Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Journal Pre-proof

An insightful recollection since the distorted key theory was born about 23 years ago

Kuo-Chen Chou

PII: S0888-7543(19)30554-3

DOI: https://doi.org/10.1016/j.ygeno.2019.09.001

Reference: YGENO 9350

To appear in: *Genomics*

Received date: 15 August 2019

Accepted date: 4 September 2019

Please cite this article as: K.-C. Chou, An insightful recollection since the distorted key theory was born about 23 years ago, *Genomics* (2019), https://doi.org/10.1016/j.ygeno.2019.09.001

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier.

Journal Pre-proof

An insightful recollection since the distorted key theory was born about 23 years ago

Kuo-Chen Chou[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

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 via this minireview that the motivation 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

Running Title: Recollection of distorted key theory

Journal Pre-proof

Introduction

Human immunodeficiency virus infection or 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 [26-39]). Now it has become very clear that the motivation 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 (**Fig.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 (**Fig.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 (**Fig.2**).

According to Fisher's "lock and key" model [40], Koshland's "induced fit" theory [40], and the "rack mechanism" [41], 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 (**Fig.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 **Fig.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 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].

For more discussion about the distorted key theory, see a comprehensive review paper [42].

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., [5, 6, 11, 13, 16, 21, 28, 30, 37, 43-46] and analyze them from the codon usage approach [47-50]).

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/.

Journal Pre-proof

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.

Conclusion and Perspectives

Just like the pseudo amino acid components [51-54] and the 5-steps rule [55], which have been stimulated by the development of the sequence bioinformatics, widely and increasingly used in proteome and genome analysis [56-151] and [150-186], it is anticipated that the “distorted key theory” that was stimulated from the structural bioinformatics [187] as well as some pioneering studies [188-195] and their follow-up studies [196-216], will be also widely and increasingly used in drug development, particularly for developing peptide drugs [217] and multi-target drugs [218].

Journal Pre-proof

Figure Legends

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 [42] 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 [42] with permission.

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 [42] with permission.

Journal Pre-proof

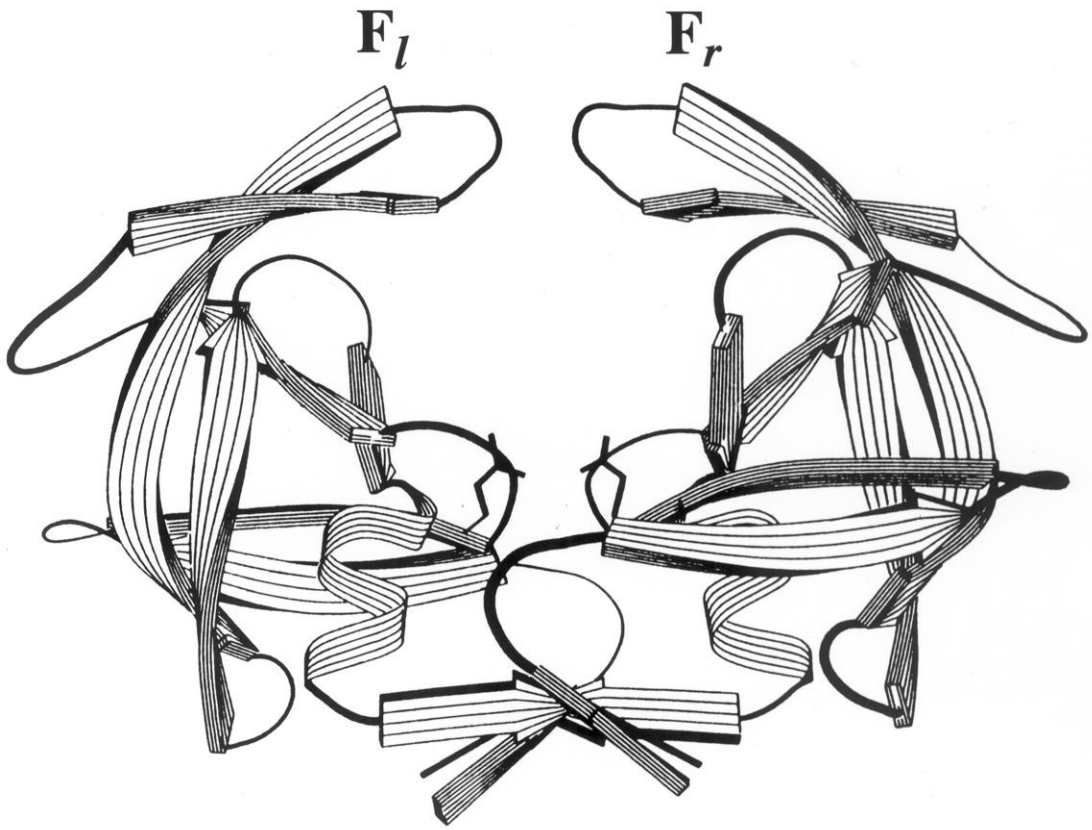

Figure 1

Journal

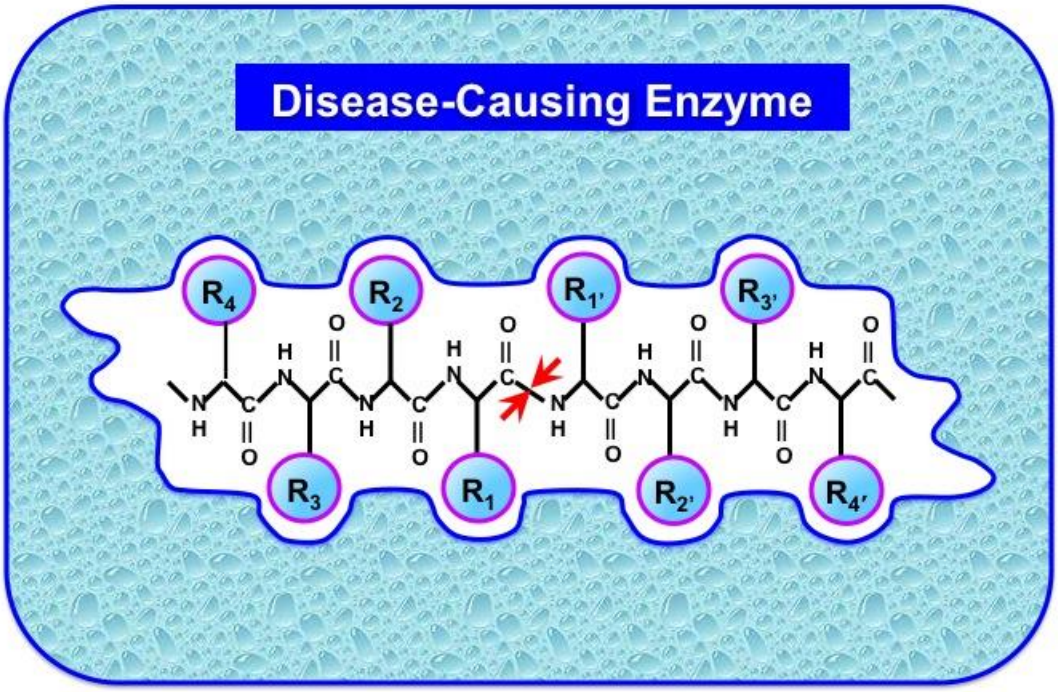

Figure 2. A schematic illustration to show a peptide in good fitting and tightly binding with the enzyme's active site before it is cleaved by the latter. Adapted from [42] with permission.

Journal

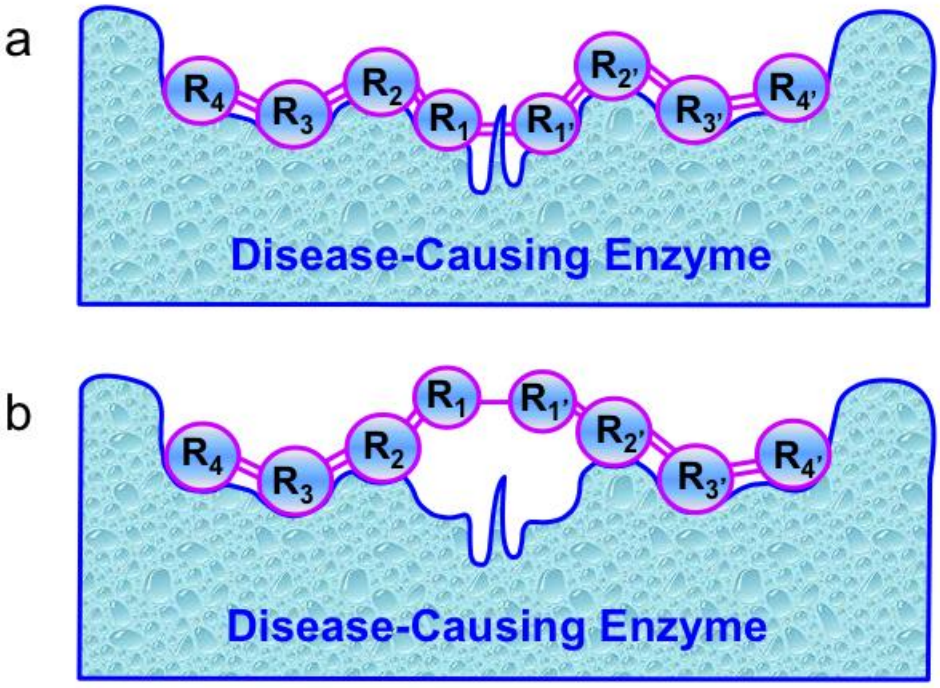

Figure 3. Schematic drawing to illustrate the "Distorted Key" theory, where panel (**a**) shows an effective binding of a cleavable peptide to the active site of a disease-causing enzyme, while panel (**b**) the same peptide has become a non-cleavable one after its scissile bond is modified although 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 [42] with permission.

References

[1] I.W. Althaus, J.J. Chou, A.J. Gonzales, M.R. Diebel, F.J. Kezdy, D.L. Romero, P.A. Aristoff, W.G. Tarpley, F. Reusser, Steady-state kinetic studies with the non-nucleoside HIV-1 reverse transcriptase inhibitor U-87201E. Journal of Biological Chemistry 268 (1993) 6119-6124.
[2] I.W. Althaus, J.J. Chou, A.J. Gonzales, M.R. Diebel, F.J. Kezdy, D.L. Romero, P.A. Aristoff, W.G. Tarpley, F. Reusser, Kinetic studies with the nonnucleoside HIV-1 reverse transcriptase inhibitor U-88204E. Biochemistry 32 (1993) 6548-6554.
[3] J.J. Chou, A formulation for correlating properties of peptides and its application to predicting human immunodeficiency virus protease-cleavable sites in proteins. Biopolymers 33 (1993) 1405-14.
[4] I.W. Althaus, A.J. Gonzales, J.J. Chou, M.R. Diebel, F.J. Kezdy, D.L. Romero, P.A. Aristoff, W.G. Tarpley, F. Reusser, The quinoline U-78036 is a potent inhibitor of HIV-1 reverse transcriptase. Journal of Biological Chemistry 268 (1993) 14875-14880.
[5] J.J. Chou, Predicting cleavability of peptide sequences by HIV protease via correlation-angle approach. Journal of Protein Chemistry 12 (1993) 291-302.
[6] K.C. Chou, A vectorized sequence-coupling model for predicting HIV protease cleavage sites in proteins. Journal of Biological Chemistry 268 (1993) 16938-16948.
[7] K.C. Chou, C.T. Zhang, Studies on the specificity of HIV protease: an application of Markov chain theory. Journal of Protein Chemistry 12 (1993) 709-724.
[8] K.C. Chou, C.T. Zhang, F.J. Kezdy, A vector approach to predicting HIV protease cleavage sites in proteins. Proteins: Structure, Function, and Genetics 16 (1993) 195-204.
[9] I.W. Althaus, J.J. Chou, A.J. Gonzales, M.R. Diebel, F.J. Kezdy, D.L. Romero, P.A. Aristoff, W.G. Tarpley, F. Reusser, Steady-state kinetic studies with the polysulfonate U-9843, an HIV reverse transcriptase inhibitor. Cellular and Molecular Life Science (Experientia) 50 (1994) 23-28.
[10] C.T. Zhang, An alternate-subsite-coupled model for predicting HIV protease cleavage sites in proteins. Protein Engineering 7 (1994) 65-73.
[11] T.B. Thompson, C. Zheng, Neural network prediction of the HIV-1 protease cleavage sites. J Theor Biol 177 (1995) 369-379.
[12] I.W. Althaus, K.M. Franks, M.R. Diebel, F.J. Kezdy, D.L. Romero, R.C. Thomas, P.A. Aristoff, W.G. Tarpley, F. Reusser, The benzylthio-pyrididine U-31,355, a potent inhibitor of HIV-1 reverse transcriptase. Biochemical Pharmacology 51 (1996) 743-750.
[13] K.C. Chou, A.L. Tomasselli, I.M. Reardon, R.L. Heinrikson, Predicting HIV protease cleavage sites in proteins by a discriminant function method. PROTEINS: Structure, Function, and Genetics 24 (1996) 51-72.
[14] Y.D. Cai, Artificial neural network model for HIV protease cleavage sites in proteins. Advances in Engineering Software 29 (1998) 119-128.
[15] Y.D. Cai, H. Yu, Using neural network for prediction of HIV protease cleavage sites in proteins. Journal of Protein Chemistry 17 (1998) 607-615.
[16] Y.D. Cai, X.J. Liu, X.B. Xu, Support Vector Machines for predicting HIV protease cleavage sites in protein. J Comput Chem 23 (2002) 267-274.
[17] S. Sirois, T. Sing, Review: HIV-1 gp120 V3 loop for structure-based drug design. Current Protein and Peptide Science 6 (2005) 413-422.

Journal Pre-proof

[18] S. Sirois, C.M. Tsoukas, D.Q. Wei, C. Boucher, G.E. Hatzakis, Selection of Molecular Descriptors with Artificial Intelligence for the Understanding of HIV-1 Protease Peptidomimetic Inhibitors-activity. Medicinal Chemistry 1 (2005) 173-184.
[19] W.N. Gao, D.Q. Wei, Y. Li, H. Gao, W.R. Xu, A.X. Li, Agaritine and its derivatives are potential inhibitors against HIV proteases. Medicinal Chemistry 3 (2007) 221-226.
[20] S. Sirois, M. Touaibia, R. Roy, Review: Glycosylation of HIV-1 gp120 V3 loop: towards the rational design of a synthetic carbohydrate vaccine. Current Medicinal Chemistry 14 (2007) 3232-3242.
[21] H.B. Shen, HIVcleave: a web-server for predicting HIV protease cleavage sites in proteins. Analytical Biochemistry 375 (2008) 388-390.
[22] J. Dev, D. Park, Q. Fu, J. Chen, H.J. Ha, F. Ghantous, T. Herrmann, W. Chang, Z. Liu, G. Frey, M.S. Seaman, B. Chen, J.J. Chou, Structural Basis for Membrane Anchoring of HIV-1 Envelope Spike. Science 353 (2016) 172-175.
[23] B. Chen, J.J. Chou, Structure of the transmembrane domain of HIV-1 envelope glycoprotein. FEBS J 284 (2017) 1171-1177.
[24] A. Piai, J. Dev, Q. Fu, J.J. Chou, Stability and Water Accessibility of the Trimeric Membrane Anchors of the HIV-1 Envelope Spikes. J Am Chem Soc 139 (2017) 18432-18435.
[25] Q. Fu, M.M. Shaik, Y. Cai, F. Ghantous, A. Piai, H. Peng, S. Rits-Volloch, Z. Liu, S.C. Harrison, M.S. Seaman, B. Chen, J.J. Chou, Structure of the membrane proximal external region of HIV-1 envelope glycoprotein. Proc Natl Acad Sci U S A 115 (2018) E8892-E8899.
[26] K. Anand, J. Ziebuhr, P. Wadhwani, J.R. Mesters, R. Hilgenfeld, Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. Science 300 (2003) 1763-1767.
[27] K.C. Chou, D.Q. Wei, W.Z. Zhong, Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design against SARS. (Erratum: ibid., 2003, Vol.310, 675). Biochem Biophys Res Comm (BBRC) 308 (2003) 148-151.
[28] Q.S. Du, S.Q. Wang, D.Q. Wei, Y. Zhu, H. Guo, S. Sirois, Polyprotein Cleavage Mechanism of SARS CoV Mpro and Chemical Modification of Octapeptide. Peptides 25 (2004) 1857-1864.
[29] S. Sirois, D.Q. Wei, Q.S. Du, Virtual Screening for SARS-CoV Protease Based on KZ7088 Pharmacophore Points. J. Chem. Inf. Comput. Sci. 44 (2004) 1111-1122.
[30] Q.S. Du, S. Wang, D.Q. Wei, S. Sirois, Molecular modelling and chemical modification for finding peptide inhibitor against SARS CoV Mpro. Analytical Biochemistry 337 (2005) 262-270.
[31] Q.S. Du, S.Q. Wang, Z.Q. Jiang, W.N. Gao, Y.D. Li, D.Q. Wei, Application of bioinformatics in search for cleavable peptides of SARS-CoV Mpro and chemical modification of octapeptides. Medicinal Chemistry 1 (2005) 209-213.
[32] M. Wang, J.S. Yao, Z.D. Huang, Z.J. Xu, G.P. Liu, H.Y. Zhao, X.Y. Wang, J. Yang, Y.S. Zhu, A new nucleotide-composition based fingerprint of SARS-CoV with visualization analysis. Medicinal Chemistry 1 (2005) 39-47.
[33] K.C. Chou, D.Q. Wei, Q.S. Du, S. Sirois, W.Z. Zhong, Review: Progress in computational approach to drug development against SARS. Current Medicinal Chemistry 13 (2006) 3263-3270.
[34] L. Gao, Y.S. Ding, H. Dai, S.H. Shao, Z.D. Huang, A novel fingerprint map for detecting SARS-CoV. Journal of Pharmaceutical and Biomedical Analysis 41 (2006) 246-250.

[35] D.Q. Wei, R. Zhang, Q.S. Du, W.N. Gao, Y. Li, H. Gao, S.Q. Wang, X. Zhang, A.X. Li, S. Sirois, Anti-SARS drug screening by molecular docking. Amino Acids 31 (2006) 73-80.
[36] R. Zhang, D.Q. Wei, Q.S. Du, Molecular modeling studies of peptide drug candidates against SARS. Medicinal Chemistry 2 (2006) 309-314.
[37] Q.S. Du, H. Sun, Inhibitor design for SARS coronavirus main protease based on "distorted key theory". Medicinal Chemistry 3 (2007) 1-6.
[38] S.Q. Wang, Q.S. Du, K. Zhao, A.X. Li, D.Q. Wei, Virtual screening for finding natural inhibitor against cathepsin-L for SARS therapy. Amino Acids 33 (2007) 129-135.
[39] K.C. Chou, D.Q. Wei, Q.S. Du, S. Sirois, H.B. Shen, W.Z. Zhong, 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, pp. 1-23.
[40] D. Voet, J.G. Voet, C.W. Pratt, Fundamentals of Biochemistry, John Wiley & Sons, New York, 2002.
[41] K.C. Chou, N.Y. Chen, The biological functions of low-frequency phonons. Scientia Sinica 20 (1977) 447-457.
[42] K.C. Chou, Review: Prediction of human immunodeficiency virus protease cleavage sites in proteins. Analytical Biochemistry 233 (1996) 1-14.
[43] K.C. Chou, C.T. Zhang, A new approach to predicting protein folding types. Journal of Protein Chemistry 12 (1993) 169-178.
[44] G.Z. Liang, S.Z. Li, A new sequence representation as applied in better specificity elucidation for human immunodeficiency virus type 1 protease. Biopolymers 88 (2007) 401-412.
[45] L. Nanni, A. Lumini, Using ensemble of classifiers for predicting HIV protease cleavage sites in proteins. Amino Acids (2008) Accepted Mar-27-2008.
[46] Y.R. Gan, H. Huang, Y.D. Huang, C.M. Rao, Y. Zhao, J.S. Liu, L. Wu, D.Q. Wei, Synthesis and activity of an octapeptide inhibitor designed for SARS coronavirus main proteinase. Peptides 27 (2006) 622-5.
[47] K.C. Chou, C.T. Zhang, Diagrammatization of codon usage in 339 HIV proteins and its biological implication. AIDS Research and Human Retroviruses 8 (1992) 1967-1976.
[48] J.J. Chou, C.T. Zhang, A joint prediction of the folding types of 1490 human proteins from their genetic codons. Journal of Theoretical Biology 161 (1993) 251-262.
[49] C.T. Zhang, Graphic analysis of codon usage strategy in 1490 human proteins. Journal of Protein Chemistry 12 (1993) 329-335.
[50] C.T. Zhang, Analysis of codon usage in 1562 E. Coli protein coding sequences. Journal of Molecular Biology 238 (1994) 1-8.
[51] K.C. Chou, Prediction of protein cellular attributes using pseudo amino acid composition. PROTEINS: Structure, Function, and Genetics (Erratum: ibid., 2001, Vol.44, 60) 43 (2001) 246-255.
[52] K.C. Chou, Using amphiphilic pseudo amino acid composition to predict enzyme subfamily classes. Bioinformatics 21 (2005) 10-19.
[53] B. Liu, F. Liu, X. Wang, J. Chen, L. Fang, Pse-in-One: a web server for generating various modes of pseudo components of DNA, RNA, and protein sequences. Nucleic Acids Research 43 (2015) W65-W71.

Journal Pre-proof

[54] B. Liu, H. Wu, Pse-in-One 2.0: An improved package of web servers for generating various modes of pseudo components of DNA, RNA, and protein sequences. Natural Science 9 (2017) 67-91.
[55] K.C. Chou, Some remarks on protein attribute prediction and pseudo amino acid composition (50th Anniversary Year Review, 5-steps rule). Journal of Theoretical Biology 273 (2011) 236-247.
[56] X.B. Zhou, C. Chen, Z.C. Li, X.Y. Zou, Using Chou's amphiphilic pseudo amino acid composition and support vector machine for prediction of enzyme subfamily classes. Journal of Theoretical Biology 248 (2007) 546–551.
[57] Y.S. Ding, T.L. Zhang, Using Chou's pseudo amino acid composition to predict subcellular localization of apoptosis proteins: an approach with immune genetic algorithm-based ensemble classifier. Pattern Recognition Letters 29 (2008) 1887-1892.
[58] Y. Fang, Y. Guo, Y. Feng, M. Li, Predicting DNA-binding proteins: approached from Chou's pseudo amino acid composition and other specific sequence features. Amino Acids 34 (2008) 103-109.
[59] 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 & Peptide Letters 15 (2008) 392-396.
[60] 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 34 (2008) 669-675.
[61] F.M. Li, Q.Z. Li, Predicting protein subcellular location using Chou's pseudo amino acid composition and improved hybrid approach. Protein & Peptide Letters 15 (2008) 612-616.
[62] H. Lin, The modified Mahalanobis discriminant for predicting outer membrane proteins by using Chou's pseudo amino acid composition. Journal of Theoretical Biology 252 (2008) 350-356.
[63] H. Lin, H. Ding, F.B. Feng-Biao Guo, A.Y. Zhang, J. Huang, Predicting subcellular localization of mycobacterial proteins by using Chou's pseudo amino acid composition. Protein & Peptide Letters 15 (2008) 739-744.
[64] L. Nanni, A. Lumini, Genetic programming for creating Chou's pseudo amino acid based features for submitochondria localization. Amino Acids 34 (2008) 653-660.
[65] G.Y. Zhang, B.S. Fang, Predicting the cofactors of oxidoreductases based on amino acid composition distribution and Chou's amphiphilic pseudo amino acid composition. Journal of Theoretical Biology 253 (2008) 310-315.
[66] G.Y. Zhang, H.C. Li, J.Q. Gao, B.S. Fang, Predicting lipase types by improved Chou's pseudo amino acid composition. Protein & Peptide Letters 15 (2008) 1132-1137.
[67] S.W. Zhang, W. Chen, F. Yang, Q. Pan, Using Chou's pseudo amino acid composition to predict protein quaternary structure: a sequence-segmented PseAAC approach. Amino Acids 35 (2008) 591-598.
[68] 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 34 (2008) 565-572.
[69] C. Chen, L. Chen, X. Zou, P. Cai, Prediction of protein secondary structure content by using the concept of Chou's pseudo amino acid composition and support vector machine. Protein & Peptide Letters 16 (2009) 27-31.

Journal Pre-proof

[70] H. Ding, L. Luo, H. Lin, Prediction of cell wall lytic enzymes using Chou's amphiphilic pseudo amino acid composition. Protein & Peptide Letters 16 (2009) 351-355.
[71] D.N. Georgiou, T.E. Karakasidis, J.J. Nieto, A. Torres, Use of fuzzy clustering technique and matrices to classify amino acids and its impact to Chou's pseudo amino acid composition. Journal of Theoretical Biology 257 (2009) 17-26.
[72] Z.C. Li, X.B. Zhou, Z. Dai, X.Y. Zou, Prediction of protein structural classes by Chou's pseudo amino acid composition: approached using continuous wavelet transform and principal component analysis. Amino Acids 37 (2009) 415-425.
[73] H. Lin, H. Wang, H. Ding, Y.L. Chen, Q.Z. Li, Prediction of Subcellular Localization of Apoptosis Protein Using Chou's Pseudo Amino Acid Composition. Acta Biotheoretica 57 (2009) 321-330.
[74] 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. Analytical Biochemistry 390 (2009) 68-73.
[75] 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. Journal of Theoretical Biology 259 (2009) 366–372.
[76] M. Esmaeili, H. Mohabatkar, S. Mohsenzadeh, Using the concept of Chou's pseudo amino acid composition for risk type prediction of human papillomaviruses. Journal of Theoretical Biology 263 (2010) 203-209.
[77] 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 & Peptide Letters 17 (2010) 559-567.
[78] H. Mohabatkar, Prediction of cyclin proteins using Chou's pseudo amino acid composition. Protein & Peptide Letters 17 (2010) 1207-1214.
[79] J.D. Qiu, J.H. Huang, S.P. Shi, R.P. Liang, Using the concept of Chou's pseudo amino acid composition to predict enzyme family classes: an approach with support vector machine based on discrete wavelet transform. Protein & Peptide Letters 17 (2010) 715-722.
[80] S.S. Sahu, G. Panda, A novel feature representation method based on Chou's pseudo amino acid composition for protein structural class prediction. Computational Biology and Chemistry 34 (2010) 320-327.
[81] 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. Journal of Theoretical Biology 267 (2010) 1-6.
[82] J. Guo, N. Rao, G. Liu, Y. Yang, G. Wang, Predicting protein folding rates using the concept of Chou's pseudo amino acid composition. Journal of Computational Chemistry 32 (2011) 1612-1617.
[83] J. Lin, Y. Wang, Using a novel AdaBoost algorithm and Chou's pseudo amino acid composition for predicting protein subcellular localization. Protein & Peptide Letters 18 (2011) 1219-1225.
[84] J. Lin, Y. Wang, X. Xu, A novel ensemble and composite approach for classifying proteins based on Chou's pseudo amino acid composition. African Journal of Biotechnology 10 (2011) 16963-16968.
[85] H. Mohabatkar, M. Mohammad Beigi, A. Esmaeili, Prediction of GABA(A) receptor proteins using the concept of Chou's pseudo amino acid composition and support vector machine. Journal of Theoretical Biology 281 (2011) 18-23.

Journal Pre-proof

[86] B.M. Mohammad, M. Behjati, H. Mohabatkar, Prediction of metalloproteinase family based on the concept of Chou's pseudo amino acid composition using a machine learning approach. Journal of Structural and Functional Genomics 12 (2011) 191-197.
[87] J.D. Qiu, S.B. Suo, X.Y. Sun, S.P. Shi, R.P. Liang, 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 30 (2011) 129-134.
[88] D. Zou, Z. He, J. He, Y. Xia, Supersecondary structure prediction using Chou's pseudo amino acid composition. Journal of Computational Chemistry 32 (2011) 271-278.
[89] J.Z. Cao, W.Q. Liu, H. Gu, Predicting Viral Protein Subcellular Localization with Chou's Pseudo Amino Acid Composition and Imbalance-Weighted Multi-Label K-Nearest Neighbor Algorithm. Protein and Peptide Letters 19 (2012) 1163-1169.
[90] C. Chen, Z.B. Shen, X.Y. Zou, Dual-Layer Wavelet SVM for Predicting Protein Structural Class Via the General Form of Chou's Pseudo Amino Acid Composition. Protein & Peptide Letters 19 (2012) 422-429.
[91] P. Du, X. Wang, C. Xu, Y. Gao, PseAAC-Builder: A cross-platform stand-alone program for generating various special Chou's pseudo amino acid compositions. Analytical Biochemistry 425 (2012) 117-119.
[92] G.L. Fan, Q.Z. Li, Predict mycobacterial proteins subcellular locations by incorporating pseudo-average chemical shift into the general form of Chou's pseudo amino acid composition. Journal of Theoretical Biology 304 (2012) 88-95.
[93] G.L. Fan, Q.Z. Li, Predicting protein submitochondria locations by combining different descriptors into the general form of Chou's pseudo amino acid composition. Amino Acids 43 (2012) 545-555.
[94] L.Q. Li, Y. Zhang, L.Y. Zou, Y. Zhou, X.Q. Zheng, Prediction of Protein Subcellular Multi-Localization Based on the General form of Chou's Pseudo Amino Acid Composition. Protein & Peptide Letters 19 (2012) 375-387.
[95] L. Liu, X.Z. Hu, X.X. Liu, Y. Wang, S.B. Li, Predicting Protein Fold Types by the General Form of Chou's Pseudo Amino Acid Composition: Approached from Optimal Feature Extractions. Protein & Peptide Letters 19 (2012) 439-449.
[96] L. Nanni, S. Brahnam, A. Lumini, Wavelet images and Chou's pseudo amino acid composition for protein classification. Amino Acids 43 (2012) 657-65.
[97] L. Nanni, A. Lumini, D. Gupta, A. Garg, 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 9 (2012) 467-475.
[98] X.H. Niu, X.H. Hu, F. Shi, J.B. Xia, Predicting Protein Solubility by the General Form of Chou's Pseudo Amino Acid Composition: Approached from Chaos Game Representation and Fractal Dimension. Protein & Peptide Letters 19 (2012) 940-948.
[99] L.Y. Ren, Y.S. Zhang, I. Gutman, Predicting the Classification of Transcription Factors by Incorporating their Binding Site Properties into a Novel Mode of Chou's Pseudo Amino Acid Composition. Protein & Peptide Letters 19 (2012) 1170-1176.
[100] X.W. Zhao, Z.Q. Ma, M.H. Yin, Predicting protein-protein interactions by combing various sequence- derived features into the general form of Chou's Pseudo amino acid composition. Protein & Peptide Letters 19 (2012) 492-500.

Journal Pre-proof

[101] Zia-ur-Rehman, A. Khan, Identifying GPCRs and their Types with Chou's Pseudo Amino Acid Composition: An Approach from Multi-scale Energy Representation and Position Specific Scoring Matrix. Protein & Peptide Letters 19 (2012) 890-903.
[102] Y.K. Chen, K.B. Li, Predicting membrane protein types by incorporating protein topology, domains, signal peptides, and physicochemical properties into the general form of Chou's pseudo amino acid composition. Journal of Theoretical Biology 318 (2013) 1-12.
[103] G.L. Fan, Q.Z. Li, Discriminating bioluminescent proteins by incorporating average chemical shift and evolutionary information into the general form of Chou's pseudo amino acid composition. Journal of Theoretical Biology 334 (2013) 45-51.
[104] D.N. Georgiou, T.E. Karakasidis, A.C. Megaritis, A short survey on genetic sequences, Chou's pseudo amino acid composition and its combination with fuzzy set theory. The Open Bioinformatics Journal 7 (2013) 41-48.
[105] M.K. Gupta, R. Niyogi, M. Misra, 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 24 (2013) 597-609.
[106] C. Huang, J. Yuan, 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 113 (2013) 50-57.
[107] C. Huang, J.Q. Yuan, 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 246 (2013) 327-34.
[108] C. Huang, J.Q. Yuan, Predicting protein subchloroplast locations with both single and multiple sites via three different modes of Chou's pseudo amino acid compositions. Journal of Theoretical Biology 335 (2013) 205-12.
[109] M. Khosravian, F.K. Faramarzi, M.M. Beigi, M. Behbahani, H. Mohabatkar, Predicting Antibacterial Peptides by the Concept of Chou's Pseudo amino Acid Composition and Machine Learning Methods. Protein & Peptide Letters 20 (2013) 180-186.
[110] H. Lin, C. Ding, L.-F. Yuan, W. Chen, H. Ding, Z.-Q. Li, F.-B. Guo, J. Huang, N.-N. Rao, 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 Biomethmatics 6 (2013) 1350003.
[111] B. Liu, X. Wang, Q. Zou, Q. Dong, Q. Chen, Protein remote homology detection by combining Chou's pseudo amino acid composition and profile-based protein representation. Molecular Informatics 32 (2013) 775-782.
[112] H. Mohabatkar, M.M. Beigi, K. Abdolahi, S. Mohsenzadeh, Prediction of Allergenic Proteins by Means of the Concept of Chou's Pseudo Amino Acid Composition and a Machine Learning Approach. Medicinal Chemistry 9 (2013) 133-137.
[113] Y.F. Qin, L. Zheng, J. Huang, Locating apoptosis proteins by incorporating the signal peptide cleavage sites into the general form of Chou's Pseudo amino acid composition. International Journal of Quantum Chemistry 113 (2013) 1660-1667.
[114] A.N. Sarangi, M. Lohani, R. Aggarwal, 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 20 (2013) 781-95.
[115] S. Wan, M.W. Mak, S.Y. Kung, GOASVM: A subcellular location predictor by incorporating term-frequency gene ontology into the general form of Chou's pseudo amino acid composition. Journal of Theoretical Biology 323 (2013) 40-48.

Journal Pre-proof

[116] X. Wang, G.Z. Li, W.C. Lu, Virus-ECC-mPLoc: a multi-label predictor for predicting the subcellular localization of virus proteins with both single and multiple sites based on a general form of Chou's pseudo amino acid composition. Protein & Peptide Letters 20 (2013) 309-317.
[117] N. Xiaohui, L. Nana, X. Jingbo, C. Dingyan, P. Yuehua, X. Yang, W. Weiquan, W. Dongming, W. Zengzhen, Using the concept of Chou's pseudo amino acid composition to predict protein solubility: An approach with entropies in information theory. Journal of Theoretical Biology 332 (2013) 211-217.
[118] P. Du, S. Gu, Y. Jiao, PseAAC-General: Fast building various modes of general form of Chou's pseudo amino acid composition for large-scale protein datasets. International Journal of Molecular Sciences 15 (2014) 3495-3506.
[119] Z. Hajisharifi, M. Piryaiee, M. Mohammad Beigi, M. Behbahani, H. Mohabatkar, Predicting anticancer peptides with Chou's pseudo amino acid composition and investigating their mutagenicity via Ames test. Journal of Theoretical Biology 341 (2014) 34-40.
[120] C. Jia, X. Lin, Z. Wang, 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 15 (2014) 10410-23.
[121] L. Kong, L. Zhang, J. Lv, 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 344 (2014) 12-18.
[122] L. Nanni, S. Brahnam, A. Lumini, Prediction of protein structure classes by incorporating different protein descriptors into general Chou's pseudo amino acid composition. J Theor Biol 360 (2014) 109-116.
[123] J. Zhang, P. Sun, X. Zhao, Z. Ma, PECM: Prediction of extracellular matrix proteins using the concept of Chou's pseudo amino acid composition. Journal of Theoretical Biology 363 (2014) 412-418.
[124] L. Zhang, X. Zhao, L. Kong, 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 355 (2014) 105-10.
[125] Y.C. Zuo, Y. Peng, L. Liu, W. Chen, L. Yang, G.L. Fan, Predicting peroxidase subcellular location by hybridizing different descriptors of Chou's pseudo amino acid patterns. Anal Biochem 458 (2014) 14-9.
[126] F. Ali, M. Hayat, Classification of membrane protein types using Voting Feature Interval in combination with Chou's Pseudo Amino Acid Composition. J Theor Biol 384 (2015) 78-83.
[127] G.L. Fan, X.Y. Zhang, Y.L. Liu, Y. Nang, H. Wang, DSPMP: Discriminating secretory proteins of malaria parasite by hybridizing different descriptors of Chou's pseudo amino acid patterns. J Comput Chem 36 (2015) 2317-27.
[128] C. Huang, J.Q. Yuan, Simultaneously Identify Three Different Attributes of Proteins by Fusing their Three Different Modes of Chou's Pseudo Amino Acid Compositions. Protein Pept Lett 22 (2015) 547-56.
[129] Z.U. Khan, M. Hayat, M.A. Khan, Discrimination of acidic and alkaline enzyme using Chou's pseudo amino acid composition in conjunction with probabilistic neural network model. J Theor Biol 365 (2015) 197-203.

Journal Pre-proof

[130] R. Kumar, A. Srivastava, B. Kumari, M. Kumar, Prediction of beta-lactamase and its class by Chou's pseudo amino acid composition and support vector machine. J Theor Biol 365 (2015) 96-103.
[131] B. Liu, J. Chen, X. Wang, Protein remote homology detection by combining Chou's distance-pair pseudo amino acid composition and principal component analysis. Mol Genet Genomics 290 (2015) 1919-31.
[132] X. Wang, W. Zhang, Q. Zhang, G.Z. Li, MultiP-SChlo: multi-label protein subchloroplast localization prediction with Chou's pseudo amino acid composition and a novel multi-label classifier. Bioinformatics 31 (2015) 2639-45.
[133] K. Ahmad, M. Waris, M. Hayat, Prediction of Protein Submitochondrial Locations by Incorporating Dipeptide Composition into Chou's General Pseudo Amino Acid Composition. J Membr Biol 249 (2016) 293-304.
[134] M. Behbahani, H. Mohabatkar, M. Nosrati, 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 411 (2016) 1-5.
[135] G.L. Fan, Y.L. Liu, H. Wang, Identification of thermophilic proteins by incorporating evolutionary and acid dissociation information into Chou's general pseudo amino acid composition. J Theor Biol 407 (2016) 138-142.
[136] Y.S. Jiao, P.F. Du, 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 402 (2016) 38-44.
[137] H. Tang, W. Chen, H. Lin, Identification of immunoglobulins using Chou's pseudo amino acid composition with feature selection technique. Mol Biosyst 12 (2016) 1269-1275.
[138] C. Xu, D. Sun, S. Liu, Y. Zhang, Protein Sequence Analysis by Incorporating Modified Chaos Game and Physicochemical Properties into Chou's General Pseudo Amino Acid Composition. J Theor Biol 406 (2016) 105-115.
[139] H.L. Zou, X. Xiao, Predicting the Functional Types of Singleplex and Multiplex Eukaryotic Membrane Proteins via Different Models of Chou's Pseudo Amino Acid Compositions. J Membr Biol 249 (2016) 23-9.
[140] H.L. Zou, X. Xiao, Classifying Multifunctional Enzymes by Incorporating Three Different Models into Chou's General Pseudo Amino Acid Composition (doi:10.1007/s00232-016-9904-3). J Membr Biol 249 (2016) 561-7.
[141] Y. Liang, S. Zhang, Predict protein structural class by incorporating two different modes of evolutionary information into Chou's general pseudo amino acid composition. J Mol Graph Model 78 (2017) 110-117.
[142] M. Rahimi, M.R. Bakhtiarizadeh, A. Mohammadi-Sangcheshmeh, OOgenesis_Pred: A sequence-based method for predicting oogenesis proteins by six different modes of Chou's pseudo amino acid composition. J Theor Biol 414 (2017) 128-136.
[143] P. Tripathi, P.N. Pandey, 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 424 (2017) 49-54.
[144] B. Yu, L. Lou, S. Li, Y. Zhang, W. Qiu, X. Wu, M. Wang, B. Tian, Prediction of protein structural class for low-similarity sequences using Chou's pseudo amino acid composition and wavelet denoising. J Mol Graph Model 76 (2017) 260-273.

Journal Pre-proof

[145] M. Arif, M. Hayat, Z. Jan, 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 442 (2018) 11-21.
[146] Z. Ju, S.Y. Wang, Prediction of citrullination sites by incorporating k-spaced amino acid pairs into Chou's general pseudo amino acid composition. Gene 664 (2018) 78-83.
[147] J. Mei, Y. Fu, J. Zhao, Analysis and prediction of ion channel inhibitors by using feature selection and Chou's general pseudo amino acid composition. J Theor Biol 456 (2018) 41-48.
[148] J. Mei, J. Zhao, Prediction of HIV-1 and HIV-2 proteins by using Chou's pseudo amino acid compositions and different classifiers. Sci Rep 8 (2018) 2359.
[149] J. Mei, J. Zhao, Analysis and prediction of presynaptic and postsynaptic neurotoxins by Chou's general pseudo amino acid composition and motif features. J Theor Biol 427 (2018) 147-153.
[150] M. Awais, W. Hussain, Y.D. Khan, N. Rasool, S.A. Khan, 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 doi:10.1109/TCBB.2019.2919025 (2019).
[151] A. Ehsan, M.K. Mahmood, Y.D. Khan, O.M. Barukab, S.A. Khan, iHyd-PseAAC (EPSV): Identify hydroxylation sites in proteins by extracting enhanced position and sequence variant feature via Chou's 5-step rule and general pseudo amino acid composition. Current Genomics 20 (2019) 124-133.
[152] X. Du, Y. Diao, H. Liu, S. Li, MsDBP: Exploring DNA-binding Proteins by Integrating Multi-scale Sequence Information via Chou's 5-steps Rule. Journal of Proteome Research 18 (2019) 3119-3132.
[153] W. Hussain, S.D. Khan, N. Rasool, S.A. Khan, 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 568 (2019) 14-23.
[154] W. Hussain, Y.D. Khan, N. Rasool, S.A. Khan, 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 468 (2019) 1-11.
[155] Z. Ju, S.Y. Wang, 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 doi:10.1016/j.ygeno.2019.05.027 (2019).
[156] M. Kabir, S. Ahmad, M. Iqbal, M. Hayat, 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 doi:10.1016/j.ygeno.2019.02.006 (2019).
[157] N.Q.K. Le, 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 doi:10.1007/s00438-019-01570-y (2019).
[158] N.Q.K. Le, E.K.Y. Yapp, Q.T. Ho, N. Nagasundaram, Y.Y. Ou, H.Y. Yeh, iEnhancer-5Step: Identifying enhancers using hidden information of DNA sequences via Chou's 5-step rule and word embedding. Anal Biochem 571 (2019) 53-61.
[159] N.Q.K. Le, E.K.Y. Yapp, Y.Y. Ou, H.Y. Yeh, iMotor-CNN: Identifying molecular functions of cytoskeleton motor proteins using 2D convolutional neural network via Chou's 5-step rule. Anal Biochem 575 (2019) 17-26.

Journal Pre-proof

[160] Y. Liang, S. Zhang, Identifying DNase I hypersensitive sites using multi-features fusion and F-score features selection via Chou's 5-steps rule. Biophysical Chemistry 253 (2019) 106227.
[161] Q. Ning, Z. Ma, X. Zhao, 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 470 (2019) 43-49.
[162] Salman, M. Khan, N. Iqbal, T. Hussain, S. Afzal, 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) Doi: 10.1007/s10989-019-09887-3 (2019).
[163] M. Tahir, H. Tayara, K.T. Chong, 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 189 (2019) 96-101.
[164] X. Cheng, X. Xiao, pLoc-mPlant: predict subcellular localization of multi-location plant proteins via incorporating the optimal GO information into general PseAAC. Molecular BioSystems 13 (2017) 1722-1727.
[165] X. Cheng, X. Xiao, pLoc-mVirus: predict subcellular localization of multi-location virus proteins via incorporating the optimal GO information into general PseAAC. Gene (Erratum: ibid., 2018, Vol.644, 156-156) 628 (2017) 315-321.
[166] X. Cheng, S.G. Zhao, W.Z. Lin, X. Xiao, pLoc-mAnimal: predict subcellular localization of animal proteins with both single and multiple sites. Bioinformatics 33 (2017) 3524-3531.
[167] X. Xiao, X. Cheng, S. Su, Q. Nao, pLoc-mGpos: Incorporate key gene ontology information into general PseAAC for predicting subcellular localization of Gram-positive bacterial proteins. Natural Science 9 (2017) 331-349.
[168] X. Cheng, X. Xiao, pLoc-mEuk: Predict subcellular localization of multi-label eukaryotic proteins by extracting the key GO information into general PseAAC. Genomics 110 (2018) 50-58.
[169] X. Cheng, X. Xiao, pLoc-mGneg: Predict subcellular localization of Gram-negative bacterial proteins by deep gene ontology learning via general PseAAC. Genomics 110 (2018) 231-239.
[170] X. Cheng, X. Xiao, pLoc-mHum: predict subcellular localization of multi-location human proteins via general PseAAC to winnow out the crucial GO information. Bioinformatics 34 (2018) 1448-1456.
[171] X. Cheng, X. Xiao, pLoc_bal-mGneg: predict subcellular localization of Gram-negative bacterial proteins by quasi-balancing training dataset and general PseAAC. Journal of Theoretical Biology 458 (2018) 92-102.
[172] X. Cheng, X. Xiao, pLoc_bal-mPlant: predict subcellular localization of plant proteins by general PseAAC and balancing training dataset Curr Pharm Des 24 (2018) 4013-4022.
[173] K.C. Chou, X. Cheng, X. Xiao, pLoc_bal-mHum: predict subcellular localization of human proteins by PseAAC and quasi-balancing training dataset Genomics doi:10.1016/j.ygeno.2018.08.007 (2018).
[174] K.C. Chou, X. Cheng, X. Xiao, pLoc_bal-mEuk: predict subcellular localization of eukaryotic proteins by general PseAAC and quasi-balancing training dataset. Med Chem 15 (2018) 472-485.

Journal Pre-proof

[175] X. Xiao, X. Cheng, G. Chen, Q. Mao, 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.
[176] X. Cheng, W.Z. Lin, X. Xiao, pLoc_bal-mAnimal: predict subcellular localization of animal proteins by balancing training dataset and PseAAC. Bioinformatics 35 (2019) 398-406.
[177] X. Xiao, X. Cheng, G. Chen, Q. Mao, pLoc_bal-mGpos: predict subcellular localization of Gram-positive bacterial proteins by quasi-balancing training dataset and PseAAC. Genomics 111 (2019) 886-892.
[178] K.C. Chou, Impacts of pseudo amino acid components and 5-steps rule to proteomics and proteome analysis. Current Topics in Medicinak Chemistry (CTMC) in press (2019).
[179] K.C. Chou, An insightful 10-year recollection since the emergence of the 5-steps rule. Current Pharmaceutical Design in press (2019).
[180] K.C. Chou, An insightful 20-year recollection since the birth of pseudo amino acid components. Applied Biochemistry and Biotechnology (ABAB) in press (2019).
[181] K.C. Chou, An insightful recollection for predicting protein subcellular locations in multi-label systems. Genomics in press (2019).
[182] K.C. Chou, An insightful recollection since the birth of Gordon Life Science Institute about 17 years ago. International Journal of Peptide Research and Therapeutics (IJPRT) in press (2019).
[183] K.C. Chou, Recent progresses in predicting protein subcellular localization with artificial intelligence tools developed via the 5-steps rule. Genomics in press (2019).
[184] O. Barukab, Y.D. Khan, S.A. Khan, iSulfoTyr-PseAAC: Identify tyrosine sulfation sites by incorporating statistical moments via Chou's 5-steps rule pseudo components Current Genomics Accepted (2019).
[185] I. Nazari, M. Tahir, H. Tayari, K.T. Chong, 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) doi: 10.1016/j.chemolab.2019.103811 (2019).
[186] K.C. Chou, Artificial intelligence (AI) tools constructed via the 5-steps rule for predicting post-translational modifications. Trends in Artificial Inttelengence (TIA) in press (2019).
[187] K.C. Chou, Review: Structural bioinformatics and its impact to biomedical science. Current Medicinal Chemistry 11 (2004) 2105-2134.
[188] K.C. Chou, S. Forsen, Diffusion-controlled effects in reversible enzymatic fast reaction system: Critical spherical shell and proximity rate constants. Biophysical Chemistry 12 (1980) 255-263.
[189] K.C. Chou, S. Forsen, Graphical rules for enzyme-catalyzed rate laws. Biochemical Journal 187 (1980) 829-835.
[190] S. Forsen, G.Q. Zhou, Three schematic rules for deriving apparent rate constants. Chemica Scripta 16 (1980) 109-113.
[191] T.T. Li, S. Forsen, The critical spherical shell in enzymatic fast reaction systems. Biophysical Chemistry 12 (1980) 265-269.
[192] T.T. Li, S. Forsen, The flow of substrate molecules in fast enzyme-catalyzed reaction systems. Chemica Scripta 16 (1980) 192-196.

Journal Pre-proof

[193] R.E. Carter, S. Forsen, A new graphical method for deriving rate equations for complicated mechanisms. Chemica Scripta 18 (1981) 82-86.
[194] N.Y. Chen, S. Forsen, The biological functions of low-frequency phonons: 2. Cooperative effects. Chemica Scripta 18 (1981) 126-132.
[195] K.C. Chou, S. Forsen, Graphical rules of steady-state reaction systems. Canadian Journal of Chemistry 59 (1981) 737-755.
[196] K.C. Chou, Low-frequency vibrations of helical structures in protein molecules. Biochemical Journal 209 (1983) 573-580.
[197] K.C. Chou, Identification of low-frequency modes in protein molecules. Biochemical Journal 215 (1983) 465-469.
[198] K.C. Chou, Biological functions of low-frequency vibrations (phonons). 3. Helical structures and microenvironment. Biophysical Journal 45 (1984) 881-889.
[199] K.C. Chou, The biological functions of low-frequency phonons. 4. Resonance effects and allosteric transition. Biophysical Chemistry 20 (1984) 61-71.
[200] K.C. Chou, Low-frequency vibrations of DNA molecules. Biochemical Journal 221 (1984) 27-31.
[201] K.C. Chou, Low-frequency motions in protein molecules: beta-sheet and beta-barrel. Biophysical Journal 48 (1985) 289-297.
[202] K.C. Chou, Prediction of a low-frequency mode in bovine pancreatic trypsin inhibitor molecule. International Journal of Biological Macromolecules 7 (1985) 77-80.
[203] K.C. Chou, Y.S. Kiang, The biological functions of low-frequency phonons: 5. A phenomenological theory. Biophysical Chemistry 22 (1985) 219-235.
[204] K.C. Chou, Origin of low-frequency motion in biological macromolecules: A view of recent progress of quasi-continuity model. Biophysical Chemistry 25 (1986) 105-116.
[205] K.C. Chou, The biological functions of low-frequency phonons: 6. A possible dynamic mechanism of allosteric transition in antibody molecules. Biopolymers 26 (1987) 285-295.
[206] K.C. Chou, Review: Low-frequency collective motion in biomacromolecules and its biological functions. Biophysical Chemistry 30 (1988) 3-48.
[207] K.C. Chou, G.M. Maggiora, The biological functions of low-frequency phonons: 7. The impetus for DNA to accommodate intercalators. British Polymer Journal 20 (1988) 143-148.
[208] K.C. Chou, Low-frequency resonance and cooperativity of hemoglobin. Trends in Biochemical Sciences 14 (1989) 212-213.
[209] K.C. Chou, G.M. Maggiora, B. Mao, Quasi-continuum models of twist-like and accordion-like low-frequency motions in DNA. Biophysical Journal 56 (1989) 295-305.
[210] K.C. Chou, C.T. Zhang, G.M. Maggiora, Solitary wave dynamics as a mechanism for explaining the internal motion during microtubule growth. Biopolymers 34 (1994) 143-153.
[211] H. Liu, M. Wang, Low-frequency Fourier spectrum for predicting membrane protein types. Biochem Biophys Res Commun (BBRC) 336 (2005) 737-739.
[212] G. Gordon, Designed Electromagnetic Pulsed Therapy: Clinical Applications. Journal of Cellular Physiology 212 (2007) 579-582.
[213] G. Gordon, Extrinsic electromagnetic fields, low frequency (phonon) vibrations, and control of cell function: a non-linear resonance system. Journal of Biomedical Science and Engineering (JBiSE) 1 (2008) 152-156

Journal Pre-proof

[214] A. Madkan, M. Blank, E. Elson, M.S. Geddis, R. Goodman, Steps to the clinic with ELF EMF Natural Science 1 (2009) 157-165.
[215] J. Lu, B. Niu, L. Liu, W.C. Lu, Y.D. Cai, Prediction of small molecules' metabolic pathways based on functional group composition. Protein & Peptide Letters 16 (2009) 969-976.
[216] K.C. Chou, Proposing pseudo amino acid components is an important milestone for proteome and genome analyses. International Journal for Peptide Research and Therapeutics (IJPRT) DOI: 10.1007/s10989-019-09910-7 (2019).
[217] K.C. Chou, Progresses in predicting post-translational modification. International Journal of Peptide Research and Therapeutics (IJPRT) DOI: 10.1007/s10989-019-09893-5 (2019).
[218] K.C. Chou, Advance in predicting subcellular localization of multi-label proteins and its implication for developing multi-target drugs. Current Medicinal Chemistry doi: 10.2174/0929867326666190507082559 (2019).

Journal Pre-proof

The author claims no conflicting interest whatsoever for this paper.

Journal Pre-proof