P., Schaduengrat, N., Nantasenamat, C., Piacham, T. and Shoombuatong, W. (2020) iQSP: A Sequence-Based Tool for the Prediction and Analysis of Quorum Sensing Peptides via Chou's 5-Steps Rule and Informative Physicochemical Properties. *International Journal of Molecular Sciences*, **21**, 75. <https://doi.org/10.3390/ijms21010075>
9. Roy, T. and Bhattacharjee, P. (2020) A LabVIEW-Based Real-Time Modeling Approach via Chou's 5-Steps Rule for Detection of Abnormalities in Cancer Cells. *Gene Reports*, Article ID: 100788. <https://doi.org/10.1016/j.genrep.2020.100788>
10. Vishnoi, S., Garg, P. and Arora, P. (2020) Physicochemical n-Grams Tool: A Tool for Protein Physicochemical Descriptor Generation via Chou's 5-Step Rule. *Chemical Biology & Drug Design*, **95**, 79-86. <https://doi.org/10.1111/cbdd.13617>
11. Ju, Z. and Wang, S.Y. (2020) Prediction of Lysine Formylation Sites Using the Composition of k-Spaced Amino Acid Pairs via Chou's 5-Steps Rule and General Pseudo Components. *Genomics*, **112**, 859-866. <https://doi.org/10.1016/j.ygeno.2019.05.027>
12. Khan, S., Khan, M., Iqbal, N., Khan, S.A. and Chou, K.C. (2020) Prediction of piRNAs and Their Function Based on Discriminative Intelligent Model Using Hybrid Features into Chou's PseKNC. *Chemometrics and Intelligent Laboratory (CHEMOLAB)*, **203**, Article ID: 104056. <https://doi.org/10.1016/j.chemolab.2020.104056>
13. Song, C. and Yang, B. (2020) Use Chou's 5-Step Rule to Classify Protein Modification Sites with Neural Network. *Scientific Programming*, **2020**, Article ID: 8894633. <https://doi.org/10.1155/2020/8894633>
14. Song, C. and Yang, B. (2020) Use Chou's 5-Step Rule to Classify Protein Modification Sites with Neural Network. *Scientific Programming*, **2020**, Article ID: 8894633. <https://doi.org/10.1155/2020/8894633>
15. Song, C. and Yang, B. (2020) Use Chou's 5-Step Rule to Classify Protein Modification Sites with Neural Network. *Scientific Programming*, **2020**, Article ID: 8894633. <https://doi.org/10.1155/2020/8894633>
16. Lu, W., Song, Z., Ding, Y., Wu, H., Cao, Y. and Zhang, Y.L. (2020) Use Chou's 5-Step Rule to Predict DNA-Binding Proteins with Evolutionary Information. *BioMed Research International*, **2020**, Article ID: 6984045. <https://doi.org/10.1155/2020/6984045>
17. Pandey, R.P., Kumar, S., Ahmad, S., Vibhuti, A., Raj, V.S., Verma, A.K., Sharma, P. and Leal, E. (2020) Use Chou's 5-Steps Rule to Evaluate Protective Efficacy Induced by Antigenic Proteins of Mycobacterium Tuberculosis Encapsulated in Chitosan Nanoparticles. *Life Sciences*, **256**, Article ID: 117961. <https://doi.org/10.1016/j.lfs.2020.117961>
18. Lin, W., Xiao, X., Qiu, W. and Chou, K.C. (2020) Use Chou's 5-Steps Rule to Predict Remote Homology Proteins by Merging Grey Incidence Analysis and Domain Similarity Analysis. *Natural Science*, **12**, 181-198. <https://doi.org/10.4236/ns.2020.123016>
19. Nguyen, D., Ho-Quang, T., Nguyen Quoc Khanh, L., Dinh-Phan, V. and Ou, Y.Y. (2020) Use Chou's 5-Steps Rule with Different Word Embedding Types to Boost Performance of Electron Transport Protein Prediction Model. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. <https://doi.org/10.1109/TCBB.2020.3010975>
20. Bouziane, H. and Chouarfia, A. (2020) Use of Chou's 5-Steps Rule to Predict the Subcellular Localization of Gram-Negative and Gram-Positive Bacterial Proteins by Multi-Label Learning Based on Gene Ontology Annotation and Profile Alignment. *Journal of Integrative Bioinformatics*. <https://doi.org/10.1515/jib-2019-0091>
21. Chen, Y. and Fan, X. (2020) Use of Chou's 5-Steps Rule to Reveal Active Compound and Mechanism of Shuangshen Pingfei San on Idiopathic Pulmonary Fibrosis. *Current Molecular Medicine*, **20**, 220-230. <https://doi.org/10.2174/1566524019666191011160543>

22. Dobosz, R., Mucko, J. and Gawinecki, R. (2020) Using Chou's 5-Step Rule to Evaluate the Stability of Tautomers: Susceptibility of 2-[(Phenylimino)-methyl]-cyclohexane-1,3-diones to Tautomerization Based on the Calculated Gibbs Free Energies. *Energies*, **13**, 183. <https://doi.org/10.3390/en13010183>
23. Vundavilli, H., Datta, A., Sima, C., Hua, J., Lopes, R. and Bittner, M. (2020) Using Chou's 5-Steps Rule to Model Feedback in Lung Cancer. *IEEE Journal of Biomedical and Health Informatics*, **21**, 1-24. <https://doi.org/10.1109/JBHI.2019.2958042>
24. Akmal, M.A., Hussain, W., Rasool, N., Khan, Y.D., Khan, S.A. and Chou, K.C. (2020) Using Chou's 5-Steps Rule to Predict O-Linked Serine Glycosylation Sites by Blending Position Relative Features and Statistical Moment. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **2020**, 1-12. <https://doi.org/10.1109/TCBB.2020.2968441>
25. Mohabatkar, H., Ebrahimi, S. and Moradi, M. (2020) Using Chou's Five-Steps Rule to Classify and Predict Glutathione S-Transferases with Different Machine Learning Algorithms and Pseudo Amino Acid Composition. *International Journal of Peptide Research and Therapeutics*, **2020**, 1-17. <https://doi.org/10.1007/s10989-020-10087-7>
26. Behbahani, M., Nosrati, M., Moradi, M. and Mohabatkar, H. (2020) Using Chou's General Pseudo Amino Acid Composition to Classify Laccases from Bacterial and Fungal Sources via Chou's Five-Step Rule. *Applied Biochemistry and Biotechnology*, **190**, 1035-1048. <https://doi.org/10.1007/s12010-019-03141-8>
27. Du, L., Meng, Q., Jiang, H. and Li, Y. (2020) Using Evolutionary Information and Multi-Label Linear Discriminant Analysis to Predict the Subcellular Location of Multi-Site Bacterial Proteins via Chou's 5-Steps Rule. *IEEE Access*, **8**, 56452-56461. <https://doi.org/10.1109/ACCESS.2020.2982160>
28. Zhao, X.F., Min, Z., Wei, X. and Ju, Y. (2020) Using the Chou's 5-Steps Rule, Transient Overexpression Technique, Subcellular Location, and Bioinformatic Analysis to verify the Function of *Vitis vinifera* O-Methyltransferase 3 (VvOMT3) Protein. *Plant Physiology and Biochemistry*, **151**, 621-629. <https://doi.org/10.1016/j.plaphy.2020.04.015>
29. Dutta, A., Dalmia, A., R, A., Singh, K.K. and Anand, A. (2020) Using the Chou's 5-Steps Rule to Predict Splice Junctions with Interpretable Bidirectional Long Short-Term Memory Networks. *Computers in Biology and Medicine*, **116**, Article ID: 103558. <https://doi.org/10.1016/j.combiomed.2019.103558>
30. Butt, A.H. and Khan, Y.D. (2018) Prediction of S-Sulfenylation Sites Using Statistical Moments Based Features via Chou's 5-Step Rule. *International Journal of Peptide Research and Therapeutics (IJPR)*, **17**, 711-738.
31. Awais, M., Hussain, W., Khan, Y.D., Rasool, N., Khan, S.A. and Chou, K.C. (2019) iPhosH-PseAAC: Identify Phosphohistidine Sites in Proteins by Blending Statistical Moments and Position Relative Features According to the Chou's 5-Step Rule and General Pseudo Amino Acid Composition. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **19**, 1-21. <https://doi.org/10.1109/TCBB.2019.2919025>
32. Barukab, O., Khan, Y.D., Khan, S.A. and Chou, K.C. (2019) iSulfoTyr-PseAAC: Identify Tyrosine Sulfation Sites by Incorporating Statistical Moments via Chou's 5-Steps Rule and Pseudo Components. *Current Genomics*, **20**, 306-320. <https://doi.org/10.2174/1389202920666190819091609>
33. Butt, A.H. and Khan, Y.D. (2019) Prediction of S-Sulfenylation Sites Using Statistical Moments Based Features via Chou's 5-Step Rule. *International Journal of Peptide Research and Therapeutics (IJPR)*, **19**, 71-82.
34. Chen, Y. and Fan, X. (2019) Use Chou's 5-Steps Rule to Reveal Active Compound and Mechanism of Shuangsheng Pingfei San on Idiopathic Pulmonary Fibrosis. *Current Molecular Medicine*, **19**, 511-563.
35. Du, X., Diao, Y., Liu, H. and Li, S. (2019) MsDBP: Exploring DNA-Binding Proteins by Integrating Multi-Scale Sequence Information via Chou's 5-Steps Rule. *Journal of Proteome Research*, **18**, 3119-3132. <https://doi.org/10.1021/acs.jproteome.9b00226>
36. Dutta, A., Dalmia, A., R, A., Singh, K.K. and Anand, A. (2019) Using the Chou's 5-Steps Rule to Predict Splice

Junctions with Interpretable Bidirectional Long Short-Term Memory Networks. *Computers in Biology and Medicine*, **116**, Article ID: 103558. <https://doi.org/10.1016/j.compbiomed.2019.103558>

37. Ehsan, A., Mahmood, M.K., Khan, Y.D., Barukab, O.M., Khan, S.A. and Chou, K.C. (2019) iHyd-PseAAC (EPSV): Identify Hydroxylation Sites in Proteins by Extracting Enhanced Position and Sequence Variant Feature via Chou's 5-Step Rule and General Pseudo Amino Acid Composition. *Current Genomics*, **20**, 124-133. <https://doi.org/10.2174/1389202920666190325162307>
38. Hussain, W., Khan, S.D., Rasool, N., Khan, S.A. and Chou, K.C. (2019) SPalmitoylC-PseAAC: A Sequence-Based Model Developed via Chou's 5-Steps Rule and General PseAAC for Identifying S-Palmitoylation Sites in Proteins. *Analytical Biochemistry*, **568**, 14-23. <https://doi.org/10.1016/j.ab.2018.12.019>
39. Hussain, W., Khan, Y.D., Rasool, N., Khan, S.A. and Chou, K.C. (2019) SPrenylC-PseAAC: A Sequence-Based Model Developed via Chou's 5-Steps Rule and General PseAAC for Identifying S-Prenylation Sites in Proteins. *Journal of Theoretical Biology*, **468**, 1-11. <https://doi.org/10.1016/j.jtbi.2019.02.007>
40. Jun, Z. and Wang, S.Y. (2019) Identify Lysine Neddylolation Sites Using Bi-Profile Bayes Feature Extraction via the Chou's 5-Steps Rule and General Pseudo Components. *Current Genomics*, **20**, 592-601. <https://doi.org/10.2174/1389202921666191223154629>
41. Khan, S., Khan, M., Iqbal, N., Hussain, T., Khan, S.A. and Chou, K.C. (2019) A Two-Level Computation Model Based on Deep Learning Algorithm for Identification of piRNA and Their Functions via Chou's 5-Steps Rule. *Human Genetics*, **19**, 756-799. <https://doi.org/10.1007/s10989-019-09887-3>
42. Khan, Z.U., Ali, F., Khan, I.A., Hussain, Y. and Pi, D. (2019) iRSpot-SPI: Deep Learning-Based Recombination Spots Prediction by Incorporating Secondary Sequence Information Coupled with Physio-Chemical Properties via Chou's 5-Step Rule and Pseudo Components. *Chemometrics and Intelligent Laboratory Systems (CHEMOLAB)*, **189**, 169-180. <https://doi.org/10.1016/j.chemolab.2019.05.003>
43. Lan, J., Liu, J., Liao, C., Merkle, D.J., Han, Q. and Li, J. (2019) A Study for Therapeutic Treatment against Parkinson's Disease via Chou's 5-Steps Rule. *Current Topics in Medicinal Chemistry*, **19**, 2318-2333. <https://doi.org/10.2174/1568026619666191019111528>
44. Le, N.Q.K. (2019) iN6-methylat (5-Step): Identifying DNA N(6)-methyladenine Sites in Rice Genome Using Continuous Bag of Nucleobases via Chou's 5-Step Rule. *Molecular Genetics and Genomics. MGG*, **294**, 1173-1182. <https://doi.org/10.1007/s00438-019-01570-y>
45. Le, N.Q.K., Yapp, E.K.Y., Ho, Q.T., Nagasundaram, N., Ou, Y.Y. and Yeh, H.Y. (2019) iEnhancer-5Step: Identifying Enhancers Using Hidden Information of DNA Sequences via Chou's 5-Step Rule and Word Embedding. *Analytical Biochemistry*, **571**, 53-61. <https://doi.org/10.1016/j.ab.2019.02.017>
46. Le, N.Q.K., Yapp, E.K.Y., Ou, Y.Y. and Yeh, H.Y. (2019) iMotor-CNN: Identifying Molecular Functions of Cytoskeleton Motor Proteins Using 2D Convolutional Neural Network via Chou's 5-Step Rule. *Analytical Biochemistry*, **575**, 17-26. <https://doi.org/10.1016/j.ab.2019.03.017>
47. Liang, R., Xie, J., Zhang, C., Zhang, M., Huang, H., Huo, H., Cao, X. and Niu, B. (2019) Identifying Cancer Targets Based on Machine Learning Methods via Chou's 5-Steps Rule and General Pseudo Components. *Current Topics in Medical Chemistry*, **19**, 2301-2317. <https://doi.org/10.2174/1568026619666191016155543>
48. Liang, Y. and Zhang, S. (2019) Identifying DNase I Hypersensitive Sites Using Multi-Features Fusion and F-Score Features Selection via Chou's 5-Steps Rule. *Biophysical Chemistry*, **253**, Article ID: 106227. <https://doi.org/10.1016/j.bpc.2019.106227>
49. Liu, Z., Dong, W., Jiang, W. and He, Z. (2019) csDMA: An Improved Bioinformatics Tool for Identifying DNA 6 mA Modifications via Chou's 5-Step Rule. *Scientific Reports*, **9**, Article No. 13109. <https://doi.org/10.1038/s41598-019-49430-4>

50. Malebary, S.J., Rehman, M.S.U. and Khan, Y.D. (2019) iCrotoK-PseAAC: Identify Lysine Crotonylation Sites by Blending Position Relative Statistical Features According to the Chou's 5-Step Rule. *PLoS ONE*, **14**, e0223993. <https://doi.org/10.1371/journal.pone.0223993>
51. Nazari, I., Tahir, M., Tayari, H. and Chong, K.T. (2019) iN6-Methyl (5-step): Identifying RNA N6-Methyladenosine Sites Using Deep Learning Mode via Chou's 5-Step Rules and Chou's General PseKNC. *Chemometrics and Intelligent Laboratory Systems (CHEMOLAB)*, **19**, 1-39. <https://doi.org/10.1016/j.chemolab.2019.103811>
52. Ning, Q., Ma, Z. and Zhao, X. (2019) dForml(KNN)-PseAAC: Detecting Formylation Sites from Protein Sequences Using K-Nearest Neighbor Algorithm via Chou's 5-Step Rule and Pseudo Components. *Journal of Theoretical Biology*, **470**, 43-49. <https://doi.org/10.1016/j.jtbi.2019.03.011>
53. Tahir, M., Tayara, H. and Chong, K.T. (2019) iDNA6mA (5-Step Rule): Identification of DNA N6-Methyladenine Sites in the Rice Genome by Intelligent Computational Model via Chou's 5-Step Rule. *CHEMOLAB*, **189**, 96-101. <https://doi.org/10.1016/j.chemolab.2019.04.007>
54. Wiktorowicz, A., Wit, A., Dziewierz, A., Rzeszutko, L., Dudek, D. and Kleczynski, P. (2019) Calcium Pattern Assessment in Patients with Severe Aortic Stenosis via the Chou's 5-Steps Rule. *Current Pharmaceutical Design*, **25**, 6-31. <https://doi.org/10.2174/1381612825666190930101258>
55. Yang, L., Lv, Y., Wang, S., Zhang, Q., Pan, Y., Su, D., Lu, Q. and Zuo, Y. (2019) Identifying FL11 Subtype by Characterizing Tumor Immune Microenvironment in Prostate Adenocarcinoma via Chou's 5-Steps Rule. *Genomics*, **112**, 1500-1515. <https://doi.org/10.1016/j.ygeno.2019.08.021>
56. Zhang, S. and Xue, T. (2020) Use Chou's 5-Steps Rule to Identify DNase I Hypersensitive Sites via Dinucleotide Property Matrix and Extreme Gradient Boosting. *Molecular Genetics and Genomics: MGG*. <https://doi.org/10.1007/s00438-020-01711-8>