Review Article

Distorted Key Theory and Its Implication for Drug Development

Chou KC[1,2]*

[1]Gordon Life Science Institute, Boston, Massachusetts 02478, United States of America
[2]Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, 610054, China

*Corresponding author: Chou KC, Gordon Life Science Institute, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China

Abstract

During the last three decades or so, many efforts have been made to study the protein cleavage sites by some disease-causing enzyme, such as HIV (Human Immunodeficiency Virus) protease and SARS (Severe Acute Respiratory Syndrome) coronavirus main proteinase. It has become increasingly clear that the motivation of driving the aforementioned studies is quite wise, and that the results acquired through these studies are very rewarding, particularly for developing peptide drugs.

Keywords: HIV; SARS; Protein cleavage sites; Lock and key; Induced fit theory; Rack mechanism; Peptide drugs

Introduction

Human immunodeficiency virus infection and acquired immune deficiency syndrome (HIV/AIDS) is caused by infection with the Human Immunodeficiency Virus (HIV). Severe Acute Respiratory Syndrome (SARS) is a viral respiratory disease of zoonotic origin caused by the SARS coronavirus (SARS-CoV). Over the last two decades or so, many efforts have been made to study the protein cleavage sites by HIV protease (see, e.g., [1-25] and SARS coronavirus main proteinase [26-40]). Now it has become very clear that the motivation of driving the aforementioned studies is quite wise, and that the results acquired through these studies are very rewarding. Without losing generality, below let us consider the case of HIV protease.

Discussion

Functioning as a dimmer of two identical subunits, HIV protease has a crab-like shape (Figure. 1). Its catalytic cleft is gated by a pair of flaps (or pincers if viewed as a crab). When the enzyme is in an inhibitor-free state, the pincer-gate is open, allowing substrates to enter the catalytic cleft (Figure. 1); when in an inhibitor-binding state, the pincer-gate is closed, blocking the entrance.

As a member of the aspartyl proteases, HIV protease is highly substrate-selective and cleavage-specific. Its susceptible sites in a protein extend to an octapeptide region (Figure. 2).

According to Fisher's "lock and key" model [41], Koshland's "induced fit" theory and the "rack mechanism" [42], 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. 2). 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. 3), 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
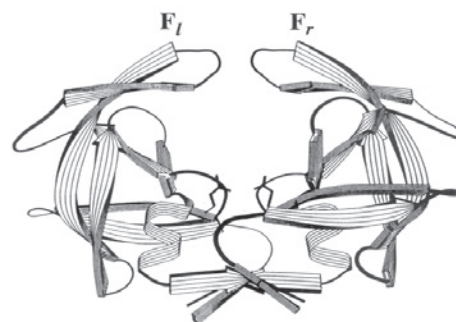

Figure 1: Functioning as a dimmer of two identical subunits, HIV protease has a crab-like shape. Its catalytic cleft is gated by a pair of flaps (or pincers if viewed as a crab). When the enzyme is in an inhibitor-free state, the pincer-gate is open, allowing substrates to enter the catalytic cleft; when in an inhibitor-binding state, the pincer-gate is closed, blocking the entrance. Adapted from [43] with permission.
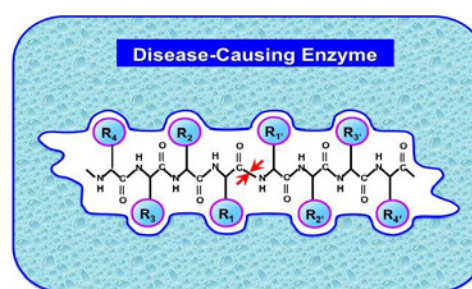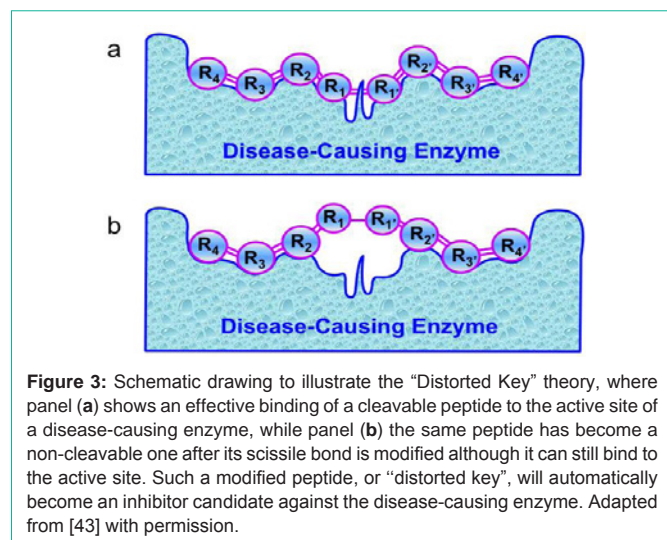

Figure 2: A schematic illustration to show a peptide in good fitting and tightly binding with the enzyme's active site before it is cleaved by the latter. Adapted from [43] with permission.

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 [28].

Figure 3: 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 [43] with permission.

For more discussion about the distorted key theory, see a comprehensive review paper [43].

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., [2,5-8,10,11,13-16,20,21] and analyze them from the codon usage approach [1,44,45]).

Furthermore, a web-server called "HIVcleave" [21] 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. To underscore the remarkable and awesome role in driving proteome and genome analyses, see [46-52], where a series of insightful recollections from different aspects have been clearly given.

Conclusion and Perspectives

Just like the pseudo amino acid components [53-56] and the 5-steps rule [57], which have been stimulated by the development of the sequence bioinformatics, widely and increasingly used in proteome and genome analysis [54, 57-230] and the guidelines for analyzing various biological systems [229-249], respectively, it is anticipated that the "distorted key theory" that was stimulated from the structural bioinformatics [250] as well as some pioneering studies [251-255] and their follow-up studies [42,44,256-282], will be also widely and increasingly used in drug development, particularly for developing peptide drugs [47] and multi-target drugs [46].

References

1. Chou KC, Zhang CT. Diagrammatization of codon usage in 339 HIV proteins and its biological implication. AIDS Research and Human Retroviruses. 1992; 12: 1967-1976.

2. Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Kezdy FJ, Romero DL, *et al.* Steady-state kinetic studies with the non-nucleoside HIV-1 reverse transcriptase inhibitor U-87201E. Journal of Biological Chemistry. 1993; 6119-6124.

3. Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Kezdy FJ, Romero DL, *et al.* Kinetic studies with the nonnucleoside HIV-1 reverse transcriptase inhibitor U-88204E. Biochemistry. 1993; 26: 6548-6554.

4. Althaus IW, Gonzales AJ, Chou JJ, Diebel MR, Kezdy FJ, Romero DL, *et al.* The quinoline U-78036 is a potent inhibitor of HIV-1 reverse transcriptase. Journal of Biological Chemistry. 1993; 20: 14875-14880.

5. Chou JJ. Predicting cleavability of peptide sequences by HIV protease via correlation-angle approach. Journal of Protein Chemistry. 1993; 3: 291-302.

6. Chou KC. A vectorized sequence-coupling model for predicting HIV protease cleavage sites in proteins. Journal of Biological Chemistry. 1993; 23: 16938-16948.

7. Chou KC, Zhang CT. Studies on the specificity of HIV protease: an application of Markov chain theory. Journal of Protein Chemistry. 1993; 6: 709-724.

8. Chou KC, Zhang CT, Kezdy FJ. A vector approach to predicting HIV protease cleavage sites in proteins. Proteins: Structure, Function, and Genetics. 1993; 17: 195-204.

9. Althaus IW, Chou JJ, Gonzales AJ, Diebel MR, Kezdy FJ, Romero DL, *et al.* Steady-state kinetic studies with the polysulfonate U-9843, an HIV reverse transcriptase inhibitor. Cellular and Molecular Life Science (Experientia). 1994; 1: 23-28.

10. Zhang CT. An alternate-subsite-coupled model for predicting HIV protease cleavage sites in proteins. Protein Engineering. 1994; 1: 65-73.

11. Thompson TB, Zheng C. Neural network prediction of the HIV-1 protease cleavage sites. J Theor Biol. 1995; 4: 369-379.

12. Althaus IW, Franks KM, Diebel MR, Kezdy FJ, Romero DL, Thomas RC, *et al.* The benzylthio-pyriditine U-31,355, a potent inhibitor of HIV-1 reverse transcriptase. Biochemical Pharmacology. 1996; 6: 743-750.

13. Chou KC, Tomasselli AL, Reardon IM, Heinrikson RL. Predicting HIV protease cleavage sites in proteins by a discriminant function method. PROTEINS: Structure, Function, and Genetics. 1996; 1: 51-72.

14. Cai YD. Artificial neural network model for HIV protease cleavage sites in proteins. Advances in Engineering Software. 1998; 2: 119-128.

15. Cai YD, Yu H. Using neural network for prediction of HIV protease cleavage sites in proteins. Journal of Protein Chemistry. 1998; 7: 607-615.

16. Cai YD, Liu XJ, Xu XB. Support Vector Machines for predicting HIV protease cleavage sites in protein. J Comput Chem. 2002; 2: 267-274.

17. Sirois S, Sing T. Review: HIV-1 gp120 V3 loop for structure-based drug design. Current Protein and Peptide Science. 2005; 5: 413-422.

18. Sirois S, Tsoukas CM, Wei DQ, Boucher C, Hatzakis GE. Selection of Molecular Descriptors with Artificial Intelligence for the Understanding of HIV-1 Protease Peptidomimetic Inhibitors-activity. Medicinal Chemistry. 2005; 2: 173-184.

19. Gao WN, Wei DQ, Li Y, Gao H, Xu WR, Li AX. Agaritine and its derivatives are potential inhibitors against HIV proteases. Medicinal Chemistry. 2007; 3: 221-226.

20. Sirois S, Touaibia M, Roy R. Review: Glycosylation of HIV-1 gp120 V3 loop: towards the rational design of a synthetic carbohydrate vaccine. Current Medicinal Chemistry. 2007; 30: 3232-3242.

21. Shen HB. HIVcleave: a web-server for predicting HIV protease cleavage sites in proteins. Analytical Biochemistry. 2008; 2: 388-390.

22. Dev J, Park D, Fu Q, Chen J, Ha HJ, Ghantous F, *et al.* Structural Basis for Membrane Anchoring of HIV-1 Envelope Spike. Science. 2016; 6295: 172-175.

23. Chen B, Chou JJ. Structure of the transmembrane domain of HIV-1 envelope glycoprotein. FEBS J. 2017; 8: 1171-1177.

24. Piai A, Dev J, Fu Q, Chou JJ. Stability and Water Accessibility of the Trimeric Membrane Anchors of the HIV-1 Envelope Spikes. J Am Chem Soc. 2017: 51: 18432-18435.

25. Fu Q, Shaik MM, Cai Y, Ghantous F, Piai A, Peng H, *et al.* Structure of the

membrane proximal external region of HIV-1 envelope glycoprotein. Proc Natl Acad Sci USA. 2018; 38: E8892-E8899.

26. Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. Science. 2003; 5626: 1763-1767.

27. Chou KC, Wei DQ, Zhong WZ. Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS. Biochem Biophys Res Comm (BBRC). 2003; 1: 148-151.

28. Du QS, Wang SQ, Wei DQ, Zhu Y, Guo H, Sirois S. Polyprotein Cleavage Mechanism of SARS CoV Mpro and Chemical Modification of Octapeptide. Peptides. 2004; 11: 1857-1864.

29. Sirois S, Wei DQ, Du QS. Virtual Screening for SARS-CoV Protease Based on KZ7088 Pharmacophore Points. J. Chem. Inf. Comput. 2004; 3: 1111-1122.

30. Du QS, Wang S, Wei DQ, Sirois S. Molecular modelling and chemical modification for finding peptide inhibitor against SARS CoV Mpro. Analytical Biochemistry. 2005; 2: 262-270.

31. Du QS, Wang SQ, Jiang ZQ, Gao WN, Li YD, Wei DQ. Application of bioinformatics in search for cleavable peptides of SARS-CoV Mpro and chemical modification of octapeptides. Medicinal Chemistry. 2005; 3: 209-213.

32. Wang M, Yao JS, Huang ZD, Xu ZJ, Liu GP, Zhao HY. A new nucleotide-composition based fingerprint of SARS-CoV with visualization analysis. Medicinal Chemistry. 2005; 1: 39-47.

33. Wei DQ, Gan YR, Du QS. A Polypeptide and Its Derivatives as Inhibitors Against SARS, Patent Application No: CN 1560074A, January. 2005.

34. Chou KC, Wei DQ, Du QS, Sirois S, Zhong WZ. Review: Progress in computational approach to drug development against SARS. Current Medicinal Chemistry. 2006; 27: 3263-3270.

35. Gao L, Ding YS, Dai H, Shao SH, Huang ZD. A novel fingerprint map for detecting SARS-CoV. Journal of Pharmaceutical and Biomedical Analysis. 2006; 1: 246-250.

36. Wei DQ, Zhang R, Du QS, Gao WN, Li Y. Gao H, *et al*. Anti-SARS drug screening by molecular docking. Amino Acids. 2006; 1: 73-80.

37. Zhang R, Wei DQ, Du QS. Molecular modeling studies of peptide drug candidates against SARS. Medicinal Chemistry. 2006; 3: 309-314.

38. Du QS, Sun H. Inhibitor design for SARS coronavirus main protease based on "distorted key theory". Medicinal Chemistry. 2007; 1: 1-6.

39. Wang SQ, Du QS, Zhao K, Li AX, Wei DQ. Virtual screening for finding natural inhibitor against cathepsin-L for SARS therapy. Amino Acids. 2007; 1: 129-135.

40. Chou KC, Wei DQ, Du QS, Sirois S, Shen HB, Zhong WZ. Study of inhibitors against SARS coronavirus by computational approaches. in: U. Lendeckel, and N.M. Hooper, (Eds.), Proteases in Biology and Disease: Viral proteases and antiviral protease inhibitor therapy, Springer Science, Media B.V. 2009; 1-23.

41. Voet D, Voet JG, Pratt CW. Fundamentals of Biochemistry, John Wiley & Sons, New York, 2002.

42. Chou KC, Chen NY. The biological functions of low-frequency phonons. Scientia Sinica. 1977; 4: 447-457.

43. Chou KC. Review: Prediction of human immunodeficiency virus protease cleavage sites in proteins. Analytical Biochemistry. 1996; 1-14.

44. Zhang CT. Graphic analysis of codon usage strategy in 1490 human proteins. Journal of Protein Chemistry. 1993; 3: 329-335.

45. Zhang CT. Analysis of codon usage in 1562 E. Coli protein coding sequences. Journal of Molecular Biology. 1994; 1: 1-8.

46. Chou KC. Advance in predicting subcellular localization of multi-label proteins and its implication for developing multi-target drugs. . Current Medicinal Chemistry. 2019; 4918-4943.

47. Chou KC. Progresses in predicting post-translational modification. International Journal of Peptide Research and Therapeutics. 2019.

48. Chou KC. An insightful recollection since the distorted key theory was born about 23 years ago. Genomics. 2019.

49. Chou KC. Proposing pseudo amino acid components is an important milestone for proteome and genome analyses. International Journal for Peptide Research and Therapeutics. 2019.

50. Chou KC. Two kinds of metrics for computational biology. 2019.

51. Chou KC. An insightful recollection for predicting protein subcellular locations in multi-label systems. Genomics. 2019; 19: 30460-30164.

52. Chou KC. Artificial Intelligence (AI) tools constructed via the 5-steps rule for predicting post-translational modifications. Trends in Artificial Inttelengence (TIA). 2019; 60-74.

53. Chou KC. Prediction of protein cellular attributes using pseudo amino acid composition. PROTEINS: Structure, Function, and Genetics. 2001; 44: 246-255.

54. Chou KC. Using amphiphilic pseudo amino acid composition to predict enzyme subfamily classes. Bioinformatics. 2005; 1: 10-19.

55. Liu B, Liu F, Wang X, Chen J, Fang L. Pse-in-One: a web server for generating various modes of pseudo components of DNA, RNA, and protein sequences. Nucleic Acids Research. 2015; 43: W65-W71.

56. Liu B, Wu H. 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; 4: 67-91.

57. 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; 1: 236-247.

58. 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.

59. Cai YD. Nearest neighbour algorithm for predicting protein subcellular location by combining functional domain composition and pseudo amino acid composition. Biochem Biophys Res Comm. 2003; 2: 407-411.

60. Chou KC, Cai YD. Predicting protein quaternary structure by pseudo amino acid composition. PROTEINS: Structure, Function, and Genetics. 2003; 2: 282-289.

61. Chou KC, Cai YD. Prediction and classification of protein subcellular location: sequence-order effect and pseudo amino acid composition. Journal of Cellular Biochemistry. 2003; 6: 1250-1260.

62. Pan YX, 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; 4: 395-402.

63. Chou KC, Cai YD. Predicting subcellular localization of proteins by hybridizing functional domain composition and pseudo amino acid composition. Journal of Cellular Biochemistry. 2004; 6: 1197-1203.

64. Wang M, Yang J, Liu GP, Xu ZJ. Weighted-support vector machines for predicting membrane protein types based on pseudo amino acid composition. Protein Engineering, Design, and Selection. 2004; 6: 509-516.

65. Cai YD. Predicting enzyme subclass by functional domain composition and pseudo amino acid composition. Journal of Proteome Research. 2005: 3: 967-971.

66. Cai YD, Zhou GP. Predicting enzyme family classes by hybridizing gene product composition and pseudo amino acid composition. Journal of Theoretical Biology. 2005; 1: 145-149.

67. Gao Y, Shao SH, Xiao X, Ding YS, Huang YS, Huang ZD. Using pseudo amino acid composition to predict protein subcellular location: approached with Lyapunov index, Bessel function, and Chebyshev filter. Amino Acids. 2005; 4: 373-376.

68. Liu H, Yang J, Wang M, Xue L. Using Fourier spectrum analysis and pseudo amino acid composition for prediction of membrane protein types. The Protein Journal. 2005; 6: 385-389.

69. Shen HB. Using optimized evidence-theoretic K-nearest neighbor classifier and pseudo amino acid composition to predict membrane protein types. Biochemical & Biophysical Research Communications (BBRC). 2005; 1: 288-292.

70. Shen HB. Predicting protein subnuclear location with optimized evidence-theoretic K-nearest classifier and pseudo amino acid composition. Biochem Biophys Res Comm. (BBRC). 2005; 3: 752-756.

71. Cai YD. Predicting membrane protein type by functional domain composition and pseudo amino acid composition. Journal of Theoretical Biology. 2006; 2: 395-400.

72. Chen C, Tian YX, Zou XY, Cai PX, Mo JY. Using pseudo amino acid composition and support vector machine to predict protein structural class. J Theor Biol. 2006; 3: 444-448.

73. 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; 1: 116-121.

74. 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.

75. 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; 2: 252-260.

76. Shen HB, Yang J. Fuzzy KNN for predicting membrane protein types from pseudo amino acid composition. Journal of Theoretical Biology. 2006; 1: 9-13.

77. Wang SQ, Yang J. Using stacked generalization to predict membrane protein types based on pseudo amino acid composition. Journal of Theoretical Biology. 2006; 4: 941-946.

78. Xiao X, Shao SH, Ding YS, Huang ZD. Using cellular automata images and pseudo amino acid composition to predict protein subcellular location. Amino Acids. 2006; 1: 49-54.

79. Xiao X, Shao SH, Huang ZD. Using pseudo amino acid composition to predict protein structural classes: approached with complexity measure factor. Journal of Computational Chemistry. 2006; 4: 478-482.

80. 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; 4: 461-468.

81. Zhou GP, Cai YD. Predicting protease types by hybridizing gene ontology and pseudo amino acid composition. PROTEINS: Structure, Function, and Bioinformatics. 2006; 3: 681-684.

82. Chen YL, Li QZ. Prediction of apoptosis protein subcellular location using improved hybrid approach and pseudo amino acid composition. Journal of Theoretical Biology. 2007; 2: 377–381.

83. Ding YS, Zhang TL. Prediction of protein structure classes with pseudo amino acid composition and fuzzy support vector machine network. Protein & Peptide Letters. 2007; 8: 811-815.

84. Lin H, Li QZ. Predicting conotoxin superfamily and family by using pseudo amino acid composition and modified Mahalanobis discriminant. Biochem Biophys Res Commun. 2007; 2: 548-551.

85. 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; 1463-1466.

86. 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; 1610-1615.

87. 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; 1: 69-74.

88. Zhang TL, Ding YS. Using pseudo amino acid composition and binary-tree support vector machines to predict protein structural classes. Amino Acids. 2007; 4: 623-629.

89. 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; 3: 546–551.

90. 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; 1: 111-117.

91. 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; 13: 1887-1892.

92. 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; 1: 103-109.

93. Gu Q, Ding Y, Zhang T. Prediction of G-protein-coupled receptor classes with pseudo amino acid composition. IEEE Xplore. 2008.

94. 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; 4: 392-396.

95. Jiang X, Wei R, Zhao Y, Zhang T. 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; 4: 669-675.

96. Li FM, Li QZ. Using pseudo amino acid composition to predict protein subnuclear location with improved hybrid approach. Amino Acids. 2008; 1: 119-125.

97. Li FM, Li QZ. Predicting protein subcellular location using Chou's pseudo amino acid composition and improved hybrid approach. Protein & Peptide Letters. 2008; 6: 612-616.

98. Lin H. The modified Mahalanobis discriminant for predicting outer membrane proteins by using Chou's pseudo amino acid composition. Journal of Theoretical Biology. 2008; 2: 350-356.

99. Lin H, Ding H, Feng-Biao Guo FB, Zhang AY, Huang J. Predicting subcellular localization of ycobacterial proteins by using Chou's pseudo amino acid composition. Protein & Peptide Letters. 2008; 7: 739-744.

100. Shen HB. PseAAC: a flexible web-server for generating various kinds of protein pseudo amino acid composition. Analytical Biochemistry. 2008; 2: 386-388.

101. 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: 321-327.

102. Xiao X, Lin WZ. Using grey dynamic modeling and pseudo amino acid composition to predict protein structural classes. Journal of Computational Chemistry. 2008; 29: 2018-2024.

103. Xiao X, Wang P. Predicting protein structural classes with pseudo amino acid composition: an approach using geometric moments of cellular automaton image. J Theor Biol. (2008); 254: 691-696.

104. Zhang GY, Fang BS. Predicting the cofactors of oxidoreductases based on amino acid composition distribution and Chou's amphiphilic pseudo amino acid composition. J Theor Biol. 2008; 253: 310-315.

105. Zhang GY, Li HC, Gao JQ, Fang BS. Predicting lipase types by improved Chou's pseudo amino acid composition. Protein & Peptide Letters. 2008; 15: 1132-1137.

106. 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: 591-598.

107. 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: 565-572.

108. Zhang TL, Ding YS. Prediction protein structural classes with pseudo amino acid composition: approximate entropy and hydrophobicity pattern. Journal of Theoretical Biology. 2008; 250: 186-193.

109. 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.

110. Chou KC. Pseudo amino acid composition and its applications in bioinformatics, proteomics and system biology. Current Proteomics. 2009; 6: 262-274.

111. 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: 351-355.

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

113. Gao QB, Jin ZC, Ye XF, Wu C, He J. Prediction of nuclear receptors with optimal pseudo amino acid composition. Analytical Biochemisry. 2009; 387: 54-59.

114. 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: 415-425.

115. Lin H, Wang H, Ding H, Chen YL, Li QZ. Prediction of Subcellular Localization of Apoptosis Protein Using Chou's Pseudo Amino Acid Composition. Acta Biotheoretica. 2009; 57: 321-330.

116. Qiu JD, Huang JH, iang R.P, 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. Anal Biochem. 2009; 390: 68-73.

117. Xiao X, Wang P. Predicting protein quaternary structural attribute by hybridizing functional domain composition and pseudo amino acid composition. Journal of Applied Crystallography. 2009; 42: 169-173.

118. 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: 366–372.

119. Esmaeili 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: 203-209.

120. Gao QB, Ye XF, Jin ZC, He J. Improving discrimination of outer membrane proteins by fusing different forms of pseudo amino acid composition. AnalBiochem. 2010; 398: 52-59.

121. 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: 500-504.

122. 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: 559-567.

123. Kandaswamy KK, Pugalenthi 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: 1473-1479.

124. 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 Pept Lett. 2010; 17: 1263-1269.

125. Mohabatkar H. Prediction of cyclin proteins using Chou's pseudo amino acid

126. 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 Pept Lett. 2010; 17: 1466-1472.

127. 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 Pept Lett. 2010; 17: 715-722.

128. 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: 320-327.

129. Wang YC, Wang XB, Yang ZX, Deng NY. Prediction of enzyme subfamily class via pseudo amino acid composition by incorporating the conjoint triad feature. Protein & Peptide Letters. 2010; 17: 1441-1449.

130. 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: 62-67.

131. 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.

132. 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 Pept Lett. 2011; 18: 58-63.

133. 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: 1612-1617.

134. 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: 10-17.

135. 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: 552-558.

136. 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: 217-223.

137. Jingbo X, Silan Z, Feng S, Huijuan X, Xuehai H, Xiaohui N. Using the concept of pseudo amino acid composition to predict resistance gene against Xanthomonas oryzae pv. oryzae in rice: An approach from chaos games representation. J Theor Biol. 2011; 284: 16-23.

138. 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: 64-69.

139. 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: 1219-1225.

140. 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: 16963-16968.

141. 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: 1244-1250.

142. 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: 132-137.

143. 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. J Theor Biol. 2011; 281: 18-23.

144. 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: 191-197.

145. 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: 1132-1138.

146. 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: 129-134.

147. 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: 625-633.

148. Shu M, Cheng X, Zhang Y, Wang Y, Lin, Wang L, Lin Z. Predicting the Activity of ACE Inhibitory Peptides with a Novel Mode of Pseudo Amino Acid Composition. Protein & Peptide Letters. 2011; 18: 1233-1243.

149. 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 Pept Lett. 2011; 18: 684-689.

150. 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: 480-487.

151. Xiao X. Using pseudo amino acid composition to predict protein attributes via cellular automata and other approaches. Current Bioinformatics. 2011; 6: 251-260.

152. Xiao X, Wang P. GPCR-2L: Predicting G protein-coupled receptors and their types by hybridizing two different modes of pseudo amino acid compositions. Molecular Biosystems. 2011; 7: 911-919.

153. Zia UrR., Khan A. Prediction of GPCRs with Pseudo Amino Acid Composition: Employing Composite Features and Grey Incidence Degree Based Classification. Protein Pept Lett.. 2011; 18: 872-878.

154. Zou D, He Z, He J, Xia Y. Supersecondary structure prediction using Chou's pseudo amino acid composition. Journal of Computational Chemistry. 2011; 32: 271-278.

155. Cao JZ, Liu WQ, Gu H. 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: 1163-1169.

156. 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: 422-429.

157. 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.

158. 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: 117-119.

159. 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.

160. Fan GL, Li QZ. Predicting protein submitochondria locations by combining different descriptors into the general form of Chou's pseudo amino acid composition. Amino Acids. 2012; 43: 545-555.

161. 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: 73-77.

162. 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: 375-387.

163. Lin WZ, Fang JA, Xiao Xb. 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.

164. 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: 439-449.

165. Nanni L, Brahnam S, Lumini A. Wavelet images and Chou's pseudo amino acid composition for protein classification. Amino Acids. 2012; 43: 657-665.

166. 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 Biolology and Bioinformatics. 2012; 9: 467-475.

167. 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: 940-948.

168. 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: 1170-1176.

169. 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: 564-574.

170. 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: 1619-1625.

171. 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.

172. Zia-ur-Rehman, Khan A. 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.

173. 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.

174. 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.

175. 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: 41-48.

176. 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; 23: 597-609.

177. 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. Biosystems. 2013; 113: 50-57.

178. 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: 327-334.

179. 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: 205-212.

180. Khosravian 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: 180-186.

181. 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.

182. Liu Bg, 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: 775-782.

183. Mohabatkar H, Beigi MM, Abdolahi K, 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: 133-137.

184. 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: 1660-1667.

185. 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. Protein Pept Lett. 2013; 20: 781-95.

186. 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: 40-48.

187. 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: 309-317.

188. Xiaohui N, Nana L, Jingbo X, Dingyan C, Yuehua P, X. Yang, *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: 211-217.

189. Xu Y, Ding J, Wu LY. 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.

190. 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: 3495-3506.

191. 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.

192. 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: 10410-10423.

193. 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: 12-18.

194. Liu B, Xu J, Lan X, Xu R, Zhou J, Wang X. 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.

195. Mondal S, Pai P. Chou's pseudo amino acid composition improves sequence-based antifreeze protein prediction. J Theor Biol. 2014; 356: 30-35.

196. 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: 109-116.

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

198. Xu Y, Wen X, Shao XJ, Deng NY. 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: 7594-7610.

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

200. 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: 412-418.

201. 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: 105-110.

202. 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.

203. 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; 47: 1485-1493.

204. Huang C, 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: 547-556.

205. 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.

206. 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.

207. 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: 1919-1931.

208. 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: 2639-2645.

209. Xu R, Zhou J, Liu B, He YA, Zou Q, Wang X. Identification of DNA-binding proteins by incorporating evolutionary information into pseudo amino acid composition via the top-n-gram approach. Journal of Biomolecular Structure & Dynamics (JBSD). 2015; 33: 1720-1730.

210. 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. Mol Biosyst. 2015; 11: 558-563.

211. 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.

212. 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.

213. 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: 138-142.

214. 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 Biomol Struct Dyn (JBSD). 2016; 34: 1946-1961.

215. 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.

216. Tang H, Chen W, Lin H. Identification of immunoglobulins using Chou's pseudo amino acid composition with feature selection technique. Mol Biosyst. 2016; 12: 1269-1275.

217. 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: 105-115.

218. 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: 23-29.

219. 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: 551-557.

220. 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: 110-117.

221. Rahimi M, Bakhtiarizadeh 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: 128-136.

222. 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.

223. Yu B, Lou L, Li S, Zhang Y, Qiu W, Wu X, et al. Tian. 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: 260-273.

224. 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.

225. Ju Z, Wang SY. Prediction of citrullination sites by incorporating k-spaced amino acid pairs into Chou's general pseudo amino acid composition. 2018; 664: 78-83.

226. 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: 41-48.

227. Mei J, Zhao J. Prediction of HIV-1 and HIV-2 proteins by using Chou's pseudo amino acid compositions and different classifiers. 2018; 8: 2359.

228. 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; 447: 147-153.

229. 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 Comput Biol Bioinform. 2019.

230. Ehsan A, 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: 124-133.

231. Butt AH, Khan YD. Prediction of S-Sulfenylation Sites Using Statistical Moments Based Features via Chou's 5-Step Rule. International Journal of Peptide Research and Therapeutics (IJPRT). 2018.

232. Barukab O, Khan YD, Khan SA, Chou KC. iSulfoTyr-PseAAC: Identify tyrosine sulfation sites by incorporating statistical moments via Chou's 5-steps rule and pseudo components Current Genomics. 2019; 20.

233. Butt AH, Khan YD. Prediction of S-Sulfenylation Sites Using Statistical Moments Based Features via Chou's 5-Step Rule. International Journal of Peptide Research and Therapeutics (IJPRT). 2019; 1-11.

234. 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: 3119-3132.

235. Hussain W, Khan YD, Khan SD, Rasool N, Khan SA, Chou KC. SPalmitoylC-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.

236. 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.

237. 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.

238. 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.

239. Khan ZU, Ali F, Khan IA, Hussain Y, Pi D. iRSpot-SPI: Deep learning-based recombination spots prediction byincorporating secondary sequence information coupled withphysio-chemical properties via Chou's 5-step rule and pseudo components. Chemometrics and Intelligent Laboratory Systems (CHEMOLAB). 2019; 189: 169-180.

240. 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.

241. 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.

242. Le NQK, Yapp EKY, Ou YY, Yeh HY. iMotor-CNN: Identifying molecular functions of cytoskeleton motor proteins using 2D convolutional neural network via Chou's 5-step rule. Anal Biochem. 2019; 575: 17-26.

243. Liang Y, Zhang S. Identifying DNase I hypersensitive sites using multi-features fusion and F-score features selection via Chou's 5-steps rule. Biophysical Chemistry. 2019; 253: 106227.

244. 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; 193.

245. 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: 43-49.

246. 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; 1-15.

247. Tahir M, Tayara H, Chong HT. 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.

248. Vishnoi S, Garg P, Arora P. Physicochemical n-Grams Tool: A tool for protein physicochemical descriptor generation via Chou's 5-steps rule. Chem Biol Drug Des. 2019.

249. Yang L, Lv Y, Wang S, Zhang Q, Pan Y, Su D, et al. Identifying FL11 subtype by characterizing tumor immune microenvironment in prostate adenocarcinoma via Chou's 5-steps rule. Genomics. 2019.

250. Chou KC. Review: Structural bioinformatics and its impact to biomedical science. Current Medicinal Chemistry. 2004; 11: 2105-2134.

251. Chou KC, Jiang SP. Studies on the rate of diffusion-controlled reactions of enzymes. Scientia Sinica. 1974; 27: 664-680.

252. Chou KC, Kuo CK, Li TT. The quantitative relations between diffusion-controlled reaction rate and characteristic parameters in enzyme-substrate reaction system: 2. Charged substrates. Scientia Sinica.1975; 18: 367-380.

253. Li TT, KC Chou KC. The quantitative relations between diffusion-controlled reaction rate and characteristic parameters in enzyme-substrate reaction system: 1. Neutral substrates. Scientia Sinica. 1976; 19: 117-136.

254. Chou KC, Forsen S. Diffusion-controlled effects in reversible enzymatic fast reaction system: Critical spherical shell and proximity rate constants.

Biophysical Chemistry. 1980; 12: 255-263.

255. Chou KC, Zhou GP. Role of the protein outside active site on the diffusion-controlled reaction of enzyme. Journal of American Chemical Society. 1982; 104: 1409-1413.

256. Chou KC, Chen NY, Forsen S. The biological functions of low-frequency phonons: 2. Cooperative effects. Chemica Scripta. 1981; 18: 126-132.

257. Chou KC. Low-frequency vibrations of helical structures in protein molecules. Biochemical Journal. 1983; 209: 573-580.

258. Chou KC. Identification of low-frequency modes in protein molecules. Biochemical Journal. 1983; 215: 465-469.

259. Chou KC. Biological functions of low-frequency vibrations (phonons). 3. Helical structures and microenvironment. Biophysical Journal. 1984; 45: 881-889.

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

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

262. Chou KC. Low-frequency motions in protein molecules: beta-sheet and beta-barrel. Biophysical Journal. 1985; 48: 289-297.

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

264. Chou KC, Kiang YS. The biological functions of low-frequency phonons: 5. A phenomenological theory. Biophysical Chemistry.1985; 22: 219-235.

265. Chou KC. Origin of low-frequency motion in biological macromolecules: A view of recent progress of quasi-continuity model. Biophysical Chemistry. 1986; 25: 105-116.

266. Chou KC. The biological functions of low-frequency phonons: 6. A possible dynamic mechanism of allosteric transition in antibody molecules. Biopolymers. 1987; 26: 285-295.

267. Chou KC. Review: Low-frequency collective motion in biomacromolecules and its biological functions. Biophysical Chemistry. 1988; 30: 3-48.

268. Chou KC, Maggiora GM. The biological functions of low-frequency phonons: 7. The impetus for DNA to accommodate intercalators. British Polymer Journal. 1988; 20: 143-148.

269. Chou KC. Low-frequency resonance and cooperativity of hemoglobin. Trends in Biochemical Sciences. 1989; 14: 212-213.

270. 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.

271. Liu H, Wang M, Chou KC. Low-frequency Fourier spectrum for predicting membrane protein types. Biochem Biophys Res Commun (BBRC). 2005; 336: 737-739.

272. Chou KC, Jiang SP, Liu WM, Fee CH. Graph theory of enzyme kinetics: 1. Steady-state reaction system. Scientia Sinica. 1979; 22: 341-358.

273. Chou KC, Forsen S. Graphical rules for enzyme-catalyzed rate laws. Biochemical Journal. 1980; 187: 829-835.

274. Chou KC. A new graphical rule for rate laws of enzyme reactions with branched pathways. Canadian Journal of Biochemistry. 1981; 59: 757-761.

275. Chou KC, Carter RE, Forsen S. A new graphical method for deriving rate equations for complicated mechanisms. Chemica Scripta. 1981; 18: 82-86.

276. Chou KC, Forsen S. Graphical rules of steady-state reaction systems. Canadian Journal of Chemistry. 1981; 59: 737-755.

277. Chou KC. Advances in graphical methods of enzyme kinetics. Biophysical Chemistry. 1983; 17: 51-55.

278. Chou KC. Graphic rules in steady and non-steady enzyme kinetics. Journal of Biological Chemistry. 1989; 264: 12074-12079.

279. 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.

280. Chou KC. Graphic rule for non-steady-state enzyme kinetics and protein folding kinetics. Journal of Mathematical Chemistry. 1993; 12: 97-108.

281. Chou KC. Graphic rule for drug metabolism systems. Current Drug Metabolism. 2010; 11: 369-378.

282. Wu ZC, Xiao X. 2D-MH: A web-server for generating graphic representation of protein sequences based on the physicochemical properties of their constituent amino acids. Journal of Theoretical Biology. 2010; 267: 29-34.