

# The Development of Gordon Life Science Institute: Its Driving Force and Accomplishments

**Kuo-Chen Chou**

Gordon Life Science Institute, Boston, MA 02478, USA

**Correspondence to:** Kuo-Chen Chou, [kcchou@gordonlifescience.org](mailto:kcchou@gordonlifescience.org), [kcchou38@gmail.com](mailto:kcchou38@gmail.com)

**Keywords:** Sweden, Cradle, San Diego, Boston, PseAAC and PseKNC, Disorted Key Theory, Wenxiang Diagram, Multi-Label System Prediction, 5-Steps Rule

**Received:** February 29, 2020

**Accepted:** March 30, 2020

**Published:** April 2, 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

Established in 2004, Gordon Life Science Institute is the first Internet Research Institute in the world. It is a non-profit institute, a gift to science. Those scientists, who are really loving science more than anything else and have shown fantastic creativity in science, can become the membership of such Institute. Their driving force is not funding but firmly belief that scientists will do much better science if they do not have to spend a lot of time for funding application, and that great scientific findings in history were often discovered by those who were without funding at all but driven by profound imagination and curiosity. Summarized in this review are also the accomplishments of the Gordon Life Science Institute and its future perspective.

## 1. INTRODUCTION

The Gordon Life Science Institute was founded by Professor Dr. Kuo-Chen Chou right after he was retired from Pfizer Global Research and Development in 2003. Its birthplace or cradle was at San Diego of California, USA. Its mission is to develop and apply new mathematical tools and physical concepts for understanding biological phenomena. For a briefing about its history and philosophy, click

<https://gordonlifescience.org/GordonLifeScience.html>.

The Institute's name reflects an interesting historical story. After the Cultural Revolution, China started to open its door, the founder was invited by Professor Sture Forsén, the then Chairman of Nobel Prize Committee, to work in Chemical Center of Lund University as a Visiting Professor. In order for Swedish people easier to pronounce his name, Professor Chou used "Gordon" as his name in Sweden. About a quarter of century later, the same name was used for the Institute, implying that "Reform and Opening" and "Free Communication" can stimulate a lot of great creativities.

The current liaison site of Gordon Life Science Institute is in Boston of Massachusetts, USA; [gls@gordonlifescience.org](mailto:gls@gordonlifescience.org).

## 2. TASKS AND CAMPUS

As an Internet Institute, it is without physical boundaries. Its members do not have to work in a same campus or building. Located around the world, they shall freely collaborate, exchange ideas, and share information and findings via various kinds of “Online Communication” methods. In some sense, it is also a very efficient practice to reduce the risk of suffering from the “New Coronavirus” that have endangered the entire world during the period of 2019-2020. Actually, during that pandemic period all the other Research Institutes and Universities in the entire world must be forced to being closed. In contrast to this, seven very powerful predictors (e.g., “pLoc\_Deep-mEuk”, “pLoc\_Deep-mHum”, and “pLoc\_Dep-mVirus”) were generated from the Gordon Life Science Institute and have become very useful tools to fight against the Coronavirus. Furthermore, its members can focus completely on science without having to cope with the troubles in obtaining visas and in paying for relocation expenses, among many others.

The Gordon Life Science Institute is a non-profit organization. It is a gift to human beings and science. Its founding principle is to pursue the excellence in science: anyone who has proved his/her creativity in science can become a member regardless of his/her age, occupation, and nationality. Accordingly, the Institute has provided an ideal organization or society for those scientists who are indeed loving science more than anything else.

Members of the Institute have firmly believed that science will be more fantastic if scientists do not have to spend a lot of time on funding applications and that great scientific findings in history were often made by those who were without any funding but driven by profound imagination and curiosity. They are absolutely concurring with the famous statement of **Albert Einstein**, “Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution”.

## 3. ACHIEVEMENTS

Up to March 2019, the Institute has 26 members. Among them 5 have been selected by Thompson Reuter and Clarivate Analytics as the “Highly Cited Researcher”: 1) Xuan Xiao (2018), 2) Hao Lin (2018), 3) Wei Chen (2018), 4) Hong-Bin Shen (2014 and 2015), and 5) Kuo-Chen Chou for continuously 5 years (2014, 2015, 2016, 2017, and 2018).

Listed below are just some represented works produced by the Gordon Life Science Institute.

### 3.1. Special PseAAC Has Been Extended to the General One

With the avalanche of biological sequences in the post-genomic age, one of the most critical problems in computational biology is how to formulate a biological sequence with a vector or discrete model, yet still keep considerable sequence-order information or key pattern characteristic. This is because all the existing machine-learning algorithms can only handle vectors as elaborated in a comprehensive review [1]. However, a vector defined in a discrete model may completely lose all the sequence-pattern information. To avoid completely losing the sequence-pattern information for proteins, the pseudo amino acid composition [2] or PseAAC [3] was proposed. Ever since then, it has been widely used in nearly all the areas of computational proteomics [4-67].

In the meantime, four powerful open access soft-wares, called “PseAAC” [68], “PseAAC-Builder” [66], “propy” [69], and “PseAAC-General” [57], were established: the former three are for generating various modes of Chou’s special PseAAC [70]; while the 4th one for those of Chou’s general PseAAC [71], including not only all the special modes of feature vectors for proteins but also the higher level feature vectors such as “Functional Domain” mode (see Eqs.9-10 of [71]), “Gene Ontology” mode (see Eqs.11-12 of [71]), and “Sequential Evolution” or “PSSM” mode (see Eqs.13-14 of [71]).

### 3.2. Extension of PseAAC to PseKNC

Encouraged by the successes of using PseAAC to deal with protein/peptide sequences, the concept of PseKNC (Pseudo K-tuple Nucleotide Composition) [72] was developed for generating various feature

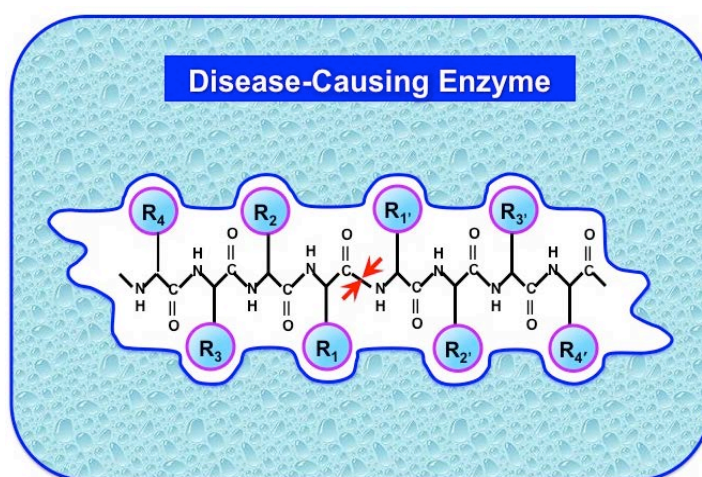
vectors for DNA/RNA sequences that have proved very useful as well [73]. Particularly, in 2015 a very powerful web-server called “Pse-in-One” [74] and its updated version “Pse-in-One2.0” [75] have been established that can be used to generate any desired feature vectors for protein/peptide and DNA/RNA sequences according to the need of users’ studies.

### 3.3. Distorted Key Theory for Peptide Drugs

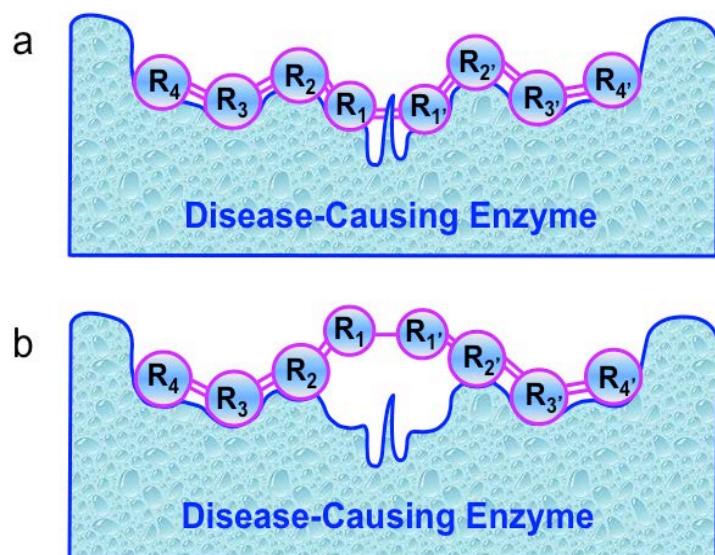
According to Fisher’s “lock and key” model [76], Koshland’s “induced fit” theory [76], and the “rack mechanism” [77], the prerequisite condition for a peptide to be cleaved by the disease-causing enzyme is a good fit and tightly binding with the enzyme’s active site (Figure 1). However, such a peptide, after a modification on its scissile bond with some simple chemical procedure, will no longer be cleavable by the enzyme but it can still tightly bind to its active site. An illustration about the distorted key theory is given in Figure 2, where panel (a) shows an effective binding of a cleavable peptide to the active site of HIV protease, while panel (b) the peptide has become a non-cleavable one after its scissile bond is modified although it can still bind to the active site. Such a modified peptide, or “distorted key”, will automatically become an inhibitor candidate against HIV protease. Even for non-peptide inhibitors, the information derived from the cleavable peptides can also provide useful insights about the key binding groups and fitting conformation in the sense of microenvironment. Besides, peptide drugs usually have no toxicity in vivo under the physiological concentration [78]. For more discussion about the distorted key theory, see a comprehensive review paper [79]. It was based on such a distorted key theory that many investigators were enthusiastic to develop various methods for predicting the protein cleavage sites by disease-causing enzymes (see, e.g., [80]). Furthermore, a web-server called “HIVcleave” [81] has been established for predicting HIV protease cleavage sites in proteins. Its website address is at <http://chou.med.harvard.edu/bioinf/HIV/>.

### 3.4. Introduction of Wenxiang Diagram

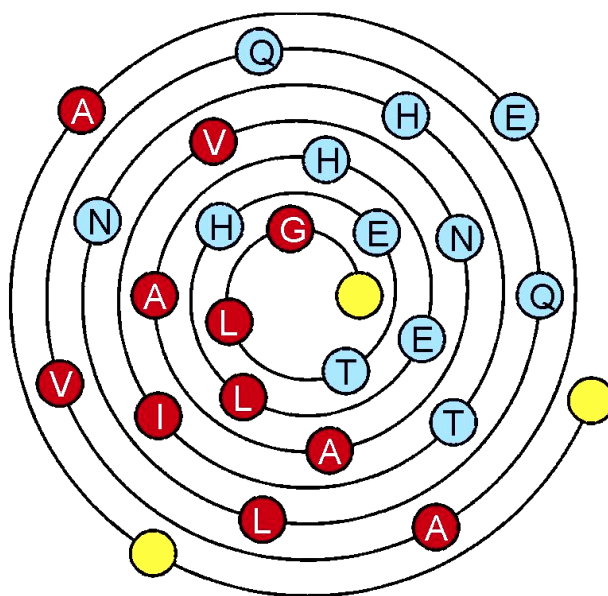
Using graphic approaches to study biological and medical systems can provide an intuitive vision and useful insights for helping analyze complicated relations therein, as indicated by many previous studies on a series of important biological topics, (see, e.g., [82]). The “wenxiang” diagram (Figure 3) [83, 84] is a special kind of graphical approach, which is very useful for in-depth studying protein-protein interaction mechanism [85, 86]. Also, the wenxiang diagram has also been used to study drug-metabolism system [87]. The name of “wenxiang” came from that its shape looks quite like the Chinese wenxiang (蚊香), a



**Figure 1.** A schematic illustration to show a peptide in good fitting and tightly binding with the enzyme’s active site before it is cleaved by the latter. Adapted from [79] with permission.



**Figure 2.** Schematic drawing to illustrate the “Distorted Key” theory, where panel (a) shows an effective binding of a cleavable peptide to the active site of a disease-causing enzyme, while panel (b) the same peptide has become a non-cleavable one after its scissile bond is modified although it can still bind to the active site. Such a modified peptide, or “distorted key”, will automatically become an inhibitor candidate against the disease-causing enzyme. Adapted from [79] with permission.



**Figure 3.** Schematic drawing to show the “wenxiang diagram”. Adapted from [83] with permission.

coil-like incense widely used in China to repel mosquitos. In the wenxiang graphs each residue is represented by a circle with a letter to indicate its code: a hydrophobic residue is denoted by a filled circle



with a white code symbol, a hydrophilic residue is denoted by an open circle with a black code symbol, whereas the invalid residue is denoted by a yellow-filled circle.

### 3.5. Predictors for Multi-Label Systems

Information of subcellular localization for a protein is indispensable for revealing its biological function. Therefore, one of the fundamental goals in molecular cell biology and proteomics is to determine the subcellular locations of proteins in an entire cell. Before 2007, most efforts in this regard were focused on the single-label system by assuming that each of the constitute proteins in a cell had one, and only one, subcellular location (see, e.g., [88-92]). However, with more experimental data uncovered, it has been found that many proteins may simultaneously occur or move between two or more location sites in a cell and hence need multiple labels to mark them. Proteins with multiple locations are also called multiplex proteins [93, 94], which are often the special targets for drug development (see, e.g., [94-96]). In studying the multi-label systems, we need two kinds of metrics to measure performance quality of a predictor: one is for the accuracy of global prediction and the other for the accuracy of local prediction [97]. As a show-case, let us consider the multi-label predictor of pLoc\_bal-mVirus [98], which was developed for studying the 6 organelles or subcellular locations (Figure 4) in a virus cell. 1) Click the link at [http://www.jci-bioinfo.cn/pLoc\\_bal-mVirus/](http://www.jci-bioinfo.cn/pLoc_bal-mVirus/), you'll see the top page of the predictor prompted on your computer screen (Figure 5). 2) You can either type or copy/paste the sequences of query virus proteins into the input box at the center of Figure 5. The input sequence should be in the FASTA format. You can click the Example button right above the input box to see the sequences in FASTA format. 3) Click on the Submit button to see the predicted result; e.g., if you use the four protein sequences in the Example window as the input, after 10 seconds or so, you will see a new screen (Figure 5) occurring. On its upper part are listed the names of the subcellular locations numbered from (1) to (6) covered by the current predictor. On its lower part are the predicted results: the query protein "P01115" of example-1 corresponds to "2", meaning it belonging to "Host cell membrane" only; the query protein "P03495" of example-2 corresponds to "4, 5", meaning it belonging to "Host cytoplasm" and "Host nucleus"; the query protein "P89873" of example-3 corresponds to "4, 5, 6", meaning it belonging to "Host cytoplasm", "Host nucleus", and "Secreted". All these results are perfectly consistent with experimental observations.

**pLoc\_bal-mVirus: predict subcellular localization of virus proteins by balancing training dataset and general PseAAC**  
| [Read Me](#) | [Supporting information](#) | [Citation](#) |

---

**Enter query sequences**

Enter the sequences of query proteins in FASTA format ([Example](#)): the number of proteins is limited at 10 or less for each submission.

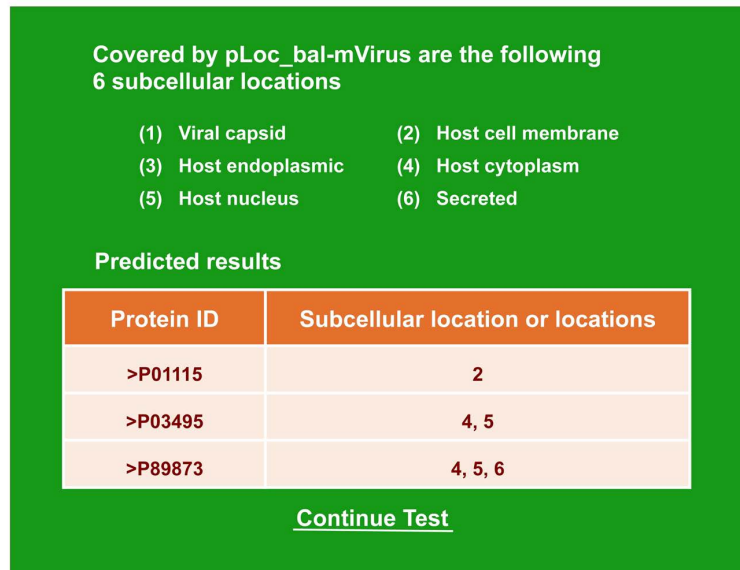
**Or, upload a file for batch prediction**

Enter your e-mail address and upload the batch input file ([Batch-example](#)). The predicted result will be sent to you by e-mail once completed; it usually takes 1 minute or so for each protein sequence

Upload file:    
Your Email:

**Figure 4.** A semi screenshot for the top page of pLoc\_bal-mVirus. Adapted from [98] with permission.



**Figure 5.** A semi screenshot for the webpage obtained by following Step 3 of Section 3.5. Adapted from [98] with permission.

### 3.6. Five-Steps Rule

The Institute was the birth-place or cradle of the famous 5-steps rule [71], which has been used in nearly all the areas of computational biology (see, e.g., [5, 11, 13, 16, 28, 32, 99-111]), material science [112], and even the commercial science (e.g., the bank account systems). The only difference between them is how to formulate the statistical samples or events with an effective mathematical expression that can truly reflect their intrinsic correlation with the target to be predicted. It just likes the case of many machine-learning algorithms. They can be widely used in nearly all the areas of statistical analysis.

Working in such Institute filled with this kind of philosophy and atmosphere, the scientists would be more prone to be stimulated by the master piece papers from the then Chairman of Nobel Prize Committee Sture Forsen (see, e.g., [113, 114]), so as to drive them substantially more creative and productive.

## 4. CONCLUSION AND PERSPECTIVE

In contrast to the conventional institutes, Gordon Life Science Institute has the following unique advantages: it can 1) attract those scientists who are really loving science more than anything else; 2) maximize their creativity in science and minimize the distraction or disturbance caused by the relocation and various followed-up tedious things; 3) provide them with an ideal environment to completely focus on doing science; 4) drive their motivation by profound imagination and curiosity; and 5) create a rich and thick scientific atmosphere to produce their scientific results more truthful, wonderful, and awesome.

Accordingly, it would not be surprised to see that a total of five members of Gordon Life Scientist have been selected by Clarivate Analytics as Highly Cited Researcher or HCR (see Section 3), indicating that for the ratio of HCR per member, the “Gordon Life Science Institute” has already exceeded the “Broad Institute of MIT and Harvard, USA”, becoming the very top in the world.

It is anticipated that more significant accomplishments will be achieved by the Gordon Life Science Institute for many years to come, as indicated by a series of very recent papers [98, 115-136].

## ETHICAL APPROVAL STATEMENT

This article does not contain any studies with human or animal participants.

## CONFLICTS OF INTEREST

No conflict of interest whatsoever is claimed by the author.

## REFERENCES

1. Chou, K.C. (2015) Impacts of Bioinformatics to Medicinal Chemistry. *Medicinal Chemistry*, **11**, 218-234. <https://doi.org/10.2174/1573406411666141229162834>
2. Chou, K.C. (2001) Prediction of Protein Cellular Attributes Using Pseudo Amino Acid Composition. *Proteins: Structure, Function, and Genetics*, **43**, 246-255. (Erratum: *Ibid.*, 2001, Vol. 44, 60) <https://doi.org/10.1002/prot.1035>
3. Chou, K.C. (2005) Using Amphiphilic Pseudo Amino Acid Composition to Predict Enzyme Subfamily Classes. *Bioinformatics*, **21**, 10-19. <https://doi.org/10.1093/bioinformatics/bth466>
4. 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>
5. 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>
6. 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>
7. Zhang, S., Yang, K., Lei, Y. and Song, K. (2019) iRSpot-DTS: Predict Recombination Spots by Incorporating the Dinucleotide-Based Spare-Cross Covariance Information into Chou's Pseudo Components. *Genomics*, **111**, 1760-1770. <https://doi.org/10.1016/j.ygeno.2018.11.031>
8. Zhang, L. and Kong, L. (2019) iRSpot-PDI: Identification of Recombination Spots by Incorporating Dinucleotide Property Diversity Information into Chou's Pseudo Components. *Genomics*, **111**, 457-464. <https://doi.org/10.1016/j.ygeno.2018.03.003>
9. Tahir, M., Hayat, M. and Khan, S.A. (2019) iNuc-ext-PseTNC: An Efficient Ensemble Model for Identification of Nucleosome Positioning by Extending the Concept of Chou's PseAAC to Pseudo-Tri-Nucleotide Composition. *Molecular genetics and genomics: MGG*, **294**, 199-210. <https://doi.org/10.1007/s00438-018-1498-2>
10. Nosrati, M., Mohabatkar, H. and Behbahani, M. (2019) 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://www.sciencedirect.com/science/article/pii/S1567576919321277> <https://doi.org/10.1016/j.intimp.2019.106020>
11. 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>
12. 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>
13. Jun, Z. and Wang, S.Y. (2019) Identify Lysine Neddylatation 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>

14. 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>
15. Behbahani, M., Nosrati, M., Moradi, M. and Mohabatkar, H. (2019) 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>
16. 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. <http://www.eurekaselect.com/174277/article>  
<https://doi.org/10.2174/1389202920666190819091609>
17. 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*. <https://www.ncbi.nlm.nih.gov/pubmed/31144645>  
<https://doi.org/10.1109/TCBB.2019.2919025>
18. Al Maruf, M.A. and Shatabda, S. (2019) iRSpot-SF: Prediction of Recombination Hotspots by Incorporating Sequence Based Features into Chou's Pseudo Components. *Genomics*, **111**, 966-972. <https://doi.org/10.1016/j.ygeno.2018.06.003>
19. Zhao, W., Wang, L., Zhang, T.X., Zhao, Z.N. and Du, P.F. (2018) A Brief Review on Software Tools in Generating Chou's Pseudo-Factor Representations for All Types of Biological Sequences. *Protein & Peptide Letters*, **25**, 822-829. <https://doi.org/10.2174/0929866525666180905111124>
20. Zhang, S., Yang, K., Lei, Y. and Song, K. (2018) iRSpot-DTS: Predict Recombination Spots by Incorporating the Dinucleotide-Based Sparse-Cross Covariance Information into Chou's Pseudo Components. *Genomics*, **11**, 457-464.
21. Mei, J. and Zhao, J. (2018) Prediction of HIV-1 and HIV-2 Proteins by Using Chou's Pseudo Amino Acid Compositions and Different Classifiers. *Scientific Reports*, **8**, Article No. 2359. <https://doi.org/10.1038/s41598-018-20819-x>
22. Ju, Z. and Wang, S.Y. (2018) Prediction of Citrullination Sites by Incorporating k-Spaced Amino Acid Pairs into Chou's General Pseudo Amino Acid Composition. *Gene*, **664**, 78-83. <https://doi.org/10.1016/j.gene.2018.04.055>
23. Cui, X., Yu, Z., Yu, B., Wang, M., Tian, B. and Ma, Q. (2018) UbiSitePred: A Novel Method for Improving the Accuracy of Ubiquitination Sites Prediction by Using LASSO to Select the Optimal Chou's Pseudo Components. *Chemometrics and Intelligent Laboratory Systems (CHEMOLAB)*, **184**, 28-43. <https://doi.org/10.1016/j.chemolab.2018.11.012>
24. Al Maruf, M.A. and Shatabda, S. (2018) iRSpot-SF: Prediction of Recombination Hotspots by Incorporating Sequence Based Features into Chou's Pseudo Components. *Genomics*, **111**, 966-972. <https://doi.org/10.1016/j.ygeno.2018.06.003>
25. Yu, B., Lou, L., Li, S., Zhang, Y., Qiu, W., Wu, X., Wang, M. and Tian, B. (2017) Prediction of Protein Structural Class for Low-Similarity Sequences Using Chou's Pseudo Amino Acid Composition and Wavelet Denoising. *Journal of Molecular Graphics and Modelling*, **76**, 260-273. <https://doi.org/10.1016/j.jmgm.2017.07.012>
26. Liang, Y. and Zhang, S. (2017) Predict Protein Structural Class by Incorporating Two Different Modes of Evolutionary Information into Chou's General Pseudo Amino Acid Composition. *Journal of Molecular Graphics and*



*Modelling*, **78**, 110-117. <https://doi.org/10.1016/j.jmngm.2017.10.003>

27. 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>
28. 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>
29. Xiao, X., Cheng, X., Chen, G., Mao, Q. and Chou, K.C. (2019) pLoc\_bal-mVirus: Predict Subcellular Localization of Multi-Label Virus Proteins by Chou's General PseAAC and IHTS Treatment to Balance Training Dataset. *Medicinal Chemistry*, **15**, 496-509. <https://doi.org/10.2174/1573406415666181217114710>
30. 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>
31. Javed, F. and Hayat, M. (2019) Predicting Subcellular Localization of Multi-Label Proteins by Incorporating the Sequence Features into Chou's PseAAC. *Genomics*, **111**, 1325-1332. <https://doi.org/10.1016/j.ygeno.2018.09.004>
32. 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>
33. Mousavizadegan, M. and Mohabatkar, H. (2018) Computational Prediction of Antifungal Peptides via Chou's PseAAC and SVM. *Journal of Bioinformatics and Computational Biology*, **16**, Article ID: 1850016. <https://doi.org/10.1142/S0219720018500166>
34. Javed, F. and Hayat, M. (2019) Predicting Subcellular Localization of Multi-Label Proteins by Incorporating the Sequence Features into Chou's PseAAC. *Genomics*, **111**, 1325-1332. <https://doi.org/10.1016/j.ygeno.2018.09.004>
35. Ghauri, A.W., Khan, Y.D., Rasool, N., Khan, S.A. and Chou, K.C. (2018) pNitro-Tyr-PseAAC: Predict Nitrotyrosine Sites in Proteins by Incorporating Five Features into Chou's General PseAAC. *Current Pharmaceutical Design*, **24**, 4034-4043. <https://doi.org/10.2174/1381612825666181127101039>
36. Fu, X., Zhu, W., Li, B., Cai, L., Peng, L. and Yang, J. (2018) Improved DNA-Binding Protein Identification by Incorporating Evolutionary Information into the Chou's PseAAC. *IEEE Access*, **6**, 66545-66556. <https://doi.org/10.1109/ACCESS.2018.2876656>
37. Butt, A.H., Rasool, N. and Khan, Y.D. (2018) Predicting Membrane Proteins and Their Types by Extracting Various Sequence Features into Chou's General PseAAC. *Molecular Biology Reports*, **45**, 2295-2306. <https://doi.org/10.1007/s11033-018-4391-5>
38. Yu, B., Li, S., Qiu, W.Y., Chen, C., Chen, R.X., Wang, L., Wang, M.H. and Zhang, Y. (2017) Accurate Prediction of Subcellular Location of Apoptosis Proteins Combining Chou's PseAAC and PsePSSM Based on Wavelet Denoising. *Oncotarget*, **8**, 107640-107665. <https://doi.org/10.18632/oncotarget.22585>
39. Xu, C., Ge, L., Zhang, Y., Dehmer, M. and Gutman, I. (2017) Prediction of Therapeutic Peptides by Incorporating q-Wiener Index into Chou's General PseAAC. *Journal of Biomedical Informatics*, **75**, 63-69. <https://doi.org/10.1016/j.jbi.2017.09.011>
40. Qiu, W.R., Zheng, Q.S., Sun, B.Q. and Xiao, X. (2017) Multi-iPPseEvo: A Multi-Label Classifier for Identifying Human Phosphorylated Proteins by Incorporating Evolutionary Information into Chou's General PseAAC via Grey System Theory. *Molecular Informatics*, **36**, UNSP 1600085. <https://doi.org/10.1002/minf.201600085>
41. Meher, P.K., Sahu, T.K., Saini, V. and Rao, A.R. (2017) Predicting Antimicrobial Peptides with Improved Accuracy by Incorporating the Compositional, Physico-Chemical and Structural Features into Chou's General

PseAAC. *Scientific Reports*, **7**, Article No. 42362. <https://doi.org/10.1038/srep42362>

42. Ju, Z. and He, J.J. (2017) Prediction of Lysine Crotonylation Sites by Incorporating the Composition of k-Spaced Amino Acid Pairs into Chou's General PseAAC. *Journal of Molecular Graphics and Modelling*, **77**, 200-204. <https://doi.org/10.1016/j.jmgm.2017.08.020>
43. Ju, Z. and He, J.J. (2017) Prediction of Lysine Propionylation Sites Using Biased SVM and Incorporating Four Different Sequence Features into Chou's PseAAC. *Journal of Molecular Graphics and Modelling*, **76**, 356-363. <https://doi.org/10.1016/j.jmgm.2017.07.022>
44. Tiwari, A.K. (2016) Prediction of G-Protein Coupled Receptors and Their Subfamilies by Incorporating Various Sequence Features into Chou's General PseAAC. *Computer Methods and Programs in Biomedicine*, **134**, 197-213. <https://doi.org/10.1016/j.cmpb.2016.07.004>
45. Tahir, M. and Hayat, M. (2016) iNuc-STNC: A Sequence-Based Predictor for Identification of Nucleosome Positioning in Genomes by Extending the Concept of SAAC and Chou's PseAAC. *Molecular BioSystems*, **12**, 2587-2593. <https://doi.org/10.1039/C6MB00221H>
46. Kabir, M. and Hayat, M. (2016) iRSpot-GAEnsC: Identifying Recombination Spots via Ensemble Classifier and Extending the Concept of Chou's PseAAC to Formulate DNA Samples. *Molecular Genetics and Genomics*, **291**, 285-296. <https://doi.org/10.1007/s00438-015-1108-5>
47. Zhang, S.L. (2015) Accurate Prediction of Protein Structural Classes by Incorporating PSSS and PSSM into Chou's General PseAAC. *Chemometrics and Intelligent Laboratory Systems (CHEMOLAB)*, **142**, 28-35. <https://doi.org/10.1016/j.chemolab.2015.01.004>
48. Zhang, M., Zhao, B. and Liu, X. (2015) Predicting Industrial Polymer Melt Index via Incorporating Chaotic Characters into Chou's General PseAAC. *Chemometrics and Intelligent Laboratory Systems (CHEMOLAB)*, **146**, 232-240. <https://doi.org/10.1016/j.chemolab.2015.05.028>
49. Sharma, R., Dehzangi, A., Lyons, J., Paliwal, K., Tsunoda, T. and Sharma, A. (2015) Predict Gram-Positive and Gram-Negative Subcellular Localization via Incorporating Evolutionary Information and Physicochemical Features into Chou's General PseAAC. *IEEE Transactions on NanoBioscience*, **14**, 915-926. <https://doi.org/10.1109/TNB.2015.2500186>
50. Sanchez, V., Peinado, A.M., Perez-Cordoba, J.L. and Gomez, A.M. (2015) A New Signal Characterization and Signal-Based Chou's PseAAC Representation of Protein Sequences. *Journal of Bioinformatics and Computational Biology*, **13**, Article ID: 1550024. <https://doi.org/10.1142/S0219720015500249>
51. Mandal, M., Mukhopadhyay, A. and Maulik, U. (2015) Prediction of Protein Subcellular Localization by Incorporating Multiobjective PSO-Based Feature Subset Selection into the General Form of Chou's PseAAC. *Medical & Biological Engineering & Computing*, **53**, 331-344. <https://doi.org/10.1007/s11517-014-1238-7>
52. Liu, B., Xu, J., Fan, S., Xu, R., Zhou, J. and Wang, X. (2015) PseDNA-Pro: DNA-Binding Protein Identification by Combining Chou's PseAAC and Physicochemical Distance Transformation. *Molecular Informatics*, **34**, 8-17. <https://doi.org/10.1002/minf.201400025>
53. Ahmad, S., Kabir, M. and Hayat, M. (2015) Identification of Heat Shock Protein Families and J-Protein Types by Incorporating Dipeptide Composition into Chou's General PseAAC. *Computer Methods and Programs in Biomedicine*, **122**, 165-174. <https://doi.org/10.1016/j.cmpb.2015.07.005>
54. Zhang, J., Zhao, X., Sun, P. and Ma, Z. (2014) PSNO: Predicting Cysteine S-Nitrosylation Sites by Incorporating Various Sequence-Derived Features into the General Form of Chou's PseAAC. *International Journal of Molecular Sciences*, **15**, 11204-11219. <https://doi.org/10.3390/ijms150711204>
55. Li, L., Yu, S., Xiao, W., Li, Y., Li, M., Huang, L., Zheng, X., Zhou, S. and Yang, H. (2014) Prediction of Bacterial Protein Subcellular Localization by Incorporating Various Features into Chou's PseAAC and a Backward Fea-

- ture Selection Approach. *Biochimie*, **104**, 100-107. <https://doi.org/10.1016/j.biochi.2014.06.001>
56. Hayat, M. and Iqbal, N. (2014) Discriminating Protein Structure Classes by Incorporating Pseudo Average Chemical Shift to Chou's General PseAAC and Support Vector Machine. *Computer Methods and Programs in Biomedicine*, **116**, 184-192. <https://doi.org/10.1016/j.cmpb.2014.06.007>
  57. Du, P., Gu, S. and Jiao, Y. (2014) PseAAC-General: Fast Building Various Modes of General Form of Chou's Pseudo Amino Acid Composition for Large-Scale Protein Datasets. *International Journal of Molecular Sciences*, **15**, 3495-3506. <https://doi.org/10.3390/ijms15033495>
  58. Xie, H.L., Fu, L. and Nie, X.D. (2013) Using Ensemble SVM to Identify Human GPCRs N-Linked Glycosylation Sites Based on the General Form of Chou's PseAAC. *Protein Engineering, Design and Selection*, **26**, 735-742. <https://doi.org/10.1093/protein/gzt042>
  59. Pacharawongsakda, E. and Theeramunkong, T. (2013) Predict Subcellular Locations of Singleplex and Multiplex Proteins by Semi-Supervised Learning and Dimension-Reducing General Mode of Chou's PseAAC. *IEEE Transactions on Nanobioscience*, **12**, 311-320. <https://doi.org/10.1109/TNB.2013.2272014>
  60. Fan, G.-L., Li, Q.-Z. and Zuo, Y.-C. (2013) Predicting Acidic and Alkaline Enzymes by Incorporating the Average Chemical Shift and Gene Ontology Informations into the General Form of Chou's PseAAC. *Process Biochemistry*, **48**, 1048-1053. <https://doi.org/10.1016/j.procbio.2013.05.012>
  61. Chang, T.H., Wu, L.C., Lee, T.Y., Chen, S.P., Huang, H.D. and Horng, J.T. (2013) EuLoc: A Web-Server for Accurately Predict Protein Subcellular Localization in Eukaryotes by Incorporating Various Features of Sequence Segments into the General Form of Chou's PseAAC. *Journal of Computer-Aided Molecular Design*, **27**, 91-103. <https://doi.org/10.1007/s10822-012-9628-0>
  62. Sun, X.Y., Shi, S.P., Qiu, J.D., Suo, S.B., Huang, S.Y. and Liang, R.P. (2012) Identifying Protein Quaternary Structural Attributes by Incorporating Physicochemical Properties into the General Form of Chou's PseAAC via Discrete Wavelet Transform. *Molecular BioSystems*, **8**, 3178-3184. <https://doi.org/10.1039/c2mb25280e>
  63. Qin, Y.F., Wang, C.H., Yu, X.Q., Zhu, J., Liu, T.G. and Zheng, X.Q. (2012) Predicting Protein Structural Class by Incorporating Patterns of Over-Represented k-Mers into the General Form of Chou's PseAAC. *Protein & Peptide Letters*, **19**, 388-397. <https://doi.org/10.2174/092986612799789350>
  64. Liao, B., Xiang, Q. and Li, D. (2012) Incorporating Secondary Features into the General form of Chou's PseAAC for Predicting Protein Structural Class. *Protein & Peptide Letters*, **19**, 1133-1138. <https://doi.org/10.2174/092986612803217051>
  65. Hayat, M. and Khan, A. (2012) Discriminating Outer Membrane Proteins with Fuzzy K-Nearest Neighbor Algorithms Based on the General Form of Chou's PseAAC. *Protein & Peptide Letters*, **19**, 411-421. <https://doi.org/10.2174/092986612799789387>
  66. Du, P., Wang, X., Xu, C. and Gao, Y. (2012) PseAAC-Builder: A Cross-Platform Stand-Alone Program for Generating Various Special Chou's Pseudo Amino Acid Compositions. *Analytical Biochemistry*, **425**, 117-119. <https://doi.org/10.1016/j.ab.2012.03.015>
  67. Zhang, S.W., Chen, W., Yang, F. and Pan, Q. (2008) Using Chou's Pseudo Amino Acid Composition to Predict Protein Quaternary Structure: A Sequence-Segmented PseAAC Approach. *Amino Acids*, **35**, 591-598. <https://doi.org/10.1007/s00726-008-0086-x>
  68. Shen, H.B. and Chou, K.C. (2008) PseAAC: A Flexible Web-Server for Generating Various Kinds of Protein Pseudo Amino Acid Composition. *Analytical Biochemistry*, **373**, 386-388. <https://doi.org/10.1016/j.ab.2007.10.012>
  69. Cao, D.S., Xu, Q.S. and Liang, Y.Z. (2013) Propy: A Tool to Generate Various Modes of Chou's PseAAC. *Bioinformatics*, **29**, 960-962. <https://doi.org/10.1093/bioinformatics/btt072>

70. Chou, K.C. (2009) Pseudo Amino Acid Composition and Its Applications in Bioinformatics, Proteomics and System Biology. *Current Proteomics*, **6**, 262-274. <https://doi.org/10.2174/157016409789973707>
71. Chou, K.C. (2011) Some Remarks on Protein Attribute Prediction and Pseudo Amino Acid Composition (50th Anniversary Year Review, 5-Steps Rule). *Journal of Theoretical Biology*, **273**, 236-247. <https://doi.org/10.1016/j.jtbi.2010.12.024>
72. Chen, W., Lei, T.Y., Jin, D.C., Lin, H. and Chou, K.C. (2014) PseKNC: A Flexible Web-Server for Generating Pseudo K-Tuple Nucleotide Composition. *Analytical Biochemistry*, **456**, 53-60. <https://doi.org/10.1016/j.ab.2014.04.001>
73. 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)*, **193**, Article ID: 103811. <https://doi.org/10.1016/j.chemolab.2019.103811>
74. Liu, B., Liu, F., Wang, X., Chen, J., Fang, L. and Chou, K.C. (2015) Pse-in-One: A Web Server for Generating Various Modes of Pseudo Components of DNA, RNA, and Protein Sequences. *Nucleic Acids Research*, **43**, W65-W71. <https://doi.org/10.1093/nar/gkv458>
75. Liu, B., Wu, H. and Chou, K.C. (2017) Pse-in-One 2.0: An Improved Package of Web Servers for Generating Various Modes of Pseudo Components of DNA, RNA, and Protein Sequences. *Natural Science*, **9**, 67-91. <https://doi.org/10.4236/ns.2017.94007>
76. Voet, D., Voet, J.G. and Pratt, C.W. (2002) *Fundamentals of Biochemistry*. John Wiley & Sons, New York.
77. Chou, K.C. and Chen, N.Y. (1977) The Biological Functions of Low-Frequency Phonons. *Scientia Sinica*, **20**, 447-457.
78. Du, Q.S., Wang, S.Q., Wei, D.Q., Zhu, Y., Guo, H., Sirois, S. and Chou, K.C. (2004) Polyprotein Cleavage Mechanism of SARS CoV Mpro and Chemical Modification of Octapeptide. *Peptides*, **25**, 1857-1864. <https://doi.org/10.1016/j.peptides.2004.06.018>
79. Chou, K.C. (1996) Review: Prediction of Human Immunodeficiency Virus Protease Cleavage Sites in Proteins. *Analytical Biochemistry*, **233**, 1-14. <https://doi.org/10.1006/abio.1996.0001>
80. Liang, G.Z. and Li, S.Z. (2007) A New Sequence Representation as Applied in Better Specificity Elucidation for Human Immunodeficiency Virus Type 1 Protease. *Biopolymers*, **88**, 401-412. <https://doi.org/10.1002/bip.20669>
81. Shen, H.B. and Chou, K.C. (2008) HIVcleave: A Web-Server for Predicting HIV Protease Cleavage Sites in Proteins. *Analytical Biochemistry*, **375**, 388-390. <https://doi.org/10.1016/j.ab.2008.01.012>
82. Chou, K.C., Jiang, S.P., Liu, W.M. and Fee, C.H. (1979) Graph Theory of Enzyme Kinetics: 1. Steady-State Reaction System. *Scientia Sinica*, **22**, 341-358.
83. Chou, K.C., Zhang, C.T. and Maggiora, G.M. (1997) Disposition of Amphiphilic Helices in Heteropolar Environments. *Proteins: Structure, Function and Genetics*, **28**, 99-108. [https://doi.org/10.1002/\(SICI\)1097-0134\(199705\)28:1<99::AID-PROT10>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1097-0134(199705)28:1<99::AID-PROT10>3.0.CO;2-C)
84. Chou, K.C., Lin, W.Z. and Xiao, X. (2011) Wenxiang: A Web-Server for Drawing Wenxiang Diagrams. *Natural Science*, **3**, 862-865. <https://doi.org/10.4236/ns.2011.310111>
85. Zhou, G.P. (2011) The Disposition of the LZCC Protein Residues in Wenxiang Diagram Provides New Insights into the Protein-Protein Interaction Mechanism. *Journal of Theoretical Biology*, **284**, 142-148. <https://doi.org/10.1016/j.jtbi.2011.06.006>
86. Zhou, G.P., Chen, D., Liao, S. and Huang, R.B. (2016) Recent Progresses in Studying Helix-Helix Interactions in Proteins by Incorporating the Wenxiang Diagram into the NMR Spectroscopy. *Current Topics in Medicinal Chemistry*, **16**, 581-590. <https://doi.org/10.2174/1568026615666150819104617>

87. Zhong, W.Z., Lalovic, B. and Ahan, J. (2009) Characterization of *in Vitro* and *in Vivo* Metabolism of AG-024322, a Novel Cyclin-Dependent Kinase (CDK) Inhibitor. *Health*, **1**, 249-262. <https://doi.org/10.4236/health.2009.14041>
88. Andrade, M.A., O'Donoghue, S.I. and Rost, B. (1998) Adaptation of Protein Surfaces to Subcellular Location. *Journal of Molecular Biology*, **276**, 517-525. <https://doi.org/10.1006/jmbi.1997.1498>
89. Reinhardt, A. and Hubbard, T. (1998) Using Neural Networks for Prediction of the Subcellular Location of Proteins. *Nucleic Acids Research*, **26**, 2230-2236. <https://doi.org/10.1093/nar/26.9.2230>
90. Yuan, Z. (1999) Prediction of Protein Subcellular Locations Using Markov Chain Models. *FEBS Letters*, **451**, 23-26. [https://doi.org/10.1016/S0014-5793\(99\)00506-2](https://doi.org/10.1016/S0014-5793(99)00506-2)
91. Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000) Predicting Subcellular Localization of Proteins Based on Their N-Terminal Amino Acid Sequence. *Journal of Molecular Biology*, **300**, 1005-1016. <https://doi.org/10.1006/jmbi.2000.3903>
92. Nakai, K. (2000) Protein Sorting Signals and Prediction of Subcellular Localization. *Advances in Protein Chemistry*, **54**, 277-344. [https://doi.org/10.1016/S0065-3233\(00\)54009-1](https://doi.org/10.1016/S0065-3233(00)54009-1)
93. Chou, K.C. and Shen, H.B. (2007) Recent Progresses in Protein Subcellular Location Prediction. *Analytical Biochemistry*, **370**, 1-16. <https://doi.org/10.1016/j.ab.2007.07.006>
94. Chou, K.C. and Shen, H.B. (2008) Cell-PLoc: A Package of Web Servers for Predicting Subcellular Localization of Proteins in Various Organisms. *Nature Protocols*, **3**, 153-162. <https://doi.org/10.1038/nprot.2007.494>
95. Zhong, W.Z. and Zhou, S.F. (2014) Molecular Science for Drug Development and Biomedicine. *International Journal of Molecular Sciences*, **15**, 20072-20078. <https://doi.org/10.3390/ijms151120072>
96. Chou, K.C. (2019) Advance in Predicting Subcellular Localization of Multi-Label Proteins and Its Implication for Developing Multi-Target Drugs. *Current Medicinal Chemistry*, **26**, 4918-4943. <http://www.eurekaselect.com/172010/article>  
<https://doi.org/10.2174/0929867326666190507082559>
97. Chou, K.C. (2019) Two Kinds of Metrics for Computational Biology. *Genomics*. <https://www.sciencedirect.com/science/article/pii/S0888754319304604?via%3Dihub>
98. Chou, K.C. (2020) The pLoc\_bal-mVirus Is a Powerful Artificial Intelligence Tool for Predicting the Subcellular Localization of Virus Proteins According to Their Sequence Information Alone. *Journal of Genetics and Genomics*, **4**.
99. 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*, **20**, 220-230. <https://doi.org/10.2174/1566524019666191011160543>
100. 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>
101. Dutta, A., Dalmia, A.A.R., 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.combiomed.2019.103558>
102. 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*. <https://doi.org/10.1007/s10989-019-09887-3>  
<https://link.springer.com/article/10.1007%2Fs10989-019-09887-3>
103. Lan, J., Liu, J., Liao, C., Merkle, D.J., Han, Q. and Li, J. (2019) A Study for Therapeutic Treatment against Par-



kinson's Disease via Chou's 5-Steps Rule. *Current Topics in Medicinal Chemistry*, **19**, 2318-2333.

<http://www.eurekaselect.com/175887/article>

<https://doi.org/10.2174/1568026619666191019111528>

104. 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>
105. 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**, 3769-3775. <https://doi.org/10.2174/1381612825666190930101258>
106. 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>
107. 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*. <https://doi.org/10.1109/TCBB.2020.2968441>
108. 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>
109. 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>
110. Dutta, A., Dalmia, A.A.R., 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>
111. 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*. (In Press) <https://doi.org/10.1109/JBHI.2019.2958042>
112. Zhai, X., Chen, M. and Lu, W. (2018) Accelerated Search for Perovskite Materials with Higher Curie Temperature Based on the Machine Learning Methods. *Computational Materials Science*, **151**, 41-48. <https://doi.org/10.1016/j.commatsci.2018.04.031>
113. Chou, K.C. and Forsen, S. (1980) Diffusion-Controlled Effects in Reversible Enzymatic Fast Reaction System: Critical Spherical Shell and Proximity Rate Constants. *Biophysical Chemistry*, **12**, 255-263. [https://doi.org/10.1016/0301-4622\(80\)80002-0](https://doi.org/10.1016/0301-4622(80)80002-0)
114. Chou, K.C. and Forsen, S. (1980) Graphical Rules for Enzyme-Catalyzed Rate Laws. *Biochemical Journal*, **187**, 829-835. <https://doi.org/10.1042/bj1870829>
115. 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>
116. Chou, K.C. (2019) Showcase to Illustrate How the Web-Server iDNA6mA-PseKNC Is Working. *Journal of Pathology Research Reviews & Reports*, **1**, 1-15.
117. Chou, K.C. (2019) The pLoc\_bal-mPlant Is a Powerful Artificial Intelligence Tool for Predicting the Subcellular

Localization of Plant Proteins Purely Based on Their Sequence Information. *International Journal of Nutrition Sciences*, **4**, 1-4. <https://doi.org/10.15761/JSCRM.1000138>

118. Chou, K.C., Cheng, X. and Xiao, X. (2019) pLoc\_bal-mEuk: Predict Subcellular Localization of Eukaryotic Proteins by General PseAAC and Quasi-Balancing Training Dataset. *Medicinal Chemistry*, **15**, 472-485. <https://doi.org/10.2174/1573406415666181218102517>
119. Chou, K.C. (2019) Showcase to Illustrate How the Web-Server iNitro-Tyr Is Working. *Glo J of Com Sci and Infor Tec.*, **2**, 1-16.s
120. Chou, K.C. (2020) Using Similarity Software to Evaluate Scientific Paper Quality Is a Big Mistake. *Natural Science*, **12**, 42-58 <https://doi.org/10.4236/ns.2020.123008>
121. Chou, K.C. (2019) The pLoc\_bal-mAnimal Is a Powerful Artificial Intelligence Tool for Predicting the Subcellular Localization of Animal Proteins Based on Their Sequence Information Alone. *Scientific Journal of Biometrics & Biostatistics*, **2**, 1-13.
122. Chou, K.C. (2020) Showcase to Illustrate How the Webserver pLoc\_bal-mEuk Is Working. *Biomedical Journal of Scientific & Technical Research*. <https://doi.org/10.18483/ijSci.2247>
123. Chou, K.C. (2020) The pLoc\_bal-mGneg Predictor Is a Powerful Web-Server for Identifying the Subcellular Localization of Gram-Negative Bacterial Proteins Based on Their Sequences Information Alone. *ijSci*, **9**, 27-34. <https://doi.org/10.18483/ijSci.2248>
124. Chou, K.C. (2020) How the Artificial Intelligence Tool iRNA-2methyl Is Working for RNA 2'-Omethylation Sites. *Journal of Medical Care Research and Review*, **3**, 348-366.
125. Chou, K.-C. (2020) Showcase to Illustrate How the Web-Server iKcr-PseEns Is Working. *Journal of Medical Care Research and Review*, **3**, 331-347. <https://doi.org/10.18483/ijSci.2247>
126. Chou, K.C. (2019) How the Artificial Intelligence Tool iSNO-PseAAC Is Working in Predicting the Cysteine S-Nitrosylation Sites in Proteins. *Journal of Stem Cells and Regenerative Medicine*, **4**, 1-9. <https://doi.org/10.15761/JSCRM.1000137>
127. Chou, K.C. (2020) Showcase to Illustrate How the Web-Server iRNA-Methyl Is Working. *Journal of Molecular Genetics*, **3**, 1-7.
128. Chou, K.C. (2020) How the Artificial Intelligence Tool iRNA-PseU Is Working in Predicting the RNA Pseudouridine Sites. *Biomedical Journal of Scientific & Technical Research*. <https://doi.org/10.15761/JSCRM.1000137>
129. Chou, K.C. (2020) Showcase to Illustrate How the Web-Server iSNO-AAPair Is Working. *Journal of Genetics & Genomics*, **4**. <https://doi.org/10.18483/ijSci.2247>
130. Chou, K.C. (2020) The pLoc\_bal-mHum Is a Powerful Web-Serve for Predicting the Subcellular Localization of Human Proteins Purely Based on Their Sequence Information. *Advances in Bioengineering and Biomedical Science Research*, **3**, 1-5.
131. Chou, K.C. (2020) Showcase to Illustrate How the Web-Server iPTM-mLys Is Working. *Infotext Journal of Infectious Diseases and Therapy*, **1**, 1-16.
132. Chou, K.C. (2020) The pLoc\_bal-mGpos Is a Powerful Artificial Intelligence Tool for Predicting the Subcellular Localization of Gram-Positive Bacterial Proteins According to Their Sequence Information Alone. *Glo J of Com Sci and Infor Tec*, **2**, 1-13.
133. Chou, K.C. (2020) Showcase to Illustrate How the Web-Server iPreny-PseAAC Is Working. *Glo J of Com Sci and Infor Tec.*, **2**, 1-15.
134. Chou, K.C. (2020) Some Illuminating Remarks on Molecular Genetics and Genomics, as Well as Drug Development. *Molecular Genetics and Genomics*, **295**, 261-274. <https://doi.org/10.1007/s00438-019-01634-z>

135. Chou, K.C. (2020) The Problem of Elsevier Series Journals Online Submission by Using Artificial Intelligence. *Natural Science*, **12**, 37-38. <https://doi.org/10.4236/ns.2020.122006>
136. Chou, K.C. (2020) The Most Important Ethical Concerns in Science. *Natural Science*, **12**, 35-36. <https://doi.org/10.4236/ns.2020.122005>