

Intriguing Story about the Birth of Gordon Life Science Institute and its Development and Driving Force

Chou CK^{1,2*}

¹Gordon Life Science Institute, Boston, Massachusetts 02478, United States of America

²Department of Physics, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, 610054, China

***Corresponding author:** Kuo-Chen Chou, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, 610054, China, E-mail: kcchou@gordonlifescience.org

Citation: Chou CK (2019) Intriguing Story about the Birth of Gordon Life Science Institute and its Development and Driving Force. *Journal of Clinical Cancer Research* V1(1): 1-22.

Received Date: Nov 08, 2019 **Accepted Date:** Nov 27, 2019 **Published Date:** Dec 06, 2019

1. Abstract

Gordon Life Science Institute is the first Internet Research Institute ever established in the world. It is a non-profit institute. Those scientists who are really dedicated themselves to science and loving science more than anything else can become its member. In the friendly door-opened Institute, they can maximize their time and energy to engage in their scientific creativity. They have also believed that science would be more truthful and wonderful if scientists do not have to spend a lot of time on funding application, and that great scientific findings and creations in history were often made by those who were least supported or funded but driven by interesting imagination and curiosity. Recollected in this minireview is its establishing and developing processes, as well as its philosophy and accomplishments.

2. Keywords: Reform and opening; Free communication; Sweden; Cradle; San Diego; Boston; Door-opening

3. Introduction

The Gordon Life Science Institute was established in 2003 at San Diego of California, USA. Its founder is Professor Dr. Kuo-Chen Chou. Its mission is to develop and apply new mathematical tools and physical concepts for understanding biological phenomena. For more detailed about the growth process of Gordon Life Science Institute as well as its novel philosophy, see an article at <https://gordonlifescience.org/GordonLifeScience.html>.

The Gordon Life Science Institute is a newly emerging academic organization in the Age of Information and Internet. Founded by Professor Dr. Kuo-Chen Chou at San Diego of California, right after he was retired from Pfizer Global Research and Development in 2003. Its mission is to develop and apply new mathematical tools and physical concepts for understanding biological phenomena.

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. To make 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, meaning that "Reform and Opening" (改革开放) and "Free Communication" (自由交换信息) 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.

4. Mission and Organization

The Institute has no physical boundaries. Its members do not have to work in a same building or campus. Distributed over different countries of the world, they shall freely collaborate, exchange ideas, and share information and findings via a variety of modern communication methods. This versatile system allows the members to focus completely on science

without having to cope with 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 science and human beings. 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 society or organization for those scientists who are really dedicated themselves to science and loving science more than anything else. In the friendly door-opened Institute, these scientists can maximize their time and energy to engage in their scientific creativity.

Members of the Institute believe science would be more truthful and wonderful if scientists do not have to spend a lot of time on funding application. We also note that great scientific findings and creations in history were often made by those who were least supported or funded but driven by interesting imagination and curiosity.

5. Accomplishments

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) Kuo-Chen Chou for continuously 5 years (2014, 2015, 2016, 2017, and 2018), (2) Hong-Bin Shen (2014 and 2015), (3) Wei Chen (2018), (4) Hao Lin (2018), and (5) Xuan Xiao (2018).

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

5.1. Extension of special PseAAC to the general one

With the explosive growth of biological sequences in the post-genomic era, one of the most challenging problems in computational biology is how to express a biological sequence with a discrete model or a vector, yet still keep considerable sequence-order information or key pattern characteristic. This is because all the existing machine-learning algorithms (such as “Optimization” algorithm [1], “Covariance Discriminant” or “CD” algorithm [2,3], “Nearest Neighbor” or “NN” algorithm [4], and “Support Vector Machine” or “SVM” algorithm [4,5]) can only handle vectors as elaborated in a comprehensive review [6]. 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 [7] or PseAAC [8] was proposed. Ever since then, it has been widely used in nearly all the areas of computational proteomics [3,9-61][58-60,62-266]. Because it has been widely and increasingly used, four powerful open access soft-wares, called ‘PseAAC’ [267],

‘PseAAC-Builder’ [268], ‘propy’ [269], and ‘PseAAC-General’ [270], were established: the former three are for generating various modes of Chou’s special PseAAC [271]; while the 4th one for those of Chou’s general PseAAC [272], 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 [272]), “Gene Ontology” mode (see Eqs.11-12 of [272]), and “Sequential Evolution” or “PSSM” mode (see Eqs.13-14 of [272]). For more information about the PseAAC, please visit an insightful Wikipedia article at https://en.wikipedia.org/wiki/Pseudo_amino_acid_composition.

5.2. 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) [273] was developed for generating various feature vectors for DNA/RNA sequences that have proved very useful as well [273-290]. Particularly, in 2015 a very powerful web-server called ‘Pse-in-One’ [291] and its updated version ‘Pse-in-One2.0’ [292] 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”. For more information about the PseKNC, please visit an insightful Wikipedia article at https://en.wikipedia.org/wiki/Pseudo_K-tuple_nucleotide_composition.

5.3. Distorted key theory for peptide drugs

According to Fisher’s “lock and key” model [293], Koshland’s “induced fit” theory [293], and the “rack mechanism” [294], 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 [295]. For more discussion about the distorted key theory, see a comprehensive review

paper [296]. 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., [295,297-302]). Furthermore, a web-server called "HIVcleave" [299] has been established for predicting HIV protease cleavage sites in proteins. Its website address is at <http://chou.med.harvard.edu/bioinf/HIV/>. For more discussions about the "distorted key theory", see an insightful Wikipedia article at https://en.wikipedia.org/wiki/Chou%27s_distorted_key_theory_for_peptide_drugs.

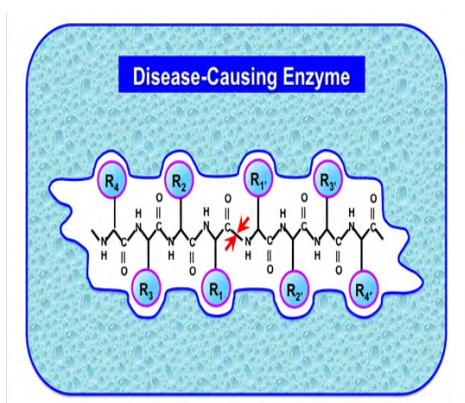


Figure 1: A schematic illustration to show a peptide in good fitting and tightly binding with the enzyme's active site before it is cleaved by the latter. Adapted from [296] with permission.

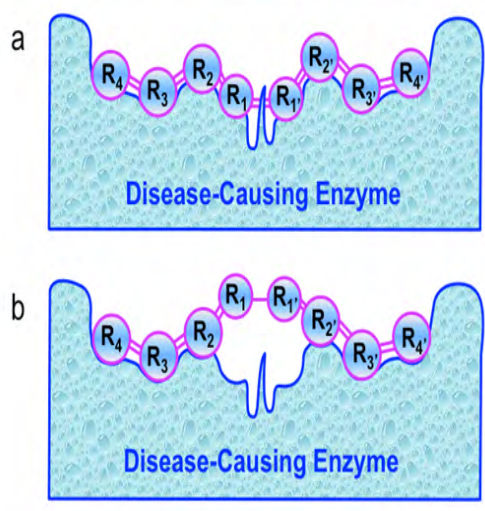


Figure 2: Schematic drawing to illustrate the "Distorted Key" theory, where panel

(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 [296] with permission.

5.4. 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., [303]). Its importance can also be seen in an insightful Wikipedia article at https://en.wikipedia.org/wiki/Graph_theory_in_enzymatic_kinetics.

Wenxiang diagram[304] is a special kind of graphical approach, which is very useful for in-depth studying protein-protein interaction mechanism [305, 306]. For more about the wenxiang diagram, see an insightful Wikipedia article at https://en.wikipedia.org/wiki/Wenxiang_diagram.

5.5. 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., [307-311]). 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 [312,313], which are often the special targets for drug development [313-319]). Therefore, how to deal with this kind of multi-label systems is a critical challenge. To take the challenge, the Institute has developed the following four series of predictors: (1) [313,320-326]; (2) [327-332]; (3) [203,204,215,224-226,333]; (4) [227-230,254,265,266]. All these predictors have yielded very high success rates, both globally and locally, as summarized in a comprehensive review paper [334]. For more about protein subcellular localization prediction, see an insightful Wikipedia article at https://en.wikipedia.org/wiki/Protein_subcellular_localization_prediction.

5.6. Five-Steps Rule

The Institute was the birth place of the famous 5-steps rule [272], which has been used in nearly all the areas of computational biology [203,204,215,224-230,233,251,254-256,259-261,264,265,277,279,288,333-368]), material science [369], and even the commercial science (e.g., the bank account systems). The only difference between the 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 like the case of many machine-learning algorithms. They can be widely used in nearly all the areas of statistical analysis. For more about the 5-steps rule, see an insightful Wikipedia article at https://en.wikipedia.org/wiki/5-step_rules

Working in such Institute filled with this kind of philosophy and atmosphere, the scientists would be more prone to be

stimulated by the eight pioneering papers from the then Chairman of Nobel Prize Committee Sture Forsen [370-377] and many of their follow-up papers [172,189,305,348,378-425], so as to render them substantially more creative and productive.

6. Conclusion and Perspective

In comparison with 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 interesting imagination and curiosity; and (5) guide their scientific results more truthful and wonderful.

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 (<https://hcr.clarivate.com/resources/archived-lists/>), 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 top in the world.

For more remarkable and awesome role of the Institute in stimulating the development of computational biology and drug development, see [334,348,418-420,423,424,426,427], where a series of insightful recollections at different angles or from various aspects have been clearly presented.

It is expected that more significant accomplishments will be achieved by the Gordon Life Science Institute for many years to come.

References

- ZhangCT, ChouKC. An optimization approach to predicting protein structural class from amino acid composition. *Protein Science*. 1992; 1: 401-8.
- ChouKC, ElrodDW. Bioinformatical analysis of G-protein-coupled receptors. *Journal of Proteome Research*. 2002; 1: 429-433.
- ChouKC, CaiYD. Prediction and classification of protein subcellular location: sequence-order effect and pseudo amino acid composition. *Journal of Cellular Biochemistry (Addendum, ibid.* 2004, 91, 1085). 2003; 90: 1250-60.
- Hu L, Huang T, Shi X, Lu WC, Cai YD, Chou KC. Predicting functions of proteins in mouse based on weighted protein-protein interaction network and protein hybrid properties PLoS ONE. 2011; 6: 14556.
- Cai YD, Feng KY, Lu WC, Chou KC. Using LogitBoost classifier to predict protein structural classes. *Journal of Theoretical Biology*. 2006; 238: 172-6.
- Chou KC. Impacts of bioinformatics to medicinal chemistry. *Medicinal Chemistry* 11: 218-34.
- Chou KC. Prediction of protein cellular attributes using pseudo amino acid composition. *PROTEINS: Structure, Function, and Genetics (Erratum: ibid., 2001, Vol.44, 60)*. 2001; 43: 246-55.
- Chou KC. Using amphiphilic pseudo amino acid composition to predict enzyme subfamily classes. *Bioinformatics*. 2005; 21: 10-9.
- Guo ZM. 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. 2002.
- Chou KC, Cai YD. Predicting protein quaternary structure by pseudo amino acid composition. *PROTEINS: Structure, Function, and Genetics*. 2003; 53: 282-9.
- PanYX, Zhang ZZ, Guo ZM, Feng GY, Huang ZD, He L. Application of pseudo amino acid composition for predicting protein subcellular location: stochastic signal processing approach. *Journal of Protein Chemistry*. 2003; 22: 395-402.
- Chou KC, Cai YD. Predicting subcellular localization of proteins by hybridizing functional domain composition and pseudo amino acid composition. *Journal of Cellular Biochemistry*. 2004; 91: 1197-203.
- Wang M, Yang J, Liu G, Xu ZJ, Chou KC. Weighted-support vector machines for predicting membrane protein types based on pseudo amino acid composition. *Protein Engineering, Design, and Selection*. 2004; 17: 509-16.
- Cai YD, Chou KC. Predicting enzyme subclass by functional domain composition and pseudo amino acid composition. *Journal of Proteome Research*. 2005; 4: 967-71.
- Cai YD, Zhou GP, Chou KC. Predicting enzyme family classes by hybridizing gene product composition and pseudo amino acid composition. *Journal of Theoretical Biology*. 2005; 234: 145-9.
- Gao Y, Shao SH, Xiao X, Ding YS, Huang YS, Huang ZD, Chou KC. Using pseudo amino acid composition to predict protein subcellular location: approached with Lyapunov index, Bessel function, and Chebyshev filter. *Amino Acids*. 2005; 28: 373-6.
- Liu H, Yang J, Wang M, Xue L, Chou KC. Using Fourier spectrum analysis and pseudo amino acid composition for prediction of membrane protein types. *The Protein Journal*. 2005; 24: 385-9.

18. Shen HB, Chou KC. Using optimized evidence-theoretic K-nearest neighbor classifier and pseudo amino acid composition to predict membrane protein types. *Biochemical & Biophysical Research Communications (BBRC)*. 2005; 334: 288-92.
19. Shen HB, Chou KC. Predicting protein subnuclear location with optimized evidence-theoretic K-nearest classifier and pseudo amino acid composition. *BiochemBiophys Res Comm. (BBRC)*. 2005; 337: 752-6.
20. Cai YD, Chou KC. Predicting membrane protein type by functional domain composition and pseudo amino acid composition. *Journal of Theoretical Biology*. 2006; 238: 395-400.
21. Chen C, Tian YX, Zou XY, Cai PX, Mo JX. Using pseudo amino acid composition and support vector machine to predict protein structural class. *J Theor Biol*. 2006; 243: 444-8.
22. Chen C, Zhou X, Tian Y, Zou X, Cai P, Predicting protein structural class with pseudo amino acid composition and support vector machine fusion network. *Analytical Biochemistry*. 2006; 357: 116-21.
23. Du P, Li Y. Prediction of protein submitochondria locations by hybridizing pseudo amino acid composition with various physicochemical features of segmented sequence. *BMC Bioinformatics*. 2006; 7: 518.
24. Mondal S, Bhavna R, Mohan Babu R, Ramakumar S. Pseudo amino acid composition and multi-class support vector machines approach for conotoxin superfamily classification. *J Theor Biol*. 2006; 243: 252-60.
25. Shen HB, Yang J, Chou KC. Fuzzy KNN for predicting membrane protein types from pseudo amino acid composition. *Journal of Theoretical Biology*. 2006 240: 9-13.
26. Wang SQ, Yang J, Chou KC. Using stacked generalization to predict membrane protein types based on pseudo amino acid composition. *Journal of Theoretical Biology*. 2006; 242: 941-6.
27. Xiao X, Shao SH, Ding YS, Huang ZD, Chou KC. Using cellular automata images and pseudo amino acid composition to predict protein subcellular location. *Amino Acids*. 2006; 30: 49-54.
28. Xiao X, Shao SH, Huang ZD, Chou KC. Using pseudo amino acid composition to predict protein structural classes: approached with complexity measure factor. *Journal of Computational Chemistry*. 2006; 27: 478-82.
29. Zhang SW, Pan Q, Zhang HC, Shao ZC, Shi JY. Prediction protein homo-oligomer types by pseudo amino acid composition: Approached with an improved feature extraction and naive Bayes feature fusion. *Amino Acids*. 2006; 30: 461-8.
30. Zhou GP, Cai YD. Predicting protease types by hybridizing gene ontology and pseudo amino acid composition. *PROTEINS: Structure, Function, and Bioinformatics*. 2006; 63: 681-4.
31. Chen YL, Li QZ. Prediction of apoptosis protein subcellular location using improved hybrid approach and pseudo amino acid composition. *Journal of Theoretical Biology*. 2007; 248: 377-81.
32. Ding YS, Zhang TL, Chou KC. Prediction of protein structure classes with pseudo amino acid composition and fuzzy support vector machine network. *Protein & Peptide Letters*. 2007; 14: 811-5.
33. Lin H, Li QZ. Predicting conotoxin superfamily and family by using pseudo amino acid composition and modified Mahalanobis discriminant. *BiochemBiophys Res Commun*. 2007; 354: 548-51.
34. Lin H, Li QZ. Using Pseudo Amino Acid Composition to Predict Protein Structural Class: Approached by Incorporating 400 Dipeptide Components. *Journal of Computational Chemistry*. 2007; 28: 1463-6.
35. Mundra P, Kumar M, Kumar KK, Jayaraman VK, Kulkarni BD. Using pseudo amino acid composition to predict protein subnuclear localization: Approached with PSSM. *Pattern Recognition Letters*. 2007; 28: 1610-5.
36. Shi JY, Zhang SW, Pan Q, Cheng YM, Xie J. Prediction of protein subcellular localization by support vector machines using multi-scale energy and pseudo amino acid composition. *Amino Acids*. 2007 33: 69-74.
37. Zhang TL, Ding YS. Using pseudo amino acid composition and binary-tree support vector machines to predict protein structural classes. *Amino Acids*. 2007; 33: 623-9.
38. Zhou XB, Chen C, Li ZC, Zou XY. Using Chou's amphiphilic pseudo amino acid composition and support vector machine for prediction of enzyme subfamily classes. *Journal of Theoretical Biology*. 2007; 248: 546-51.
39. Diao Y, Ma D, Wen Z, Yin J, Xiang J, Li M. Using pseudo amino acid composition to predict transmembrane regions in protein: cellular automata and Lempel-Ziv complexity. *Amino Acids*. 2008; 34: 111-7.
40. Ding YS, Zhang TL. 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*. 2008; 29: 1887-92.
41. Fang Y, Guo Y, Feng Y, Li M. Predicting DNA-binding proteins: approached from Chou's pseudo amino acid composition and other specific sequence features. *Amino Acids*. 2008; 34: 103-9.
42. Gu Q, Ding Y, Zhang Y. Prediction of G-protein-coupled receptor classes with pseudo amino acid composition. *IEEE Xplore(iCBBE)*,

Shanghai, China). 2008.

43. Jiang X, Wei R, Zhang TL, Gu Q. Using the concept of Chou's pseudo amino acid composition to predict apoptosis proteins subcellular location: an approach by approximate entropy. *Protein & Peptide Letters*. 2008; 15: 392-6.

44. Jiang X, Wei R, Zhao Y, Zhang Y. Using Chou's pseudo amino acid composition based on approximate entropy and an ensemble of AdaBoost classifiers to predict protein subnuclear location. *Amino Acids*. 2008; 34: 669-75.

45. Li FM, Li QZ. Using pseudo amino acid composition to predict protein subnuclear location with improved hybrid approach. *Amino Acids*. 2008; 34: 119-25.

46. Li FM, Li QZ. Predicting protein subcellular location using Chou's pseudo amino acid composition and improved hybrid approach. *Protein & Peptide Letters*. 2008; 15: 612-6.

47. Lin H. The modified Mahalanobis discriminant for predicting outer membrane proteins by using Chou's pseudo amino acid composition. *Journal of Theoretical Biology*. 2008; 252: 350-6.

48. Lin H, Ding H, Feng-Biao Guo FB, Zhang AY, Huang J. Predicting subcellular localization of mycobacterial proteins by using Chou's pseudo amino acid composition. *Protein & Peptide Letters*. 2008; 15: 739-44.

49. Shi JY, Zhang SW, Pan Q, Zhou GP. Using Pseudo Amino Acid Composition to Predict Protein Subcellular Location: Approached with Amino Acid Composition Distribution. *Amino Acids*. 2008; 35: 321-7.

50. Xiao X, Lin WZ, Chou KC. Using grey dynamic modeling and pseudo amino acid composition to predict protein structural classes. *Journal of Computational Chemistry*. 2008; 29: 2018-24.

51. Xiao X, Wang P, Chou KC. Predicting protein structural classes with pseudo amino acid composition: an approach using geometric moments of cellular automaton image. *Journal of Theoretical Biology*. 2008; 254: 691-6.

52. Zhang GY, Fang BS. Predicting the cofactors of oxidoreductases based on amino acid composition distribution and Chou's amphiphilic pseudo amino acid composition. *Journal of Theoretical Biology*. 2008; 253: 310-5.

53. Zhang GY, Li HC, Gao JQ, Fang BS. Predicting lipase types by improved Chou's pseudo amino acid composition. *Protein & Peptide Letters*. 2008; 15: 1132-7.

54. Zhang SW, Chen W, Yang F, Pan Q. Using Chou's pseudo amino acid composition to predict protein quaternary structure: a sequence-

segmented PseAAC approach. *Amino Acids*. 2008; 35: 591-8.

55. Zhang SW, Zhang YL, Yang HF, Zhao CH, Pan Q. 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*. 2008; 34: 565-72.

56. Zhang TL, Ding YS, Chou KC. Prediction protein structural classes with pseudo amino acid composition: approximate entropy and hydrophobicity pattern. *Journal of Theoretical Biology*. 2008; 250: 186-93.

57. Chen C, Chen L, Zou X, Cai P. Prediction of protein secondary structure content by using the concept of Chou's pseudo amino acid composition and support vector machine. *Protein & Peptide Letters*. 2009; 16: 27-31.

58. Ding H, Luo L, Lin H. Prediction of cell wall lytic enzymes using Chou's amphiphilic pseudo amino acid composition. *Protein & Peptide Letters*. 2009; 16: 351-5.

59. Du P, Cao S, Li Y. SubChlo: predicting protein subchloroplast locations with pseudo amino acid composition and the evidence-theoretic K-nearest neighbor (ET-KNN) algorithm. *Journal of Theoretical Biology*. 2009; 261: 330-335.

60. Gao QB, Jin ZC, Ye XF, Wu C, He J. Prediction of nuclear receptors with optimal pseudo amino acid composition. *Analytical Biochemistry*. 2009; 387: 54-9.

61. Georgiou DN, Karakasidis TE, Nieto JJ, Torres A. 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*. 2009; 257: 17-26.

62. Li ZC, Zhou XB, Dai Z, Zou XY. Prediction of protein structural classes by Chou's pseudo amino acid composition: approached using continuous wavelet transform and principal component analysis. *Amino Acids*. 2009; 37: 415-25.

63. Lin H, Wang H, Ding H, Chen YL, Li QL. Prediction of Subcellular Localization of Apoptosis Protein Using Chou's Pseudo Amino Acid Composition. *Acta Biotheoretica*. 2009; 57: 321-30.

64. Qiu JD, Huang JH, Liang RP, Lu XQ. 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*. 2009; 390: 68-73.

65. Xiao X, Wang P, Chou KC. Predicting protein quaternary structural attribute by hybridizing functional domain composition and pseudo amino acid composition. *Journal of Applied Crystallography*. 2009; 42:

169-73.

66. Zeng YH, Guo YZ, Xiao RQ, Yang L, Yu LZ, Li ML. Using the augmented Chou's pseudo amino acid composition for predicting protein submitochondria locations based on auto covariance approach. *Journal of Theoretical Biology*. 2009; 259: 366-72.

67. Esmaili M, Mohabatkar H, Mohsenzadeh S. Using the concept of Chou's pseudo amino acid composition for risk type prediction of human papillomaviruses. *Journal of Theoretical Biology*. 2010; 263: 203-9.

68. Gao QB, Ye XF, Jin ZC, He J. Improving discrimination of outer membrane proteins by fusing different forms of pseudo amino acid composition. *Analytical Biochemistry*. 2010; 398: 52-9.

69. Gu Q, Ding Y, Zhang T, Shen Y. Prediction of G-protein-coupled receptor classes with pseudo amino acid composition. *Journal of biomedical engineering*. 2010; 27: 500-4.

70. Gu Q, Ding YS, Zhang TL. 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*. 2010; 17: 559-67.

71. Kandaswamy KK, Pugalenti G, Moller S, Hartmann E, Kalies KU, Suganthan PN, et al. 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*. 2010; 17: 1473-9.

72. Liu T, Zheng X, Wang C, Wang J. Prediction of Subcellular Location of Apoptosis Proteins using Pseudo Amino Acid Composition: An Approach from Auto Covariance Transformation. *Protein & Peptide Letters*. 2010; 17: 1263-9.

73. Mohabatkar H. Prediction of cyclin proteins using Chou's pseudo amino acid composition. *Protein & Peptide Letters*. 2010; 17: 1207-14.

74. Nanni L, Brahnam S, Lumini A. High performance set of PseAAC and sequence based descriptors for protein classification. *Journal of Theoretical Biology*. 2010; 266: 1-10.

75. Niu XH, Li NN, Shi F, Hu XH, Xia JB, Xiong HJ. Predicting protein solubility with a hybrid approach by pseudo amino Acid composition. *Protein and Peptide Letters*. 2010; 17: 1466-72.

76. Qiu JD, Huang JH, Shi SP, Liang RP. 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*. 2010; 17: 715-22.

77. Sahu SS, Panda G. A novel feature representation method

based on Chou's pseudo amino acid composition for protein structural class prediction. *Computational Biology and Chemistry*. 2010; 34: 320-7.

78. Wang YC, Wang XB, YangZX, Deng NY. Prediction of enzyme subfamily class via pseudo amino acid composition by incorporating the conjoint triad feature. *Protein & Peptide Letters*. 2010; 17: 1441-9.

79. Wu J, Li ML, Yu LZ, Wang C. An ensemble classifier of support vector machines used to predict protein structural classes by fusing auto covariance and pseudo amino acid composition. *Protein J*. 2010; 29: 62-7.

80. Yu L, Guo Y, Li Y, Li G, Li M, Luo J, et al. SecretP: Identifying bacterial secreted proteins by fusing new features into Chou's pseudo amino acid composition. *Journal of Theoretical Biology*. 2010; 267: 1-6.

81. Ding H, Liu L, Guo FB, Huang J, Lin H. Identify Golgi protein types with modified mahalanobis discriminant algorithm and pseudo amino acid composition. *Protein & Peptide Letters*. 2011; 18: 58-63.

82. Guo J, Rao N, Liu G, Yang Y, Wang G. Predicting protein folding rates using the concept of Chou's pseudo amino acid composition. *Journal of Computational Chemistry*. 2011; 32: 1612-7.

83. Hayat M, Khan A. Predicting membrane protein types by fusing composite protein sequence features into pseudo amino acid composition. *Journal of Theoretical Biology*. 2011; 271: 10-7.

84. Hu L, Zheng L, Wang Z, Li B, Liu L. Using pseudo amino Acid composition to predict protease families by incorporating a series of protein biological features. *Protein and Peptide Letters*. 2011; 18: 552-8.

85. Huang Y, Yang L, Wang T. Phylogenetic analysis of DNA sequences based on the generalized pseudo amino acid composition. *Journal of Theoretical Biology*. 2011; 269: 217-23.

86. Jingbo X, Silan Z, Feng S, Huijuan X, Xuehai H, Xiaohui N, et al. Using the concept of pseudo amino acid composition to predict resistance gene against *Xanthomonas oryzae* in rice: An approach from chaos games representation. *Journal of Theoretical Biology*. 2011; 284: 16-23.

87. Liao B, Jiang JB, Zeng QG, Zhu W. Predicting Apoptosis Protein Subcellular Location with PseAAC by Incorporating Tripeptide Composition. *Protein & Peptide Letters*. 2011; 18: 1086-92.

88. Lin H, Ding H. Predicting ion channels and their types by the dipeptide mode of pseudo amino acid composition. *Journal of*

Theoretical Biology. 2011; 269: 64-9.

89. Lin J, Wang Y. Using a novel AdaBoost algorithm and Chou's pseudo amino acid composition for predicting protein subcellular localization. *Protein & Peptide Letters*. 2011; 18: 1219-25.

90. Lin J, Wang Y, Xu X. A novel ensemble and composite approach for classifying proteins based on Chou's pseudo amino acid composition. *African Journal of Biotechnology*. 2011; 10: 16963-8.

91. Liu XL, Lu JL, Hu XH. Predicting Thermophilic Proteins with Pseudo Amino Acid Composition: Approached from Chaos Game Representation and Principal Component Analysis. *Protein & Peptide Letters*. 2011; 18: 1244-50.

92. Mahdavi A, Jahandideh S. Application of density similarities to predict membrane protein types based on pseudo amino acid composition. *Journal of Theoretical Biology*. 2011; 276: 132-7.

93. Mohabatkar H, Mohammad Beigi M, Esmaeili A. Prediction of GABA(A) receptor proteins using the concept of Chou's pseudo amino acid composition and support vector machine. *Journal of Theoretical Biology*. 2011; 281: 18-23.

94. Mohammad BM, Behjati M, Mohabatkar H. 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*. 2011; 12: 191-7.

95. Qiu JD, Sun XY, Suo SB, Shi SP, Huang SY, Liang RP, et al. Predicting homo-oligomers and hetero-oligomers by pseudo amino acid composition: an approach from discrete wavelet transformation. *Biochimie*. 2011; 93: 1132-8.

96. Qiu JD, Suo SB, Sun XY, Shi SP, Liang RP. OligoPred: A web-server for predicting homo-oligomeric proteins by incorporating discrete wavelet transform into Chou's pseudo amino acid composition. *Journal of Molecular Graphics & Modelling*. 2011; 30: 129-34.

97. Shi R, Xu C. Prediction of rat protein subcellular localization with pseudo amino acid composition based on multiple sequential features. *Protein and Peptide Letters*. 2011; 18: 625-33.

98. Shu M, Cheng X, Zhang Y, Wang Y, Lin Y, Wang L, et al. Predicting the Activity of ACE Inhibitory Peptides with a Novel Mode of Pseudo Amino Acid Composition. *Protein & Peptide Letters*. 2011; 18: 1233-43.

99. Wang D, Yang L, Fu Z, Xia J. Prediction of thermophilic protein with pseudo amino acid composition: an approach from combined feature selection and reduction. *Protein & Peptide Letters*. 2011; 18: 684-9.

100. Wang W, Geng XB, Dou Y, Liu T, Zheng X. Predicting

protein subcellular localization by pseudo amino acid composition with a segment-weighted and features-combined approach. *Protein and Peptide Letters*. 2011; 18: 480-7.

101. Xiao X, Chou KC. Using pseudo amino acid composition to predict protein attributes via cellular automata and other approaches. *Current Bioinformatics*. 2011; 6: 251-60.

102. Xiao X, Wang P, Chou KC. GPCR-2L: Predicting G protein-coupled receptors and their types by hybridizing two different modes of pseudo amino acid compositions. *Molecular Biosystems*. 2011; 7: 911-9.

103. Zia Ur R, Khan A. Prediction of GPCRs with Pseudo Amino Acid Composition: Employing Composite Features and Grey Incidence Degree Based Classification. *Protein & Peptide Letters*. 2011; 18: 872-8.

104. Zou D, He Z, He J, Xia Y. Supersecondary structure prediction using Chou's pseudo amino acid composition. *Journal of Computational Chemistry*. 2011; 32: 271-8.

105. Cao JZ, Liu WQ, GuH. 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*. 2012; 19: 1163-9.

106. Chen C, Shen ZB, Zou XY. Dual-Layer Wavelet SVM for Predicting Protein Structural Class Via the General Form of Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*. 2012; 19: 422-9.

107. Chen YL, Li QZ, Zhang LQ. Using increment of diversity to predict mitochondrial proteins of malaria parasite: integrating pseudo amino acid composition and structural alphabet. *Amino Acids*. 2012; 42: 1309-16.

108. Fan GL, Li QZ. 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*. 2012; 304: 88-95.

109. Fan GL, Li QF. Predicting protein submitochondria locations by combining different descriptors into the general form of Chou's pseudo amino acid composition. *Amino Acids*. 2012; 43: 545-55.

110. Gao QB, Zhao H, Ye X, He J. Prediction of pattern recognition receptor family using pseudo amino acid composition. *Biochemical and Biophysical Research Communications*. 2012; 417: 73-7.

111. Hayat M, Khan A. Discriminating Outer Membrane

Proteins with Fuzzy K-Nearest Neighbor Algorithms Based on the General Form of Chou's PseAAC. *Protein & Peptide Letters*. 2012; 19: 411-21.

112. Li LQ, Zhang Y, Zou LY, Zhou Y, Zheng XQ. Prediction of Protein Subcellular Multi-Localization Based on the General form of Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*. 2012; 19: 375-87.

113. Liao B, Xiang Q, Li D. Incorporating Secondary Features into the General form of Chou's PseAAC for Predicting Protein Structural Class. *Protein & Peptide Letters*. 2012; 19: 1133-8.

114. Lin WZ, Fang JA, Xiao X, Chou KC. Predicting Secretory Proteins of Malaria Parasite by Incorporating Sequence Evolution Information into Pseudo Amino Acid Composition via Grey System Model. *PLoS One*. 2012; 7: e49040.

115. Liu L, Hu XZ, Liu XX, Wang Y, Li SB. Predicting Protein Fold Types by the General Form of Chou's Pseudo Amino Acid Composition: Approached from Optimal Feature Extractions. *Protein & Peptide Letters*. 2012; 19: 439-49.

116. Mei S. Multi-kernel transfer learning based on Chou's PseAAC formulation for protein submitochondria localization. *Journal of Theoretical Biology*. 2012; 293: 121-30.

117. Mei S. Predicting plant protein subcellular multi-localization by Chou's PseAAC formulation based multi-label homolog knowledge transfer learning. *Journal of Theoretical Biology*. 2012; 310: 80-7.

118. Nanni L, Brahnam S, Lumini A. Wavelet images and Chou's pseudo amino acid composition for protein classification. *Amino Acids*. 2012; 43: 657-65.

119. Nanni L, Lumini A, Gupta D, Garg A. 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*. 2012; 9: 467-75.

120. Niu XH, Hu XH, Shi F, Xia JB. 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*. 2012; 19: 940-8.

121. Qin YF, Wang CH, Yu XQ, Zhu J, Liu TG, Zheng XQ. Predicting Protein Structural Class by Incorporating Patterns of Over-Represented k-mers into the General form of Chou's PseAAC. *Protein & Peptide Letters*. 2012; 19: 388-97.

122. Ren LY, Zhang YS, Gutman I. 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. 2012; 19: 1170-6.

123. Sun XY, Shi SP, Qiu JD, Suo SB, Huang SY, Liang RP. Identifying protein quaternary structural attributes by incorporating physicochemical properties into the general form of Chou's PseAAC via discrete wavelet transform. *Molecular BioSystems*. 2012; 8: 3178-84.

124. Wang J, Li Y, Wang Q, You X, Man J, Wang C, et al. ProClusEnsem: predicting membrane protein types by fusing different modes of pseudo amino acid composition. *Comput Biol Med*. 2012; 42: 564-74.

125. Yu X, Zheng X, Liu T, Dou Y, Wang J. Predicting subcellular location of apoptosis proteins with pseudo amino acid composition: approach from amino acid substitution matrix and auto covariance transformation. *Amino Acids*. 2012; 42: 1619-25.

126. Zhao XW, Ma ZQ, Yin MH. Predicting protein-protein interactions by combing various sequence- derived features into the general form of Chou's Pseudo amino acid composition. *Protein & Peptide Letters*. 2012; 19: 492-500.

127. Zia-ur-Rehman, A. Khan, 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*. 2012; 19: 890-903.

128. Chang TH, Wu LC, Lee TY, Chen SP, Huang HD, Horng JT. 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*. 2013; 27: 91-103.

129. Chen YK, Li KB. 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*. 2013; 318: 1-12.

130. Fan GL, Li QZ, Zuo YC. 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*. 2013; 48: 1048-53.

131. Fan GL, Li QZ. 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*. 2013; 334: 45-51.

132. Georgiou DN, Karakasidis TE, Megaritis AC. A short survey on genetic sequences, Chou's pseudo amino acid composition and its combination with fuzzy set theory. *The Open Bioinformatics Journal*.

2013; 7: 41-8.

133. Gupta MK, Niyogi R, Misra M. An alignment-free method to find similarity among protein sequences via the general form of Chou's pseudo amino acid composition. *SAR QSAR Environ Res*. 2013; 24: 597-609.

134. Huang C, Yuan J. 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*. 2013; 113: 50-7.

135. Huang C, Yuan JQ. A multilabel model based on Chou's pseudo amino acid composition for identifying membrane proteins with both single and multiple functional types. *J Membr Biol*. 2013; 246: 327-34.

136. Huang C, Yuan JQ. 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*. 2013; 335: 205-12.

137. Khosravi M, Faramarzi FK, Beigi MM, Behbahani M, Mohabatkar H. Predicting Antibacterial Peptides by the Concept of Chou's Pseudo Amino Acid Composition and Machine Learning Methods. *Protein & Peptide Letters*. 2013; 20: 180-6.

138. Lin H, Ding C, Yuan LF, Chen W, Ding H, Li ZQ, et al. 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*. 2013; 6: 1350003.

139. Liu B, Wang X, Zou Q, Dong Q, Chen Q. Protein remote homology detection by combining Chou's pseudo amino acid composition and profile-based protein representation. *Molecular Informatics*. 2013; 32: 775-82.

140. Mohabatkar H, Beigi MM, Abdolahi A, Mohsenzadeh S. Prediction of Allergenic Proteins by Means of the Concept of Chou's Pseudo Amino Acid Composition and a Machine Learning Approach. *Medicinal Chemistry*. 2013; 9: 133-7.

141. Pacharawongsakda E, Theeramunkong T. 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*. 2013; 12: 311-20.

142. Qin YF, Zheng L, Huang J. 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*. 2013; 113: 1660-7.

143. Sarangi AN, Lohani M, Aggarwal R. Prediction of Essential

Proteins in Prokaryotes by Incorporating Various Physico-chemical Features into the General form of Chou's Pseudo Amino Acid Composition. *ProteinPept Lett*. 2013; 20: 781-95.

144. Wan S, Mak MW, Kung SY. 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*. 2013; 323: 40-8.

145. Wang X, Li GZ, Lu WC. 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*. 2013; 20: 309-17.

146. Xiao X, Min JL, Wang P, Chou KC. iCDI-PseFpt: Identify the channel-drug interaction in cellular networking with PseAAC and molecular fingerprints. *Journal of Theoretical Biology*. 2013; 337: 71-9.

147. Xiaohui N, Nana L, Jingbo X, Dingyan C, Yuehua P, Yang X, et al. 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*. 2013; 332: 211-7.

148. Xie HL, Fu L, Nie XD. Using ensemble SVM to identify human GPCRs N-linked glycosylation sites based on the general form of Chou's PseAAC. *Protein EngDes Sel*. 2013; 26: 735-42.

149. Xu Y, Ding J, Wu LY, Chou KC. iSNO-PseAAC: Predict cysteine S-nitrosylation sites in proteins by incorporating position specific amino acid propensity into pseudo amino acid composition. *PLoS ONE*. 2013; 8: e55844.

150. Xu Y, Shao XJ, Wu LY, Deng NJ, Chou KC. iSNO-AAPair: incorporating amino acid pairwise coupling into PseAAC for predicting cysteine S-nitrosylation sites in proteins. *PeerJ*. 2013; 1: e171.

151. Hajisharifi Z, Piryaiee M, Mohammad Beigi M, Behbahani M, Mohabatkar H. Predicting anticancer peptides with Chou's pseudo amino acid composition and investigating their mutagenicity via Ames test. *Journal of Theoretical Biology*. 2014; 341: 34-40.

152. Han GS, Yu ZG, Anh V. 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. *J Theor Biol*. 2014; 344: 31-9.

153. Hayat M, Iqbal N. Discriminating protein structure classes by incorporating Pseudo Average Chemical Shift to Chou's general PseAAC and Support Vector Machine. *Comput Methods Programs Biomed*. 2014; 116: 184-92.

154. Jia C, Lin X, Wang Z. Prediction of Protein S-Nitrosylation

Sites Based on Adapted Normal Distribution Bi-Profile Bayes and Chou's Pseudo Amino Acid Composition. *Int J Mol Sci.* 2014; 15: 10410-23.

155. Kong L, Zhang L, Lv J. Accurate prediction of protein structural classes by incorporating predicted secondary structure information into the general form of Chou's pseudo amino acid composition. *J Theor Biol.* 2014; 344: 12-8.

156. L. Li, S. Yu, W. Xiao, Y. Li, M. Li, L. Huang, X. Zheng, S. Zhou, H. Yang, Prediction of bacterial protein subcellular localization by incorporating various features into Chou's PseAAC and a backward feature selection approach. *Biochimie* 104 (2014) 100-7.

157. Liu B, Xu J, Lan X, Xu R, Zhou J, Wang X, et al. 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.* 2014; 9: e106691.

158. Mondal S, Pai PP. Chou's pseudo amino acid composition improves sequence-based antifreeze protein prediction. *J Theor Biol.* 2014; 356: 30-5.

159. Nanni L, Brahnam S, Lumini A. Prediction of protein structure classes by incorporating different protein descriptors into general Chou's pseudo amino acid composition. *J Theor Biol.* 2014; 360: 109-16.

160. Qiu WR, Xiao X, Chou KC. iRSpot-TNCPseAAC: Identify recombination spots with trinucleotide composition and pseudo amino acid components. *Int J MolSci (IJMS).* 2014; 15: 1746-66.

161. Qiu WR, Xiao X, Lin WZ, Chou KC. iMethyl-PseAAC: Identification of Protein Methylation Sites via a Pseudo Amino Acid Composition Approach. *Biomed Res Int (BMRI).* 2014; 2014: 947416.

162. Xu Y, Wen X, Shao XJ, Deng NY, Chou KC. 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).* 2014; 15: 7594-610.

163. Xu Y, Wen X, Wen LS, Wu LY, Deng NY, Chou KC. iNitro-Tyr: Prediction of nitrotyrosine sites in proteins with general pseudo amino acid composition. *PLoS ONE.* 2014; 9: e105018.

164. Zhang J, Sun P, Zhao X, Ma Z. PECM: Prediction of extracellular matrix proteins using the concept of Chou's pseudo amino acid composition. *Journal of Theoretical Biology.* 2014; 363: 412-8.

165. Zhang J, Zhao X, Sun P, Ma Z. PSNO: Predicting Cysteine S-Nitrosylation Sites by Incorporating Various Sequence-Derived Features into the General Form of Chou's PseAAC. *Int JMol Sci.* 2014;

15: 11204-19.

166. Zhang L, Zhao X, Kong L. Predict protein structural class for low-similarity sequences by evolutionary difference information into the general form of Chou's pseudo amino acid composition. *J Theor Biol.* 2014; 355: 105-10.

167. Ahmad S, Kabir M, Hayat M. Identification of Heat Shock Protein families and J-protein types by incorporating Dipeptide Composition into Chou's general PseAAC. *Comput Methods Programs Biomed.* 2015; 122: 165-74.

168. Ali F, Hayat M. Classification of membrane protein types using Voting Feature Interval in combination with Chou's Pseudo Amino Acid Composition. *J Theor Biol.* 2015 384: 78-83.

169. Chen L, Chu C, Huang T, Kong X, Cai YD. Prediction and analysis of cell-penetrating peptides using pseudo amino acid composition and random forest models. *Amino Acids.* 2015.

170. Dehzangi, Heffernan R, Sharma A, Lyons J, Paliwal K, Sattar A. Gram-positive and Gram-negative protein subcellular localization by incorporating evolutionary-based descriptors into Chou's general PseAAC. *J Theor Biol.* 2015; 364: 284-94.

171. Huang, Yuan JQ. Simultaneously Identify Three Different Attributes of Proteins by Fusing their Three Different Modes of Chou's Pseudo Amino Acid Compositions. *Protein Pept Lett.* 2015; 22: 547-56.

172. Jia J, Liu Z, Xiao X, Chou KC. iPPI-Esml: an ensemble classifier for identifying the interactions of proteins by incorporating their physicochemical properties and wavelet transforms into PseAAC. *J Theor Biol.* 2015; 377: 47-56.

173. Ju Z, Cao JZ, Gu H. iLM-2L: A two-level predictor for identifying protein lysine methylation sites and their methylation degrees by incorporating K-gap amino acid pairs into Chou's general PseAAC. *J Theor Biol.* 2015; 385: 50-7.

174. Khan ZU, Hayat M, Khan MA. Discrimination of acidic and alkaline enzyme using Chou's pseudo amino acid composition in conjunction with probabilistic neural network model. *J Theor Biol.* 2015; 365: 197-203.

175. Kumar R, Srivastava A, Kumari B, Kumar M. Prediction of beta-lactamase and its class by Chou's pseudo amino acid composition and support vector machine. *J Theor Biol.* 2015; 365: 96-103.

176. Liu B, Chen J, Wang X. Protein remote homology detection by combining Chou's distance-pair pseudo amino acid composition and principal component analysis. *Mol Genet Genomics.* 2015; 290: 1919-31.

177. Liu B, Xu J, Fan S, Xu R, Jiyun Zhou J, et al. PseDNA-Pro: DNA-binding protein identification by combining Chou's PseAAC and physicochemical distance transformation. *Molecular Informatics*. 2015; 34: 8-17.
178. Mandal M, Mukhopadhyay A, Maulik U. Prediction of protein subcellular localization by incorporating multiobjective PSO-based feature subset selection into the general form of Chou's PseAAC. *Med BiolEngComput*. 2015; 53: 331-44.
179. Sanchez V, Peinado AM, Perez-Cordoba JL, Gomez AM. A new signal characterization and signal-based Chou's PseAAC representation of protein sequences. *J BioinformComput Biol*. 2015; 13: 1550024.
180. Sharma R, Dehzangi A, Lyons J, Paliwal K, Tsunoda T, Sharma A. Predict Gram-Positive and Gram-Negative Subcellular Localization via Incorporating Evolutionary Information and Physicochemical Features Into Chou's General PseAAC. *IEEE Trans Nanobioscience*. 2015; 14: 915-26.
181. Wang X, Zhang W, Zhang Q, Li GZ. MultiP-SChlo: multi-label protein subchloroplast localization prediction with Chou's pseudo amino acid composition and a novel multi-label classifier. *Bioinformatics*. 2015; 31: 2639-45.
182. Xu R, Zhou J, Liu B, He YA, Zou Q, Wang X, et al. Identification of DNA-binding proteins by incorporating evolutionary information into pseudo amino acid composition via the top-n-gram approach. *Journal of BiomolecularStructure & Dynamics (JBSD)*. 2015; 33: 1720-30.
183. Zhang M, Zhao B, Liu X. Predicting industrial polymer melt index via incorporating chaotic characters into Chou's general PseAAC. *Chemometrics and Intelligent Laboratory Systems (CHEMOLAB)*. 2015; 146: 232-40.
184. Zhang SL. Accurate prediction of protein structural classes by incorporating PSSS and PSSM into Chou's general PseAAC. *Chemometrics and Intelligent Laboratory Systems (CHEMOLAB)*. 2015; 142: 28-35.
185. Zhu PP, Li WC, Zhong ZJ, Deng EZ, Ding H, Chen W, et al. Predicting the subcellular localization of mycobacterial proteins by incorporating the optimal tripeptides into the general form of pseudo amino acid composition. *MolBiosyst*. 2015; 11: 558-63.
186. Ahmad K, Waris M, Hayat M. Prediction of Protein Submitochondrial Locations by Incorporating Dipeptide Composition into Chou's General Pseudo Amino Acid Composition. *J Membr Biol*. 2016; 249: 293-304.
187. Behbahani M, Mohabatkar H, Nosrati M. Analysis and comparison of lignin peroxidases between fungi and bacteria using three different modes of Chou's general pseudo amino acid composition. *J Theor Biol*. 2016; 411: 1-5.
188. Fan GL, Liu YL, Wang H. Identification of thermophilic proteins by incorporating evolutionary and acid dissociation information into Chou's general pseudo amino acid composition. *J Theor Biol*. 2016; 407: 138-42.
189. Jia J, Liu Z, Xiao X, Liu B, Chou KC. Identification of protein-protein binding sites by incorporating the physicochemical properties and stationary wavelet transforms into pseudo amino acid composition (iPPBS-PseAAC). *J BiomolStructDyn (JBSD)*. 2016; 34: 1946-1.
190. Jia J, Liu Z, Xiao X, Liu B, Chou KC. pSuc-Lys: Predict lysine succinylation sites in proteins with PseAAC and ensemble random forest approach. *Journal of Theoretical Biology*. 2016; 394: 223-30.
191. Jia J, Liu Z, Xiao X, Liu B, Chou KC. iCar-PseCp: identify carbonylation sites in proteins by Monto Carlo sampling and incorporating sequence coupled effects into general PseAAC. *Oncotarget*. 2016; 7: 34558-70.
192. Jia J, Zhang L, Liu Z, Xiao X, Chou KC. pSumo-CD: Predicting sumoylation sites in proteins with covariance discriminant algorithm by incorporating sequence-coupled effects into general PseAAC. *Bioinformatics*. 2016; 32: 3133-41.
193. Jiao YS, Du PF. Prediction of Golgi-resident protein types using general form of Chou's pseudo amino acid compositions: Approaches with minimal redundancy maximal relevance feature selection. *J Theor Biol*. 2016; 402: 38-44.
194. Ju Z, Cao JZ, Gu H. Predicting lysine phosphoglycerylation with fuzzy SVM by incorporating k-spaced amino acid pairs into Chou's general PseAAC. *J Theor Biol*. 2016; 397: 145-50.
195. Kabir M, Hayat M. iRSpot-GAEnsC: identifying recombination spots via ensemble classifier and extending the concept of Chou's PseAAC to formulate DNA samples. *Molecular Genetics and Genomics*. 2016; 291: 285-96.
196. Qiu WR, Sun BQ, Xiao X, Xu ZC, Chou KC. iHyd-PseCp: Identify hydroxyproline and hydroxylysine in proteins by incorporating sequence-coupled effects into general PseAAC. *Oncotarget*. 2016; 7: 44310-21.
197. Tahir M, Hayat M. iNuc-STNC: a sequence-based predictor for identification of nucleosome positioning in genomes by extending the concept of SAAC and Chou's PseAAC. *MolBiosyst*. 2016; 12: 2587-93.
198. Tang H, Chen W, Lin H. Identification of immunoglobulins

using Chou's pseudo amino acid composition with feature selection technique. *MolBiosyst.* 2016; 12: 1269-75.

199. Tiwari AK. Prediction of G-protein coupled receptors and their subfamilies by incorporating various sequence features into Chou's general PseAAC. *Comput Methods Programs Biomed.* 2016; 134: 197-213.

200. Xu C, Sun D, Liu S, Zhang Y. Protein Sequence Analysis by Incorporating Modified Chaos Game and Physicochemical Properties into Chou's General Pseudo Amino Acid Composition. *J Theor Biol.* 2016; 406: 105-15.

201. Zou HL, Xiao X. Predicting the Functional Types of Singleplex and Multiplex Eukaryotic Membrane Proteins via Different Models of Chou's Pseudo Amino Acid Compositions. *J Membr Biol.* 2016; 249: 23-9.

202. Zou HL, Xiao X. Classifying Multifunctional Enzymes by Incorporating Three Different Models into Chou's General Pseudo Amino Acid Composition. *J Membr Biol.* 2016; 249: 561-7.

203. Cheng X, Xiao X, Chou KC. pLoc-mPlant: predict subcellular localization of multi-location plant proteins via incorporating the optimal GO information into general PseAAC. *Molecular BioSystems.* 2017; 13: 1722-7.

204. Cheng X, Xiao X, Chou KC. pLoc-mVirus: predict subcellular localization of multi-location virus proteins via incorporating the optimal GO information into general PseAAC. *Gene.* 2017; 628: 315-21.

205. Ju Z, He JJ. Prediction of lysine propionylation sites using biased SVM and incorporating four different sequence features into Chou's PseAAC. *J Mol Graph Model.* 2017; 76: 356-63.

206. Ju Z, He JJ. Prediction of lysine crotonylation sites by incorporating the composition of k-spaced amino acid pairs into Chou's general PseAAC. *J Mol Graph Model.* 2017; 77: 200-4.

207. Khan M, Hayat M, Khan SA, Iqbal N. Unb-DPC: Identify mycobacterial membrane protein types by incorporating un-biased dipeptide composition into Chou's general PseAAC. *J Theor Biol.* 2017; 415: 13-9.

208. Liang Y, Zhang S. Predict protein structural class by incorporating two different modes of evolutionary information into Chou's general pseudo amino acid composition. *J Mol Graph Model.* 2017; 78: 110-7.

209. Liu LM, Xu Y, Chou KC. iPGK-PseAAC: identify lysine phosphoglycerlation sites in proteins by incorporating four different tiers of amino acid pairwise coupling information into the general PseAAC. *Med Chem.* 2017; 13: 552-9.

210. Meher PK, Sahu TK, Saini V, Rao AR. Predicting antimicrobial peptides with improved accuracy by incorporating the compositional, physico-chemical and structural features into Chou's general PseAAC. *Sci Rep.* 2017; 7: 42362.

211. Qiu WR, Sun BQ, Xiao X, Xu D, Chou KC. iPhos-PseEvo: Identifying human phosphorylated proteins by incorporating evolutionary information into general PseAAC via grey system theory. *Molecular Informatics.* 2017; 36: 1600010.

212. Qiu WR, Zheng QS, Sun BQ, Xiao X. 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.* 2017; 36: 1600085.

213. Rahimi M, Bakhtiarzadeh MR, Mohammadi-Sangcheshmeh A. Oogenesis_Pred: A sequence-based method for predicting oogenesis proteins by six different modes of Chou's pseudo amino acid composition. *J Theor Biol.* 2017; 414: 128-36.

214. Tripathi P, Pandey PN. A novel alignment-free method to classify protein folding types by combining spectral graph clustering with Chou's pseudo amino acid composition. *J Theor Biol.* 2017; 424: 49-54.

215. X. Xiao, X. Cheng, S. Su, Q. Nao, K.C. Chou, pLoc-mGpos: Incorporate key gene ontology information into general PseAAC for predicting subcellular localization of Gram-positive bacterial proteins. *Natural Science.* 2017; 9: 331-49.

216. Xu C, Ge L, Zhang Y, Dehmer M, Gutman I. Prediction of therapeutic peptides by incorporating q-Wiener index into Chou's general PseAAC. *J Biomed Inform.* 2017.

217. Xu Y, Li C, Chou KC. iPreNy-PseAAC: identify C-terminal cysteine prenylation sites in proteins by incorporating two tiers of sequence couplings into PseAAC. *Med Chem.* 2017; 13: 544-51.

218. Yu B, Li S, Qiu WY, Chen C, Chen RX, Wang L, et al. Accurate prediction of subcellular location of apoptosis proteins combining Chou's PseAAC and PsePSSM based on wavelet denoising. *Oncotarget.* 2017; 8: 107640-65.

219. Yu B, Lou L, Li S, Zhang Y, Qiu W, Wu X, et al. Prediction of protein structural class for low-similarity sequences using Chou's pseudo amino acid composition and wavelet denoising. *J Mol Graph Model.* 2017; 76: 260-73.

220. Ahmad J, Hayat M. MFSC: Multi-voting based Feature Selection for Classification of Golgi Proteins by Adopting the General form of Chou's PseAAC components. *J Theor Biol.* 2018; 463: 99-109.

221. Akbar S, Hayat M. iMethyl-STTNC: Identification of N(6)-methyladenosine sites by extending the Idea of SAAC into Chou's

- PseAAC to formulate RNA sequences. *J Theor Biol.* 2018; 455: 205-11.
222. Arif M, Hayat M, Jan Z. iMem-2LSAAC: A two-level model for discrimination of membrane proteins and their types by extending the notion of SAAC into Chou's pseudo amino acid composition. *J Theor Biol.* 2018; 442: 11-21.
223. Butt AH, Rasool N, Khan YD. Predicting membrane proteins and their types by extracting various sequence features into Chou's general PseAAC. *MolBiol Rep.* 2018.
224. Cheng X, Xiao X, Chou KC. pLoc-mEuk: Predict subcellular localization of multi-label eukaryotic proteins by extracting the key GO information into general PseAAC. *Genomics.* 2018; 110: 50-8.
225. Cheng X, Xiao X, Chou KC. pLoc-mGneg: Predict subcellular localization of Gram-negative bacterial proteins by deep gene ontology learning via general PseAAC. *Genomics.* 2018; 110: 231-9.
226. Cheng X, Xiao X, Chou KC. pLoc-mHum: predict subcellular localization of multi-location human proteins via general PseAAC to winnow out the crucial GO information. *Bioinformatics.* 2018; 34: 1448-56.
227. Cheng X, Xiao X, Chou KC. pLoc_bal-mGneg: predict subcellular localization of Gram-negative bacterial proteins by quasi-balancing training dataset and general PseAAC. *Journal of Theoretical Biology.* 2018; 458: 92-102.
228. Cheng X, Xiao X, Chou KC. pLoc_bal-mPlant: predict subcellular localization of plant proteins by general PseAAC and balancing training dataset *Curr Pharm Des.* 2018; 24: 4013-22.
229. Chou KC, Cheng X, Xiao X. pLoc_bal-mHum: predict subcellular localization of human proteins by PseAAC and quasi-balancing training dataset *Genomics.* 2018.
230. Chou KC, Cheng X, Xiao X. pLoc_bal-mEuk: predict subcellular localization of eukaryotic proteins by general PseAAC and quasi-balancing training dataset. *Med Chem.* 2019; 15: 472-85.
231. Contreras-Torres E. Predicting structural classes of proteins by incorporating their global and local physicochemical and conformational properties into general Chou's PseAAC. *J Theor Biol.* 2018; 454: 139-45.
232. Fu X, Zhu W, Liso B, Cai L, Peng L, Yang J. Improved DNA-binding protein identification by incorporating evolutionary information into the Chou's PseAAC. *IEEE Access.* 2018; 20.
233. Ghauri AW, Khan YD, Rasool N, Khan SA, Chou KC. pNitro-Tyr-PseAAC: Predict nitrotyrosine sites in proteins by incorporating five features into Chou's general PseAAC. *Curr Pharm Des.* 2018; 24: 4034-43.
234. Javed F, Hayat M. Predicting subcellular localizations of multi-label proteins by incorporating the sequence features into Chou's PseAAC. *Genomics.* 2018.
235. Ju Z, Wang SY. Prediction of citrullination sites by incorporating k-spaced amino acid pairs into Chou's general pseudo amino acid composition. *Gene.* 2018; 664: 78-83.
236. Khan YD, Rasool N, Hussain W, Khan SA, Chou KC. iPhosT-PseAAC: Identify phosphothreonine sites by incorporating sequence statistical moments into PseAAC. *Analytical Biochemistry.* 2018; 550: 109-16.
237. Khan YD, Rasool N, Hussain W, Khan SA, Chou KC. iPhosY-PseAAC: identify phosphotyrosine sites by incorporating sequence statistical moments into PseAAC. *MolBiol Rep.* 2018; 45: 2501-9.
238. Krishnan MS. Using Chou's general PseAAC to analyze the evolutionary relationship of receptor associated proteins (RAP) with various folding patterns of protein domains. *J Theor Biol.* 2018; 445: 62-74.
239. Liang Y, Zhang S. Identify Gram-negative bacterial secreted protein types by incorporating different modes of PSSM into Chou's general PseAAC via Kullback-Leibler divergence. *J Theor Biol.* 2018; 454: 22-9.
240. Mei J, Fu Y, Zhao J. Analysis and prediction of ion channel inhibitors by using feature selection and Chou's general pseudo amino acid composition. *J Theor Biol.* 2018; 456: 41-8.
241. Mei J, Zhao J. Prediction of HIV-1 and HIV-2 proteins by using Chou's pseudo amino acid compositions and different classifiers. *Sci Rep.* 2018; 8: 2359.
242. Mei J, Zhao J. Analysis and prediction of presynaptic and postsynaptic neurotoxins by Chou's general pseudo amino acid composition and motif features. *J Theor Biol.* 2018; 427: 147-53.
243. Mousavizadegan M, Mohabatkar H. Computational prediction of antifungal peptides via Chou's PseAAC and SVM. *J BioinformComput Biol.* 2018; 1850016.
244. Rahman SM, Shatabda S, Saha S, Kaykobad M, Sohel Rahman M. DPP-PseAAC: A DNA-binding Protein Prediction model using Chou's general PseAAC. *J Theor Biol.* 2018; 452: 22-34.
245. E.S. Sankari, D.D. Manimegalai, Predicting membrane protein types by incorporating a novel feature set into Chou's general PseAAC. *J Theor Biol.* 2018; 455: 319-28.
246. Srivastava, Kumar R, Kumar M. BlaPred: predicting and

classifying beta-lactamase using a 3-tier prediction system via Chou's general PseAAC. *J Theor Biol.* 2018; 457: 29-36.

247. Zhang S, Duan X. Prediction of protein subcellular localization with oversampling approach and Chou's general PseAAC. *J Theor Biol.* 2018; 437: 239-50.

248. Zhang S, Liang Y. Predicting apoptosis protein subcellular localization by integrating auto-cross correlation and PSSM into Chou's PseAAC. *J Theor Biol.* 2018; 457: 163-9.

249. Adilina S, Farid DM, Shatabda S. Effective DNA binding protein prediction by using key features via Chou's general PseAAC. *J Theor Biol.* 2019; 460: 64-78.

250. Ahmad J, Hayat M. MFSC: Multi-voting based feature selection for classification of Golgi proteins by adopting the general form of Chou's PseAAC components. *J Theor Biol.* 2019; 463: 99-109.

251. Awais M, Hussain W, Khan YD, Rasool N, Khan SA, Chou KC.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 Trans ComputBiolBioinform.* 2019.

252. Butt AH, Rasool N, Khan YD. Prediction of antioxidant proteins by incorporating statistical moments based features into Chou's PseAAC. *Journal of Theoretical Biology.* 2019; 473: 1-8.

253. Chen G, Cao M, Yu J, Guo X, Shi S. Prediction and functional analysis of prokaryote lysine acetylation site by incorporating six types of features into Chou's general PseAAC. *J Theor Biol.* 2019; 461: 92-101.

254. Cheng X, Lin WZ, Xiao X, Chou KC.pLoc_bal-mAnimal: predict subcellular localization of animal proteins by balancing training dataset and PseAAC. *Bioinformatics.* 2019; 35: 398-406.

255. Ehsan, Mahmood MK, Khan YD, Barukab OM, Khan SA, Chou KC.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.* 2019; 20: 124-33.

256. Hussain W, Khan SD, Rasool N, Khan SA, Chou KC.SPalmitylC-PseAAC: A sequence-based model developed via Chou's 5-steps rule and general PseAAC for identifying S-palmitoylation sites in proteins. *Anal Biochem.* 2019; 568: 14-23.

257. Hussain W, Khan YD, Rasool N, Khan SA, Chou KC.SPrenylC-PseAAC: A sequence-based model developed via Chou's 5-steps rule and general PseAAC for identifying S-prenylation sites in proteins. *J Theor Biol.* 2019; 468: 1-11.

258. Jia J, Li X, Qiu W, Xiao X, Chou KC.iPPI-PseAAC(CGR): Identify protein-protein interactions by incorporating chaos game representation into PseAAC. *Journal of Theoretical Biology.* 2019; 460: 195-203.

259. Kabir M, Ahmad S, Iqbal M, Hayat M. 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.* 2019.

260. Khan YD, Jamil M, Hussain W, Rasool N, Khan SA, Chou KC.pSSbond-PseAAC: Prediction of disulfide bonding sites by integration of PseAAC and statistical moments. *J Theor Biol.* 2019; 463: 47-55.

261. Ning Q, Ma Z, Zhao X.dForml(KNN)-PseAAC: Detecting formylation sites from protein sequences using K-nearest neighbor algorithm via Chou's 5-step rule and pseudo components. *J Theor Biol.* 2019; 470: 43-9.

262. Shen Y, Tang J, Guo F. Identification of protein subcellular localization via integrating evolutionary and physicochemical information into Chou's general PseAAC. *J Theor Biol.* 2019; 462: 230-9.

263. Tahir M, Hayat M, Khan SA.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. *Mol Genet Genomics.* 2019; 294: 199-210.

264. Wang L, Zhang R, Mu Y. Fu-SulfPred: Identification of Protein S-sulfenylation Sites by Fusing Forests via Chou's General PseAAC. *J Theor Biol.* 461: 51-8.

265. Xiao X, Cheng X, Chen G, Mao Q, Chou KC.pLoc_bal-mGpos: predict subcellular localization of Gram-positive bacterial proteins by quasi-balancing training dataset and PseAAC. *Genomics.* 2019; 111: 886-92.

266. Xiao X, Cheng X, Chen G, Mao Q, Chou KC.pLoc_bal-mVirus: Predict Subcellular Localization of Multi-Label Virus Proteins by Chou's General PseAAC and IHTS Treatment to Balance Training Dataset. *Med Chem.* 2018; 15: 496-509.

267. Shen HB, Chou KC.PseAAC: a flexible web-server for generating various kinds of protein pseudo amino acid composition. *Analytical Biochemistry.* 2008; 373: 386-8.

268. Du P, Wang X, Xu C, Gao Y, PseAAC-Builder: A cross-platform stand-alone program for generating various special Chou's pseudo amino acid compositions. *Analytical Biochemistry.* 2012; 425: 117-9.

269. Cao DS, Xu QS, Liang YZ. propy: a tool to generate various modes of Chou's PseAAC. *Bioinformatics*. 2013; 29: 960-2.
270. Du P, Gu S, Jiao Y. 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*. 2014; 15: 3495-506.
271. Chou KC. Pseudo amino acid composition and its applications in bioinformatics, proteomics and system biology. *Current Proteomics*. 2009; 6: 262-74.
272. Chou KC. Some remarks on protein attribute prediction and pseudo amino acid composition (50th Anniversary Year Review, 5-steps rule). *Journal of Theoretical Biology*. 2011; 273: 236-47.
273. Chen W, Lei TY, Jin DC, Lin H, Chou KC. PseKNC: a flexible web-server for generating pseudo K-tuple nucleotide composition. *Analytical Biochemistry*. 2014; 456: 53-60.
274. Chen W, Lin H, Chou KC. Pseudo nucleotide composition or PseKNC: an effective formulation for analyzing genomic sequences. *MolBioSyst*. 2015; 11: 2620-34.
275. Liu B, Yang F, Huang DS, Chou KC. iPromoter-2L: a two-layer predictor for identifying promoters and their types by multi-window-based PseKNC. *Bioinformatics*. 2018; 34: 33-40.
276. Guo SH, Deng EZ, Xu LQ, Ding H, Lin H, Chen W, et al. iNuc-PseKNC: a sequence-based predictor for predicting nucleosome positioning in genomes with pseudo k-tuple nucleotide composition. *Bioinformatics*. 2014; 30: 1522-9.
277. Lin H, Deng EZ, Ding H, Chen W, Chou KC. iPro54-PseKNC: a sequence-based predictor for identifying sigma-54 promoters in prokaryote with pseudo k-tuple nucleotide composition. *Nucleic Acids Research*. 2014; 42: 12961-72.
278. Chen W, Zhang X, Brooker J, Lin H, Zhang L, Chou KC. PseKNC-General: a cross-platform package for generating various modes of pseudo nucleotide compositions. *Bioinformatics*. 2015; 31: 119-20.
279. Liu B, Fang L, Long R, Lan X, Chou KC. iEnhancer-2L: a two-layer predictor for identifying enhancers and their strength by pseudo k-tuple nucleotide composition. *Bioinformatics*. 2016; 32: 362-9.
280. Xiao X, Ye HX, Liu Z, Jia JH, Chou KC. iROS-gPseKNC: predicting replication origin sites in DNA by incorporating dinucleotide position-specific propensity into general pseudo nucleotide composition. *Oncotarget*. 2016; 7: 34180-9.
281. Feng P, Ding H, Yang H, Chen W, Lin H, Chou KC. iRNA-PseColl: Identifying the occurrence sites of different RNA modifications by incorporating collective effects of nucleotides into PseKNC. *Molecular Therapy - Nucleic Acids*. 2017; 7: 155-63.
282. Qiu WR, Jiang SY, Sun BQ, Xiao X, Cheng X, Chou KC. iRNA-2methyl: identify RNA 2'-O-methylation sites by incorporating sequence-coupled effects into general PseKNC and ensemble classifier. *Medicinal Chemistry*. 2017; 13: 734-43.
283. Liu B, Weng F, Huang DS, Chou KC. iRO-3wPseKNC: Identify DNA replication origins by three-window-based PseKNC. *Bioinformatics*. 2018; 34: 3086-93.
284. Sabooh MF, Iqbal N, Khan M, Khan M, Maqbool MF. Identifying 5-methylcytosine sites in RNA sequence using composite encoding feature into Chou's PseKNC. *J Theor Biol*. 2018; 452: 1-9.
285. Su ZD, Huang Y, Zhang ZY, Zhao YW, Wang D, Chen W, et al. iLoc-lncRNA: predict the subcellular location of lncRNAs by incorporating octamer composition into general PseKNC. *Bioinformatics*. 2018; 34: 4196-204.
286. Xiao X, Xu ZC, Qiu WR, Wang P, Ge HT, Chou KC. iPSW(2L)-PseKNC: A two-layer predictor for identifying promoters and their strength by hybrid features via pseudo K-tuple nucleotide composition. *Genomics*. 2018.
287. Yang H, Qiu WR, Liu G, Guo FB, Chen W, Chou KC, et al. iRSpot-Pse6NC: Identifying recombination spots in *Saccharomyces cerevisiae* by incorporating hexamer composition into general PseKNC. *International Journal of Biological Sciences*. 2018; 14: 883-91.
288. Feng P, Yang H, Ding H, Lin H, Chen W, Chou KC. iDNA6mA-PseKNC: Identifying DNA N(6)-methyladenosine sites by incorporating nucleotide physicochemical properties into PseKNC. *Genomics*. 2019; 111: 96-102.
289. Nazari I, Tahir M, Tayari H, Chong KT. 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)*. 2019.
290. Tahir M, Tayara H, Chong KT. iRNA-PseKNC(2methyl): Identify RNA 2'-O-methylation sites by convolution neural network and Chou's pseudo components. *J Theor Biol*. 2019; 465: 1-6.
291. Liu B, Liu F, Wang X, Chen J, Fang L, Chou KC. Pse-in-One: a web server for generating various modes of pseudo components of DNA, RNA, and protein sequences. *Nucleic Acids Research*. 2015; 43: 65-71.
292. Liu B, Wu H, Chou KC. 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*. 2017; 9: 67-91.

293. Voet D, Voet JG, Pratt CW. *Fundamentals of Biochemistry*, John Wiley & Sons, New York. 2002.
294. Chou KC, Chen NY. The biological functions of low-frequency phonons. *ScientiaSinica*. 1977; 20: 447-57.
295. Du QS, Wang SQ, Wei DQ, Zhu Y, Guo H, Sirois S, et al. Polyprotein Cleavage Mechanism of SARS CoVMpro and Chemical Modification of Octapeptide. *Peptides*. 2004; 25: 1857-64.
296. Chou KC. Review: Prediction of human immunodeficiency virus protease cleavage sites in proteins. *Analytical Biochemistry*. 1996; 233: 1-14.
297. Liang GZ, Li SZ. A new sequence representation as applied in better specificity elucidation for human immunodeficiency virus type 1 protease. *Biopolymers*. 2007; 88: 401-12.
298. Nanni L, Lumini A. Using ensemble of classifiers for predicting HIV protease cleavage sites in proteins. *Amino Acids*. 2008.
299. Shen HB, Chou KC. HIVcleave: a web-server for predicting HIV protease cleavage sites in proteins. *Analytical Biochemistry*. 2008;375: 388-90.
300. Du QS, Sun H, Chou KC. Inhibitor design for SARS coronavirus main protease based on "distorted key theory". *Medicinal Chemistry*. 2007; 3: 1-6.
301. Du QS, Wang S, Wei DQ, Sirois S, Chou KC. Molecular modelling and chemical modification for finding peptide inhibitor against SARS CoVMpro. *Analytical Biochemistry*. 2005; 337: 262-70.
302. Gan YR, Huang H, Huang YD, Rao CM, Zhao Y, Liu JS, et al. Synthesis and activity of an octapeptide inhibitor designed for SARS coronavirus main proteinase. *Peptides*. 2006; 27: 622-5.
303. Chou KC, Jiang SP, Liu WM, Fee CH. Graph theory of enzyme kinetics: 1. Steady-state reaction system. *ScientiaSinica*. 1979; 22: 341-58.
304. Chou KC, Zhang CT, Maggiora GM. Disposition of amphiphilic helices in heteropolar environments. *PROTEINS: Structure, Function, and Genetics*. 1997; 28: 99-108.
305. Zhou GP. The disposition of the LZCC protein residues in wenxiang diagram provides new insights into the protein-protein interaction mechanism. *Journal of Theoretical Biology*. 2011; 284: 142-8.
306. Zhou GP. The Structural Determinations of the Leucine Zipper Coiled-Coil Domains of the cGMP-Dependent Protein Kinase I alpha and its Interaction with the Myosin Binding Subunit of the Myosin Light Chains Phosphase. *Proteins & Peptide Letters*. 2011; 18: 966-78.
307. Andrade MA, O'Donoghue SI, Rost B. Adaptation of protein surfaces to subcellular location. *J. Mol. Biol.* 1998; 276: 517-25.
308. Reinhardt, Hubbard T. Using neural networks for prediction of the subcellular location of proteins. *Nucleic Acids Research*. 1998; 26: 2230-6.
309. Yuan Z. Prediction of protein subcellular locations using Markov chain models. *FEBS Letters*. 1999; 451: 23-6.
310. Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of Molecular Biology*. 2000; 300: 1005-16.
311. Nakai K. Protein sorting signals and prediction of subcellular localization. *Advances in Protein Chemistry*. 2000; 54: 277-344.
312. Chou KC, Shen HB. Recent progresses in protein subcellular location prediction. *Analytical Biochemistry*. 2007; 370: 1-16.
313. Chou KC, Shen HB. Cell-PLoc: A package of Web servers for predicting subcellular localization of proteins in various organisms. *Nature Protocols*. 2008; 3: 153-62.
314. Zhong WZ, Zhou SF. Molecular science for drug development and biomedicine. *International Journal of Molecular Sciences*. 2014; 15: 20072-8.
315. Du QS, Huang RB, Wang SQ, Chou KC. Designing inhibitors of M2 proton channel against H1N1 swine influenza virus. *PLoS ONE*. 2010; 5: e9388.
316. Wang SQ, Cheng XC, Dong WL, Wang RL, Chou KC. Three new powerful Oseltamivir derivatives for inhibiting the neuraminidase of influenza virus. *BiochemBiophys Res Commun*. 2010; 401: 188-91.
317. Li XB, Wang SQ, Xu WR, Wang RL, Chou KC. Novel Inhibitor Design for Hemagglutinin against H1N1 Influenza Virus by Core Hopping Method. *PLoS One*. 2011; 6: e28111.
318. Ma Y, Wang SQ, Xu WR, Wang RL, Chou KC. Design novel dual agonists for treating type-2 diabetes by targeting peroxisome proliferator-activated receptors with core hopping approach. *PLoS One*. 2012; 7: e38546.
319. Liu L, Ma Y, Wang RL, Xu WR, Wang SQ, Chou KC. Find novel dual-agonist drugs for treating type 2 diabetes by means of cheminformatics. *Drug Design, Development and Therapy*. 2013; 7: 279-87.
320. Chou KC, Shen HB. Hum-PLoc: A novel ensemble classifier for predicting human protein subcellular localization. *Biochem. Biophys. Res. Commun. (BBRC)*. 2006; 347: 150-7.

321. Chou KC, Shen HB. Addendum to "Hum-PLoc: A novel ensemble classifier for predicting human protein subcellular localization". *Biochem. Biophys. Res. Commun. (BBRC)*. 2006; 348: 1479.
322. Shen HB, Chou KC. Gpos-PLoc: an ensemble classifier for predicting subcellular localization of Gram-positive bacterial proteins. *Protein Engineering, Design, and Selection*. 2007; 20: 39-46.
323. Shen HB, Chou KC. Virus-PLoc: A fusion classifier for predicting the subcellular localization of viral proteins within host and virus-infected cells. *Biopolymers*. 2007; 85: 233-40.
324. Shen HB, Chou KC. Nuc-PLoc: A new web-server for predicting protein subnuclear localization by fusing PseAA composition and PsePSSM. *Protein Engineering, Design & Selection*. 2007; 20: 561-7.
325. Shen HB, Yang J, Chou KC. Euk-PLoc: an ensemble classifier for large-scale eukaryotic protein subcellular location prediction. *Amino Acids*. 2007; 33: 57-67.
326. Chou KC, Shen HB. Cell-PLoc 2.0: An improved package of web-servers for predicting subcellular localization of proteins in various organisms. *Natural Science*. 2010; 2: 1090-103.
327. Chou KC, Wu ZC, Xiao X. iLoc-Euk: A Multi-Label Classifier for Predicting the Subcellular Localization of Singleplex and Multiplex Eukaryotic Proteins. *PLoS One*. 2011; 6: e18258.
328. Wu ZC, Xiao X, Chou KC. iLoc-Plant: a multi-label classifier for predicting the subcellular localization of plant proteins with both single and multiple sites. *Molecular BioSystems*. 2011; 7: 3287-97.
329. Xiao X, Wu ZC, Chou KC. iLoc-Virus: A multi-label learning classifier for identifying the subcellular localization of virus proteins with both single and multiple sites. *Journal of Theoretical Biology*. 2011; 284: 42-51.
330. Chou KC, Wu ZC, Xiao X. iLoc-Hum: Using accumulation-label scale to predict subcellular locations of human proteins with both single and multiple sites. *Molecular Biosystems*. 2012; 8: 629-41.
331. Wu ZC, Xiao X, Chou KC. iLoc-Gpos: A Multi-Layer Classifier for Predicting the Subcellular Localization of Singleplex and Multiplex Gram-Positive Bacterial Proteins. *Protein & Peptide Letters*. 2012; 19: 4-14.
332. Lin WZ, Fang JA, Xiao X, Chou KC. iLoc-Animal: A multi-label learning classifier for predicting subcellular localization of animal proteins *Molecular BioSystems*. 2013; 9: 634-44.
333. Cheng X, Zhao SG, Lin WZ, Xiao X, Chou KC. pLoc-mAnimal: predict subcellular localization of animal proteins with both single and multiple sites. *Bioinformatics*. 2017; 33: 3524-31.
334. Chou KC. Advance in predicting subcellular localization of multi-label proteins and its implication for developing multi-target drugs. *Current Medicinal Chemistry*. 2019; 26: 4918-43.
335. Chen W, Feng PM, Lin H, Chou KC. iRSpot-PseDNC: identify recombination spots with pseudo dinucleotide composition *Nucleic Acids Research*. 2013; 41: e68.
336. Feng PM, Chen W, Lin H, Chou KC. iHSP-PseRAAAC: Identifying the heat shock protein families using pseudo reduced amino acid alphabet composition. *Analytical Biochemistry*. 2013; 442: 118-25.
337. Chen W, Feng PM, Deng EZ, Lin H, Chou KC. iTIS-PseTNC: a sequence-based predictor for identifying translation initiation site in human genes using pseudo trinucleotide composition. *Analytical Biochemistry*. 2014; 462: 76-83.
338. Ding H, Deng EZ, Yuan LF, Liu L, Lin H, Chen W, Chou KC. iCTX-Type: A sequence-based predictor for identifying the types of conotoxins in targeting ion channels. *BioMed Research International (BMRI)*. 2014; 2014: 286419.
339. Liu B, Fang L, Liu F, Wang X, Chen J, Chou KC. Identification of real microRNA precursors with a pseudo structure status composition approach. *PLoS ONE*. 2015; 10: e0121501.
340. Liu Z, Xiao X, Qiu WR, Chou KC. iDNA-Methyl: Identifying DNA methylation sites via pseudo trinucleotide composition. *Analytical Biochemistry*. 2015; 474: 69-77.
341. Xiao X, Min JL, Lin WZ, Liu Z, Cheng X, Chou KC. iDrug-Target: predicting the interactions between drug compounds and target proteins in cellular networking via the benchmark dataset optimization approach. *J BiomolStructDyn (JBSD)*. 2015; 33: 2221-33.
342. Jia J, Liu Z, Xiao X, Liu B, Chou KC. iSuc-PseOpt: Identifying lysine succinylation sites in proteins by incorporating sequence-coupling effects into pseudo components and optimizing imbalanced training dataset. *Anal Biochem*. 2016; 497: 48-56.
343. Chen W, Feng P, Yang H, Ding H, Lin H, Chou KC. iRNA-AI: identifying the adenosine to inosine editing sites in RNA sequences. *Oncotarget*. 2017; 8: 4208-17.
344. Chen W, Ding H, Zhou X, Lin H, Chou KC. iRNA(m6A)-PseDNC: Identifying N6-methyladenosine sites using pseudo dinucleotide composition. *Analytical Biochemistry*. 2018; 561-562: 59-65.
345. Chen W, Feng P, Yang H, Ding H, Lin H, Chou KC. iRNA-3typeA: identifying 3-types of modification at RNA's adenosine sites.

Molecular Therapy: Nucleic Acid. 2018; 11: 468-74.

346. Li JX, Wang SQ, Du QS, Wei H, Li XM, Meng JZ, et al. Simulated protein thermal detection (SPTD) for enzyme thermostability study and an application example for pullulanase from *Bacillus deramificans*. *Curr Pharm Des*. 2018; 24: 4023-33.

347. Qiu WR, Sun BQ, Xiao X, Xu ZC, Jia JH, Chou KC. iKcr-PseEns: Identify lysine crotonylation sites in histone proteins with pseudo components and ensemble classifier. *Genomics*. 2018; 110: 239-46.

348. Chou KC. Progresses in predicting post-translational modification. *International Journal of Peptide Research and Therapeutics (IJPRT)*. 2019.

349. Du X, Diao Y, Liu H, Li S. MsDBP: Exploring DNA-binding Proteins by Integrating Multi-scale Sequence Information via Chou's 5-steps Rule. *Journal of Proteome Research*. 2019; 18: 3119-32.

350. Ju Z, Wang SY. 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*. 2019.

351. Khan YD, Batool A, Rasool N, Khan A, Chou KC. Prediction of nitrosocysteine sites using position and composition variant features. *Letters in Organic Chemistry*. 2019; 16: 283-93.

352. Le NQK. iN6-methylat (5-step): identifying DNA N(6)-methyladenine sites in rice genome using continuous bag of nucleobases via Chou's 5-step rule. *Mol Genet Genomics*. 2019.

353. Le NQK, Yapp EKY, Ho QT, Nagasundaram N, Ou YY, Yeh HY. iEnhancer-5Step: Identifying enhancers using hidden information of DNA sequences via Chou's 5-step rule and word embedding. *Anal Biochem*. 2019; 571: 53-61.

354. Lu Y, Wang S, Wang J, Zhou G, Zhang Q, Zhou X, et al. An Epidemic Avian Influenza Prediction Model Based on Google Trends. *Letters in Organic Chemistry*. 2019; 16: 303-10.

355. Romero-Molina S, Ruiz-Blanco YB, Harms M, Münch J, Sanchez-Garcia E. PPI-Detect: A Support Vector Machine Model for Sequence-Based Prediction of Protein - Protein Interactions. *Journal of Computational Chemistry*. 2019.

356. Salman, Khan M, Iqbal N, Hussain T, Afzal S, Chou KC. A two-level computation model based on deep learning algorithm for identification of piRNA and their functions via Chou's 5-steps rule. *International Journal of Peptide Research and Therapeutics (IJPRT)*. 2019.

357. Tahir M, Tayara H, Chong KT. iDNA6mA (5-step rule): Identification of DNA N6-methyladenine sites in the rice genome by

intelligent computational model via Chou's 5-step rule. *CHEMOLAB*. 2019; 189: 96-101.

358. Song J, Li F, Takemoto K, Haffari G, Akutsu T, Chou KC, et al. PREvail, an integrative approach for inferring catalytic residues using sequence, structural and network features in a machine learning framework. *Journal of Theoretical Biology*. 2018; 443: 125-37.

359. Chen Z, Liu X, Li F, Li C, Marquez-Lago T, Leier A, et al. Large-scale comparative assessment of computational predictors for lysine post-translational modification sites. *Brief in Bioinform*. 2018.

360. Chen Z, Zhao PY, Li F, Leier A, Marquez-Lago TT, Wang Y, et al. iFeature: a python package and web server for features extraction and selection from protein and peptide sequences. *Bioinformatics*. 2018; 34: 2499-502.

361. Li F, Li C, Marquez-Lago TT, Leier A, Akutsu T, Purcell AW, et al. Quokka: a comprehensive tool for rapid and accurate prediction of kinase family-specific phosphorylation sites in the human proteome. *Bioinformatics*. 2018; 34: 4223-31.

362. Li F, Wang Y, Li C, Marquez-Lago TT, Leier A, Rawlings ND, et al. Twenty years of bioinformatics research for protease-specific substrate and cleavage site prediction: a comprehensive revisit and benchmarking of existing methods. *Brief in Bioinform*. 2018.

363. Song J, Li F, Leier A, Marquez-Lago TT, Akutsu T, Haffari G, et al. PROSPEROUS: high-throughput prediction of substrate cleavage sites for 90 proteases with improved accuracy. *Bioinformatics*. 2018; 34: 684-7.

364. Song J, Wang Y, Li F, Akutsu T, Rawlings ND, Webb GI, et al. iProt-Sub: a comprehensive package for accurately mapping and predicting protease-specific substrates and cleavage sites. *Brief in Bioinform*. 2018; 20: 638-58.

365. Wang J, Li J, Yang B, Xie R, Marquez-Lago TT, Leier A, et al. Bastion3: a two-layer approach for identifying type III secreted effectors using ensemble learning. *Bioinformatics*. 2018; 35: 2017-28.

366. Wang J, Yang B, Leier A, Marquez-Lago TT, Hayashida M, Rucker A, et al. Bastion6: a bioinformatics approach for accurate prediction of type VI secreted effectors. *Bioinformatics*. 2018; 34: 2546-55.

367. Zhang S, Yang K, Lei Y, Song K. iRSpot-DTS: Predict recombination spots by incorporating the dinucleotide-based spare-cross covariance information into Chou's pseudo components. *Genomics*. 2018.

368. Zhang Y, Xie R, Wang J, Leier A, Marquez-Lago TT, Akutsu T, et al. Computational analysis and prediction of lysine malonylation sites by exploiting informative features in an integrative machine-

learning framework. *Brief in Bioinform.* 2018.

369. Zhai X, Chen M, Lu W. Accelerated search for perovskite materials with higher Curie temperature based on the machine learning methods. *Computational Materials Science.* 2018; 151: 41-8.

370. Chou KC, Forsen S. Diffusion-controlled effects in reversible enzymatic fast reaction system: Critical spherical shell and proximity rate constants. *Biophysical Chemistry.* 1980; 12: 255-63.

371. Chou KC, Forsen S. Graphical rules for enzyme-catalyzed rate laws. *Biochemical Journal.* 1980; 187: 829-35.

372. Chou KC, Forsen S, Zhou GQ. Three schematic rules for deriving apparent rate constants. *ChemicaScripta.* 1980; 16: 109-13.

373. Chou KC, Li TT, Forsen S. The critical spherical shell in enzymatic fast reaction systems. *Biophysical Chemistry.* 1980; 12: 265-9.

374. Li TT, Chou KC, Forsen S. The flow of substrate molecules in fast enzyme-catalyzed reaction systems. *ChemicaScripta.* 1980; 16: 192-6.

375. Chou KC, Carter RT, Forsen S. A new graphical method for deriving rate equations for complicated mechanisms. *ChemicaScripta.* 1981; 18: 82-6.

376. Chou KC, Chen NY, Forsen S. The biological functions of low-frequency phonons: 2. Cooperative effects. *ChemicaScripta.* 1981; 18: 126-32.

377. Chou KC, Forsen S. Graphical rules of steady-state reaction systems. *Canadian Journal of Chemistry.* 1981; 59: 737-55.

378. Chou KC. Low-frequency vibrations of helical structures in protein molecules. *Biochemical Journal.* 1983; 209: 573-80.

379. Chou KC. Identification of low-frequency modes in protein molecules. *Biochemical Journal.* 1983; 215: 465-9.

380. Zhou GP, Deng MH. An extension of Chou's graphic rules for deriving enzyme kinetic equations to systems involving parallel reaction pathways. *Biochemical Journal.* 1984; 222: 169-76.

381. Chou KC. Biological functions of low-frequency vibrations (phonons).3. Helical structures and microenvironment. *Biophysical Journal.* 1984; 45: 881-9.

382. Chou KC. The biological functions of low-frequency phonons. 4. Resonance effects and allosteric transition. *Biophysical Chemistry.* 1984; 20: 61-71.

383. Chou KC. Low-frequency vibrations of DNA molecules. *Biochemical Journal.* 1984; 221: 27-31.

384. Chou KC. Low-frequency motions in protein molecules:

beta-sheet and beta-barrel. *Biophysical Journal.* 1985; 48: 289-97.

385. Chou KC. Prediction of a low-frequency mode in bovine pancreatic trypsin inhibitor molecule. *International Journal of Biological Macromolecules.* 1985; 7: 77-80.

386. Chou KC, Kiang YS. The biological functions of low-frequency phonons: 5. A phenomenological theory. *Biophysical Chemistry.* 1985; 22: 219-35.

387. Chou KC. Origin of low-frequency motion in biological macromolecules: A view of recent progress of quasi-continuity model. *Biophysical Chemistry.* 1986; 25: 105-16.

388. Chou KC. The biological functions of low-frequency phonons: 6. A possible dynamic mechanism of allosteric transition in antibody molecules. *Biopolymers.* 1987; 26: 285-95.

389. Chou KC. Review: Low-frequency collective motion in bi-macromolecules and its biological functions. *Biophysical Chemistry.* 1988; 30: 3-48.

390. Chou KC, Maggiora GM. The biological functions of low-frequency phonons: 7. The impetus for DNA to accommodate intercalators. *British Polymer Journal.* 1988; 20: 143-8.

391. Chou KC. Low-frequency resonance and cooperativity of hemoglobin. *Trends in Biochemical Sciences.* 1989; 14: 212-3.

392. Chou KC, Maggiora GM, Mao B. Quasi-continuum models of twist-like and accordion-like low-frequency motions in DNA. *Biophysical Journal.* 1989; 56: 295-305.

393. Chou KC. Graphic rules in steady and non-steady enzyme kinetics. *Journal of Biological Chemistry.* 1989; 264: 12074-9.

394. Chou KC. Review: Applications of graph theory to enzyme kinetics and protein folding kinetics. Steady and non-steady state systems. *Biophysical Chemistry.* 1990; 35: 1-24.

395. Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Chou KC, Kezdy FJ, et al. Steady-state kinetic studies with the non-nucleoside HIV-1 reverse transcriptase inhibitor U-87201E. *Journal of Biological Chemistry.* 1993; 268: 6119-24.

396. Althaus IW, Gonzales AJ, Chou JJ, Diebel MR, Chou KC, Kezdy FJ, et al. The quinoline U-78036 is a potent inhibitor of HIV-1 reverse transcriptase. *Journal of Biological Chemistry.* 1993; 268: 14875-80.

397. Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Chou KC, Kezdy FJ, et al. Kinetic studies with the nonnucleoside HIV-1 reverse transcriptase inhibitor U-88204E. *Biochemistry.* 1993; 32: 6548-54.

398. Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Chou KC, Kezdy FJ, et al. Steady-state kinetic studies with the polysulfonate

- U-9843, an HIV reverse transcriptase inhibitor. *Cellular and Molecular Life Science (Experientia)*. 1994; 50: 23-8.
399. Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Chou KC, Kezdy FJ, et al. Kinetic studies with the non-nucleoside human immunodeficiency virus type-1 reverse transcriptase inhibitor U-90152e. *Biochemical Pharmacology*. 1994;47: 2017-28.
400. Chou KC, Kezdy FJ, Reusser F. Review: Kinetics of processive nucleic acid polymerases and nucleases. *Analytical Biochemistry*. 1994; 221: 217-30.
401. Chou KC, Zhang CT, Maggiora GM. Solitary wave dynamics as a mechanism for explaining the internal motion during microtubule growth. *Biopolymers*. 1994; 34: 143-53.
402. Althaus IW, Chou KC, Franks KM, Diebel MR, Kezdy FJ, Romero DL, et al. The benzylthio-pyrididine U-31,355, a potent inhibitor of HIV-1 reverse transcriptase. *Biochemical Pharmacology*. 1996; 51: 743-50.
403. Liu H, Wang M, Chou KC. Low-frequency Fourier spectrum for predicting membrane protein types. *BiochemBiophys Res Commun (BBRC)*. 2005; 336: 737-9.
404. Gordon G. Designed Electromagnetic Pulsed Therapy: Clinical Applications. *Journal of Cellular Physiology*. 2007; 212: 579-82.
405. Andraos J. Kinetic plasticity and the determination of product ratios for kinetic schemes leading to multiple products without rate laws: new methods based on directed graphs. *Canadian Journal of Chemistry*. 2008; 86: 342-57.
406. Chou KC, Shen HB. FoldRate: A web-server for predicting protein folding rates from primary sequence. *The Open Bioinformatics Journal*. 2009; 3: 31-50
407. Shen HB, Song JN, Chou KC. Prediction of protein folding rates from primary sequence by fusing multiple sequential features *Journal of Biomedical Science and Engineering (JBiSE)*. 2009; 2: 136-43.
408. Wang JF, Chou KC. Insight into the molecular switch mechanism of human Rab5a from molecular dynamics simulations. *BiochemBiophys Res Commun (BBRC)*. 2009; 390: 608-12.
409. Gordon G. Extrinsic electromagnetic fields, low frequency (phonon) vibrations, and control of cell function: a non-linear resonance system. *Journal of Biomedical Science and Engineering (JBiSE)*. 2008; 1: 152-6
410. Madkan, Blank M, Elson E, Chou KC, Geddis MS, Goodman R. Steps to the clinic with ELF EMF *Natural Science*. 2009; 1: 157-65.
411. Chou KC. Graphic rule for drug metabolism systems. *Current Drug Metabolism*. 2010; 11: 369-78.
412. Chou KC, Lin WZ, Xiao X. Wenxiang: a web-server for drawing wenxiang diagrams *Natural Science*. 2011; 3: 862-5
413. Lian P, Wei DQ, Wang JF, Chou KC. An allosteric mechanism inferred from molecular dynamics simulations on phospholamban-pentamer lipid membranes. *PLoS ONE*. 2011; 6: e18587.
414. Liao QH, Gao QZ, Wei J, Chou KC. Docking and Molecular Dynamics Study on the Inhibitory Activity of Novel Inhibitors on Epidermal Growth Factor Receptor (EGFR). *Medicinal Chemistry*. 2011; 7: 24-31.
415. Li J, Wei DQ, Wang JF, Yu ZT, Chou KC. Molecular Dynamics Simulations of CYP2E1. *Medicinal Chemistry*. 2012; 8: 208-21.
416. Wang JF, Chou KC. Recent advances in computational studies on influenza A virus M2 proton channel. *Mini Reviews in Medicinal Chemistry*. 2012; 12: 971-8.
417. Zhang T, Wei DQ, Chou KC. A Pharmacophore Model Specific to Active Site of CYP1A2 with a Novel Molecular Modeling Explorer and CoMFA. *Medicinal Chemistry*. 2012; 8: 198-207.
418. Chou KC. Proposing pseudo amino acid components is an important milestone for proteome and genome analyses. *International Journal for Peptide Research and Therapeutics (IJPRT)*. 2019.
419. Chou KC. Artificial intelligence (AI) tools constructed via the 5-steps rule for predicting post-translational modifications. *Trends in Artificial Intelligence (TIA)*. 2019; 3: 60-74.
420. Chou KC. Impacts of pseudo amino acid components and 5-steps rule to proteomics and proteome analysis. *Current Topics in Medicinal Chemistry (CTMC) (Special Issue ed. G.P Zhou)*. 2019.
421. Chou KC. An insightful 10-year recollection since the emergence of the 5-steps rule. *Current Pharmaceutical Design (Special Issue ed. Bing Nui)* in press. 2019.
422. Chou KC. An insightful 20-year recollection since the birth of pseudo amino acid components. *Computers in Biology and Medicine* in press. 2019.
423. Chou KC. An insightful recollection for predicting protein subcellular locations in multi-label systems. *Genomics*. 2019.
424. Chou KC. An insightful recollection since the distorted key theory was born about 23 years ago. *Genomics*. 2019.
425. Chou KC. Recent progresses in predicting protein

subcellular localization with artificial intelligence tools developed via the 5-steps rule. Medicinal Chemistry Submitted. 2019.

426. Chou KC. Two kinds of metrics for computational biology. Genomics. 2019.

427. Chou KC. Distorted Key Theory and Its Implication for Drug Development. Current Genomics. 2020; 17.