

Pse-in-One 2.0: An Improved Package of Web Servers for Generating Various Modes of Pseudo Components of DNA, RNA, and Protein Sequences

Bin Liu^{1,2,3}, Hao Wu¹, Kuo-Chen Chou^{3,4,5}

¹School of Computer Science and Technology, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China; ²Key Laboratory of Network Oriented Intelligent Computation, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China; ³Gordon Life Science Institute, Boston, Massachusetts, USA; ⁴Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China; ⁵Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University, Jeddah, KSA

Correspondence to: Bin Liu, bliu@hit.edu.cn; Kuo-Chen Chou, kcchou@gordonlifescience.org

Keywords: Pseudo Components, DNA Sequences, RNA Sequences, Protein Sequences

Received: April 10, 2017

Accepted: April 25, 2017

Published: April 28, 2017

Copyright © 2017 by authors 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

Pse-in-One 2.0 is a package of web-servers evolved from Pse-in-One (Liu, B., Liu, F., Wang, X., Chen, J. Fang, L. & Chou, K.C. Nucleic Acids Research, 2015, 43:W65-W71). In order to make it more flexible and comprehensive as suggested by many users, the updated package has incorporated 23 new pseudo component modes as well as a series of new feature analysis approaches. It is available at <http://bioinformatics.hitsz.edu.cn/Pse-in-One2.0/>. Moreover, to maximize the convenience of users, provided is also the stand-alone version called “Pse-in-One-Analysis”, by which users can significantly speed up the analysis of massive sequences.

1. INTRODUCTION

With the avalanche of biological sequences generated in the post-genomic age, one of the most challenging problems in computational biology today is how to effectively formulate the sequence of a biological sample (such as DNA, RNA or protein) with a discrete model or a vector that can effectively reflect its sequence pattern information or capture its key features concerned. This is because almost all the existing machine-learning algorithms, such as “Neural Network” or NN algorithm [1-3] “Support Vector Machine” or SVM algorithm [4-12] “Nearest Neighbor” or NN algorithm [13, 14] and “Random Forest” algorithm [15-22] can only handle vectors but not sequence samples as elucidated in a review paper [23]. Unfortunately, if using the sequential model, *i.e.*, the model in which all the samples are represented by their original sequences, it is hardly able to train a machine learning model that can cover all the possible cases concerned, as elaborated in [24].

To avoid completely losing the sequence-order information for proteins, the quasi-sequence-order

approach [25] or PseAAC (pseudo amino acid composition) approach [26-28] or Chou's PseAAC [29-32] was proposed. Ever since the concept of Chou's PseAAC [29-31] was proposed, it has rapidly penetrated into many biomedicine and drug development areas [33, 34] and nearly all the areas of computational proteomics [35-212].

Because it has been widely and increasingly used, and also because it would be a trend and future direction to establish user-friendly and publically accessible web-servers for various analysis methods as pointed out in [213], four powerful web-servers were established; they are "PseAAC" [214], "PseAAC-Builder" [29], "propy" [30] and "PseAAC-General" [32]. The former three are for generating various modes of Chou's special PseAAC; while the 4th one for those of Chou's general PseAAC [28, 108] 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 [108], "Gene Ontology" mode (see Eqs.11-12 of 108), and "Sequential Evolution" or "PSSM" mode (see Eqs.13-14 of [108]).

Encouraged by the successes of using PseAAC to deal with protein/peptide sequences, the concept of PseAAC has been extended to cover DNA/RNA sequences as well via the PseKNC (Pseudo K-tuple Nucleotide Composition) approach [215-223]. Meanwhile, four publically accessible web-servers [215, 217, 224, 225] were developed for generating various pseudo components or feature vectors for DNA/RNA sequences as well.

Particularly, recently a very powerful web-server called Pse-in-One [226] has been established that can be used to generate any desired pseudo components or feature vectors for protein/peptide and DNA/RNA sequences according to the need of users' studies.

Since then some novel pseudo component modes have been proposed for dealing with various problems in proteomics and genome analysis [7, 10-12, 18-22, 221, 227-263]. In order to incorporate these new and important developments into the Pse-in-One package, an updated version called "Pse-in-One 2.0" has been established.

2. RESULTS AND DISCUSSION

Compared with the original one, the updated version has the following new features and functions.

2.1. Modes of Pseudo Components

Added in are a total of 23 new pseudo component modes, of which 6 for DNA sequences (Table 1), 8 for RNA sequences (Table 2), and 9 for protein sequences (Table 3). These new modes reflect the recent developments of the pseudo components, particularly in extending the coverage scope to those features derived from (1) RNA secondary structures, and (2) the multiple sequence alignments and profiles. As a consequence, **Pse-in-One 2.0** covers a total of 51 different features, of which 20 for DNA sequences, 14 for RNA sequences, and 17 for protein sequences. The overall structure can be reflected via the following three sub web-servers.

PseDAC-General is for generating the feature vectors of DNA sequences. It contains three categories: nucleotide composition, nucleotide autocorrelation, and pseudo nucleotide composition. Of the 6 new modes, 3 are added into the first category, including IDKmer [224], Mismatch [264], and Subsequence[265]; while the other 3 are added to the second category, including Moran autocorrelation, Geary autocorrelation, and Normalized Moreau-Broto autocorrelation [217].

PseRAC-General is aimed to generate the feature vectors for RNA sequences, and it has four categories, of which the "predicted structure composition" is a newly added category for extracting the structure-based features of RNA sequences, in which the following 3 new modes are incorporated: Triplet [266], PseSSC [267] and PseDPC [10]. Triplet is an early approach to use the structure information of RNA sequences and has shown better performance for microRNA identification in comparison with other sequence-based approaches. PseSSC and PseDPC can be used to incorporate the global or long-range structure-order information so as to remarkably improve the prediction quality in identifying the pre-miRNAs. Of the other 5 new modes, 2 are added into the nucleic acid composition category, *i.e.*,

Table 1. List of the 6 new modes for DNA sequences.

Category	Mode
Nucleic acid composition	1) increment of diversity (IDKmer) [226, 270, 291]
	2) The occurrences of kmers, allowing at most m mismatches (Mismatch) [264, 265, 292]
	3) The occurrences of kmers, allowing non-contiguous matches (Subsequence) [265, 292, 293]
Autocorrelation	4) Moran autocorrelation (MAC) [268, 294]
	5) Geary autocorrelation (GAC) [217, 295]
	6) Normalized Moreau-Broto autocorrelation (NMBAC) [217, 296]

Table 2. List of the 8 new modes for RNA sequences.

Category	Mode
Nucleic acid composition	1) The occurrences of kmers, allowing at most m mismatches (Mismatch) [264, 265, 292]
	2) The occurrences of kmers, allowing non-contiguous matches (Subsequence) [265, 292, 293]
Autocorrelation	3) Moran autocorrelation (MAC) [217, 294]
	4) Geary autocorrelation (GAC) [217, 295]
	5) Normalized Moreau-Broto autocorrelation (NMBAC) [217, 296]
Predicted structure composition	6) Local structure-sequence triplet element (Triplet) [266]
	7) Pseudo-structure status composition (PseSSC) [226]
	8) Pseudo-distance structure status pair composition (PseDPC) [10]

Table 3. List of the 8 new modes for protein sequences.

Category	Mode
Amino acid composition	1) Distance-based Residue (DR) [271]
	2) PseAAC of Distance-Pairs and Reduced Alphabet (Distance Pair) [271]
Autocorrelation	3) Physicochemical distance transformation (PDT) [270]
	4) Select and combine the n most frequent amino acids according to their frequencies (Top-n-gram) [269]
	5) Profile-based Physicochemical distance transformation (PDT-Pofile) [270]
Profile-based features	6) Distance-based Top-n-gram (DT) [271]
	7) Profile-based Auto covariance (AC-PSSM) [272]
	8) Profile-based Cross covariance (CC-PSSM) [272]
	9) Profile-based Auto-cross covariance (ACC-PSSM) [272]

Mismatch [264] and Subsequence [265]; and 3 are added into the autocorrelation category, *i.e.*, Moran autocorrelation, Geary autocorrelation, and Normalized Moreau-Broto autocorrelation [268].

PseAAC-General is designed to generate the feature vectors for protein sequences. For this sub web-server, we have created a special category called “profile-based” category, into which 6 new modes are added; they are “top-n-gram” [269], “PDT-Profile” [270], “DT” [271], “AC-PSSM”, “CC-PSSM” and “ACC-PSSM” [272]. The top-n-gram combines the n most frequent amino acids in each amino acid frequency profile; PDT-Profile is the abbreviation for “Profile-based physicochemical distance transformation” and it is similar to PDT except that PDT-Profile extracts the evolutionary information from the frequency profile; DT is the abbreviation for “distance-based Top-n-gram” and this method extends Top-n-gram by considering the distances between Top-n-gram pairs; AC-PSSM, CC-PSSM and ACC-PSSM incorporate the position-specific score matrix (PSSM) into the methods of AC, CC and ACC [272, 273]. These profile-based methods can significantly improve the protein remote homology detection [7, 8], protein fold recognition and so forth. Moreover, added into the amino acid composition category are 3 new modes: they are “DR” [274], “Distance Pair” [271], and “PDT” [270]. DR is the abbreviation for “Distance-based Residue”. It is sequence-based method, in which the generated feature vector for protein sequence is based on the distance between residue pairs and has shown better performance for protein remote homology detection. “Distance Pair” method incorporates the amino acid distance pair coupling information and the amino acid reduced alphabet profile into the general pseudo amino acid composition (PseAAC) [108] vector, which is very useful for analysing DNA-binding proteins [15, 170, 189, 275]. PDT is the abbreviation for “physicochemical distance transformation”, which can incorporate considerable sequence-order information or important patterns of protein/peptide sequences into Pseudo components [28], which is very useful for conducting various proteome analyses [17, 23, 215-217, 224, 225, 231, 235, 276-289] and genome analysis as well [216, 218, 220, 223, 229, 255, 277, 290].

For more information about the three sub-webservers, see [Supporting information S1](#).

2.2. New Facility

Added into the updated version is also a new facility called “**Pse-in-One-Analysis**”, by which the feature vectors for the input DNA, RNA, or protein sequences can be automatically generated according to the selected modes and parameters. And the results will be sent to the users via their e-mail addresses. The users can also see the result by revisiting the link concerned. Moreover, provided are also the feature vector visualization and the predicted RNA secondary structure visualization functions, which are very useful for the feature analysis and interpretation. See [Supporting Information S2](#) for detailed information in this regard.

The stand-alone version of **Pse-in-One 2.0** is available. Users can easily download it into their own computer for conducting high throughput analysis of massive biological sequences.

By means of the new facility **Pse-in-One-Analysis** or **Pse-Analysis** [254], all the tedious jobs in developing a predictor, such as selecting optimal features and parameters as well as evaluating anticipated prediction quality, can be automatically fulfilled by the computer as elaborated in [254]. It will save scientists a lot of time, one big step forward to realize the dream of using robots or computers to conduct genome/proteome analyses.

2.3. New Kits

Newly provided in **Pse-in-One 2.0** are also some useful kits, including automatic notification of results by e-mail, RNA secondary structure visualization, etc. Meanwhile, some bugs have been fixed to make the web-server work more smoothly and fully consistent.

A flowchart of Pse-in-One 2.0 is given in [Figure 1](#).

3. CONCLUSION

Evolved from the original **Pse-in-One** package [226], **Pse-in-One 2.0** is much more flexible and

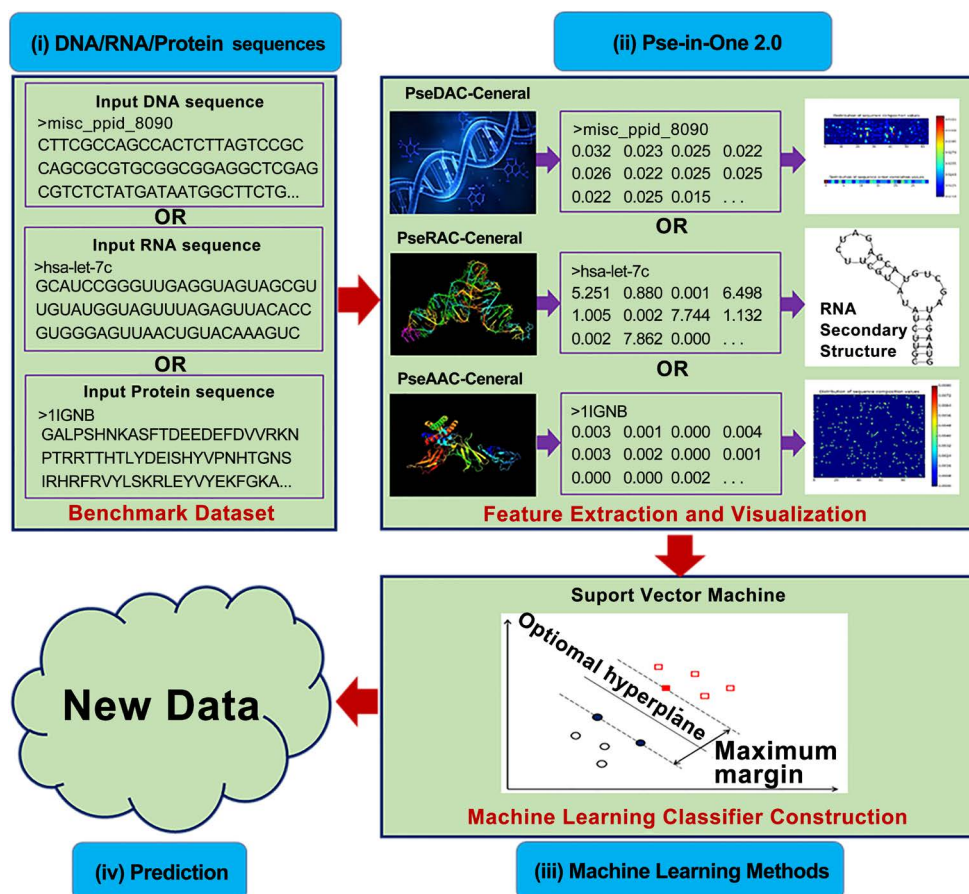


Figure 1. The flowchart of Pse-in-One 2.0. The first two steps are implemented in Pse-in-One 2.0 webserver. The last two steps are implemented in Pse-in-One-Analysis. The output of the webserver can be directly used as the input of Pse-in-One-Analysis package.

powerful than the former. In comparison with the 2015 version that has been widely used in bioinformatics and computational biology and biomedicine within a very short period of time, the new version is even more powerful for conducting various genome analyses and proteome analyses. Science is rapidly developing, particularly in life science. Once having new and important developments, the future version for the **Pse-in-One** series will be announced via a publication or web-page.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (No. 61672184), the Natural Science Foundation of Guangdong Province (2014A030313695), Guangdong Natural Science Funds for Distinguished Young Scholars (2016A030306008), and Scientific Research Foundation in Shenzhen (Grant No. JCYJ 20150626110425228).

SUPPORTING INFORMATION

Supporting Information S1. Pse-in-One 2.0 Description. This document describes totally 51 different modes, including 20 modes for DNA sequences, 14 modes for RNA sequences, and 17 modes for protein sequences. The document is available at http://bioinformatics.hitsz.edu.cn/Pse-in-One2.0/static/download/Pse-in-One%202.0_description.pdf.

Supporting Information S2. Pse-in-One 2.0 User's Manual. This document describes how to use **Pse-in-One-Analysis**, including installation, function description, commands and methods description. The document is available at http://bioinformatics.hitsz.edu.cn/Pse-in-One2.0/static/download/Pse-in-One%202.0_manual.pdf.

REFERENCES

1. Feng, K.Y., Cai, Y.D. and Chou, K.C. (2005) Boosting Classifier for Predicting Protein Domain Structural Class. *Biochemical & Biophysical Research Communications (BBRC)*, **334**, 213-217. <https://doi.org/10.1016/j.bbrc.2005.06.075>
2. Cai, Y.D. and Chou, K.C. (1999) Artificial Neural Network for Predicting Alpha-Turn Types. *Analytical Biochemistry*, **268**, 407-409. <https://doi.org/10.1006/abio.1998.2992>
3. Thompson, T.B., Chou, K.C. and Zheng, C. (1995) Neural Network Prediction of the HIV-1 Protease Cleavage Sites. *Journal of Theoretical Biology*, **177**, 369-379. <https://doi.org/10.1006/jtbi.1995.0254>
4. Feng, P.M., Chen, W., Lin, H. and Chou, K.C. (2013) iHSP-PseRAAAC: Identifying the Heat Shock Protein Families Using Pseudo Reduced Amino Acid Alphabet Composition. *Analytical Biochemistry*, **442**, 118-125. <https://doi.org/10.1016/j.ab.2013.05.024>
5. Chen, W., Feng, P.M., Lin, H. and Chou, K.C. (2013) iRSpot-PseDNC: Identify Recombination Spots with Pseudo Dinucleotide Composition. *Nucleic Acids Research*, **41**, e68. <https://doi.org/10.1093/nar/gks1450>
6. Xiao, X., Wang, P. and Chou, K.C. (2012) iNR-PhysChem: A Sequence-Based Predictor for Identifying Nuclear Receptors and Their Subfamilies via Physical-Chemical Property Matrix. *PLoS ONE*, **7**, e30869. <https://doi.org/10.1371/journal.pone.0030869>
7. Chen, J., Long, R., Wang, X.L., Liu, B., Chou, K.C. (2016) dRHP-PseRA: Detecting Remote Homology Proteins Using Profile-Based Pseudo Protein Sequence and Rank Aggregation. *Scientific Reports*, **6**, 32333. <https://doi.org/10.1038/srep32333>
8. Liu, B., Zhang, D., Xu, R., Xu, J., Wang, X., Chen, Q., Dong, Q. and Chou, K.C. (2014) Combining evolutionary Information Extracted from Frequency Profiles with Sequence-Based Kernels for Protein Remote Homology Detection. *Bioinformatics*, **30**, 472-479. <https://doi.org/10.1093/bioinformatics/btt709>
9. Liu, B., Fang, L., Wang, S., Wang, X., Li, H. and Chou, K.C. (2015) Identification of microRNA Precursor with the Degenerate K-Tuple or Kmer Strategy. *Journal of Theoretical Biology*, **385**, 153-159. <https://doi.org/10.1016/j.jtbi.2015.08.025>
10. Liu, B., Fang, L., Liu, F., Wang, X. and Chou, K.C. (2016) iMiRNA-PseDPC: microRNA Precursor Identification with a Pseudo Distance-Pair Composition Approach. *Journal of Biomolecular Structure and Dynamics (JBSD)*, **34**, 223-235. <https://doi.org/10.1080/07391102.2015.1014422>
11. Liu, B., Fang, L., Long, R., Lan, X. and Chou, K.C. (2016) iEnhancer-2L: A Two-Layer Predictor for Identifying Enhancers and Their Strength by Pseudo k-Tuple Nucleotide Composition. *Bioinformatics*, **32**, 362-369. <https://doi.org/10.1093/bioinformatics/btv604>
12. Liu, B., Wang, S., Long, R. and Chou, K.C. (2017) iRSpot-EL: Identify Recombination Spots with an Ensemble Learning Approach. *Bioinformatics*, **33**, 35-41. <https://doi.org/10.1093/bioinformatics/btw539>
13. Cai, Y.D. and Chou, K.C. (2004) Predicting Subcellular Localization of Proteins in a Hybridization Space. *Bioinformatics*, **20**, 1151-1156. <https://doi.org/10.1093/bioinformatics/bth054>
14. Chou, K.C. and Cai, Y.D. (2006) Prediction of Protease Types in a Hybridization Space. *Biochem Biophys Res Comm (BBRC)*, **339**, 1015-1020.
15. Lin, W.Z., Fang, J.A., Xiao, X. and Chou, K.C. (2011) iDNA-Prot: Identification of DNA Binding Proteins Using

Random Forest with Grey Model. *PLoS ONE*, **6**, e24756.

16. Kandaswamy, K.K., Chou, K.C., Martinetz, T., Moller, S., Suganthan, P.N., Sridharan, S. and Pugalenthi, G. (2011) AFP- Pred: A Random Forest Approach for Predicting Antifreeze Proteins from Sequence-Derived Properties. *Journal of Theoretical Biology*, **270**, 56-62. <https://doi.org/10.1016/j.jtbi.2010.10.037>
17. Jia, J., Liu, Z., Xiao, X. and Chou, K.C. (2015) iPPI-Esml: An Ensemble Classifier for Identifying the Interactions of Proteins by Incorporating Their Physicochemical Properties and Wavelet Transforms into PseAAC. *Journal of Theoretical Biology*, **377**, 47-56. <https://doi.org/10.1016/j.jtbi.2015.04.011>
18. Jia, J., Liu, Z., Xiao, X., Liu, B. and Chou, K.C. (2016) iSuc-PseOpt: Identifying Lysine Succinylation Sites in Proteins by Incorporating Sequence-Coupling Effects into Pseudo Components and Optimizing Imbalanced Training Dataset. *Analytical Biochemistry*, **497**, 48-56. <https://doi.org/10.1016/j.ab.2015.12.009>
19. Jia, J., Liu, Z., Xiao, X., Liu, B. and Chou, K.C. (2016) pSuc-Lys: Predict Lysine Succinylation Sites in Proteins with PseAAC and Ensemble Random Forest Approach. *Journal of Theoretical Biology*, **394**, 223-230. <https://doi.org/10.1016/j.jtbi.2016.01.020>
20. Jia, J., Liu, Z., Xiao, X., Liu, B. and Chou, K.C. (2016) iCar-PseCp: Identify Carbonylation Sites in Proteins by Monto Carlo Sampling and Incorporating Sequence Coupled Effects into General PseAAC. *Oncotarget*, **7**, 34558-34570.
21. Qiu, W.R., Sun, B.Q., Xiao, X., Xu, D. and Chou, K.C. (2016) iPhos-PseEvo: Identifying Human Phosphorylated Proteins by Incorporating Evolutionary Information into General PseAAC via Grey System Theory. *Molecular Informatics*. <https://doi.org/10.1002/minf.201600010>
22. Qiu, W.R., Sun, B.Q., Xiao, X., Xu, Z.C. and Chou, K.C. (2016) iPTM-mLys: Identifying Multiple Lysine PTM Sites and Their Different Types. *Bioinformatics*, **32**, 3116-3123. <https://doi.org/10.1093/bioinformatics/btw380>
23. Chou, K.C. (2015) Impacts of Bioinformatics to Medicinal Chemistry *Medicinal Chemistry*, **11**, 218-234. <https://doi.org/10.2174/1573406411666141229162834>
24. Chou, K.C. and Shen, H.B. (2007) Review: Recent Progresses in Protein Subcellular Location Prediction. *Analytical Biochemistry*, **370**, 1-16. <https://doi.org/10.1016/j.ab.2007.07.006>
25. Chou, K.C. (2000) Prediction of Protein Subcellular Locations by Incorporating Quasi-Sequence-Order Effect. *Biochem Biophys Res Comm (BBRC)*, **278**, 477-483.
26. Chou, K.C. (2001) Prediction of Protein Cellular Attributes Using Pseudo Amino Acid Composition. *PROTEINS: Structure, Function, and Genetics (Erratum: ibid.)*, **43**, 246-255.
27. 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>
28. 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>
29. 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>
30. 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.
31. Lin, S.X. and Lapointe, J. (2013) Theoretical and Experimental Biology in One—A Symposium in Honour of Professor Kuo-Chen Chou's 50th Anniversary and Professor Richard Giegé's 40th Anniversary of Their Scientific Careers. *J. Biomedical Science and Engineering (JBISE)*, **6**, 435-442. <https://doi.org/10.4236/jbise.2013.64054>

32. 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>
33. Zhong, W.Z. and Zhou, S.F. (2014) Molecular Science for Drug Development and Biomedicine. *International Journal of Molecular Sciences*, **15**, 20072-20078.
34. Zhou, G.P., Zhong, W.Z. (2016) Perspectives in Medicinal Chemistry. *Current Topics in Medicinal Chemistry*, **16**, 381-382. <https://doi.org/10.2174/156802661604151014114030>
35. Guo, Z.M. (2002) Prediction of Membrane Protein Types by Using Pattern Recognition Method Based on Pseudo Amino Acid Composition. Master Thesis, Bio-X Life Science Research Center, Shanghai Jiaotong University, Shanghai.
36. Chou, K.C. and Cai, Y.D. (2003) Predicting Protein Quaternary Structure by Pseudo Amino Acid Composition. *Proteins: Structure, Function, and Genetics*, **53**, 282-289. <https://doi.org/10.1002/prot.10500>
37. Chou, K.C. and Cai, Y.D. (2003) Prediction and Classification of Protein Subcellular Location: Sequence-Order Effect and Pseudo Amino Acid Composition. *Journal of Cellular Biochemistry (Addendum, ibid. 2004, 91, 1085)*, **90**, 1250-1260.
38. Pan, Y.X., Zhang, Z.Z., Guo, Z.M., Feng, G.Y., Huang, Z.D. and He, L. (2003) Application of Pseudo Amino Acid Composition for Predicting Protein Subcellular Location: Stochastic Signal Processing Approach. *Journal of Protein Chemistry*, **22**, 395-402. <https://doi.org/10.1023/A:1025350409648>
39. Chou, K.C. and Cai, Y.D. (2004) Predicting Subcellular Localization of Proteins by Hybridizing Functional Domain Composition and Pseudo Amino Acid Composition. *Journal of Cellular Biochemistry*, **91**, 1197-1203. <https://doi.org/10.1002/jcb.10790>
40. Wang, M., Yang, J., Liu, G.P., Xu, Z.J. and Chou, K.C. (2004) Weighted-Support Vector Machines for Predicting Membrane Protein Types Based on Pseudo Amino Acid Composition. *Protein Engineering, Design, and Selection*, **17**, 509-516. <https://doi.org/10.1093/protein/gzh061>
41. Cai, Y.D. and Chou, K.C. (2005) Predicting Enzyme Subclass by Functional Domain Composition and Pseudo Amino Acid Composition. *Journal of Proteome Research*, **4**, 967-971. <https://doi.org/10.1021/pr0500399>
42. Cai, Y.D., Zhou, G.P. and Chou, K.C. (2005) Predicting Enzyme Family Classes by Hybridizing Gene Product Composition and Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **234**, 145-149. <https://doi.org/10.1016/j.jtbi.2004.11.017>
43. Gao, Y., Shao, S.H., Xiao, X., Ding, Y.S., Huang, Y.S., Huang, Z.D. and Chou, K.C. (2005) Using Pseudo Amino Acid Composition to Predict Protein Subcellular Location: Approached with Lyapunov Index, Bessel Function, and Chebyshev Filter. *Amino Acids*, **28**, 373-376. <https://doi.org/10.1007/s00726-005-0206-9>
44. Liu, H., Yang, J., Wang, M., Xue, L. and Chou, K.C. (2005) Using Fourier Spectrum Analysis and Pseudo Amino Acid Composition for Prediction of Membrane Protein Types. *The Protein Journal*, **24**, 385-389. <https://doi.org/10.1007/s10930-005-7592-4>
45. Shen, H.B. and Chou, K.C. (2005) Using Optimized Evidence-Theoretic K-Nearest Neighbor Classifier and Pseudo Amino Acid Composition to Predict Membrane Protein Types. *Biochemical & Biophysical Research Communications (BBRC)*, **334**, 288-292.
46. Shen, H.B. and Chou, K.C. (2005) Predicting Protein Subnuclear Location with Optimized Evidence-Theoretic K-Nearest Classifier and Pseudo Amino Acid Composition. *Biochemical and Biophysical Research Communications (BBRC)*, **337**, 752-756.
47. Cai, Y.D. and Chou, K.C. (2006) Predicting Membrane Protein Type by Functional Domain Composition and Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **238**, 395-400.

<https://doi.org/10.1016/j.jtbi.2005.05.035>

48. Chen, C., Tian, Y.X., Zou, X.Y., Cai, P.X. and Mo, J.Y. (2006) Using Pseudo Amino Acid Composition and Support Vector Machine to Predict Protein Structural Class. *Journal of Theoretical Biology*, **243**, 444-448. <https://doi.org/10.1016/j.jtbi.2006.06.025>
49. Chen, C., Zhou, X., Tian, Y., Zou, X. and Cai, P. (2006) Predicting Protein Structural Class with Pseudo Amino Acid Composition and Support Vector Machine Fusion Network. *Analytical Biochemistry*, **357**, 116-121. <https://doi.org/10.1016/j.ab.2006.07.022>
50. Du, P. and Li, Y. (2006) Prediction of Protein Submitochondria Locations by Hybridizing Pseudo Amino Acid Composition with Various Physicochemical Features of Segmented Sequence. *BMC Bioinformatics*, **7**, 518. <https://doi.org/10.1186/1471-2105-7-518>
51. Mondal, S., Bhavna, R., Mohan Babu, R. and Ramakumar, S. (2006) Pseudo Amino Acid Composition and Multi-Class Support Vector Machines Approach for Conotoxin Superfamily Classification. *Journal of Theoretical Biology*, **243**, 252-260. <https://doi.org/10.1016/j.jtbi.2006.06.014>
52. Shen, H.B., Yang, J. and Chou, K.C. (2006) Fuzzy KNN for Predicting Membrane Protein Types from Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **240**, 9-13. <https://doi.org/10.1016/j.jtbi.2005.08.016>
53. Wang, S.Q., Yang, J. and Chou, K.C. (2006) Using Stacked Generalization to Predict Membrane Protein Types Based on Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **242**, 941-946. <https://doi.org/10.1016/j.jtbi.2006.05.006>
54. Xiao, X., Shao, S.H., Ding, Y.S., Huang, Z.D. and Chou, K.C. (2006) Using Cellular Automata Images and Pseudo Amino Acid Composition to Predict Protein Subcellular Location. *Amino Acids*, **30**, 49-54. <https://doi.org/10.1007/s00726-005-0225-6>
55. Xiao, X., Shao, S.H., Huang, Z.D. and Chou, K.C. (2006) Using Pseudo Amino Acid Composition to Predict Protein Structural Classes: Approached with Complexity Measure Factor. *Journal of Computational Chemistry*, **27**, 478-482. <https://doi.org/10.1002/jcc.20354>
56. Zhang, S.W., Pan, Q., Zhang, H.C., Shao, Z.C. and Shi, J.Y. (2006) Prediction Protein Homo-Oligomer Types by Pseudo Amino Acid Composition: Approached with an Improved Feature Extraction and Naive Bayes Feature Fusion. *Amino Acids*, **30**, 461-468.
57. Zhou, G.P. and Cai, Y.D. (2006) Predicting Protease Types by Hybridizing Gene Ontology and Pseudo Amino Acid Composition. *PROTEINS: Structure, Function, and Bioinformatics*, **63**, 681-684. <https://doi.org/10.1002/prot.20898>
58. Chen, Y.L. and Li, Q.Z. (2007) Prediction of Apoptosis Protein Subcellular Location Using Improved Hybrid Approach and Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **248**, 377-381. <https://doi.org/10.1016/j.jtbi.2007.05.019>
59. Ding, Y.S., Zhang, T.L. and Chou, K.C. (2007) Prediction of Protein Structure Classes with Pseudo Amino Acid Composition and Fuzzy Support Vector Machine Network. *Protein & Peptide Letters*, **14**, 811-815. <https://doi.org/10.2174/092986607781483778>
60. Lin, H. and Li, Q.Z. (2007) Predicting Conotoxin Superfamily and Family by Using Pseudo Amino Acid Composition and Modified Mahalanobis Discriminant. *Biochemical and Biophysical Research Communications*, **354**, 548-551. <https://doi.org/10.1016/j.bbrc.2007.01.011>
61. Lin, H. and Li, Q.Z. (2007) Using Pseudo Amino Acid Composition to Predict Protein Structural Class: Approached by Incorporating 400 Dipeptide Components. *Journal of Computational Chemistry*, **28**, 1463-1466. <https://doi.org/10.1002/jcc.20554>
62. Mundra, P., Kumar, M., Kumar, K.K., Jayaraman, V.K. and Kulkarni, B.D. (2007) Using Pseudo Amino Acid

Composition to Predict Protein Subnuclear Localization: Approached with PSSM. *Pattern Recognition Letters*, **28**, 1610-1615. <https://doi.org/10.1016/j.patrec.2007.04.001>

63. Shi, J.Y., Zhang, S.W., Pan, Q., Cheng, Y.-M. and Xie, J. (2007) Prediction of Protein Subcellular Localization by Support Vector Machines Using Multi-Scale Energy and Pseudo Amino Acid Composition. *Amino Acids*, **33**, 69-74. <https://doi.org/10.1007/s00726-006-0475-y>
64. Zhang, T.L. and Ding, Y.S. (2007) Using Pseudo Amino Acid Composition and Binary-Tree Support Vector Machines to Predict Protein Structural Classes. *Amino Acids*, **33**, 623-629. <https://doi.org/10.1007/s00726-007-0496-1>
65. Zhou, X.B., Chen, C., Li, Z.C. and Zou, X.Y. (2007) Using Chou's Amphiphilic Pseudo Amino Acid Composition and Support Vector Machine for Prediction of Enzyme Subfamily Classes. *Journal of Theoretical Biology*, **248**, 546-551. <https://doi.org/10.1016/j.jtbi.2007.06.001>
66. Diao, Y., Ma, D., Wen, Z., Yin, J., Xiang, J. and Li, M. (2008) Using Pseudo Amino Acid Composition to Predict Transmembrane Regions in Protein: Cellular Automata and Lempel-Ziv Complexity. *Amino Acids*, **34**, 111-117. <https://doi.org/10.1007/s00726-007-0550-z>
67. Ding, Y.S. and Zhang, T.L. (2008) Using Chou's Pseudo Amino Acid Composition to Predict Subcellular Localization of Apoptosis Proteins: An Approach with Immune Genetic Algorithm-Based Ensemble Classifier. *Pattern Recognition Letters*, **29**, 1887-1892. <https://doi.org/10.1016/j.patrec.2008.06.007>
68. Fang, Y., Guo, Y., Feng, Y. and Li, M. (2008) Predicting DNA-Binding Proteins: Approached from Chou's Pseudo Amino Acid Composition and Other Specific Sequence Features. *Amino Acids*, **34**, 103-109. <https://doi.org/10.1007/s00726-007-0568-2>
69. Jiang, X., Wei, R., Zhang, T.L. and Gu, Q. (2008) Using the Concept of Chou's Pseudo Amino Acid Composition to Predict Apoptosis Proteins Subcellular Location: An Approach By Approximate Entropy. *Protein & Peptide Letters*, **15**, 392-396. <https://doi.org/10.2174/092986608784246443>
70. Jiang, X., Wei, R., Zhao, Y. and Zhang, T. (2008) Using Chou's Pseudo Amino Acid Composition Based on Approximate Entropy and an Ensemble of AdaBoost Classifiers to Predict Protein Subnuclear Location. *Amino Acids*, **34**, 669-675. <https://doi.org/10.1007/s00726-008-0034-9>
71. Li, F.M. and Li, Q.Z. (2008) Using Pseudo Amino Acid Composition to Predict Protein Subnuclear Location with Improved Hybrid Approach. *Amino Acids*, **34**, 119-125. <https://doi.org/10.1007/s00726-007-0545-9>
72. Li, F.M. and Li, Q.Z. (2008) Predicting Protein Subcellular Location Using Chou's Pseudo Amino Acid Composition and Improved Hybrid Approach. *Protein & Peptide Letters*, **15**, 612-616. <https://doi.org/10.2174/092986608784966930>
73. Lin, H. (2008) The Modified Mahalanobis Discriminant for Predicting Outer Membrane Proteins by Using Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **252**, 350-356. <https://doi.org/10.1016/j.jtbi.2008.02.004>
74. Lin, H., Ding, H., Guo, F.-B., Zhang, A.Y. and Huang, J. (2008) Predicting Subcellular Localization of Mycobacterial Proteins by Using Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **15**, 739-744. <https://doi.org/10.2174/092986608785133681>
75. Nanni, L. and Lumini, A. (2008) Genetic Programming for Creating Chou's Pseudo Amino Acid Based Features for Submitochondria Localization. *Amino Acids*, **34**, 653-660. <https://doi.org/10.1007/s00726-007-0018-1>
76. Shi, J.Y., Zhang, S.W., Pan, Q. and Zhou, G.P. (2008) Using Pseudo Amino Acid Composition to Predict Protein Subcellular Location: Approached with Amino Acid Composition Distribution. *Amino Acids*, **35**, 321-327.
77. Xiao, X., Lin, W.Z. and Chou, K.C. (2008) Using Grey Dynamic Modeling and Pseudo Amino Acid Composition to Predict Protein Structural Classes. *Journal of Computational Chemistry*, **29**, 2018-2024.

<https://doi.org/10.1002/jcc.20955>

78. Xiao, X., Wang, P. and Chou, K.C. (2008) Predicting Protein Structural Classes with Pseudo Amino Acid Composition: An Approach Using Geometric Moments of Cellular Automaton Image. *Journal of Theoretical Biology*, **254**, 691-696. <https://doi.org/10.1016/j.jtbi.2008.06.016>
79. Zhang, G.Y. and Fang, B.S. (2008) Predicting the Cofactors of Oxidoreductases Based on Amino Acid Composition Distribution and Chou's Amphiphilic Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **253**, 310-315. <https://doi.org/10.1016/j.jtbi.2008.03.015>
80. Zhang, G.Y., Li, H.C., Gao, J.Q. and Fang, B.S. (2008) Predicting Lipase Types by Improved Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **15**, 1132-1137. <https://doi.org/10.2174/092986608786071184>
81. 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>
82. Zhang, S.W., Zhang, Y.L., Yang, H.F., Zhao, C.H. and Pan, Q. (2008) Using the Concept of Chou's Pseudo Amino Acid Composition to Predict Protein Subcellular Localization: An Approach by Incorporating Evolutionary Information and von Neumann Entropies. *Amino Acids*, **34**, 565-572. <https://doi.org/10.1007/s00726-007-0010-9>
83. Zhang, T.L., Ding, Y.S. and Chou, K.C. (2008) Prediction Protein Structural Classes with Pseudo Amino Acid Composition: Approximate Entropy and Hydrophobicity Pattern. *Journal of Theoretical Biology*, **250**, 186-193. <https://doi.org/10.1016/j.jtbi.2007.09.014>
84. Chen, C., Chen, L., Zou, X. and Cai, P. (2009) Prediction of Protein Secondary Structure Content by Using the Concept of Chou's Pseudo Amino Acid Composition and Support Vector Machine. *Protein & Peptide Letters*, **16**, 27-31. <https://doi.org/10.2174/092986609787049420>
85. Ding, H., Luo, L. and Lin, H. (2009) Prediction of Cell Wall Lytic Enzymes Using Chou's Amphiphilic Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **16**, 351-355. <https://doi.org/10.2174/092986609787848045>
86. Du, P., Cao, S. and Li, Y. (2009) SubChlo: Predicting Protein Subchloroplast Locations with Pseudo Amino Acid Composition and the Evidence-Theoretic K-Nearest Neighbor (ET-KNN) Algorithm. *Journal of Theoretical Biology*, **261**, 330-335.
87. Gao, Q.B., Jin, Z.C., Ye, X.F., Wu, C. and He, J. (2009) Prediction of Nuclear Receptors with Optimal Pseudo Amino Acid Composition. *Analytical Biochemistry*, **387**, 54-59. <https://doi.org/10.1016/j.ab.2009.01.018>
88. Georgiou, D.N., Karakasidis, T.E., Nieto, J.J. and Torres, A. (2009) Use of Fuzzy Clustering Technique and Matrices to Classify Amino Acids and Its Impact to Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **257**, 17-26. <https://doi.org/10.1016/j.jtbi.2008.11.003>
89. Li, Z.C., Zhou, X.B., Dai, Z. and Zou, X.Y. (2009) Prediction of Protein Structural Classes by Chou's Pseudo Amino Acid Composition: Approached Using Continuous Wavelet Transform and Principal Component Analysis. *Amino Acids*, **37**, 415-425. <https://doi.org/10.1007/s00726-008-0170-2>
90. Lin, H., Wang, H., Ding, H., Chen, Y.L. and Li, Q.Z. (2009) Prediction of Subcellular Localization of Apoptosis Protein Using Chou's Pseudo Amino Acid Composition. *Acta Biotheoretica*, **57**, 321-330.
91. Qiu, J.D., Huang, J.H., Liang, R.P. and Lu, X.Q. (2009) Prediction of G-Protein-Coupled Receptor Classes Based on the Concept of Chou's Pseudo Amino Acid Composition: An Approach from Discrete Wavelet Transform. *Analytical Biochemistry*, **390**, 68-73. <https://doi.org/10.1016/j.ab.2009.04.009>
92. Xiao, X., Wang, P. and Chou, K.C. (2009) Predicting Protein Quaternary Structural Attribute by Hybridizing Functional Domain Composition And Pseudo Amino Acid Composition. *Journal of Applied Crystallography*,

42, 169-173. <https://doi.org/10.1107/S0021889809002751>

93. Zeng, Y.H., Guo, Y.Z., Xiao, R.Q., Yang, L., Yu, L.Z. and Li, M.L. (2009) Using the Augmented Chou's Pseudo Amino Acid Composition for Predicting Protein Submitochondria Locations Based on Auto Covariance Approach. *Journal of Theoretical Biology*, **259**, 366-372. <https://doi.org/10.1016/j.jtbi.2009.03.028>
94. Esmaeili, M., Mohabatkar, H. and Mohsenzadeh, S. (2010) Using the Concept of Chou's Pseudo Amino Acid Composition for Risk Type Prediction of Human Papilloma Viruses. *Journal of Theoretical Biology*, **263**, 203-209. <https://doi.org/10.1016/j.jtbi.2009.11.016>
95. Gao, Q.B., Ye, X.F., Jin, Z.C. and He, J. (2010) Improving Discrimination of Outer Membrane Proteins by Fusing Different Forms of Pseudo Amino Acid Composition. *Analytical Biochemistry*, **398**, 52-59. <https://doi.org/10.1016/j.ab.2009.10.040>
96. Gu, Q., Ding, Y., Zhang, T. and Shen, Y. (2010) Prediction of G-Protein-Coupled Receptor Classes with Pseudo Amino Acid Composition. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*, **27**, 500-504.
97. Gu, Q., Ding, Y.S. and Zhang, T.L. (2010) Prediction of G-Protein-Coupled Receptor Classes in Low Homology Using Chou's Pseudo Amino Acid Composition with Approximate Entropy and Hydrophobicity Patterns. *Protein & Peptide Letters*, **17**, 559-567.
98. Kandaswamy, K.K., Pugalenti, G., Moller, S., Hartmann, E., Kalies, K.U., Suganthan, P.N. and Martinetz, T. (2010) Prediction of Apoptosis Protein Locations with Genetic Algorithms and Support Vector Machines Through a New Mode of Pseudo Amino Acid Composition. *Protein and Peptide Letters*, **17**, 1473-1479.
99. Liu, T., Zheng, X., Wang, C. and Wang, J. (2010) Prediction of Subcellular Location of Apoptosis Proteins Using Pseudo Amino Acid Composition: An Approach from Auto Covariance Transformation. *Protein & Peptide Letters*, **17**, 1263-1269. <https://doi.org/10.2174/092986610792231528>
100. Mohabatkar, H. (2010) Prediction of Cyclin Proteins Using Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **17**, 1207-1214. <https://doi.org/10.2174/092986610792231564>
101. Niu, X.H., Li, N.N., Shi, F., Hu, X.H., Xia, J.B. and Xiong, H.J. (2010) Predicting Protein Solubility with a Hybrid Approach by Pseudo Amino Acid Composition. *Protein and Peptide Letters*, **17**, 1466-1472. <https://doi.org/10.2174/0929866511009011466>
102. Qiu, J.D., Huang, J.H., Shi, S.P. and Liang, R.P. (2010) Using the Concept of Chou's Pseudo Amino Acid Composition to Predict Enzyme Family Classes: An Approach with support Vector Machine Based on Discrete Wavelet Transform. *Protein & Peptide Letters*, **17**, 715-722. <https://doi.org/10.2174/092986610791190372>
103. Sahu, S.S. and Panda, G. (2010) A Novel Feature Representation Method Based on Chou's Pseudo Amino Acid Composition for Protein Structural Class Prediction. *Computational Biology and Chemistry*, **34**, 320-327.
104. Wang, Y.C., Wang, X.B., Yang, Z.X. and Deng, N.Y. (2010) Prediction of Enzyme Subfamily Class via Pseudo Amino Acid Composition by Incorporating the Conjoint Triad Feature. *Protein & Peptide Letters*, **17**, 1441-1449. <https://doi.org/10.2174/0929866511009011441>
105. Wu, J., Li, M.L., Yu, L.Z. and Wang, C. (2010) An Ensemble Classifier of Support Vector Machines Used to Predict Protein Structural Classes by Fusing Auto Covariance and Pseudo Amino Acid Composition. *The Protein Journal*, **29**, 62-67. <https://doi.org/10.1007/s10930-009-9222-z>
106. Yu, L., Guo, Y., Li, Y., Li, G., Li, M., Luo, J., Xiong, W. and Qin, W. (2010) SecretP: Identifying Bacterial Secreted Proteins by Fusing New Features into Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **267**, 1-6. <https://doi.org/10.1016/j.jtbi.2010.08.001>
107. Chou, K.C. and Shen, H.B. (2010) Cell-PLoc 2.0: An Improved Package of Web-Servers for Predicting Subcellular Localization of Proteins in Various Organisms. *Natural Science*, **2**, 1090-1103.
108. Chou, K.C. (2011) Some Remarks on Protein Attribute Prediction and Pseudo Amino Acid Composition (50th

Anniversary Year Review). *Journal of Theoretical Biology*, **273**, 236-247.

<https://doi.org/10.1016/j.jtbi.2010.12.024>

109. Ding, H., Liu, L., Guo, F.B., Huang, J. and Lin, H. (2011) Identify Golgi Protein Types with Modified Mahalanobis Discriminant Algorithm and Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **18**, 58-63.
110. Guo, J., Rao, N., Liu, G., Yang, Y. and Wang, G. (2011) Predicting Protein Folding Rates Using the Concept of Chou's Pseudo Amino Acid Composition. *Journal of Computational Chemistry*, **32**, 1612-1617.
<https://doi.org/10.1002/jcc.21740>
111. Hayat, M. and Khan, A. (2011) Predicting Membrane Protein Types by Fusing Composite Protein Sequence Features into Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **271**, 10-17.
<https://doi.org/10.1016/j.jtbi.2010.11.017>
112. Hu, L., Zheng, L., Wang, Z., Li, B. and Liu, L. (2011) Using Pseudo Amino Acid Composition to Predict Protease Families by Incorporating a Series of Protein Biological Features. *Protein and Peptide Letters*, **18**, 552-558.
<https://doi.org/10.2174/092986611795222795>
113. Huang, Y., Yang, L. and Wang, T. (2011) Phylogenetic Analysis of DNA Sequences Based on the Generalized Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **269**, 217-223.
<https://doi.org/10.1016/j.jtbi.2010.10.027>
114. Jingbo, X., Silan, Z., Feng, S., Huijuan, X., Xuehai, H., Xiaohui, N. and Zhi, L. (2011) Using the Concept of Pseudo Amino Acid Composition to predict Resistance Gene against *Xanthomonas oryzae* pv. *oryzae* in Rice: An Approach from Chaos Games Representation. *Journal of Theoretical Biology*, **284**, 16-23.
<https://doi.org/10.1016/j.jtbi.2011.06.003>
115. Lin, H. and Ding, H. (2011) Predicting Ion Channels and Their Types by the Dipeptide Mode of Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **269**, 64-69. <https://doi.org/10.1016/j.jtbi.2010.10.019>
116. Lin, J. and Wang, Y. (2011) Using a Novel AdaBoost Algorithm and Chou's Pseudo Amino Acid Composition for Predicting Protein Subcellular Localization. *Protein & Peptide Letters*, **18**, 1219-1225.
<https://doi.org/10.2174/092986611797642797>
117. Lin, J., Wang, Y. and Xu, X. (2011) A Novel Ensemble and Composite Approach for Classifying Proteins Based on Chou's Pseudo Amino Acid Composition. *African Journal of Biotechnology*, **10**, 16963-16968.
118. Liu, X.L., Lu, J.L. and Hu, X.H. (2011) Predicting Thermophilic Proteins with Pseudo Amino Acid Composition: Approached from Chaos Game Representation and Principal Component Analysis. *Protein & Peptide Letters*, **18**, 1244-1250.
119. Mahdavi, A. and Jahandideh, S. (2011) Application of Density Similarities to Predict Membrane Protein Types Based on Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **276**, 132-137.
<https://doi.org/10.1016/j.jtbi.2011.01.048>
120. Mohabatkar, H., Mohammad Beigi, M. and Esmaili, A. (2011) Prediction of GABA(A) Receptor Proteins Using the Concept of Chou's Pseudo Amino Acid Composition and Support Vector Machine. *Journal of Theoretical Biology*, **281**, 18-23. <https://doi.org/10.1016/j.jtbi.2011.04.017>
121. Mohammad Beigi, M., Behjati, M. and Mohabatkar, H. (2011) Prediction of Metalloproteinase Family Based on the Concept of Chou's Pseudo Amino Acid Composition Using a Machine Learning Approach. *Journal of Structural and Functional Genomics*, **12**, 191-197. <https://doi.org/10.1007/s10969-011-9120-4>
122. Qiu, J.D., Sun, X.Y., Suo, S.B., Shi, S.P., Huang, S.Y., Liang, R.P. and Zhang, L. (2011) Predicting Homo-Oligomers and Hetero-Oligomers by Pseudo Amino Acid Composition: An Approach from Discrete Wavelet Transformation. *Biochimie*, **93**, 1132-1138. <https://doi.org/10.1016/j.biochi.2011.03.010>
123. Qiu, J.D., Suo, S.B., Sun, X.Y., Shi, S.P. and Liang, R.P. (2011) OligoPred: A Web-Server For Predicting Ho-

- mo-Oligomeric Proteins by Incorporating Discrete Wavelet Transform into Chou's Pseudo Amino Acid Composition. *Journal of Molecular Graphics & Modelling*, **30**, 129-134. <https://doi.org/10.1016/j.jmgm.2011.06.014>
124. Shi, R. and Xu, C. (2011) Prediction of Rat Protein Subcellular Localization with Pseudo Amino Acid Composition Based on Multiple Sequential Features. *Protein and Peptide Letters*, **18**, 625-633. <https://doi.org/10.2174/092986611795222768>
125. Shu, M., Cheng, X., Zhang, Y., Wang, Y., Lin, Y., Wang, L. and Lin, Z. (2011) Predicting the Activity of ACE Inhibitory Peptides with a Novel Mode of Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **18**, 1233-1243.
126. Wang, D., Yang, L., Fu, Z. and Xia, J. (2011) Prediction of Thermophilic Protein with Pseudo Amino Acid Composition: An Approach from Combined Feature Selection and Reduction. *Protein & Peptide Letters*, **18**, 684-689. <https://doi.org/10.2174/092986611795446085>
127. Wang, W., Geng, X.B., Dou, Y., Liu, T. and Zheng, X. (2011) Predicting Protein Subcellular Localization by Pseudo Amino Acid Composition with a Segment-Weighted and Features-Combined Approach. *Protein and Peptide Letters*, **18**, 480-487. <https://doi.org/10.2174/092986611794927947>
128. Xiao, X. and Chou, K.C. (2011) Using Pseudo Amino Acid Composition to Predict Protein Attributes via Cellular Automata and Other Approaches. *Current Bioinformatics*, **6**, 251-260. <https://doi.org/10.2174/1574893611106020251>
129. Xiao, X., Wang, P. and Chou, K.C. (2011) GPCR-2L: Predicting G Protein-Coupled Receptors and Their Types by Hybridizing Two Different Modes of Pseudo Amino Acid Compositions. *Molecular Biosystems*, **7**, 911-919. <https://doi.org/10.1039/C0MB00170H>
130. Zia Ur, R. and Khan, A. (2011) Prediction of GPCRs with Pseudo Amino Acid Composition: Employing Composite Features and Grey Incidence Degree Based Classification. *Protein & Peptide Letters*, **18**, 872-878.
131. Zou, D., He, Z., He, J. and Xia, Y. (2011) Supersecondary Structure Prediction Using Chou's Pseudo Amino Acid Composition. *Journal of Computational Chemistry*, **32**, 271-278. <https://doi.org/10.1002/jcc.21616>
132. Cao, J.Z., Liu, W.Q. and Gu, H. (2012) Predicting Viral Protein Subcellular Localization with Chou's Pseudo Amino Acid Composition and Imbalance-Weighted Multi-Label K-Nearest Neighbor Algorithm. *Protein and Peptide Letters*, **19**, 1163-1169. <https://doi.org/10.2174/092986612803216999>
133. Chen, C., Shen, Z.B. and Zou, X.Y. (2012) Dual-Layer Wavelet SVM for Predicting Protein Structural Class via the General Form of Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **19**, 422-429.
134. Chen, Y.L., Li, Q.Z. and Zhang, L.Q. (2012) Using Increment of Diversity to Predict Mitochondrial Proteins of Malaria Parasite: Integrating Pseudo Amino Acid Composition and Structural Alphabet. *Amino Acids*, **42**, 1309-1316. <https://doi.org/10.1007/s00726-010-0825-7>
135. Fan, G.L. and Li, Q.Z. (2012) Predict Mycobacterial Proteins Subcellular Locations by Incorporating Pseudo-Average Chemical Shift into the General Form of Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **304**, 88-95. <https://doi.org/10.1016/j.jtbi.2012.03.017>
136. Fan, G.L. and Li, Q.Z. (2012) Predicting Protein Submitochondria Locations by Combining Different Descriptors into the General Form of Chou's Pseudo Amino Acid Composition. *Amino Acids*, **43**, 545-555. <https://doi.org/10.1007/s00726-011-1143-4>
137. Gao, Q.B., Zhao, H., Ye, X. and He, J. (2012) Prediction of Pattern Recognition Receptor Family Using Pseudo Amino Acid Composition. *Biochemical and Biophysical Research Communications*, **417**, 73-77. <https://doi.org/10.1016/j.bbrc.2011.11.057>
138. Li, L.Q., Zhang, Y., Zou, L.Y., Zhou, Y. and Zheng, X.Q. (2012) Prediction of Protein Subcellular Multi-Localization Based on the General Form of Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*

ters, **19**, 375-387.

139. Lin, W.Z., Fang, J.A., Xiao, X. and Chou, K.C. (2012) Predicting Secretory Proteins of Malaria Parasite by Incorporating Sequence Evolution Information into Pseudo Amino Acid Composition via Grey System Model. *PLoS ONE*, **7**, e49040. <https://doi.org/10.1371/journal.pone.0049040>
140. Liu, L., Hu, X.Z., Liu, X.X., Wang, Y. and Li, S.B. (2012) Predicting Protein Fold Types by the General Form of Chou's Pseudo Amino Acid Composition: Approached from Optimal Feature Extractions. *Protein & Peptide Letters*, **19**, 439-449. <https://doi.org/10.2174/092986612799789378>
141. Nanni, L., Brahnam, S. and Lumini, A. (2012) Wavelet Images and Chou's Pseudo Amino Acid Composition for Protein Classification. *Amino Acids*, **43**, 657-665. <https://doi.org/10.1007/s00726-011-1114-9>
142. Nanni, L., Lumini, A., Gupta, D. and Garg, A. (2012) Identifying Bacterial Virulent Proteins by Fusing a Set of Classifiers Based on Variants of Chou's Pseudo Amino Acid Composition and on Evolutionary Information. *IEEE-ACM Transaction on Computational Biology and Bioinformatics*, **9**, 467-475.
143. Niu, X.H., Hu, X.H., Shi, F. and Xia, J.B. (2012) Predicting Protein Solubility by the General Form of Chou's Pseudo Amino Acid Composition: Approached from Chaos Game Representation and Fractal Dimension. *Protein & Peptide Letters*, **19**, 940-948.
144. Ren, L.Y., Zhang, Y.S. and Gutman, I. (2012) Predicting the Classification of Transcription Factors by Incorporating their Binding Site Properties into a Novel Mode of Chou's Pseudo Amino Acid Composition *Protein & Peptide Letters*, **19**, 1170-1176.
145. Wang, J., Li, Y., Wang, Q., You, X., Man, J., Wang, C. and Gao, X. (2012) ProClusEnsem: Predicting Membrane Protein Types by Fusing Different Modes of Pseudo Amino Acid Composition. *Computers in Biology and Medicine*, **42**, 564-574. <https://doi.org/10.1016/j.compbiomed.2012.01.012>
146. Yu, X., Zheng, X., Liu, T., Dou, Y. and Wang, J. (2012) Predicting Subcellular Location of Apoptosis Proteins with Pseudo Amino Acid Composition: Approach from Amino Acid Substitution Matrix and Auto Covariance Transformation. *Amino Acids*, **42**, 1619-1625. <https://doi.org/10.1007/s00726-011-0848-8>
147. Zhao, X.W., Ma, Z.Q. and Yin, M.H. (2012) Predicting Protein-Protein Interactions by Combing Various Sequence-Derived Features into the General Form of Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **19**, 492-500. <https://doi.org/10.2174/092986612800191080>
148. Zia-ur-Rehman, K.A. (2012) Identifying GPCRs and Their Types with Chou's Pseudo Amino Acid Composition: An Approach from Multi-Scale Energy Representation and Position Specific Scoring Matrix. *Protein & Peptide Letters*, **19**, 890-903.
149. Chen, Y.K. and Li, K.B. (2013) Predicting Membrane Protein Types by Incorporating Protein Topology, Domains, Signal Peptides, and Physicochemical Properties into the General Form of Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **318**, 1-12. <https://doi.org/10.1016/j.jtbi.2012.10.033>
150. Fan, G.L. and Li, Q.Z. (2013) Discriminating Bioluminescent Proteins by Incorporating Average Chemical Shift and Evolutionary Information into the General Form of Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **334**, 45-51. <https://doi.org/10.1016/j.jtbi.2013.06.003>
151. Georgiou, D.N., Karakasidis, T.E. and Megaritis, A.C. (2013) A Short Survey on Genetic Sequences, Chou's Pseudo Amino Acid Composition and Its Combination with Fuzzy Set Theory. *The Open Bioinformatics Journal*, **7**, 41-48. <https://doi.org/10.2174/1875036201307010041>
152. Gupta, M.K., Niyogi, R. and Misra, M. (2013) An Alignment-Free Method to Find Similarity among Protein Sequences via the General Form of Chou's Pseudo Amino Acid Composition. *SAR and QSAR in Environmental Research*, **24**, 597-609. <https://doi.org/10.1080/1062936X.2013.773378>
153. Huang, C. and Yuan, J. (2013) Using Radial Basis Function on the General Form of Chou's Pseudo Amino Acid

Composition and PSSM to Predict Subcellular Locations of Proteins with Both Single and Multiple Sites. *Bio-systems*, **113**, 50-57. <https://doi.org/10.1016/j.biosystems.2013.04.005>

154. Huang, C. and Yuan, J.Q. (2013) A Multilabel Model Based on Chou's Pseudo Amino Acid Composition for Identifying Membrane Proteins with Both Single and Multiple Functional Types. *Journal of Membrane Biology*, **246**, 327-334. <https://doi.org/10.1007/s00232-013-9536-9>
155. Huang, C. and Yuan, J.Q. (2013) Predicting Protein Subchloroplast Locations with Both Single and Multiple Sites via Three Different Modes of Chou's Pseudo Amino Acid Compositions. *Journal of Theoretical Biology*, **335**, 205-212. <https://doi.org/10.1016/j.jtbi.2013.06.034>
156. Khosravian, M., Faramarzi, F.K., Beigi, M.M., Behbahani, M. and Mohabatkar, H. (2013) Predicting Antibacterial Peptides by the Concept of Chou's Pseudo Amino Acid Composition and Machine Learning Methods. *Protein & Peptide Letters*, **20**, 180-186.
157. Lin, H., Ding, C., Yuan, L.-F., Chen, W., Ding, H., Li, Z.-Q., Guo, F.-B., Huang, J. and Rao, N.-N. (2013) Predicting Subchloroplast Locations of Proteins Based on the General Form of Chou's Pseudo Amino Acid Composition: Approached from Optimal Tripeptide Composition. *International Journal of Biomathematics*, **6**, 1350003.
158. Lin, H., Ding, C., Yuan, L.F., Chen, W., Ding, H., Li, Z.Q., Guo, F.B., Hung, J. and Rao, N.N. (2013) Predicting Subchloroplast Locations of Proteins Based on the General Form of Chou's Pseudo Amino Acid Composition: Approached from Optimal Tripeptide Composition. *International Journal of Biomathematics*, **6**, Article Number: 1350003. <https://doi.org/10.1142/S1793524513500034>
159. Liu, B., Wang, X., Zou, Q., Dong, Q. and Chen, Q. (2013) Protein Remote Homology Detection by Combining Chou's Pseudo Amino Acid Composition and Profile-Based Protein Representation. *Molecular Informatics*, **32**, 775-782. <https://doi.org/10.1002/minf.201300084>
160. Mohabatkar, H., Beigi, M.M., Abdolahi, K. and Mohsenzadeh, S. (2013) Prediction of Allergenic Proteins by Means of the Concept of Chou's Pseudo Amino Acid Composition and a Machine Learning Approach. *Medicinal Chemistry*, **9**, 133-137.
161. Qin, Y.F., Zheng, L. and Huang, J. (2013) Locating Apoptosis Proteins by Incorporating the Signal Peptide Cleavage Sites into the General Form of Chou's Pseudo Amino Acid Composition. *International Journal of Quantum Chemistry*, **113**, 1660-1667. <https://doi.org/10.1002/qua.24383>
162. Sarangi, A.N., Lohani, M. and Aggarwal, R. (2013) Prediction of Essential Proteins in Prokaryotes by Incorporating Various Physico-Chemical Features into the General Form of Chou's Pseudo Amino Acid Composition. *Protein and Peptide Letters*, **20**, 781-795.
163. Wan, S., Mak, M.W. and Kung, S.Y. (2013) GOASVM: A Subcellular Location Predictor by Incorporating Term-Frequency Gene Ontology into the General Form of Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **323**, 40-48. <https://doi.org/10.1016/j.jtbi.2013.01.012>
164. Wang, X., Li, G.Z. and Lu, W.C. (2013) Virus-ECC-mPLoc: A Multi-Label Predictor for Predicting the Subcellular Localization of Virus Proteins with Both Single and Multiple Sites Based on a general Form of Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **20**, 309-317. <https://doi.org/10.2174/092986613804910608>
165. Xiaohui, N., Nana, L., Jingbo, X., Dingyan, C., Yuehua, P., Yang, X., Weiquan, W., Dongming, W. and Zengzhen, W. (2013) Using the Concept of Chou's Pseudo Amino Acid Composition to Predict Protein Solubility: An Approach with Entropies in Information Theory. *Journal of Theoretical Biology*, **332**, 211-217. <https://doi.org/10.1016/j.jtbi.2013.03.010>
166. Xu, Y., Ding, J., Wu, L.Y. and Chou, K.C. (2013) iSNO-PseAAC: Predict Cysteine S-Nitrosylation Sites in Proteins by Incorporating Position Specific Amino Acid Propensity into Pseudo Amino Acid Composition. *PLoS*

ONE, **8**, e55844. <https://doi.org/10.1371/journal.pone.0055844>

167. Hajisharifi, Z., Piryaiee, M., Mohammad Beigi, M., Behbahani, M. and Mohabatkar, H. (2014) Predicting Anti-cancer Peptides with Chou's Pseudo Amino Acid Composition and Investigating Their Mutagenicity via Ames Test. *Journal of Theoretical Biology*, **341**, 34-40. <https://doi.org/10.1016/j.jtbi.2013.08.037>
168. Jia, C., Lin, X. and Wang, Z. (2014) Prediction of Protein S-Nitrosylation Sites Based on Adapted Normal Distribution Bi-Profile Bayes and Chou's Pseudo Amino Acid Composition. *International Journal of Molecular Sciences*, **15**, 10410- 10423.
169. Kong, L., Zhang, L. and Lv, J. (2014) Accurate Prediction of Protein Structural Classes by Incorporating Predicted Secondary Structure Information into the General Form of Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **344**, 12-18. <https://doi.org/10.1016/j.jtbi.2013.11.021>
170. Liu, B., Xu, J., Lan, X., Xu, R., Zhou, J., Wang, X. and Chou, K.C. (2014) iDNA-Prot|dis: Identifying DNA-Binding Proteins by Incorporating Amino Acid Distance-Pairs and Reduced Alphabet Profile into the General Pseudo Amino Acid Composition. *PLoS ONE*, **9**, e106691. <https://doi.org/10.1371/journal.pone.0106691>
171. Mondal, S. and Pai, P.P. (2014) Chou's Pseudo Amino Acid Composition Improves Sequence-Based Antifreeze Protein Prediction. *Journal of Theoretical Biology*, **356**, 30-35. <https://doi.org/10.1016/j.jtbi.2014.04.006>
172. Nanni, L., Brahnam, S. and Lumini, A. (2014) Prediction of Protein Structure Classes by Incorporating Different Protein Descriptors into General Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **360**, 109-116. <https://doi.org/10.1016/j.jtbi.2014.07.003>
173. Qiu, W.R., Xiao, X. and Chou, K.C. (2014) iRSpot-TNCPseAAC: Identify Recombination Spots with Trinucleotide Composition and Pseudo Amino Acid Components. *International Journal of Molecular Sciences (IJMS)*, **15**, 1746-1766.
174. Qiu, W.R., Xiao, X., Lin, W.Z. and Chou, K.C. (2014) iMethyl-PseAAC: Identification of Protein Methylation Sites via a Pseudo Amino Acid Composition Approach. *BioMed Research International (BMRI)*, **2014**, 947416.
175. Xu, Y., Wen, X., Shao, X.J., Deng, N.Y. and Chou, K.C. (2014) iHyd-PseAAC: Predicting Hydroxyproline and Hydroxylysine in Proteins by Incorporating Dipeptide Position-Specific Propensity into Pseudo Amino Acid Composition. *International Journal of Molecular Sciences (IJMS)*, **15**, 7594-7610.
176. Xu, Y., Wen, X., Wen, L.S., Wu, L.Y., Deng, N.Y. and Chou, K.C. (2014) iNitro-Tyr: Prediction of Nitrotyrosine Sites in Proteins with General Pseudo Amino Acid Composition. *PLoS ONE*, **9**, e105018. <https://doi.org/10.1371/journal.pone.0105018>
177. Zhang, J., Sun, P., Zhao, X. and Ma, Z. (2014) PECM: Prediction of Extracellular Matrix Proteins Using the Concept of Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **363**, 412-418. <https://doi.org/10.1016/j.jtbi.2014.08.002>
178. Zhang, L., Zhao, X. and Kong, L. (2014) Predict Protein Structural Class for Low-Similarity Sequences by Evolutionary Difference Information into the General Form of Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **355**, 105-110. <https://doi.org/10.1016/j.jtbi.2014.04.008>
179. Zuo, Y.C., Peng, Y., Liu, L., Chen, W., Yang, L. and Fan, G.L. (2014) Predicting peroxidase Subcellular Location by Hybridizing Different Descriptors of Chou's Pseudo Amino Acid Patterns. *Analytical Biochemistry*, **458**, 14-19. <https://doi.org/10.1016/j.ab.2014.04.032>
180. Ding, H., Deng, E.Z., Yuan, L.F., Liu, L., Lin, H., Chen, W. and Chou, K.C. (2014) iCTX-Type: A Sequence-Based Predictor for Identifying the Types of Conotoxins in Targeting Ion Channels. *BioMed Research International (BMRI)*, **2014**, 286419.
181. Ali, F. and Hayat, M. (2015) Classification of Membrane Protein Types Using Voting Feature Interval in Combination with Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **384**, 78-83.

182. Chen, L., Chu, C., Huang, T., Kong, X. and Cai, Y.D. (2015) Prediction and Analysis of Cell-Penetrating Peptides Using Pseudo Amino Acid Composition and Random Forest Models. *Amino Acids*.
183. Fan, G.L., Zhang, X.Y., Liu, Y.L., Nang, Y. and Wang, H. (2015) DSPMP: Discriminating Secretory Proteins of Malaria Parasite by Hybridizing Different Descriptors of Chou's Pseudo Amino Acid Patterns. *Journal of Computational Chemistry*, **36**, 2317-2327. <https://doi.org/10.1002/jcc.24210>
184. Huang, C. and Yuan, J.Q. (2015) Simultaneously Identify Three Different Attributes of Proteins by Fusing their Three Different Modes of Chou's Pseudo Amino Acid Compositions. *Protein and Peptide Letters*, **22**, 547-556.
185. Khan, Z.U., Hayat, M. and Khan, M.A. (2015) Discrimination of Acidic and Alkaline Enzyme Using Chou's Pseudo Amino Acid Composition in Conjunction with Probabilistic Neural Network Model. *Journal of Theoretical Biology*, **365**, 197-203. <https://doi.org/10.1016/j.jtbi.2014.10.014>
186. Kumar, R., Srivastava, A., Kumari, B. and Kumar, M. (2015) Prediction of Beta-Lactamase and Its Class by Chou's Pseudo Amino Acid Composition and Support Vector Machine. *Journal of Theoretical Biology*, **365**, 96-103. <https://doi.org/10.1016/j.jtbi.2014.10.008>
187. Liu, B., Chen, J. and Wang, X. (2015) Protein Remote Homology Detection by Combining Chou's Distance-Pair Pseudo Amino Acid Composition and Principal Component Analysis. *Molecular Genetics and Genomics*, **290**, 1919-1931. <https://doi.org/10.1007/s00438-015-1044-4>
188. Wang, X., Zhang, W., Zhang, Q. and Li, G.Z. (2015) MultiP-SChlo: Multi-Label Protein Subchloroplast Localization Prediction with Chou's Pseudo Amino Acid Composition and a Novel Multi-Label Classifier. *Bioinformatics*, **31**, 2639-2645. <https://doi.org/10.1093/bioinformatics/btv212>
189. Xu, R., Zhou, J., Liu, B., He, Y.A., Zou, Q., Wang, X. and Chou, K.C. (2015) Identification of DNA-Binding Proteins by Incorporating Evolutionary Information into Pseudo Amino Acid Composition via the Top-n-Gram Approach. *Journal of Biomolecular Structure & Dynamics (JBSD)*, **33**, 1720-1730.
190. Zhu, P.P., Li, W.C., Zhong, Z.J., Deng, E.Z., Ding, H., Chen, W. and Lin, H. (2015) Predicting the Subcellular Localization of Mycobacterial Proteins by Incorporating the Optimal Tripeptides into the General Form of Pseudo Amino Acid Composition. *Molecular BioSystems*, **11**, 558-563. <https://doi.org/10.1039/C4MB00645C>
191. 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>
192. Dehzangi, A., Heffernan, R., Sharma, A., Lyons, J., Paliwal, K. and Sattar, A. (2015) Gram-Positive and Gram-Negative Protein Subcellular Localization by Incorporating Evolutionary-Based Descriptors into Chou's General PseAAC. *Journal of Theoretical Biology*, **364**, 284-294. <https://doi.org/10.1016/j.jtbi.2014.09.029>
193. 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>
194. 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>
195. 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**, 1550024. <https://doi.org/10.1142/S0219720015500249>
196. 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 Trans Nanobioscience*, **14**, 915-926. <https://doi.org/10.1109/TNB.2015.2500186>

197. 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.
198. 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.
199. Han, G.S., Yu, Z.G. and Anh, V. (2014) A Two-Stage SVM Method to Predict Membrane Protein Types by Incorporating Amino Acid Classifications and Physicochemical Properties into a General Form of Chou's PseAAC. *Journal of Theoretical Biology*, **344**, 31-39. <https://doi.org/10.1016/j.jtbi.2013.11.017>
200. 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. *Comput Methods Programs Biomed*, **116**, 184-192.
201. 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 Feature Selection Approach. *Biochimie*, **104**, 100-107. <https://doi.org/10.1016/j.biochi.2014.06.001>
202. 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>
203. 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>
204. 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. *Pocess Biochemistry*, **48**, 1048-1053.
205. 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>
206. 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>
207. 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.
208. 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.
209. Mei, S. (2012) Multi-Kernel Transfer Learning Based on Chou's PseAAC Formulation for Protein Submitochondria Localization. *Journal of Theoretical Biology*, **293**, 121-130. <https://doi.org/10.1016/j.jtbi.2011.10.015>
210. Mei, S. (2012) Predicting Plant Protein Subcellular Multi-Localization by Chou's PseAAC Formulation Based Multi-Label Homolog Knowledge Transfer Learning. *Journal of Theoretical Biology*, **310**, 80-87. <https://doi.org/10.1016/j.jtbi.2012.06.028>
211. 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.

212. 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>
213. Chou, K.C. and Shen, H.B. (2009) Review: Recent Advances in Developing Web-Servers for Predicting Protein Attributes. *Natural Science*, **1**, 63-92.
214. 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>
215. 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>
216. Chen, W., Feng, P.M., Deng, E.Z., Lin, H. and Chou, K.C. (2014) iTIS-PseTNC: A Sequence-Based Predictor for Identifying Translation Initiation Site in Human Genes Using Pseudo Trinucleotide Composition. *Analytical Biochemistry*, **462**, 76-83. <https://doi.org/10.1016/j.ab.2014.06.022>
217. Chen, W., Zhang, X., Brooker, J., Lin, H., Zhang, L. and Chou, K.C. (2015) PseKNC-General: A Cross-Platform Package for Generating Various Modes of Pseudo Nucleotide Compositions. *Bioinformatics*, **31**, 119-120. <https://doi.org/10.1093/bioinformatics/btu602>
218. Chen, W., Feng, P.M., Lin, H. and Chou, K.C. (2014) iSS-PseDNC: Identifying Splicing Sites Using Pseudo Dinucleotide Composition. *BioMed Research International (BMRI)*, **2014**, 623149.
219. Guo, S.H., Deng, E.Z., Xu, L.Q., Ding, H., Lin, H., Chen, W. and Chou, K.C. (2014) iNuc-PseKNC: A Sequence-Based Predictor for Predicting Nucleosome Positioning in Genomes with Pseudo k-Tuple Nucleotide Composition. *Bioinformatics*, **30**, 1522-1529. <https://doi.org/10.1093/bioinformatics/btu083>
220. Lin, H., Deng, E.Z., Ding, H., Chen, W. and Chou, K.C. (2014) iPro54-PseKNC: A Sequence-Based Predictor for Identifying Sigma-54 Promoters in Prokaryote with Pseudo k-Tuple Nucleotide Composition. *Nucleic Acids Research*, **42**, 12961-12972. <https://doi.org/10.1093/nar/gku1019>
221. Xiao, X., Ye, H.X., Liu, Z., Jia, J.H. and Chou, K.C. (2016) iROS-gPseKNC: Predicting Replication Origin Sites in DNA by Incorporating Dinucleotide Position-Specific Propensity into General Pseudo Nucleotide Composition. *Oncotarget*, **7**, 34180-34189.
222. Chen, W., Lin, H. and Chou, K.C. (2015) Pseudo Nucleotide Composition or PseKNC: An Effective Formulation for Analyzing Genomic Sequences. *Molecular BioSystems*, **11**, 2620-2634. <https://doi.org/10.1039/C5MB00155B>
223. Feng, P., Ding, H., Yang, H., Chen, W., Lin, H. and Chou, K.C. (2017) iRNA-PseColl: Identifying the Occurrence Sites of Different RNA Modifications by Incorporating Collective Effects of Nucleotides into PseKNC. *Molecular Therapy - Nucleic Acids*. <https://doi.org/10.1016/j.omtn.2017.03.006>
224. Liu, B., Liu, F., Fang, L., Wang, X. and Chou, K.C. (2015) repDNA: A Python Package to Generate Various Modes of Feature Vectors for DNA Sequences by Incorporating User-Defined Physicochemical Properties and Sequence-Order Effects. *Bioinformatics*, **31**, 1307-1309. <https://doi.org/10.1093/bioinformatics/btu820>
225. Liu, B., Liu, F., Fang, L., Wang, X. and Chou, K.C. (2016) repRNA: A Web Server for Generating Various Feature Vectors of RNA Sequences. *Molecular Genetics and Genomics*, **291**, 473-481. <https://doi.org/10.1007/s00438-015-1078-7>
226. 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>

227. Ahmad, K., Waris, M. and Hayat, M. (2016) Prediction of Protein Submitochondrial Locations by Incorporating Dipeptide Composition into Chou's General Pseudo Amino Acid Composition. *Journal of Membrane Biology*, **249**, 293-304.
228. Behbahani, M., Mohabatkar, H. and Nosrati, M. (2016) Analysis and Comparison of Lignin Peroxidases between Fungi and Bacteria Using Three Different Modes of Chou's General Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **411**, 1-5. <https://doi.org/10.1016/j.jtbi.2016.09.001>
229. Chen, W., Tang, H., Ye, J., Lin, H. and Chou, K.C. (2016) iRNA-PseU: Identifying RNA Pseudouridine Sites. *Molecular Therapy-Nucleic Acids*, **5**, e332.
230. Fan, G.L., Liu, Y.L. and Wang, H. (2016) Identification of Thermophilic Proteins by Incorporating Evolutionary and Acid Dissociation Information into Chou's General Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **407**, 138-142. <https://doi.org/10.1016/j.jtbi.2016.07.010>
231. Jia, J., Liu, Z., Xiao, X., Liu, B. and Chou, K.C. (2016) Identification of Protein-Protein Binding Sites by Incorporating the Physicochemical Properties and Stationary Wavelet Transforms into Pseudo Amino Acid Composition (iPPBS-PseAAC). *Journal of Biomolecular Structure & Dynamics (JBSD)*, **34**, 1946-1961. <https://doi.org/10.1080/07391102.2015.1095116>
232. Jia, J., Zhang, L., Liu, Z., Xiao, X. and Chou, K.C. (2016) pSumo-CD: Predicting Sumoylation Sites in Proteins with Covariance Discriminant Algorithm by Incorporating Sequence-Coupled Effects into General PseAAC. *Bioinformatics*, **32**, 3133-3141.
233. Jiao, Y.S. and Du, P.F. (2016) Prediction of Golgi-Resident Protein Types Using General Form of Chou's Pseudo Amino Acid Compositions: Approaches with Minimal Redundancy Maximal Relevance Feature Selection. *Journal of Theoretical Biology*, **402**, 38-44. <https://doi.org/10.1016/j.jtbi.2016.04.032>
234. Liu, B., Long, R. and Chou, K.C. (2016) iDHS-EL: Identifying DNase I Hypersensitive Sites by Fusing Three Different Modes of Pseudo Nucleotide Composition into an Ensemble Learning Framework *Bioinformatics*, **32**, 2411-2418. <https://doi.org/10.1093/bioinformatics/btw186>
235. Liu, Z., Xiao, X., Yu, D.J., Jia, J., Qiu, W.R. and Chou, K.C. (2016) pRNAm-PC: Predicting N-Methyladenosine Sites in RNA Sequences via Physical-Chemical Properties. *Analytical Biochemistry*, **497**, 60-67. <https://doi.org/10.1016/j.ab.2015.12.017>
236. Qiu, W.R., Sun, B.Q., Xiao, X., Xu, Z.C. and Chou, K.C. (2016) iHyd-PseCp: Identify Hydroxyproline and Hydroxylysine in Proteins by Incorporating Sequence-Coupled Effects into General PseAAC. *Oncotarget*, **7**, 44310-44321.
237. Qiu, W.R., Xiao, X., Xu, Z.H. and Chou, K.C. (2016) iPhos-PseEn: Identifying Phosphorylation Sites in Proteins by Fusing Different Pseudo Components into an Ensemble Classifier. *Oncotarget*, **7**, 51270-51283.
238. Qiu, W.R., Zheng, Q.S., Sun, B.Q. and Xiao, X. (2016) Multi-iPPseEvo: A Multi-Label Classifier for Identifying Human Phosphorylated Proteins by Incorporating Evolutionary Information into Chou's General PseAAC via Grey System Theory *Mol Inform.* <https://doi.org/10.1002/minf.201600010>
239. Rahimi, M., Bakhtiarizadeh, M.R. and Mohammadi-Sangcheshmeh, A. (2016) OOgenesis_Pred: A Sequence-Based Method for Predicting Oogenesis Proteins by Six Different Modes of Chou's Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **414**, 128-136. <https://doi.org/10.1016/j.jtbi.2016.11.028>
240. Sun, W.J., Li, J.H., Liu, S., Wu, J., Zhou, H., Qu, L.H. and Yang, J.H. (2016) RMBase: A Resource for Decoding the Landscape of RNA Modifications from High-Throughput Sequencing Data. *Nucleic Acids Research*, **44**, D259-D265. <https://doi.org/10.1093/nar/gkv1036>
241. Tang, H., Chen, W. and Lin, H. (2016) Identification of Immunoglobulins Using Chou's Pseudo Amino Acid Composition with Feature Selection Technique. *Molecular BioSystems*, **12**, 1269-1275. <https://doi.org/10.1039/C5MB00883B>

242. 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>
243. Xu, C., Sun, D., Liu, S. and Zhang, Y. (2016) Protein Sequence Analysis by Incorporating Modified Chaos Game and Physicochemical Properties into Chou's General Pseudo Amino Acid Composition. *Journal of Theoretical Biology*, **406**, 105-115.
244. Xu, Y. and Chou, K.C. (2016) Recent Progress in Predicting Posttranslational Modification Sites in Proteins. *Current Topics in Medicinal Chemistry*, **16**, 591-603. <https://doi.org/10.2174/1568026615666150819110421>
245. Zhang, C.J., Tang, H., Li, W.C., Lin, H., Chen, W. and Chou, K.C. (2016) iOri-Human: Identify Human Origin of Replication by Incorporating Dinucleotide Physicochemical Properties into Pseudo Nucleotide Composition. *Oncotarget*, **7**, 69783-69793.
246. Zou, H.L. and Xiao, X. (2016) Predicting the Functional Types of Singleplex and Multiplex Eukaryotic Membrane Proteins via Different Models of Chou's Pseudo Amino Acid Compositions. *Journal of Membrane Biology*, **249**, 23-29.
247. Zou, H.L. and Xiao, X. (2016) Classifying Multifunctional Enzymes by Incorporating Three Different Models into Chou's General Pseudo Amino Acid Composition. *Journal of Membrane Biology*, **249**, 561-567.
248. Chen, W., Ding, H., Feng, P., Lin, H. and Chou, K.C. (2016) iACP: A Sequence-Based Tool for Identifying Anticancer Peptides. *Oncotarget*, **7**, 16895-16909. <https://doi.org/10.18632/oncotarget.7815>
249. Ju, Z., Cao, J.Z. and Gu, H. (2016) Predicting Lysine Phosphoglycerylation with Fuzzy SVM by Incorporating k-Spaced Amino Acid Pairs into Chou's General PseAAC. *Journal of Theoretical Biology*, **397**, 145-150. <https://doi.org/10.1016/j.jtbi.2016.02.020>
250. 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>
251. 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. *Mol Biosyst*, **12**, 2587-2593. <https://doi.org/10.1039/C6MB00221H>
252. Chen, W., Feng, P., Yang, H., Ding, H., Lin, H. and Chou, K.C. (2017) iRNA-AI: Identifying the Adenosine to Inosine Editing Sites in RNA Sequences. *Oncotarget*, **8**, 4208-4217. <https://doi.org/10.18632/oncotarget.13758>
253. Jiao, Y.S. and Du, P.F. (2017) Predicting Protein Submitochondrial Locations by Incorporating the Positional-Specific Physicochemical Properties into Chou's General Pseudo-Amino Acid Compositions. *Journal of Theoretical Biology*, **416**, 81-87. <https://doi.org/10.1016/j.jtbi.2016.12.026>
254. Liu, B., Wu, H., Zhang, D., Wang, X. and Chou, K.C. (2017) Pse-Analysis: A Python Package for DNA/RNA and Protein/Peptide Sequence Analysis Based on Pseudo Components and Kernel Methods. *Oncotarget*, **8**, 4208-4217. <https://doi.org/10.18632/oncotarget.14524>
255. Qiu, W.R., Jiang, S.Y., Xu, Z.C., Xiao, X. and Chou, K.C. (2017) iRNAm5C-PseDNC: Identifying RNA 5-Methylcytosine Sites by Incorporating Physical-Chemical Properties into Pseudo Dinucleotide Composition. *Oncotarget*, **in press (026264R2)**.
256. 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. *Mol Inform*, **36**. <https://doi.org/10.1002/minf.201600085>
257. Rahimi, M., Bakhtiarizadeh, M.R. and Mohammadi-Sangcheshmeh, A. (2017) OOgenesis_Pred: A Sequence-Based Method for Predicting Oogenesis Proteins by Six Different Modes of Chou's Pseudo Amino Acid Com-

- position. *Journal of Theoretical Biology*, **414**, 128-136. <https://doi.org/10.1016/j.jtbi.2016.11.028>
258. Khan, M., Hayat, M., Khan, S.A. and Iqbal, N. (2017) Unb-DPC: Identify Mycobacterial Membrane Protein Types by Incorporating Un-Biased Dipeptide Composition into Chou's General PseAAC. *Journal of Theoretical Biology*, **415**, 13-19. <https://doi.org/10.1016/j.jtbi.2016.12.004>
259. Liu, L.M., Xu, Y. and Chou, K.C. (2017) iPGK-PseAAC: Identify Lysine Phosphoglycerylation Sites in Proteins by Incorporating Four Different Tiers of Amino Acid Pairwise Coupling Information into the General PseAAC. *Medicinal Chemistry*, in press.
260. 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**, 42362. <https://doi.org/10.1038/srep42362>
261. Xu, Y., Li, C. and Chou, K.C. (2017) iPreny-PseAAC: Identify C-Terminal Cysteine Prenylation Sites in Proteins by Incorporating Two Tiers of Sequence Couplings into PseAAC. *Medicinal Chemistry*, in press.
262. Zhou, S.F. and Zhong, W.Z. (2017) Drug Design and Discovery: Principles and Applications. *Molecules*, **22**.
263. Qiu, W.R., Jiang, S.Y., Sun, B.Q., Xiao, X. and Chou, K.C. (2017) iRNA-2methyl: An Ensemble Classifier for Identifying RNA 2'-O-Methylation Modification Sites by Incorporating Sequence-Coupled Effects into General PseKNC. *Genomics*, in press.
264. Leslie, C.S., Eskin, E., Cohen, A., Weston, J. and Noble, W.S. (2004) Mismatch String Kernels for Discriminative Protein Classification. *Bioinformatics*, **20**, 467-476. <https://doi.org/10.1093/bioinformatics/btg431>
265. Luo, L., Li, D., Zhang, W., Tu, S., Zhu, X. and Tian, G. (2016) Accurate Prediction of Transposon-Derived piRNAs by Integrating Various Sequential and Physicochemical Features. *PLoS ONE*, **11**, e0153268.
266. Xue, C., Li, F., He, T., Liu, G.-P., Li, Y. and Zhang, X. (2005) Classification of Real and Pseudo MicroRNA Precursors Using Local Structure-Sequence Features and Support Vector Machine. *BMC bioinformatics*, **6**, 1. <https://doi.org/10.1186/1471-2105-6-310>
267. Liu, B., Fang, L., Liu, F., Wang, X., Chen, J. and Chou, K.C. (2015) Identification of Real MicroRNA Precursors with a Pseudo Structure Status Composition Approach. *PLoS ONE*, **10**, e0121501. <https://doi.org/10.1371/journal.pone.0121501>
268. Chen, W., Zhang, X., Brooker, J., Lin, H., Zhang, L. and Chou, K.-C. (2015b) PseKNC-General: A Cross-Platform Package for Generating Various Modes of Pseudo Nucleotide Compositions. *Bioinformatics*, **31**, 119-120. <https://doi.org/10.1093/bioinformatics/btu602>
269. Liu, B., Wang, X., Lin, L., Dong, Q. and Wang, X. (2008) A Discriminative Method for Protein Remote Homology Detection and Fold Recognition Combining Top-n-Grams and Latent Semantic Analysis. *BMC Bioinformatics*, **9**, 1. <https://doi.org/10.1186/1471-2105-9-510>
270. Liu, B., Wang, X., Chen, Q., Dong, Q. and Lan, X. (2012) Using Amino Acid Physicochemical Distance Transformation for Fast Protein Remote Homology Detection. *PLoS ONE*, **7**, e46633. <https://doi.org/10.1371/journal.pone.0046633>
271. Liu, B., Xu, J., Zou, Q., Xu, R., Wang, X. and Chen, Q. (2014) Using Distances between Top-n-Gram and Residue Pairs for Protein Remote Homology Detection. *BMC Bioinformatics*, **15**, 1. <https://doi.org/10.1093/bib/bbs075>
272. Dong, Q., Zhou, S. and Guan, J. (2009) A New Taxonomy-Based Protein Fold Recognition Approach Based on Autocross-Covariance Transformation. *Bioinformatics*, **25**, 2655-2662. <https://doi.org/10.1093/bioinformatics/btp500>
273. Guo, Y., Yu, L., Wen, Z. and Li, M. (2008) Using Support Vector Machine Combined with Auto Covariance to Predict Protein-Protein Interactions from Protein Sequences. *Nucleic Acids Research*, **36**, 3025-3030.

<https://doi.org/10.1093/nar/gkn159>

274. Liu, B., Xu, J., Lan, X., Xu, R., Zhou, J., Wang, X. and Chou, K.-C. (2014) iDNA-Prot| dis: Identifying DNA-Binding Proteins by Incorporating Amino Acid Distance-Pairs and Reduced Alphabet Profile into the General Pseudo Amino Acid Composition. *PLoS ONE*, **9**, e106691. <https://doi.org/10.1371/journal.pone.0106691>
275. Xiao, X., Lin, W.Z. and Chou, K.C. (2013) Recent Advances in Predicting Protein Classification and Their Applications to Drug Development. *Current Topics in Medicinal Chemistry*, **13**, 1622-1635. <https://doi.org/10.2174/15680266113139990113>
276. Shen, H.B., Song, J.N. and Chou, K.C. (2009) Prediction of Protein Folding Rates from Primary Sequence by Fusing Multiple Sequential Features. *Journal of Biomedical Science and Engineering (JBISE)*, **2**, 136-143.
277. Chen, W., Lin, H., Feng, P.M., Ding, C., Zuo, Y.C. and Chou, K.C. (2012) iNuc-PhysChem: A Sequence-Based Predictor for Identifying Nucleosomes via Physicochemical Properties. *PLoS ONE*, **7**, e47843. <https://doi.org/10.1371/journal.pone.0047843>
278. Fan, Y.N., Xiao, X., Min, J.L. and Chou, K.C. (2014) iNR-Drug: Predicting the Interaction of Drugs with Nuclear Receptors in Cellular Networking. *International Journal of Molecular Sciences (IJMS)*, **15**, 4915-4937.
279. He, Z., Zhang, J., Shi, X.H., Hu, L.L., Kong, X., Cai, Y.D. and Chou, K.C. (2010) Predicting Drug-Target Interaction Networks Based on Functional Groups and Biological Features. *PLoS ONE*, **5**, e9603. <https://doi.org/10.1371/journal.pone.0009603>
280. Hu, L., Huang, T., Shi, X., Lu, W.C., Cai, Y.D. and Chou, K.C. (2011) Predicting Functions of Proteins in Mouse Based on Weighted Protein-Protein Interaction Network and Protein Hybrid Properties. *PLoS ONE*, **6**, e14556. <https://doi.org/10.1371/journal.pone.0014556>
281. Huang, R.B., Du, Q.S., Wang, C.H., Liao, S.M. and Chou, K.C. (2010) A Fast and Accurate Method for Predicting pKa of Residues in Proteins. *Protein Engineering, Design and Selection (PEDS)*, **23**, 35-42.
282. Huang, T., Chen, L., Cai, Y.D. and Chou, K.C. (2011) Classification and Analysis of Regulatory Pathways Using Graph Property, Biochemical and Physicochemical Property, and Functional Property. *PLoS ONE*, **6**, e25297. <https://doi.org/10.1371/journal.pone.0025297>
283. Huang, T., Niu, S., Xu, Z., Huang, Y., Kong, X., Cai, Y.D. and Chou, K.C. (2011) Predicting Transcriptional Activity of Multiple Site p53 Mutants Based on Hybrid Properties. *PLoS ONE*, **6**, e22940.
284. Huang, T., Shi, X.H., Wang, P., He, Z., Feng, K.Y., Hu, L., Kong, X., Li, Y.X., Cai, Y.D. and Chou, K.C. (2010) Analysis and Prediction of the Metabolic Stability of Proteins Based on Their Sequential Features, Subcellular Locations and Interaction Networks. *PLoS ONE*, **5**, e10972. <https://doi.org/10.1371/journal.pone.0010972>
285. Huang, T., Zhang, J., Xu, Z.P., Hu, L.L., Chen, L., Shao, J.L., Zhang, L., Kong, X.Y., Cai, Y.D. and Chou, K.C. (2012) Deciphering the Effects of Gene Deletion on Yeast Longevity Using Network and Machine Learning Approaches. *Biochimie*, **94**, 1017-1025. <https://doi.org/10.1016/j.biochi.2011.12.024>
286. Li, B.Q., Hu, L.L., Chen, L., Feng, K.Y., Cai, Y.D. and Chou, K.C. (2012) Prediction of Protein Domain with mRMR Feature Selection and Analysis. *PLoS ONE*, **7**, e39308.
287. Li, B.Q., Hu, L.L., Niu, S., Cai, Y.D. and Chou, K.C. (2012) Predict and Analyze S-Nitrosylation Modification Sites with the mRMR and IFS Approaches. *Journal of Proteomics*, **75**, 1654-1665. <https://doi.org/10.1016/j.jprot.2011.12.003>
288. Min, J.L., Xiao, X. and Chou, K.C. (2013) iEzy-Drug: A Web Server for Identifying the Interaction between Enzymes and Drugs in Cellular Networking. *BioMed Research International (BMRI)*, **2013**, 701317.
289. Pugalenti, G., Kandaswamy, K.K., Chou, K.C., Vivekanandan, S. and Kolatkar, P. (2012) RSARF: Prediction of Residue Solvent Accessibility from Protein Sequence Using Random Forest Method. *Protein & Peptide Letters*, **19**, 50-56.

290. Chen, W., Feng, P., Ding, H., Lin, H. and Chou, K.C. (2016) Using Deformation Energy to Analyze Nucleosome Positioning in Genomes. *Genomics*, **107**, 69-75. <https://doi.org/10.1016/j.ygeno.2015.12.005>
291. Chen, W., Luo, L. and Zhang, L. (2010) The Organization of Nucleosomes around Splice Sites. *Nucleic Acids Research*, **38**, 2788-2798. <https://doi.org/10.1093/nar/gkq007>
292. El-Manzalawy, Y., Dobbs, D. and Honavar, V. (2008) Predicting Flexible Length Linear B-Cell Epitopes. *Computational Systems Bioinformatics*, **7**, 121-132.
293. Lodhi, H., Saunders, C., Shawe-Taylor, J., Cristianini, N. and Watkins, C. (2002) Text Classification Using String Kernels. *Journal of Machine Learning Research*, **2**, 419-444.
294. Horne, D.S. (1988) Prediction of Protein Helix Content from an Autocorrelation Analysis of Sequence Hydrophobicities. *Biopolymers*, **27**, 451-477. <https://doi.org/10.1002/bip.360270308>
295. Sokal, R.R. and Thomson, B.A. (2006) Population Structure Inferred by Local Spatial Autocorrelation: An Example from an Amerindian Tribal Population. *American Journal of Physical Anthropology*, **129**, 121-131. <https://doi.org/10.1002/ajpa.20250>
296. Feng, Z.-P. and Zhang, C.-T. (2000) Prediction of Membrane Protein Types Based on the Hydrophobic Index of Amino Acids. *Journal of Protein Chemistry*, **19**, 269-275. <https://doi.org/10.1023/A:1007091128394>



Scientific Research Publishing

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact ns@scirp.org