Enliven Archive
Intensifying Thoughts

JOURNAL OF MATHEMATICS, STATISTICS AND COMPUTING

OPEN ACCESS
Research Article

An Insightful 20-Year Recollection Since the Birth of Pseudo Amino Acid Components

Chou KC[1,2*]

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

***Corresponding Author:**

Kuo-Chen Chou,
Professor Department of Computational Biology & Biomedicine,
Gordon Life Science Institute,Boston, Massachusetts 02478, United States of America.
Email: kcchou@gordonlifescience.org

Received: November 21, 2019
Accepted: December 12, 2019
Published: December 17, 2019

Abstract

In this paper, the author has recalled the proposal of pseudo amino acid components, and its significant and substantial impacts to proteome and genome analyses.

Keywords: PseAAC; PseKNC; 5-steps rule; Byproducts; Distorted Key Theory.

Background

Given a protein sequence P with *L* amino acid residues, its most straightforward expression is:

$$\mathbf{P} = R_1 R_2 R_3 R_4 R_5 R_6 R_7 \ldots\ldots R_L \text{ ------ (1)}$$

where R_1 is its 1st amino acid residue, R_2 its 2nd residue, and so forth. With the avalanche of protein sequences generated in the postgenomic age, scientists are facing a critical challenge; i.e., how to express a protein sequence of Eq.1 with a vector but still considerably keep its sequential order information or characteristic patterns. This is because all the existing operation engines, such as "Optimization" algorithm [1], "Correlation Angle" algorithm [2-4], "Component Coupled" algorithm [5-10], "Covariance Discriminant" or CD algorithm [7, 11-19], "Principal Component Analysis" algorithm [20], "Top-n-Gram" algorithm [21], "SLLE" algorithm [22], "Neural Network" or NN algorithm [23-25], "Ensemble Classifier" algorithm [26, 27], "Support Vector Machine" or SVM algorithm [28-30], "Random Forest" algorithm [31, 32], "Conditional Random Field" algorithm [14], "Nearest Neighbor" or NN algorithm [33, 34], "K-Nearest Neighbor" or KNN algorithm [35-37], "Optimized Evidence-Theoretic K Nearest Neighbor" or OET-KNN algorithm [38-41], and "Fuzzy K-Nearest Neighbor" algorithm [42-46], can only handle a vector but not a sequence sample. The simplest discrete model used to represent a protein sample is its amino acid (AA) composition or AAC [47]. According to the AAC-discrete model, the protein of Eq.1 can be expressed by [7].

$$\mathbf{P} = [f_1 f_2 f_3 \ldots\ldots f_{20}]^T \text{ ------ (2)}$$

where f_i (i = 1, 2, ..., 20) are the normalized occurrence frequencies of the 20 native amino acidsin P, and T the transposing operator. Many methods for predicting various protein attributes were based on the AAC-discrete model (see, e.g., [2, 4, 8-11, 15, 17, 20, 23, 48-60].

Unfortunately, as we can see from Eq.2, a vector defined in a discrete model may completely lose the sample's sequence-order information [61], and hence the prediction quality thus obtained might be limited. This is the main shortcoming of the AAC discrete model.

Birth of PseAAC

To avoid completely losing the sequence-order information, a completely different discrete model, or the so-called "pseudo amino acid composition" (PseAAC) model [61], was proposed to represent the sample of a protein, as formulated by

$$P = [p_1 \ldots p_{20} p_{20+1} \ldots p_{20+\lambda}]^T \text{ -----(3)}$$

where the first 20 elements are associated with the 20 elements in Eq.2 orthe 20 amino acid components of the protein, while the additional "λ" factors are used to incorporate some sequence-order information via various modes. Typically, these additional factors are a series of rank-different correlation factors along a protein

chain, but they can also be any combinations of other factors so long as they can reflect some sorts of sequence-order effects in one way or the other. For the convenience of users, a web-server called "PseAAC" [62] was established at http://www.csbio.sjtu.edu.cn/bioinf/PseAAC/, by which some commonly used PseAAC forms can be automatically generated. For more about PseAAC, see an insightful Wikipedia article at https://en.wikipedia.org/wiki/Pseudo_amino_acid_composition.

From PseAAC to PseKNC

Shortly afterwards, the idea of PseAAC in dealing with peptide/protein sequences was extended to PseKNC for dealing with DNA/RNA sequences [29, 63-90], which has been found very successful as well. For more about PseKNC, see an insightful Wikipedia article at https://en.wikipedia.org/wiki/Pseudo_K-tuple_nucleotide_composition.

Impacts on Attribute Prediction

Ever since their birth, PseAAC and PseKNC have remarkable impacts in nearly all the areas of computational proteomics and computational proteomics, respectively, such as predicting enzymes and their family/sub-family classification [91-95], protein subcellular location prediction [13, 96-104], apoptosis protein subcellular location prediction [105-109], mycobacterial protein subcellular location prediction [110], predicting protein subnuclear localization [111-113], predicting protein subchloroplast locations [114], predicting protein submitochondria locations [115-117], predicting membrane proteins and their types [40, 43, 118-122], discrimination of outer membrane proteins [123, 124], identifying transmembrane regions in proteins [125], identifying proteases and their types [126, 127], predicting protein solubility [128], identifying GPCRs and their classes [44, 129-133], prediction of nuclear receptors [134], prediction of cyclin proteins [135], identifying bacterial secreted proteins [136], identifying risk type of human papillomaviruses [137], prediction of cell wall lytic enzymes [138], prediction of lipases types [139], predicting conotoxin superfamily and family [140, 141], predicting the cofactors of oxidoreductases [142], predicting DNA-binding proteins [143], predict protein structural classes [19, 144-153], supersecondary structure prediction [154], protein secondary structure content prediction [155], and predicting protein quaternary structural attributes [156-161], fold pattern prediction [162-163], and others (e.g., [164]). Detailed below are just a few.

Protein structural class prediction

The knowledge of protein 3D (three-dimensional) structures or their complexes with ligands is vitally important for rational drug design. Although X-ray crystallography is a powerful tool in determining these structures, it is time-consuming and expensive, and not all proteins can be successfully crystallized. Membrane proteins are difficult to crystallize and most of them will not dissolve in normal solvents. Therefore, so far very few membrane protein structures have been determined. NMR is indeed a very powerful tool in determining the 3D structures of membrane proteins (see, e.g., [165-182]), but it is also time-consuming and costly. The structural class of a protein can provide useful clues for finding the structure of bovine somatotropin by the heuristic approach [183], and stimulate the modeling of 3D structures for many other proteins [184-202] important for drug development.

Protein subcellular location prediction

It is the PseAAC, particularly the general PseAAC [203] that have significantly improve the prediction quality for the localization of proteins in a multi-label system [26, 27, 36, 204-228].

PTM site prediction

Post-translational modification (PTM) refers to the covalent and generally enzymatic modification of proteins following protein biosynthesis. Prediction of PTM (post-translational modification) sites is currently a very hot topic. Again, it is PseAAC or PseKNC that has played the key role in this regard (see, e.g., [14, 75, 85, 229-271].

Byproducts

Meanwhile the introduction of PseAAC has also yielded some important byproducts.

Five-steps rule

Stimulated by the PseAAC and PseKNC, a very important rule, called "5-steps rule", has been established, as briefed in the Abstract and the 2nd paragraph of a comprehensive review paper [203] and detailed in its Sections 2, 3, 4, 5, and 6. As the guidelines for proteome or genome analysis, the 5-steps rule has been used nearly all the areas of computational proteomics or genomics [28, 29, 64, 70, 74, 77, 86, 215-225, 227, 228, 238, 255-257, 260, 263, 267-269, 271, 292]. For more about the5-steps rule, see an insightful Wikipedia article at https://en.wikipedia.org/wiki/5-step_rules.

Distorted key theory

Originally proposed in a comprehensive review paper [293], the "distorted key theory" has been further developed [192] and applied [294-297]. For more about the fantastic "distorted key theory, see an insightful Wikipedia article at https://en.wikipedia.org/wiki/Chou%27s_distorted_key_theory_for_peptide_drugs.

Conclusion and perspective

PseAAC is indeed a milestone for proteome analysis. Also, it has not escaped our notice that stimulated by the eight master pieces of pioneering papers from the then Chairman of Nobel Prize Committee StureForsen [298-305], many follow-up papers have been published [306-349, 257, 271, 278, 350]. They are very useful for in-depth investigation into the topic of the current paper, and we will use them in our future efforts.

References

[1]. Zhang CT, Chou KC. An optimization approach to predicting protein structural class from amino acid composition. Protein Sci. 1992 Mar;1(3):401-8.PMID: 1304347.

[2]. Chou KC, Zhang CT. A correlation coefficient method to predicting protein structural classes from amino acid compositions. Eur J Biochem. 1992 Jul 15;207(2):429-3. PMID: 1633801.

[3]. Chou KC. Prediction of protein folding types from amino acid composition by correlation angles. Amino Acids. 1994 Oct 1;6(3):231-46.

[4]. Nakashima H, Nishikawa K. Discrimination of intracellular and extracellular proteins using amino acid composition and residue-pair frequencies. J Mol Biol. 1994 Apr 22;238(1):54-61.

PMID:1633801.
[5]. K.C. Chou, C.T. Zhang, Predicting protein folding types by distance functions that make allowances for amino acid interactions. Journal of Biological Chemistry 269 (1994) 22014-22020.
[6]. K.C. Chou, Does the folding type of a protein depend on its amino acid composition? FEBS Lett. 1995 Apr 17;363(1-2):127-31. PMID: 7729532.
[7]. Chou KC. A novel approach to predicting protein structural classes in a (20-1)-D amino acid composition space. Proteins. 1995 Apr;21(4):319-44.PMID: 7567954.
[8]. Chou KC. A novel approach to predicting protein structural classes in a (20–1)-D amino acid composition space. Proteins: Structure, Function, and Bioinformatics. 1995 Apr;21(4):319-44.
[9]. Chou KC, Liu WM, Maggiora GM, Zhang CT. Prediction and classification of domain structural classes. Proteins: Structure, Function, and Bioinformatics. 1998 Apr 1;31(1):97-103.
[10]. Chou KC, Maggiora GM. Domain structural class prediction. Protein Engineering. 1998 Jul 1;11(7):523-38.
[11]. Chou KC. The convergence-divergence duality in lectin domains of the selectin family and its implications. FEBS Lett. 1995 Apr 17;363(1-2):123-6.PMID: 7537222.
[12]. Chou KC. The convergence-divergence duality in lectin domains of selectin family and its implications. FEBS letters. 1995 Apr 17;363(1-2):123-6.
[13]. Chou KC, Elrod DW. Bioinformatical analysis of G-protein-coupled receptors. J Proteome Res. 2002 Oct 1;1(5):429-33.
[14]. Chou KC, Elrod DW. Prediction of enzyme family classes. J Proteome Res. 2003 Apr 1;2(2):183-90.
[15]. Gao Y, Shao S, Xiao X, Ding Y, Huang Y, Huang Z, et al. Using pseudo amino acid composition to predict protein subcellular location: approached with Lyapunov index, Bessel function, and Chebyshev filter. Amino Acids. 2005 Jun;28(4):373-6. Epub 2005 May 17.PMID: 15889221.
[16]. Chou KC. Prediction of G-protein-coupled receptor classes. J Proteome Res. 2005 Aug 8;4(4):1413-8.
[17]. Xiao X, Wang P, Chou KC. Predicting protein structural classes with pseudo amino acid composition: an approach using geometric moments of cellular automaton image. J Theor Biol. 2008 Oct 7;254(3):691-6.PMID: 18634802.
[18]. Chen W, Lin H, Feng PM, Ding C, Zuo YC, Chou KC. iNuc-PhysChem: A Sequence-Based Predictor for Identifying Nucleosomes via Physicochemical Properties. PLoS One. 2012;7(10):e47843. PMID: 23144709.
[19]. 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 8 (2013) e55844.PMID: 23409062.
[20]. Du QS1, Jiang ZQ, He WZ, Li DP, Chou KC. Amino acid principal component analysis (AAPCA) and its applications in protein structural class prediction. J Biomol Struct Dyn. 2006 Jun;23(6):635-40.PMID: 16615809.
[21]. Xu R, Zhou J, Liu B, He Y, 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. J Biomol Struct Dyn. 2015;33(8):1720-30.PMID: 25252709.
[22]. Wang M, Yang J, Xu ZJ, Chou KC. SLLE for predicting membrane protein types. J Theor Biol. 2005 Jan 7;232(1):7-15.PMID: 15498588.
[23]. Thompson TB, Chou KC, Zheng C. Neural network prediction of the HIV-1 protease cleavage sites. J Theor Biol. 1995 Dec 21;177(4):369-79.
[24]. Cai YD, Chou KC. Artificial neural network for predicting alpha-turn types. Anal Biochem. 1999 Mar 15;268(2):407-9.PMID: 10075834.
[25]. Vitale G, Pellizzari R, Recchi C, Napolitani G, Mock M, Montecucco C. Biochem. Biophys. Res. Commun.
[26]. Shen HB, Yang J, Chou KC., Euk-PLoc: an ensemble classifier for large-scale eukaryotic protein subcellular location prediction. Amino Acids. 2007 Jul;33(1):57-67.PMID: 17235453.
[27]. 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 Feb 15;85(3):233-40.PMID: 17120237.
[28]. Xiao X, Wang P, Chou KC. iNR-PhysChem: A Sequence-Based Predictor for Identifying Nuclear Receptors and Their Subfamilies via Physical-Chemical Property Matrix. PLoS One. 2012;7(2):e30869.
[29]. 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 442 (2013) 118-25.PMID: 23756733.
[30]. Chen W, Feng PM, Lin H, Chou KC. iRSpot-PseDNC: identify recombination spots with pseudo dinucleotide composition. Nucleic Acids Res. 2013 Apr 1;41(6):e68.
[31]. Lin WZ, Fang JA, Xiao X, Chou KC. iDNA-Prot: Identification of DNA Binding Proteins Using Random Forest with Grey Model. PLoS One. 2011;6(9):e24756.PMID: 21935457.
[32]. K.K. Kandaswamy, T. Martinetz, S. Moller, P.N. Suganthan, S. Sridharan, G. Pugalenthi, AFP-Pred: A random forest approach for predicting antifreeze proteins from sequence-derived properties. J Theor Biol. 2011 Feb 7;270(1):56-62.PMID: 21056045.
[33]. Cai YD, Chou KC. Predicting subcellular localization of proteins in a hybridization space. Bioinformatics. 2004 May 1;20(7):1151-6. Epub 2004 Feb 5.PMID: 14764553.
[34]. Chou KC, Cai YD. Prediction of protease types in a hybridization space. Biochem Biophys Res Commun. 2006 Jan 20;339(3):1015-20.
[35]. Chou KC, Shen HB. Predicting eukaryotic protein subcellular location by fusing optimized evidence-theoretic K-nearest neighbor classifiers. J Proteome Res. 2006 Aug 4;5(8):1888-97.
[36]. Chou KC, Shen HB. Hum-PLoc: a novel ensemble classifier for predicting human protein subcellular localization. Biochem Biophys Res Commun. 2006 Aug 18;347(1):150-7.
[37]. Chou KC, Shen HB. Large-scale predictions of gram-negative bacterial protein subcellular locations. J Proteome Res. 2006 Dec 1;5(12):3420-8.
[38]. Chou KC, Shen HB. Euk-mPLoc: a fusion classifier for large-scale eukaryotic protein subcellular location prediction by incorporating multiple sites. J Proteome Res. 2007 May 4;6(5):1728-34.
[39]. Chou KC, Shen HB. Signal-CF: a subsite-coupled and window-fusing approach for predicting signal peptides. Biochem Biophys Res Commun. 2007 Jun 8;357(3):633-40.
[40]. Shen H, Chou KC., Using optimized evidence-theoretic K-nearest neighbor classifier and pseudo amino acid composition to predict membrane protein types. Biochem Biophys Res Commun. 2005 Aug 19;334(1):288-92.PMID: 16002049.
[41]. Shen HB, Chou KC. A top-down approach to enhance the power of predicting human protein subcellular localization: Hum-mPLoc 2.0. Analytical biochemistry. 2009 Nov 15;394(2):269-74.
[42]. 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. Mol. Biosyst. 2011;7(3):911-9.
[43]. Shen HB, Yang J, Chou KC. Fuzzy KNN for predicting membrane protein types from pseudo amino acid composition. J Theor Biol. 2006 May 7;240(1):9-13. Epub 2005 Sep 28.PMID: 16197963.
[44]. Xiao X, Min JL, Wang P, Chou KC. iGPCR-Drug: A web server for predicting interaction between GPCRs and drugs in cellular networking. PloS one. 2013 Aug 27;8(8):e72234.
[45]. Xiao X, Min JL, Wang P, Chou KC. iCDI-PseFpt: Identify the channel-drug interaction in cellular networking with PseAAC and molecular fingerprints. J Theor Biol. 2013 Nov 21;337:71-9. PMID: 23988798.
[46]. Xiao X, Wang P, Lin WZ, Jia JH, Chou KC. iAMP-2L: a two-level multi-label classifier for identifying antimicrobial peptides and their functional types. Anal. Biochem. 2013 May 15;436(2):168-

77.
[47]. Nakashima H, Nishikawa K, Ooi T.The folding type of a protein is relevant to the amino acid composition. J Biochem. 1986 Jan;99(1):153-62.PMID: 3957893.
[48]. Klein P, Delisi C. Prediction of protein structural class from the amino acid sequence. Biopolymers: Original Research on Biomolecules. 1986 Sep;25(9):1659-72.
[49]. Klein P. Prediction of protein structural class by discriminant analysis. Biochim Biophys Acta. 1986 Nov 21;874(2):205-15.
[50]. Chou PY. Prediction of protein structural classes from amino acid compositions. InPrediction of protein structure and the principles of protein conformation 1989 (pp. 549-586). Springer, Boston, MA.
[51]. Metfessel BA, Saurugger PN, Connelly DP, Rich SS. Cross-validation of protein structural class prediction using statistical clustering and neural networks. Protein Sci. 1993 Jul;2(7):1171-82.PMID: 8358300.
[52]. Cedano J, Aloy P, Perez-Pons JA, Querol E. Relation between amino acid composition and cellular location of proteins. J. Mol. Biol. 1997 Feb 28;266(3):594-600.
[53]. Zhou GP. An intriguing controversy over protein structural class prediction. Journal of protein chemistry. 1998 Nov 1;17(8):729-38.
[54]. Liu WM, Chou KC. Prediction of protein structural classes by modified Mahalanobis discriminant algorithm. J Protein Chem. 1998 Apr 1;17(3):209-17.
[55]. Chou KC A key driving force in determination of protein structural classes. Biochem Biophys Res Commun. 1999 Oct 14;264(1):216-24.PMID: 10527868.
[56]. Chou KC. Review: Prediction of protein structural classes and subcellular locations. Curr Protein Pept Sci. 2000 Sep;1(2):171-208. PMID: 12369916.
[57]. Zhou GP, Assa-Munt N. Some insights into protein structural class prediction. Proteins. 2001 Jul 1;44(1):57-9.PMID: 11354006.
[58]. Zhou GP, Doctor K. Subcellular location prediction of apoptosis proteins. Proteins: Structure, Function, and Bioinformatics. 2003 Jan 1;50(1):44-8.
[59]. Niu B, Cai YD, Lu WC, Li GZ, Chou KC. Predicting protein structural class with AdaBoost learner. Protein Pept Lett. 2006;13(5):489-92.PMID: 16800803.
[60]. Jahandideh S, Abdolmaleki P, Jahandideh M, Asadabadi EB. Novel two-stage hybrid neural discriminant model for predicting proteins structural classes. Biophys Chem. 2007 Jun;128(1):87-93. Epub 2007 Mar 13.PMID: 17467878.
[61]. Chou KC. Prediction of protein cellular attributes using pseudo-amino acid composition. Proteins: Structure, Function, and Bioinformatics. 2001 May 15;43(3):246-55.
[62]. Shen HB, Chou KC. PseAAC: a flexible web-server for generating various kinds of protein pseudo amino acid composition. Anal Biochem. 2008 Feb 15;373(2):386-8. Epub 2007 Oct 13.PMID: 17976365.
[63]. Zhou X, Li Z, Dai Z, Zou X. Predicting promoters by pseudo-trinucleotide compositions based on discrete wavelets transform. J Theor Biol. 2013 Feb 21;319:1-7.PMID: 23211833.
[64]. 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. Anal Biochem. 2014 Oct 1;462:76-83.PMID: 25016190.
[65]. Chen W, Feng PM, Lin H, Chou KC, iSS-PseDNC: identifying splicing sites using pseudo dinucleotide composition. Biomed Res Int. 2014;2014:623149.PMID: 24967386.
[66]. Chen W, Lei TY, Jin DC, Lin H, Chou KC. PseKNC: a flexible web-server for generating pseudo K-tuple nucleotide composition. Anal Biochem. 2014 Jul 1;456:53-60.PMID: 24732113.
[67]. Feng P, Chen W, Lin H. Prediction of CpG island methylation status by integrating DNA physicochemical properties. Genomics. 2014 Oct;104(4):229-33.PMID: 25172426.
[68]. Feng P, Jiang N, Liu N. Prediction of DNase I hypersensitive sites by using pseudo nucleotide compositions. ScientificWorldJournal. 2014;2014:740506.PMID: 25215331.
[69]. 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 30 (2014) 1522-1529.PMID: 24504871.
[70]. 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 Res. 2014 Dec 1;42(21):12961-72.PMID: 25361964.
[71]. Qiu WR, Xiao X, Chou KC, iRSpot-TNCPseAAC: Identify recombination spots with trinucleotide composition and pseudo amino acid components. Int J Mol Sci (IJMS) 15 (2014) 1746-1766.PMID: 24469313.
[72]. 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 Jan 1;31(1):119-20.
[73]. Liu B, Liu F, Fang L, Wang X, Chou KC. repDNA: a Python package to generate various modes of feature vectors for DNA sequences by incorporating user-defined physicochemical properties and sequence-order effects. Bioinformatics. 2015 Apr 15;31(8):1307-9.PMID: 25504848.
[74]. Liu Z, Xiao X, Qiu WR, Chou KC. iDNA-Methyl: Identifying DNA methylation sites via pseudo trinucleotide composition. Analytical biochemistry. 2015 Apr 1;474:69-77.
[75]. Chen W, Tang H, Ye J, Lin H, Chou KC. iRNA-PseU: Identifying RNA pseudouridine sites Molecular Therapy. Mol Ther Nucleic Acids. 2016;5:e332.PMID: 28427142.
[76]. Liu B, Fang L, Liu F, Wang X, Chou KC, iMiRNA-PseDPC: microRNA precursor identification with a pseudo distance-pair composition approach. J Biomol Struct Dyn. 2016;34(1):223-35. PMID: 25645238.
[77]. 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. 2015 Oct 17;32(3):362-9.
[78]. Liu B, Liu F, Fang L, Wang X, Chou KC. repRNA: a web server for generating various feature vectors of RNA sequences. Molecular Genetics and Genomics 291 (2016) 473-481. Mol Genet Genomics. 2016 Feb;291(1):473-81.PMID: 26085220.
[79]. Liu B, Long R, Chou KC. iDHS-EL: Identifying DNase I hypersensi-tivesites by fusing three different modes of pseudo nucleotide composition into an ensemble learning framework. Bioinformatics 32 (2016) 2411-2418.PMID: 27153623.
[80]. X. Xiao, H.X. Ye, Z. Liu, J.H. Jia, iROS-gPseKNC: predicting replication origin sites in DNA by incorporating dinucleotide position-specific propensity into general pseudo nucleotide composition. Oncotarget. 2016 Jun 7;7(23):34180-9.PMID: 27147572.
[81]. Liu B, Wang S, Long R, Chou KC. iRSpot-EL: identify recombination spots with an ensemble learning approach. Bioinformatics 33 (2017) 35-41.PMID: 27531102.
[82]. 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 Apr 28;9(04):67.
[83]. Liu B, Wu H, Zhang D, Wang X, Chou KC, Pse-Analysis: a python package for DNA/RNA and protein/peptide sequence analysis based on pseudo components and kernel methods. Oncotarget 8 (2017) 13338-13343.PMID: 28076851.
[84]. Liu B, Yang F, Chou KC. 2L-piRNA: A two-layer ensemble classifier for identifying piwi-interacting RNAs and their function. Molecular Therapy - Nucleic Acids 7 (2017) 267-277.PMID: 28624202.
[85]. Qiu WR, Jiang SY, Xu ZC, Xiao X, Chou KC. iRNAm5C-PseDNC: identifying RNA 5-methylcytosine sites by incorporating physical-chemical properties into pseudo dinucleotide composition. Oncotarget. 2017 Jun 20;8(25):41178-41188.PMID:

28476023.
[86]. Chen J, Guo M, Wang X, Liu B. A comprehensive review and comparison of different computational methods for protein remote homology detection. Brief Bioinform. 2018 Mar 1;19(2):231-244. PMID: 27881430.
[87]. Liu B. BioSeq-Analysis: a platform for DNA, RNA and protein sequence analysis based on machine learning approaches. Brief Bioinform. 2017 Dec 19.
[88]. Liu B, Li K, Huang DS, Chou KC. iEnhancer-EL: Identifying enhancers and their strength with ensemble learning approach. Bioinformatics. 2018 Nov 15;34(22):3835-3842.PMID: 29878118.
[89]. Liu B, Weng F, Huang DS, Chou KC. iRO-3wPseKNC: Identify DNA replication origins by three-window-based PseKNC. Bioinformatics. 2018 Sep 15;34(18):3086-3093..PMID: 29684124.
[90]. 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 Jan 1;34(1):33-40. PMID: 28968797.
[91]. 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 May 7;234(1):145-9.
[92]. Cai YD, Chou KC. Predicting enzyme subclass by functional domain composition and pseudo amino acid composition. J Proteome Res. 2005 May-Jun;4(3):967-71.PMID: 15952744.
[93]. 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. J. Theor. Biol. 2007 Oct 7;248(3):546-51.
[94]. 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 Jun;17(6):715-22.PMID: 19961429.
[95]. 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 Pept Lett. 2010 Nov;17(11):1441-9.PMID: 20666729.
[96]. Cai YD, Chou KC. Nearest neighbour algorithm for predicting protein subcellular location by combining functional domain composition and pseudon amino acid composition. Biochem Biophys Res Commun. 2003 May 30;305(2):407-11.PMID: 12745090.
[97]. 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. J Protein Chem. 2003 May;22(4):395-402.PMID: 13678304.
[98]. Chou KC, Cai YD. Prediction and classification of protein subcellular location: sequence-order effect and pseudo amino acid composition. J Cell Biochem. 2003 Dec 15;90(6):1250-60.PMID: 14635197.
[99]. Chou KC, Cai YD. Predicting subcellular localization of proteins by hybridizing functional domain composition and pseudo-amino acid composition. Journal of Cellular Biochemistry. 2004 Apr 15;91(6):1197-203.
[100]. Xiao X, Shao S, Ding Y, Huang Z, Chou KC. Using cellular automata images and pseudo amino acid composition to predict protein subcellular location. Amino acids. 2006 Feb 1;30(1):49-54.
[101]. J.Y. Shi, S.W. Zhang, Q. Pan, Y.-M. Cheng, J. Xie. Prediction of protein subcellular localization by support vector machines using multi-scale energy and pseudo amino acid composition. Amino Acids. 2007 Jul;33(1):69-74.PMID:17235454.
[102]. S.W. Zhang, Y.L. Zhang, H.F. Yang, C.H. Zhao, Q. Pan, 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 May;34(4):565-72. PMID:18074191.
[103]. J.Y. Shi, S.W. Zhang, Q. Pan, G.P. Zhou, Using Pseudo Amino Acid Composition to Predict Protein Subcellular Location: Approached with Amino Acid Composition Distribution. Amino Acids. 2008 Aug;35(2):321-7. PMID:18209947.
[104]. F.M. Li, Q.Z. Li, Predicting protein subcellular location using Chou's pseudo amino acid composition and improved hybrid approach. Protein Pept Lett. 2008;15(6):612-6. PMID:18680458.
[105]. Y.L. Chen, Q.Z. Li, Prediction of apoptosis protein subcellular location using improved hybrid approach and pseudo amino acid composition. J Theor Biol. 2007 Sep 21;248(2):377-81. PMID:17572445.
[106]. X. Jiang, R. Wei, T.L. Zhang, Q. Gu, Using the concept of Chou's pseudo amino acid composition to predict apoptosis proteins subcellular location: an approach by approximate entropy. Protein Pept Lett. 2008;15(4):392-6. PMID:18473953.
[107]. 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 Sep 1;57(3):321-30.
[108]. 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 and peptide letters. 2010 Oct 1;17(10):1263-9.
[109]. Kandaswamy KK, Pugalenthi G, Möller 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 Pept Lett. 2010 Dec;17(12):1473-9. PMID:20666727.
[110]. Lin H, Ding H, Guo FB, Zhang AY, Huang J. Predicting subcellular localization of mycobacterial proteins by using Chou's pseudo amino acid composition. Protein and peptide letters. 2008 Jul 1;15(7):739-44.
[111]. H.B. Shen, Predicting protein subnuclear location with optimized evidence-theoretic K-nearest classifier and pseudo amino acid composition. Biochem Biophys Res Commun. 2005 Nov 25;337(3):752-6. PMID:16213466.
[112]. F.M. Li, Q.Z. Li, Using pseudo amino acid composition to predict protein subnuclear location with improved hybrid approa h. Amino Acids. 2008 Jan;34(1):119-25. PMID:17514493.
[113]. X. Jiang, R. Wei, Y. Zhao, T. Zhang, 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 May;34(4):669-75. PMID:18256886.
[114]. P. Du, S. Cao, Y. Li, SubChlo: predicting protein subchloroplast locations with pseudo amino acid composition and the evidence-theoretic K-nearest neighbor (ET-KNN) algorithm. J Theor Biol. 2009 Nov 21;261(2):330-5. PMID:19679138.
[115]. P. Du, Y. Li, Prediction of protein submitochondria locations by hybridizing pseudo amino acid composition with various physicochemical features of segmented sequence. BMC Bioinformatics. 2006 Nov 30;7:518. PMID:17134515.
[116]. L. Nanni, A. Lumini, Genetic programming for creating Chou's pseudo amino acid based features for submitochondria localization. Amino Acids. 2008 May;34(4):653-60. PMID:18175047.
[117]. Y.H. Zeng, Y.Z. Guo, R.Q. Xiao, L. Yang, L.Z. Yu, M.L. Li, Using the augmented Chou's pseudo amino acid composition for predicting protein submitochondria locations based on auto covariance approach. J Theor Biol. 2009 Jul 21;259(2):366-72. PMID:19341746.
[118]. M. Wang, J. Yang, G.P. Liu, Z.J. Xu, Weighted-support vector machines for predicting membrane protein types based on pseudo amino acid composition. Protein Engineering, Protein Eng Des Sel. 2004 Jun;17(6):509-16. PMID:15314209.
[119]. Cai YD, Chou KC. Predicting membrane protein type by functional domain composition and pseudo-amino acid composition. J Theor Biol.. 2006 Jan 21;238(2):395-400.
[120]. 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 Aug 1;24(6):385-9.
[121]. S.Q. Wang, J. Yang, Using stacked generalization to predict mem-

brane protein types based on pseudo amino acid composition J Theor Biol. 2006 Oct 21;242(4):941-6. PMID:16806277.

[122]. K.C. Chou, H.B. Shen, MemType-2L: A Web server for predicting membrane proteins and their types by incorporating evolution information through Pse-PSSM. Biochem Biophys Res Commun. 2007 Aug 24;360(2):339-45. PMID:17586467.

[123]. H. Lin, The modified Mahalanobis discriminant for predicting outer membrane proteins by using Chou's pseudo amino acid composition. J Theor Biol. 2008 May 21;252(2):350-6. PMID:18355838.

[124]. Q.B. Gao, X.F. Ye, Z.C. Jin, J. He, Improving discrimination of outer membrane proteins by fusing different forms of pseudo amino acid composition. Anal Biochem. 2010 Mar 1;398(1):52-9. PMID:19874797.

[125]. Y. Diao, D. Ma, Z. Wen, J. Yin, J. Xiang, M. Li, Using pseudo amino acid composition to predict transmembrane regions in protein: cellular automata and Lempel-Ziv complexity. Amino Acids. 2008 Jan;34(1):111-7. PMID:17520325.

[126]. G.P. Zhou, Y.D. Cai, Predicting protease types by hybridizing gene ontology and pseudo amino acid composition. Proteins. 2006 May 15;63(3):681-4. PMID:16456852.

[127]. K.C. Chou, H.B. Shen, ProtIdent: A web server for identifying proteases and their types by fusing functional domain and sequential evolution information. Biochem. Biochem Biophys Res Commun. 2008 Nov 14;376(2):321-5. PMID:18774775.

[128]. 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. J Theor Biol. 2013 Sep 7;332:211-7. PMID:23524162.

[129]. J.D. Qiu, J.H. Huang, R.P. Liang, X.Q. Lu, 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 Jul 1;390(1):68-73. PMID:19364489.

[130]. Q. Gu, Y.S. Ding, T.L. Zhang, Prediction of G-Protein-Coupled Receptor Classes in Low Homology Using Chou's Pseudo Amino Acid Composition with Approximate Entropy and Hydrophobicity Patterns. Protein Pept Lett. 2010 May;17(5):559-67. PMID:19594431.

[131]. Gu Q, Ding YS, Zhang TL. Prediction of G-protein-coupled receptor classes with pseudo amino acid composition. In2008 2nd International Conference on Bioinformatics and Biomedical Engineering 2008 May 16 (pp. 876-879). IEEE.

[132]. W.Z. Lin, X. Xiao, GPCR-GIA: a web-server for identifying G-protein coupled receptors and their families with grey incidence analysis. Protein Eng Des Sel. 2009 Nov;22(11):699-705. PMID:19776029.

[133]. X. Xiao, P. Wang, GPCR-CA: A cellular automaton image approach for predicting G-protein-coupled receptor functional classes. J Comput Chem. 2009 Jul 15;30(9):1414-23. PMID:19037861.

[134]. Q.B. Gao, Z.C. Jin, X.F. Ye, C. Wu, J. He, Prediction of nuclear receptors with optimal pseudo amino acid composition. Anal Biochem. 2009 Apr 1;387(1):54-9. PMID:19454254.

[135]. H. Mohabatkar, Prediction of cyclin proteins using Chou's pseudo amino acid composition. Protein Pept Lett. 2010 Oct;17(10):1207-14. PMID:20450487.

[136]. L. Yu, Y. Guo, Y. Li, G. Li, M. Li, J. Luo, W. Xiong, W. Qin, SecretP: Identifying bacterial secreted proteins by fusing new features into Chou's pseudo amino acid composition. J Theor Biol. 2010 Nov 7;267(1):1-6. PMID:20691704.

[137]. M. Esmaeili, H. Mohabatkar, S. Mohsenzadeh, Using the concept of Chou's pseudo amino acid composition for risk type prediction of human papillomaviruses. J Theor Biol. 2010 Mar 21;263(2):203-9. PMID:19961864.

[138]. H. Ding, L. Luo, H. Lin, Prediction of cell wall lytic enzymes using Chou's amphiphilic pseudo amino acid composition. Protein Pept Lett. 2009;16(4):351-5. PMID:19356130.

[139]. G.Y. Zhang, H.C. Li, J.Q. Gao, B.S. Fang, Predicting lipase types by improved Chou's pseudo amino acid composition. Protein Pept Lett. 2008;15(10):1132-7. PMID:19075826.

[140]. S. Mondal, R. Bhavna, R. Mohan Babu, S. Ramakumar, Pseudo amino acid composition and multi-class support vector machines approach for conotoxin superfamily classification. J Theor Biol. 2006 Nov 21;243(2):252-60. PMID:16890961.

[141]. H. Lin, Q.Z. Li, Predicting conotoxin superfamily and family by using pseudo amino acid composition and modified Mahalanobis discriminant. Biochem Biophys Res Commun. 2007 Mar 9;354(2):548-51.PMID: 17239817.

[142]. 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 Jul 21;253(2):310-5.PMID: 18471832.

[143]. 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 34 (2008) 103-109. PMID: 17234492.

[144]. Xiao X, Shao SH, Huang ZD, Chou KC. Using pseudo amino acid composition to predict protein structural classes: approached with complexity measure factor. J Comput Chem. 2006 Mar;27(4):478-82.

[145]. Chen C, Tian YX, Zou XY, Cai PX, Mo JY. Using pseudo-amino acid composition and support vector machine to predict protein structural class. Journal of theoretical biology. 2006 Dec 7;243(3):444-8.

[146]. 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. Anal Biochem. 2006 Oct 1;357(1):116-21.PMID: 16920060.

[147]. Zhang TL, Ding YS. Using pseudo amino acid composition and binary-tree support vector machines to predict protein structural classes. Amino Acids. 2007 Nov;33(4):623-9.PMID: 17308864.

[148]. Lin H, Li QZ. Using Pseudo Amino Acid Composition to Predict Protein Structural Class: Approached by Incorporating 400 Dipeptide Components. J Comput Chem. 2007 Jul 15;28(9):1463-1466.PMID: 17330882.

[149]. Zhang TL, Ding YS, Chou KC. Prediction protein structural classes with pseudo amino acid composition: approximate entropy and hydrophobicity pattern. J Theor Biol. 2008 Jan 7;250(1):186-93. PMID: 17959199.

[150]. Ding YS, Zhang TL, Chou KC. Prediction of protein structure classes with pseudo amino acid composition and fuzzy support vector machine network. Protein Pept Lett. 2007;14(8):811-5. PMID: 17979824.

[151]. Xiao X, Lin WZ, Chou KC. Using grey dynamic modeling and pseudo amino acid composition to predict protein structural classes. J Comput Chem. 2008 Sep;29(12):2018-24.PMID: 18381630.

[152]. 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 Jul;37(2):415-25.PMID: 18726140.

[153]. 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 29 (2010) 62-7.PMID: 20049515.

[154]. Zou D, He Z, He J, Xia Y. Supersecondary structure prediction using Chou's pseudo amino acid composition. J Comput Chem. 2011 Jan 30;32(2):271-8.PMID: 20652881.

[155]. 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 Pept Lett. 2009;16(1):27-31.PMID: 19149669.

[156]. Chou KC, Cai YD. Predicting protein quaternary structure by pseudo amino acid composition. Proteins: Structure, Function, and Bioinformatics. 2003 Nov 1;53(2):282-9.

[157]. 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 Jun;30(4):461-8.PMID: 16773245.

[158]. 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 Oct;35(3):591-8.PMID: 18427713.

[159]. Shen HB, Chou KC. QuatIdent: A web server for identifying protein quaternary structural attribute by fusing functional domain and sequential evolution information. J Proteome Res. 2009 Mar;8(3):1577-84.PMID: 19226167.

[160]. Xiao X, Wang P, Chou KC. Quat-2L: a web-server for predicting protein quaternary structural attributes. Molecular Diversity 15 (2011) 149-155.PMID: 20148364.

[161]. Xiao X, Wang P, Chou KC. Predicting protein quaternary structural attribute by hybridizing functional domain composition and pseudo amino acid composition. Mol Divers. 2011 Feb;15(1):149-55.PMID:20148364.

[162]. Shen HB, Chou KC. Ensemble classifier for protein fold pattern recognition. Bioinform. 2006 May 3;22(14):1717-22.

[163]. Shen HB, Chou KC. Predicting protein fold pattern with functional domain and sequential evolution information. J Theor Biol. 2009 Feb 7;256(3):441-6. PMID:18996396.

[164]. 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. J Theor Biol. 2009 Mar 7;257(1):17-26. PMID:19056401.

[165]. Chou JJ, Matsuo H, Duan H, Wagner G. Solution structure of the RAIDD CARD and model for CARD/CARD interaction in caspase-2 and caspase-9 recruitment. Cell 94 (1998) 171-180. PMID:9695946.

[166]. Oxenoid K, Dong Y, Cao C, Cui T, Sancak Y, Markhard AL. et al. Architecture of the mitochondrial calcium uniporter. Nature. 2016 May;533(7602):269.

[167]. Dev J, Park D, Fu Q, Chen J, Ha HJ, Ghantous F. Structural basis for membrane anchoring of HIV-1 envelope spike. Science. 2016 Jul 8;353(6295):172-5.

[168]. Schnell JR, Chou JJ. Structure and mechanism of the M2 proton channel of influenza A virus. Nature. 2008 Jan;451(7178):591.

[169]. Berardi MJ, Shih WM, Harrison SC, Chou JJ. Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. Nature. 2011 Jul 24;476(7358):109-13. PMID:21785437.

[170]. Chou JJ, Li S, Klee CB, Bax A. Solution structure of Ca 2+–calmodulin reveals flexible hand-like properties of its domains. Nat. Struct. Mol. Biol. 2001 Nov;8(11):990.

[171]. OuYang B, Xie S, Berardi MJ, Zhao X, Dev J, Yu W. at al. Unusual architecture of the p7 channel from hepatitis C virus. Nature. 2013 Jun;498(7455):521.

[172]. Wang J, Pielak RM, McClintock MA, Chou JJ. Chou, Solution structure and functional analysis of the influenza B proton channel. Nat Struct Mol Biol. 2009 Dec;16(12):1267-71. Nat Struct Mol Biol. 2009 Dec;16(12):1267-71.

[173]. Fu Q, Fu TM, Cruz AC, Sengupta P, Thomas SK, Siegel RM. et al. Structural Basis and Functional Role of Intramembrane Trimerization of the Fas/CD95 Death Receptor. Mol Cell. 2016 Feb 18;61(4):602-613. PMID:26853147.

[174]. Chou JJ, Li H, Salvesen GS, Yuan J, Wagner G. Solution structure of BID, an intracellular amplifier of apoptotic signaling. Cell. 1999 Mar 5;96(5):615-24.

[175]. Oxenoid K, Chou JJ. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. Proc Natl Acad Sci U S A. 2005 Aug 2;102(31):10870-5. PMID:16043693.

[176]. Call ME, Schnell JR, Xu C, Lutz RA, Chou JJ, Wucherpfennig KW. The structure of the zetazeta transmembrane dimer reveals features essential for its assembly with the T cell receptor. Cell. 2006 Oct 20;127(2):355-68. PMID:17055436.

[177]. Call ME, Wucherpfennig KW, Chou JJ. The structural basis for intramembrane assembly of an activating immunoreceptor complex. Nat Immunol. 2010 Nov;11(11):1023-9. PMID:20890284.

[178]. Gagnon E, Xu C, Yang W, Chu HH, Call ME, Chou JJ. et al. Response multilayered control of T cell receptor phosphorylation. Cell. 2010 Sep 3;142(5):669-71. PMID:20813252.

[179]. Brüschweiler S, Yang Q, Run C, Chou JJ. Substrate-modulated ADP/ATP-transporter dynamics revealed by NMR relaxation dispersion. Nat Struct Mol Biol. 2015 Aug;22(8):636-41. PMID:26167881.

[180]. Cao C, Wang S, Cui T, Su XC, Chou JJ. Ion and inhibitor binding of the double-ring ion selectivity filter of the mitochondrial calcium uniporter. Proc Natl Acad Sci U S A. 2017 Apr 4;114(14):E2846-E2851. PMID:28325874.

[181]. 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 Dec 27;139(51):18432-18435. PMID:29193965.

[182]. Pan L, Fu TM, Zhao W, Zhao L, Chen W, Qiu C,et al. Higher-Order Clustering of the Transmembrane Anchor of DR5 Drives Signaling. Cell. 2019 Mar 7;176(6):1477-1489.e14. PMID:30827683.

[183]. Carlacci L, Chou KC, Maggiora GM. A heuristic approach to predicting the tertiary structure of bovine somatotropin. Biochemistry. 1991 May 7;30(18):4389-98. PMID:2021631.

[184]. Chou KC, Jones D, Heinrikson RL. Prediction of the tertiary structure and substrate binding site of caspase-8. FEBS Lett. 1997 Dec 8;419(1):49-54. PMID:9426218.

[185]. Chou KC, Tomasselli AG, Heinrikson RL. Prediction of the Tertiary Structure of a Caspase-9/Inhibitor Complex. FEBS Lett. 2000 Mar 31;470(3):249-56. PMID:10745077.

[186]. Chou KC, Carlacci L, Maggiora GM, Parodi LA, Schulz MW. An energy-based approach to packing the 7-helix bundle of bacteriorhodopsin. Protein Sci. 1992 Jun;1(6):810-27. PMID:1304922.

[187]. Chou KC. Knowledge-based model building of tetiary structures for lectin domains of the selectin family. J Protein Chem. 1996 Feb;15(2):161-8. PMID:8924200.

[188]. Chou KC, Watenpaugh KD, Heinrikson RL. A Model of the complex between cyclin-dependent kinase 5 (Cdk5) and the activation domain of neuronal Cdk5 activator. Biochem Biophys Res Commun. 1999 Jun 7;259(2):420-8.

[189]. Chou KC, Howe WJ. Prediction of the tertiary structure of the beta-secretase zymogen. Biochem Biophys Res Commun. 2002 Apr 5;292(3):702-8. PMID:11922623.

[190]. Chou KC. Coupling interaction between thromboxane A2 receptor and alpha-13 subunit of guanine nucleotide-binding protein. J Proteome Res. 2005 Sep-Oct;4(5):1681-6. PMID:16212421.

[191]. Chou KC. Modelling extracellular domains of GABA-A receptors: subtypes 1, 2, 3, and 5. Biochem Biophys Res Commun. 2004 Apr 9;316(3):636-42. PMID:15033447.

[192]. 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 Commun. 2003 Aug 15;308(1):148-51. PMID:12890493.

[193]. Chou KC. Insights from modelling the 3D structure of the extracellular domain of alpha7 nicotinic acetylcholine receptor. Biochem Biophys Res Commun. 2004 Jun 25;319(2):433-8. PMID:15178425.

[194]. Chou KC. Insights from modelling the tertiary structure of BACE2. J Proteome Res. 2004 Sep-Oct;3(5):1069-72. PMID:15473697.

[195]. Chou KC. Molecular therapeutic target for type-2 diabetes. J Proteome Res. 2004 Dec 13;3(6):1284-8.

[196]. Chou KC. Insights from modeling three-dimensional structures of the human potassium and sodium channels. J Proteome Res. 2004 Aug 9;3(4):856-61.

[197]. Chou KC. Insights from modeling the 3D structure of DNA– CB-F3b complex. J Proteome Res. 2005 Oct 10;4(5):1657-60.

[198]. Chou KC. Modeling the tertiary structure of human cathep-

sin-E. Biochem. Biochem Biophys Res Commun. 2005 May 27;331(1):56-60. PMID:15845357.

[199]. Wang SQ, Du QS, Chou KC. Study of drug resistance of chicken influenza A virus (H5N1) from homology-modeled 3D structures of neuraminidases. Biochem Biophys Res Commun. 2007 Mar 16;354(3):634-40. PMID:17266937.

[200]. Wang SQ, Du QS, Huang RB, Zhang DW, Chou KC. Insights from investigating the interaction of oseltamivir (Tamiflu) with neuraminidase of the 2009 H1N1 swine flu virus. Biochem Biophys Res Commun. 2009 Aug 28;386(3):432-6. PMID:19523442.

[201]. 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(11):e28111. PMID:22140516.

[202]. 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(6):e38546. PMID:22685582.

[203]. Chou KC. Some remarks on protein attribute prediction and pseudo amino acid composition. J. Theor. Biol. 2011 Mar 21;273(1):236-47.

[204]. Chou KC, Shen HB. Hum-PLoc: a novel ensemble classifier for predicting human protein subcellular localization. Biochem Biophys Res Commun. 2006 Aug 18;347(1):150-7.

[205]. Shen HB, Chou KC. Gpos-PLoc: an ensemble classifier for predicting subcellular localization of Gram-positive bacterial proteins. Protein Eng Des Sel. 2007 Jan;20(1):39-46. PMID:17244638.

[206]. Shen HB, Chou KC. Nuc-PLoc: A new web-server for predicting protein subnuclear localization by fusing PseAA composition and PsePSSM. Protein Eng Des Sel. 2007 Nov;20(11):561-7. PMID:17993650.

[207]. Chou KC, Shen HB. Cell-PLoc: A package of Web servers for predicting subcellular localization of proteins in various organisms. Nat Protoc. 2008;3(2):153-62. PMID:18274516.

[208]. Chou KC, Shen HB. Cell-PLoc 2.0: An improved package of web-servers for predicting subcellular localization of proteins in various organisms. development. 2010;109:1091.

[209]. 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 Mar 30;6(3):e18258. PMID:21483473.

[210]. 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. Mol Biosyst. 2011 Dec;7(12):3287-97. PMID:21984117.

[211]. 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. J Theor Biol. 2011 Sep 7;284(1):42-51. PMID:21684290.

[212]. 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. Mol Biosyst. 2012 Feb;8(2):629-41. PMID:22134333.

[213]. 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 Pept Lett. 2012 Jan;19(1):4-14. PMID:21919865.

[214]. Lin WZ, Fang JA, Xiao X, Chou KC. iLoc-Animal: A multi-label learning classifier for predicting subcellular localization of animal proteins. Mol Biosyst. 2013 Apr 5;9(4):634-44. PMID:23370050.

[215]. Cheng X, Xiao X, Chou KC. pLoc-mPlant: predict subcellular localization of multi-location plant proteins by incorporating the optimal GO information into general PseAAC. Mol Biosyst. 2017;13(9):1722-7.

[216]. 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 Sep 10;628:315-21.

[217]. 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 Nov 15;33(22):3524-3531. PMID:29036535.

[218]. Xiao X, Cheng X, Su S, Mao Q, Chou KC. pLoc-mGpos: incorporate key gene ontology information into general PseAAC for predicting subcellular localization of Gram-positive bacterial proteins. Natural Science. 2017 Sep 11;9(09):330.

[219]. 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 110 (2018) 50-58. PMID:28818512.

[220]. 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 110 (2018) 231-239. PMID:28989035.

[221]. 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 May 1;34(9):1448-1456. PMID:29106451.

[222]. 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. J Theor Biol. 2018 Dec 7;458:92-102. PMID:30201434.

[223]. 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(34):4013-4022. PMID:30451108.

[224]. Chou KC, Cheng X, Xiao X. pLoc_bal-mHum: predict subcellular localization of human proteins by PseAAC and quasi-balancing training dataset. Genomics. 2019 Dec;111(6):1274-1282. PMID:30179658.

[225]. 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(5):472-485. PMID:30569871.

[226]. 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 15 (2018) 496-509. PMID:30556503.

[227]. 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 Feb 1;35(3):398-406. PMID:30010789.

[228]. 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 Jul;111(4):886-892. PMID:29842950.

[229]. 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 Eng Des Sel. 2013 Nov;26(11):735-42. PMID:24048266.

[230]. Xu Y, Shao XJ, Wu LY, Deng NY, Chou KC. iSNO-AAPair: incorporating amino acid pairwise coupling into PseAAC for predicting cysteine S-nitrosylation sites in proteins. PeerJ. 2013 Oct 3;1:e171. PMID:24109555.

[231]. 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 Jun 10;15(6):10410-23. PMID:24918295.

[232]. 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. 2014;2014:947416. PMID:24977164.

[233]. 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. Int J Mol Sci. 2014 May

5;15(5):7594-610. PMID:24857907.
[234]. 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 Aug 14;9(8):e105018. PMID:25121969.
[235]. 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 J Mol Sci. 2014 Jun 25;15(7):11204-19. PMID:24968264.
[236]. Chen W, Feng P, Ding H, Lin H, Chou KC. iRNA-Methyl: Identifying N6-methyladenosine sites using pseudo nucleotide composition. Anal Biochem. 2015 Dec 1;490:26-33. PMID:26314792.
[237]. Qiu WR, Xiao X, Lin WZ, Chou KC. iUbiq-Lys: prediction of lysine ubiquitination sites in proteins by extracting sequence evolution information via a gray system model. J Biomol Struct Dyn. 2015 Aug 3;33(8):1731-42.
[238]. 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 Mar 15;497:48-56.
[239]. 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. J Theor Biol. 2016 Apr 7;394:223-230. PMID:26807806.
[240]. Jia J, Liu Z, Xiao X, Liu B, Chou KC. iCar-PseCp: identify carbonylation sites in proteins by Monte Carlo sampling and incorporating sequence coupled effects into general PseAAC. Oncotarget. 2016 Jun 7;7(23):34558.
[241]. 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 Oct 15;32(20):3133-3141. PMID:27354696.
[242]. 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 May 21;397:145-50.
[243]. Liu Z, Xiao X, Yu DJ, Jia J, Qiu WR, Chou KC. pRNAm-PC: Predicting N6-methyladenosine sites in RNA sequences via physical–chemical properties. Anal. Biochem. 2016 Mar 15;497:60-7.
[244]. 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 Jul 12;7(28):44310-44321. PMID:27322424.
[245]. Qiu WR, Sun BQ, Xiao X, Xu ZC, Chou KC. iPTM-mLys: identifying multiple lysine PTM sites and their different types. Bioinformatics. 2016 Oct 15;32(20):3116-3123. PMID:27334473.
[246]. Qiu WR, Xiao X, Xu ZC, Chou KC. iPhos-PseEn: identifying phosphorylation sites in proteins by fusing different pseudo components into an ensemble classifier. Oncotarget. 2016 Aug 9;7(32):51270-51283. PMID:27323404.
[247]. Xu Y, Chou KC. Recent progress in predicting posttranslational modification sites in proteins. Curr Top Med Chem. 2016;16(6):591-603. PMID:26286211.
[248]. 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. Mol Ther Nucleic Acids. 2017 Jun 16;7:155-163. PMID:28624191.
[249]. 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 Oct;77:200-204. PMID:28886434.
[250]. Liu LM, Xu Y, Chou KC. iPGK-PseAAC: identify lysine phosphoglycerylation sites in proteins by incorporating four different tiers of amino acid pairwise coupling information into the general PseAAC. Med Chem. 2017;13(6):552-559. PMID:28521678.
[251]. 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. Med Chem. 2017;13(8):734-743.PMID: 28641529.
[252]. 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. Mol Inform. 2017 May;36(5-6).PMID: 28488814.
[253]. Y. Xu, C. Li, iPreny-PseAAC: identify C-terminal cysteine prenylation sites in proteins by incorporating two tiers of sequence couplings into PseAAC. Med Chem 13 (2017) 544-551.PMID: 28488814.
[254]. 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 Oct 14;455:205-211.PMID: 300317943.
[255]. Chen W, Ding H, Zhou X, Lin H, Chou KC, iRNA(m6A)-PseDNC: Identifying N6-methyladenosine sites using pseudo dinucleotide composition. Analytical Biochemistry 561-562 (2018) 59-65. PMID: 30201554.
[256]. Chen W, Feng P, Yang H, Ding H, Lin H, Chou KC. iRNA-3typeA: identifying 3-types of modification at RNA's adenosine sites. Mol Ther Nucleic Acids. 2018 Jun 1;11:468-474.PMID: 29858081.
[257]. 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 Jan;111(1):96-102.PMID: 29360500.
[258]. 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 Jul 20;664:78-83.PMID: 29694908.
[259]. Khan YD, Rasool N, Hussain W, Khan SA, Chou KC. iPhosT-PseAAC: Identify phosphothreonine sites by incorporating sequence statistical moments into PseAAC. Anal Biochem. 2018 Jun 1;550:109-116.PMID: 29704476.
[260]. 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 Sep;110(5):239-246.PMID: 29107015.
[261]. Sabooh MF, Iqbal N, Khan M, Khan M, Maqbool HF. Maqbool. Identifying 5-methylcytosine sites in RNA sequence using composite encoding feature into Chou's PseKNC. J Theor Biol. 2018 Sep 7;452:1-9.
[262]. Khan YD, Rasool N, Hussain W, Khan SA, Chou KC. iPhosY-PseAAC: identify phosphotyrosine sites by incorporating sequence statistical moments into PseAAC. Mol Biol Rep. 2018 Dec;45(6):2501-2509.PMID: 30311130.
[263]. Hussain W, Khan YD, 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 Mar 1;568:14-23. PMID: 30593778.
[264]. Li F, Zhang Y, Purcell AW, Webb GI, Chou KC, Lithgow T,et al. Positive-unlabelled learning of glycosylation sites in the human proteome. BMC Bioinformatics. 2019 Mar 6;20(1):112. PMID: 30841845.
[265]. 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. 2019 Jan 14;461:51-58. PMID: 30365947.
[266]. Kumar VS, Vellaichamy A. Sequence and structure-based characterization of human and yeast ubiquitination sites by using Chou's sample formulation. Proteins. 2019 Aug;87(8):646-657.PMID: 30958587.

[267]. Zhang Y, Xie R, Wang J, Leier A, Marquez-Lago TT, Akutsu T. Computational analysis and prediction of lysine malonylation sites by exploiting informative features in an integrative machine-learning framework. Brief Bioinform. 2018 Aug 24. PMID: 30351377.
[268]. Chen Z, Liu X, Li F, Li C, Marquez-Lago T, Leier A. Large-scale comparative assessment of computational predictors for lysine post-translational modification sites. Brief Bioinform. 2018 Oct 4.PMID: 30285084.
[269]. 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 May 27.
[270]. 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 May 7;468:1-11.PMID: 30768975.
[271]. Chou KC. Progresses in predicting post-translational modification. International Journal of Peptide Research and Therapeutics. 2019:1-6.
[272]. Ding H, Deng EZ, Yuan LF, Liu L, Lin H, Chen W,et al. iC-TX-Type: A sequence-based predictor for identifying the types of conotoxins in targeting ion channels. Biomed Res Int. 2014;2014:286419.PMID: 24991545.
[273]. 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 Mar 30;10(3):e0121501.PMID: 25821974.
[274]. 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 Biomol Struct Dyn. 2015;33(10):2221-33.PMID: 25513722.
[275]. 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 Jan 17;8(3):4208-4217.PMID: 27926534.
[276]. 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(34):4034-4043.PMID: 30479209.
[277]. 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(34):4023-4033.PMID: 30421671.
[278]. Chou KC. Advance in predicting subcellular localization of multi-label proteins and its implication for developing multi-target drugs. Curr Med Chem. 2019 May 6. .PMID: 31060481.
[279]. Ehsan A, Mahmood MK, Khan YD, Barukab OM, Khan SA. iHyd-PseAAC (EPSV): Identifying 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(2):124-33.
[280]. Khan YD, Batool A, Rasool N, Khan SA, Chou KC. Prediction of Nitrosocysteine Sites Using Position and Composition Variant Features. Letters in Organic Chemistry. 2019 Apr 1;16(4):283-93.
[281]. 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 Feb 21;463:47-55.PMID: 30550863.
[282]. 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 Apr 1;16(4):303-10.
[283]. Khan S, Khan M, Iqbal N, Hussain T, Khan SA, 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. 2019:1-5.
[284]. 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. J Theor Biol. 2018 Apr 14;443:125-137. PMID: 29408627.
[285]. Chen Z, Zhao P, 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 Jul 15;34(14):2499-2502.PMID: 29528364.
[286]. 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 Dec 15;34(24):4223-4231.PMID: 29947803.
[287]. 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 Bioinform. 2018 Aug 29.PMID: 30184176.
[288]. Song J, Li F, Leier A, Marquez-Lago TT, Akutsu T, Haffari G, et al. ROSPERous: high-throughput prediction of substrate cleavage sites for 90 proteases with improved accuracy. Bioinformatics. 2018 Feb 15;34(4):684-687.PMID: 29069280.
[289]. 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 Bioinform. 2019 Mar 25;20(2):638-658.PMID: 29897410.
[290]. 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. 201.
[291]. Wang J, Yang B, Leier A, Marquez-Lago TT, Hayashida M, Rocker A, et al. Bastion6: a bioinformatics approach for accurate prediction of type VI secreted effectors. Bioinformatics. 2018 Aug 1;34(15):2546-2555.PMID: 29547915.
[292]. 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. 2019 Dec;111(6):1760-1770.
[293]. Chou KC. Review: Prediction of human immunodeficiency virus protease cleavage sites in proteins. Anal Biochem. 1996 Jan 1;233(1):1-14.PMID: 8789141.
[294]. Du QS, Wang SQ, Zhu Y, Wei DQ, Guo H, Sirois S, et al. Polyprotein Cleavage Mechanism of SARS CoV Mpro and Chemical Modification of Octapeptide. Peptides 25 (2004) 1857-1864. PMID: 15501516.
[295]. Du Q, Wang S, Wei D, Sirois S, Chou KC. Molecular modeling and chemical modification for finding peptide inhibitor against severe acute respiratory syndrome coronavirus main proteinase. Analytical Biochemistry. 2005 Feb 15;337(2):262-70.
[296]. 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 Apr;27(4):622-5.PMID: 16242214.
[297]. Du QS, Sun H, Chou KC. Inhibitor design for SARS coronavirus main protease based on "distorted key theory". Med Chem. 2007 Jan;3(1):1-6.PMID: 17266617.
[298]. Chou KC, Forsén S. Diffusion-controlled effects in reversible enzymatic fast reaction system: Critical spherical shell and proximity rate constants. Biophys Chem. 1980 Dec;12(3-4):255-63.PMID: 7225518.
[299]. Chou KC, Forsén S. Graphical rules for enzyme-catalyzed rate laws. Biochem J. 1980 Jun 1;187(3):829-35.PMID: 7188428.
[300]. Stanton RE, McIver Jr JW. Group theoretical selection rules for the transition states of chemical reactions. J. Am. Chem. Soc. 1975 Jun;97(13):3632-46.
[301]. Chou KC, Li TT, Forsén S. The critical spherical shell in enzymatic fast reaction systems. Biophys. Chem. 1980 Dec 1;12(3-4):265-9.

[302]. Li TT, Chou KC. The flow of substrate molecules in fast enzyme-catalyzed reaction systems. Chemica Scripta. 1980 Jan 1;16(5):192-6.
[303]. Carter RE, Forsen S. A new graphical method for deriving rate equations for complicated mechanisms. Chemica Scripta 18 (1981) 82-86.
[304]. Chou KC. Biological functions of low-frequency vibrations (phonons). III. Helical structures and microenvironment. Biophys. J.1984 May 1;45(5):881-9.
[305]. Chou KC, Forsén S. Graphical rules of steady-state reaction systems. Canadian Biochem J. 1980 Jun 1;187(3):829-35.PMID: 7188428.
[306]. Chou KC. Low-frequency vibrations of helical structures in protein molecules. Biochem. J. 1983 Mar 1;209(3):573-80.
[307]. Chou KC. Identification of low-frequency modes in protein molecules. Biochem. J. 1983 Dec 1;215(3):465-9.
[308]. 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 Aug 15;222(1):169-76.
[309]. Chou KC. Biological functions of low-frequency vibrations (phonons). III. Helical structures and microenvironment. Biophysical journal. 1984 May 1;45(5):881-9.
[310]. Chou KC. The biological functions of low-frequency vibrations (phonons): 4. Resonance effects and allosteric transition. Biophys. Chem. 1984 Aug 1;20(1-2):61-71.
[311]. Chou KC. Low-frequency vibrations of DNA molecules. Biochemical Journal. 1984 Jul 1;221(1):27-31.
[312]. Chen L, Cai YD, Shi XH, Huang T. Analysis of metabolic pathway using hybrid properties. Protein Pept.Lett. 2012 Jan 1;19(1):99-107.
[313]. Chou KC. Low-frequency motions in protein molecules. Beta-sheet and beta-barrel. Biophys J. 1985 Aug 1;48(2):289-97.
[314]. Chou KC. Prediction of a low-frequency mode in bovine pancreatic trypsin inhibitor molecule. International Journal of Biological Macromolecules. 1985 Apr 1;7(2):77-80.
[315]. Chou KC, Kiang YS. The biological functions of low-frequency phonons: 5. A phenomenological theory. Biophys Chem. 1985 Aug;22(3):219-35.PMID: 4052576.
[316]. Chou C. Origin of low-frequency motion in biological macromolecules: A view of recent progress of quasi-continuity model. Biophys Chem. 1986 Dec 15;25(2):105-16.PMID: 3101760.
[317]. Chou KC. The biological functions of low-frequency phonons: 6. A possible dynamic mechanism of allosteric transition in antibody molecules. Biopolymers. 20;20(4):447-57.
[318]. Chou KC. Review: Low-frequency collective motion in biomacromolecules and its biological functions. Biophys Chem. 1988 May;30(1):3-48.PMID: 3046672.
[319]. Chou KC, Maggiora GM. Biological functions of low-frequency vibrations (phonons). 7. The impetus for DNA to accommodate intercalators. British polymer journal. 1988;20(2):143-8.
[320]. Chou KC. Low-frequency resonance and cooperativity of hemoglobin. Trends in Biochemical Sciences. 1989 Jun 1;14(6):212.
[321]. Chou KC, Maggiora GM, Mao B. Quasi-continuum models of twist-like and accordion-like low-frequency motions in DNA. Biophys J. 1989 Aug;56(2):295-305.
[322]. Chou KC. Graphic rules in steady and non-steady state enzyme kinetics. J. Biol. Chem.1989 Jul 15;264(20):12074-9.
[323]. Chou KC. Applications of graph theory to enzyme kinetics and protein folding kinetics: steady and non-steady-state systems. Biophys. Chem. 1990 Jan 1;35(1):1-24.
[324]. Althaus IW, Chou JJ, Gonzales AJ, Deibel MR, Chou KC, Kezdy FJ, et al. Steady-state kinetic studies with the non-nucleoside HIV-1 reverse transcriptase inhibitor U-87201E. Int. J. Biol. Chem. 1993 Mar 25;268(9):6119-24.
[325]. Althaus IW, Gonzales AJ, Chou JJ, Romero DL, Deibel MR, Chou KC, et al. The quinoline U-78036 is a potent inhibitor of HIV-1 reverse transcriptase. Int. J. Biol. Chem. 1993 Jul 15;268(20):14875-80.
[326]. Althaus IW, Gonzales AJ, Chou JJ, Romero DL, Deibel MR, Chou KC, et al. The quinoline U-78036 is a potent inhibitor of HIV-1 reverse transcriptase. J. Biol. Chem. 1993 Jul 15;268(20):14875-80.
[327]. Althaus IW, Chou JJ, Gonzales AJ, LeMay RJ, Deibel MR, Chou KC, et al. Steady-state kinetic studies with the polysulfonate U-9843, an HIV reverse transcriptase inhibitor. Experientia. 1994 Feb 1;50(1):23-8.
[328]. Althaus IW, Chou JJ, Gonzales AJ, Deibel MR, Kuo-Chen C, Kezdy FJ, et al. Kinetic studies with the non-nucleoside human immunodeficiency virus type-1 reverse transcriptase inhibitor U-90152E. Biochem. Pharmacol. 1994 Jun 1;47(11):2017-28.
[329]. Chou KC, Kézdy FJ, Reusser F. Kinetics of processive nucleic acid polymerases and nucleases. Anal. Biochem. 1994 Sep 1;221(2):217-30.
[330]. Chou KC, Zhang CT, Maggiora GM. Solitary wave dynamics as a mechanism for explaining the internal motion during microtubule growth. Biopolymers. 1994 Jan;34(1):143-53.
[331]. Althaus IW, Chou KC, Lemay RJ, Franks KM, Deibel MR, Kezdy FJ, et al. The benzylthio-pyrimidine U-31,355, a potent inhibitor of HIV-1 reverse transcriptase. Biochem. Pharmacol. 1996 Mar 22;51(6):743-50.
[332]. Liu H, Wang M, Chou KC. Low-frequency Fourier spectrum for predicting membrane protein types. Biochem. Biophys. Res. Commun. 2005 Oct 28;336(3):737-9.
[333]. Gordon GA. Designed electromagnetic pulsed therapy: clinical applications. J Cell Physiol. 2007 Sep;212(3):579-82.
[334]. 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. Can. J. Chem. 2008 Apr 1;86(4):342-57.
[335]. Chou KC, Shen HB. FoldRate: A web-server for predicting protein folding rates from primary sequence. Open Bioinforma J. 2009;3(2):31-50.
[336]. Shen HB, Song JN, Chou KC. Prediction of protein folding rates from primary sequence by fusing multiple sequential features. J. Biomed. Eng. Sci. 2009 Jun 1;2:136-43.
[337]. Wang JF, Chou KC. Insight into the molecular switch mechanism of human Rab5a from molecular dynamics simulations. Biochem. Biophys. Res. Commun. 2009 Dec 18;390(3):608-12.
[338]. Gordon GA. Extrinsic electromagnetic fields, low frequency (phonon) vibrations, and control of cell function: a non-linear resonance system. J. Biomed. Eng. Sci. 2008 Oct 15;1(03):152.
[339]. Madkan A, Blank M, Elson E, Chou KC, Geddis MS, Goodman R. Steps to the clinic with ELF EMF. Natural Science. 2009 Dec 15;1(03):157.
[340]. Chou KC. Graphic rule for drug metabolism systems. Curr. Drug Metab. 2010 May 1;11(4):369-78.
[341]. Chou KC, Lin WZ, Xiao X. Wenxiang: a web-server for drawing wenxiang diagrams. Natural Science. 2011 Oct 13;3(10):862.
[342]. Lian P, Wei DQ, Wang JF, Chou KC. An allosteric mechanism inferred from molecular dynamics simulations on phospholamban pentamer in lipid membranes. PLoS One. 2011 Apr 15;6(4):e18587.PMID: 21525996.
[343]. 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). Med Chem. 2011 Jan;7(1):24-31.PMID: 21235516.
[344]. Zhou GP, The disposition of the LZCC protein residues in wenxiang diagram provides new insights into the protein-protein interaction mechanism. J Theor Biol. 2011 Sep 7;284(1):142-8.PMID: 21718705.
[345]. Li J, Wei DQ, Wang JF, Yu ZT, Chou KC. Molecular dynamics simulations of CYP2E1. Med Chem. 2012 Mar 1;8(2):208-21.
[346]. Wang JF, Chou KC. Recent advances in computational studies on influenza a virus M2 proton channel. Mini Rev Med Chem. 2012 Sep 1;12(10):971-8.
[347]. Zhang T, Wei DQ, Chou KC. A pharmacophore model specific to

active site of CYP1A2 with a novel molecular modeling explorer and CoMFA. Med Chem. 2012 Mar 1;8(2):198-207.

[348]. Jia J, Liu Z, Xiao X, Liu B, 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 377 (2015) 47-56.

[349]. 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. J Biomol Struct Dyn. 2016 Sep 1;34(9):1946-61.

[350]. Chou KC. Artificial intelligence (AI) tools constructed via the 5-steps rule for predicting post-translational modifications. Trends in Artificial Inttelengence (TIA) 3 (2019) 60-74.