J. Biomedical Science and Engineering, 2013, 6, 435-442 **JBiSE**
http://dx.doi.org/10.4236/jbise.2013.64054 Published Online April 2013 (http://www.scirp.org/journal/jbise/)

Theoretical and experimental biology in one

—A symposium in honour of Professor Kuo-Chen Chou's 50th anniversary and Professor Richard Giegé's 40th anniversary of their scientific careers

Sheng-Xiang Lin[1,2], Jacques Lapointe[3]

[1]Laboratory of Molecular Endocrinology and Oncology, Research Center of Laval University Hospital Center (CHUQ-CHUL) and PROTEO, Laval University, Quebec, Canada
[2]WHO Collaborating Center for Research in Human Reproductive Health, Shanghai, China
[3]Department of Biochemistry, Microbiology and Bioinformatics and PROTEO, Laval University, Quebec, Canada
Email: sheng-xiang.lin@crchul.ulaval.ca

Received 15 July 2012; revised 4 September 2012; accepted 25 September 2012

ABSTRACT

It has been a dream that theoretical biology can be extensively applied in experimental biology to accelerate the understanding of the sophiscated movements in living organisms. A brave assay and an excellent example were represented by enzymology, in which the well-established physico-chemistry is used to describe, to fit, to predict and to improve enzyme reactions. Before the modern bioinformatics, the developments of the combination of theoretical biology and experimental biology have been mainly limited to various classic formulations. The systematic use of graphic rules by Prof. Kuo-Chen Chou and his coworkers has significantly facilitated to deal with complicated enzyme systems. With the recent fast progress of bioinformatics, prediction of protein structures and various protein attributes have been well established by Chou and co-workers, stimulating the experimental biology. For example, their recent method for predicting protein subcellular localization (one of the important attributes of proteins) has been extensively applied by scientific colleagues, yielding many new results with thousands of citations. The research by Prof. Chou is characterized by introducing novel physical concepts as well as powerful and elegant mathematical methods into important biomedical problems, a focus throughout his career, even when facing enormous difficulties. His efforts in 50 years have greatly helped us to realize the dream to make "theoretical and experimental biology in one". Prof. Richard Giege is well known for his multi-disciplinary research combining physics, chemistry, enzymology and molecular biology. His major focus of study is on the identity of tRNAs and their interactions with aminoacyl-tRNA synthetases (aaRS), which are of critical importance to the fidelity of protein biosynthesis. He and his colleagues have carried out the first crystallization of a tRNA/aaRS complex, that between $tRNA^{Asp}$ and AspRS from yeast. The determination of the complex structure contributed significantly to understand the interaction of protein and RNA. From his fine research, they have also found other biological function of these small RNAs. He has developed in parallel appropriate methods for his research, of which the protein crystallogenesis, a name he has coined, is an excellent example. Now macromolecular crystallogenesis has become a developed science. In fact, such contribution has accelerated the development of protein crystallography, stimulating the study of macromolecular structure and function.

Keywords: Theoretical Biology; Experimental Biology; Chou's Graphic Rules in Enzymology; Prediction of Protein Attributes; Chou's Invariance Theorem; Macromolecular Crystallogenesis; tRNA Identity on Aminoacylation Specificity; Physico-Biochemistry

1. A PARADOX IN ENZYME FAST REACTIONS AND CHOU'S DIIFUSION-CONTROLLED REACTION MODEL

The upper limit of enzyme-substrate reaction was originally estimated by Prof. Manfred Eigen and his coworkers using a simple mathematical model [1]. According to their estimation, the upper limit of enzyme-substrate reaction was 10^9/Msec [1]. In 1972 a paradox occurred [2] that some enzyme-substrate reaction could reach the rate with one order of magnitude higher than the upper limit estimated by Prof. Eigen *et al.* [1]. To address such a paradox, Prof. Chou and his co-workers proposed a more rigorous and advanced model by taking into account the spatial factor and force field factor be-

tween the enzyme and its substrate [3,4]. It was found by solving complicated differential equations that the upper limit can really be one order of magnitude higher than what was originally estimated by Prof. Eigen and co-workers. Meanwhile, it was also found that the protein outside the active site may also play an important role for the upper limit [5].

For a brief discussion about this, see a Wikipedia article entitled "Diffusion-controlled reaction" by clicking the link http://en.wikipedia.org/wiki/Diffusion-controlled_reaction.

2. CHOU'S GRAPHIC RULES IN BIOLOGICAL SYSTEMS

The extensive calculation in enzyme kinetics, especially for complicated systems often hinders the experimental work of biologists in making significant progress. Using graphical approaches to study biological problems can provide an intuitive picture or useful insights for helping analyzing complicated mechanisms in these systems. The powerful and elegant graphic rules proposed by Chou and his co-workers [6-8], usually cited as "Chou's graphic rules" in the literatures, has been rigorously proved and widely used, such as in studying conformational change in liver glucokinase in a non-steady-state [9], deriving enzyme kinetic equations to systems involving parallel reaction pathways [10], studying mixtures of tight-binding enzyme inhibitors [11], developing microcosm-puter tools for steady-state enzyme kinetics [12], inhibition of HIV-1 reverse transcriptase [13-15], studying kinetic plasticity and the determination of product ratios [16], inhibition kinetics of processive nucleic acid polymerases and nucleases [17], protein folding kinetics [18], and drug metabolism systems [19]. The systematic application of graphic theory into enzymology has a significant contribution in this domain. In fact, such graphic methods provide a visually intuitive relation between calculations and reaction graphics, highlighting key points from complicated details, significantly simplifying the calculations and facilitating to check the complicated results. Chou's graphic rules 1 - 3 [6,7] were established for steady-state kinetics systems while Chou's graphic rule 4 [6] for non-steady-state kinetics systems. With the development of modern biology with a transition from molecules to cells and further to intact organisms, enzymology in cells and organisms has become more and more important [20]. Hence, it is expected that the Chou's graphic rules will be also very useful for studying the cells and even organisms.

For a brief introduction about "graph theory in enzyme kinetics", see a Wikipedia article at http://en.wikipedia.org/wiki/Graph_theory_in_enzymatic_kinetics.

3. CHOU'S PSEAAC AND ITS APPLICATIONS IN PREDICTING PROTEIN ATTRIBUTES

To develop a powerful predictor for protein systems, one of the keys is to formulate the protein samples with an effective mathematical expression or feature vector that can truly reflect their intrinsic correlation with the target to be predicted. In view of this, Professor Chou introduced the pseudo amino acid composition (PseAAC) to deal with this problem. Since the concept of PseAAC was proposed [21] in 2001, the application of PseAAC has been widely penetrated into almost all fields of identifying protein attributes, such as predicting protein secondary structure content [22], predicting super-secondary structure [23], predicting protein structural classes [24, 25], identifying protein quaternary structure attribute [26], identifying enzyme family and sub-family classes [27, 28], identifying protein subcellular localizations [29,30], identifying subcellular localization of apoptosis proteins [31-33], predicting protein subnuclear localizations [34], identifying protein sub-mitochondria locations [35-37], identifying cell wall lytic enzymes [38], identifying risk type of human papillomaviruses [39], identifying DNA-binding proteins [40], identifying G-Protein-Coupled Receptor Classes [41,42], predicting protein folding rates [43], identifying outer membrane proteins [44], identifying cyclin proteins [45], identifying GABA(A) receptor proteins [46], identifying bacterial secreted proteins [47], identifying the cofactors of oxidoreductases [48], identifying lipase types [49], classifying amino acids [50], identifying metalloproteinase family [51], among many others. For a summary about the development of various different modes of PseAAC and their applications, see a recent comprehensive review [52].

For a brief introduction about Chou's "pseudo amino acid composition", see a Wikipedia article at http://en.wikipedia.org/wiki/Pseudo_amino_acid_composition.

4. WEB-SERVER PREDICTORS

For helping experimental scientists to acquire useful information and data, Prof. Chou and his co-workers have established a series of powerful web-servers, such as Cell-PLoc [53] for identifying protein subcellular localization in various organisms, MemType-2L [54] for membrane proteins and their types, EzyPred [55] for enzymes and their family classes, ProtIdent [56] for proteases and their types, GPCR-2L [57] for GPCR and their types, iDNA-Prot [58] for DNA-binding proteins, NR-2L [59] for nuclear receptors and their types, iLoc-Euk [60] for subcellular localization of eukaryotic proteins, iLoc-Hum [61] for subcellular localization of human proteins, Plant-mPLoc [62] for subcellular localiza-

tion of plant proteins, Signal-CF [63] for predicting protein peptides, as well as a series of web-server predictors listed in Table 3 of [64].

5. LOW-FREQUENCY INTERNAL MOTION IN BIOMOLECULES AND ITS BIOLOGICAL FUNCTIONS

The concept of low-frequency phonons (or internal motion) in proteins was originally proposed in order to solve a perplexing "free-energy deficit" problem [65], which was encountered in studying the binding interaction between insulin and insulin receptor [66]. According to the inference elaborated in [65], the wave numbers of the low-frequency phonons were in the range of 10 - 100 cm^{-1}, corresponding to a frequency in the tera Hz range (3×10^{11} - 3×10^{12} Hz). In the mean time, the possible biological functions of low-frequency phonons in proteins were also discussed [65].

Subsequently, the aforementioned low-frequency modes as predicted by Professor Chou have been indeed observed later by Raman spectroscopy for a number of protein molecules [67,68] and different types of DNA [69-72]. These results have also been further confirmed by neutron scattering experiments [73].

To identify and analyze this kind of low-frequency motions in protein and DNA molecules, the quasi-continuum model was developed by Prof. Chou and his coworkers [74-78]. It has been successfully used to simulate various low-frequency collective motions in protein and DNA molecules, such as accordion-like motion, pulsation or breathing motion, as reflected by the fact that the low-frequency wave numbers thus derived were quite close to the experimental observations [74,76-78]. It was also revealed through the Chou's quasi-continuum model that the low-frequency motions in biomacromolecules originate from their two common and intrinsic characteristics; *i.e.*, they usually contain 1) a series of weak bonds, such as hydrogen bonds, and 2) a substantial mass distributed over the region of these weak bonds [79].

The most interesting is that many marvelous biological functions and their profound dynamic mechanisms, such as cooperative effects [80,81], allosteric transition [82, 83], and intercalation of drugs into DNA [84,85], can be revealed through the low-frequency collective motion or resonance in protein and DNA molecules. In this regard, some phenomenological theories [82,83,85,86] were established. Meanwhile, the solitary wave motion was also used to address the internal motion during microtubule growth [87]. A soliton is a self-reinforcing solitary wave (a wave packet or pulse) that maintains its shape while it travels at constant speed. The relationship between the solitons and the low-frequency phonons in proteins have been discussed in a recent paper [88].

Furthermore, the low-frequency internal motions in proteins as originally inferred by Professor Chou in 1977 [65] have also been clearly observed in 2001 by NMR [89], and applied in medical treatments [90-92].

As stated on the web-page of Vermont Photonics Technologies Corp., "Study of low-frequency (or Terahertz frequency) motions in biomacromolecules holds a very exciting potential that could lead to revolutionize biophysics, molecular biology, and biomedicine".

For a systematic introduction of the low-frequency collective motion in biomacromolecules and its biological functions, refer to a comprehensive review article [93]. For a brief introduction in this regard, see a Wikipedia article at http://en.wikipedia.org/wiki/Low-frequency_Collective_Motion.

6. CHOU'S INVARIENCE THEOREM

In developing methodology for protein attribute prediction, the dimension-reduced operation is often needed when calculating the covariant discriminant [21] between two feature-vectors with certain number of normalized components in order to avoid the divergence problem. However, which one of these normalized components should be removed? Will the result be different by removing a different component? To address these problems, the Chou's Invariance Theorem was developed in 1995.

According to the Chou's Invariance Theorem, the outcome of the covariant discriminant will remain exactly the same regardless of which one of the components is left out. In other words, any one of the constituent normalized components can be left out to overcome the divergence problem without changing the final result.

For more information about Chou's Invariance Theorem and its applications as well as the relevant references, see a Wikipedia article at http://en.wikipedia.org/wiki/Chou's_invariance_theorem.

Finally, for more information about Professor Chou's publications, visit his Research ID at http://www.researcherid.com/rid/A-8340-2009.

7. IDENTITY OF TRNAS IN AMINOACYLATION SPECIFICITY AND THE CORRECT EXPRESSION OF GENETIC CODE DEMONSTRATED BY PROF. GIEGE

Prof. Giege has devoted most of his efforts to study the interactions between two important biomacromolecules, RNA and proteins, taking the transfer RNA (tRNA) and aminoacyl-tRNA synthetases (aaRS) as examples. Using a multiple physico-biochemical method, he demonstrated the concept of kinetic specificity for the aminoacylation of tRNAs [94]. This was continued by extensive studies

on the characterization and properties of the determinants accounting for the identity of several tRNAs and the discovery of anti-determinants that are chemical negative marks in a tRNA preventing false aminoacylations. His laboratory published the first crystallization of a tRNA/aaRS complex, that between $tRNA^{Asp}$ and AspRS from yeast [95] in collaboration with his colleagues. He found the identity elements for yeast $tRNA^{Asp}$ charging aminoacylation [96], and engineered the structure and function of yeast $tRNA^{Asp}$ [97]. Together with Paul Schimmel, Dino Moras and Shigeyuki Yokoyama, he proposed the model of an operational RNA code for amino acids and its possible relationship to the genetic code [98]. During the period of intensive search for the determinants on the tRNA structure for their recognition by their cognate aaRS, he discovered that faithful aminoacylation of tRNAs also relied on anti-determinants [99]. In 1998, from a survey of the large amount of information gathered at that time on tRNA/aaRS recognition in various organisms, he proposed [100] universal rules, and also idiosyncrasies that distinguish individual or groups of tRNA identities (including determinants, anti-determinants and tRNA architecture). These rules allowed the manipulation of identity elements and the engineering of tRNAs with altered specificities. He and his colleagues have rationalized the conserved and variable structural features in the tRNA World, with emphasis to structural plasticity [101]. Analysis of the new structures in relation with functional data obtained in his department gave a refined understanding on aminoacylation identity and on phylogenetic structural and enzymologic species differences. From the viewpoint of theoretical biology, the structure/function understanding of tRNA and synthetase molecules allowed him to design experiments in order to verify conjectures on tRNA architecture and specificity and to predict their results. This was successful and, amongst others, functional tRNAs with triple aminoacid specificity could be produced as well as a minimalist RNA construct consisting solely of a small circular RNA hybridized with a short RNA oligomer that is chargeable with histidine by a histidyl-tRNA synthetase following canonical tRNA identity rules.

Some special functions of tRNA were also demonstrated, such as the organization of the tRNA-like domain of TYMV RNAs [102].

8. MACROMOLECULAR CRYSTALLOGENESIS

In the earlier years of Prof. Giege's career, he progressively created a research devoted to investigate the structure/function relationships in tRNA molecules, with the support of Prof. J. P. Ebel. The first bottleneck he and his colleagues should overcome, was the crystallization of such multi-domain proteins. In the case of aspatyl-tRNA synthetase/tRNA complex, it took several years to get high resolution crystals. Through this fine study, they have shown the importance of macromolecular quality for their successful crystallization. Their detailed study demonstrated that biocrystallization is a multi-parametric process including intrinsic physico-chemical, biochemical, biophysical, and biological parameters, the purity of macromolecules the conformational homogenity [103]. This was why he devoted much energy to the new field of biocrystallogenesis whose major aim is to provide conceptual means and practical tools to the biologists enabling them to overcome the delicate bottleneck of crystallization. Here, his contributions were important, notably the data he gained on the role of physical-chemical parameters (temperature, pH, pressure, gravity) on protein solubility, pre-nucleation and nucleation and crystal growth, on the diagnostic of crystallization by Dynamic Light Scattering, the development of a microfluidic crystallization system operating by counter-diffusion with possibility of in situ crystal diffraction analysis, and the direct visualization by X-ray topography and AFM of perturbations in crystal perfection and crystal growth, not only on model proteins but also more essential on the synthetases and tRNAs he was studying. As a consequence several difficult crystallization problems with synthetases could be overcome and structures solved or improved, in particular of cytosolic aspartyl-tRNA synthetases and the first one of a human mitochondrial synthetase, that specific for tyrosine. Many foreign senior scientists were visiting scientists in Prof. Giege's laboratory and about 50 researchers prepared a PhD under his supervision or were postdoctoral fellows with him, and many of them became distinguished scientists in Academia or in Industry.

His well-known book "Crystallization of Nucleic Acids and Proteins" [104] is extensively used in the courses and workshops for macromolecular crystallization all over the world.

9. PHYSICO-BIOCHEMISTRY IS ALWAYS APPLIED IN THE RESEARCH BY PROF. GIEGE

We always remember his great interests to use various physico-chemical methods to study protein-nucleic acid interaction, from the Raman spectroscopy, X-ray crystallography, Dynamic Light Scattering and many others. He initiated a research department devoted to investigate the structure/function relationship in tRNA molecules, with various physico-chemical means. Using Raman spectroscopy he provided insights on the structure and dynamics of tRNA, notably base-stacking effects for transfer of information. He developed chemical methods and used Small Angle Neutron Scattering (SANS) for studying the static and dynamic solution structure of

tRNAs and their complexes with aminoacyl-tRNA synthetases. In this regard his demonstration via SANS on the stability of tRNA/synthetase complexes in very high concentrations of ammonium sulfate was the clue to crystallize for the first time a tRNA/synthetase complex and opened new routes in the structural biology of protein/nucleic acid complexes.

10. ACKNOWLEDGEMENTS

The work was supported by a Canadian Institutes of Health Research grant (FRN: MOP89851) to S.-X. Lin, and by a Natural Sciences and Engineering Research Council of Canada (9597-2010) to J. Lapointe..

REFERENCES

[1] Eigen, M. and Hammes, G.G. (1963) Elementary steps in enzyme reactions (as studies by relaxation spectrometry). *Advances in Enzymology and Related Subjects of Biochemistry*, **25**, 1-38.

[2] Koening, S.H. and Brown, R.D. (1972) H_2CO_3 as substrate for carbonic anhydrase in the dehydration of H_2CO_3. *Proceedings of the National Academy of Sciences of USA*, **69**, 2422-2425. doi:10.1073/pnas.69.9.2422

[3] Chou, K.C. (1976) The kinetics of the combination reaction between en zyme and substrate. *Scientia Sinica*, **19**, 505-528.

[4] Chou, K.C. and Jiang, S.P. (1974) Studies on the rate of diffusion-controlled reactions of enzymes. *Scientia Sinica*, **17**, 664-680.

[5] Chou, K.C. and Zhou, G.P. (1982) Role of the protein outside active site on the diffusion-controlled reaction of enzyme. *Journal of American Chemical Society*, **104**, 1409-1413. doi:10.1021/ja00369a043

[6] Chou, K.C. (1989) Graphic rules in steady and non-steady enzyme kinetics. *Journal of Biological Chemistry*, **264**, 12074-12079.

[7] Chou, K.C. (1990) Review: Applications of graph theory to enzyme kinetics and protein folding kinetics. Steady and non-steady state systems. *Biophysical Chemistry*, **35**, 1-24. doi:10.1016/0301-4622(90)80056-D

[8] Chou, K.C., Jiang, S.P., Liu, W.M. and Fee, C.H. (1979) Graph theory of enzyme kinetics: 1. Steady-state reaction system. *Scientia Sinica*, **22**, 341-358.

[9] Lin, S.X. and Neet, K.E. (1990) Demonstration of a slow conformational change in liver glucokinase by fluorescence spectroscopy. *Journal of Biological Chemistry*, **265**, 9670-9675.

[10] Zhou, G.P. and Deng, M.H. (1984) An extension of Chou's graphic rules for deriving enzyme kinetic equations to systems involving parallel reaction pathways. *Biochemical Journal*, **222**, 169-176.

[11] Kuzmic, P., Ng, K.Y. and Heath, T.D. (1992) Mixtures of tight-binding enzyme inhibitors. Kinetic analysis by a recursive rate equation. *Analytical Biochemistry*, **200**, 68-73. doi:10.1016/0003-2697(92)90278-F

[12] Myers, D. and Palmer, G. (1985) Microcomputer tools for steady-state enzyme kinetics. *Bioinformatics*, **1**, 105-110.

[13] Althaus, I.W., Reusser, F., *et al*. (1993) Steady-state kinetic studies with the non-nucleoside HIV-1 reverse transcriptase inhibitor U-87201E. *Journal of Biological Chemistry*, **268**, 6119-6124.

[14] Althaus, I.W., Reusser, F., *et al*. (1993) The quinoline U-78036 is a potent inhibitor of HIV-1 reverse transcriptase. *Journal of Biological Chemistry*, **268**, 14875-14880.

[15] Althaus, I.W., Reusser, F., *et al*. (1993) Kinetic studies with the nonnucleoside HIV-1 reverse transcriptase inhibitor U-88204E. *Biochemistry*, **32**, 6548-6554. doi:10.1021/bi00077a008

[16] Andraos, J. (2008) Kinetic plasticity and the determination of product ratios for kinetic schemes leading to multiple products without rate laws: New methods based on directed graphs. *Canadian Journal of Chemistry*, **86**, 342-357. doi:10.1139/v08-020

[17] Chou, K.C., Kezdy, F.J. and Reusser, F. (1994) Review: Steady-state inhibition kinetics of processive nucleic acid polymerases and nucleases. *Analytical Biochemistry*, **221**, 217-230. doi:10.1006/abio.1994.1405

[18] Shen, H.B., Song, J.N. and Chou, K.C. (2009) Prediction of protein folding rates from primary sequence by fusing multiple sequential features. *Journal of Biomedical Science and Engineering*, **2**, 136-143. doi:10.4236/jbise.2009.23024

[19] Chou, K.C. (2010) Graphic rule for drug metabolism systems. *Current Drug Metabolism*, **11**, 369-378. doi:10.2174/138920010791514261

[20] Aka, J.A., Mazumdar, M., Chen, C.Q., Poirier, D. and Lin, S.X. (2010) 17beta-hydroxysteroid dehydrogenase type 1 stimulates breast cancer by dihydrotestosterone inactivation in addition to estradiol production. *Molecular Endocrinology*, **24**, 832-845. doi:10.1210/me.2009-0468

[21] Chou, K.C. (2001) Prediction of protein cellular attributes using pseudo amino acid composition. *Proteins*: *Structure*, *Function*, *and Genetics*, **43**, 246-255.

[22] Chen, C., Chen, L., Zou, X. and Cai, P. (2009) Prediction of protein secondary structure content by using the concept of Chou's pseudo amino acid composition and support vector machine. *Protein & Peptide Letters*, **16**, 27-31. doi:10.2174/092986609787049420

[23] Zou, D., He, Z., He, J. and Xia, Y. (2011) Supersecondary structure prediction using Chou's pseudo amino acid composition. *Journal of Computational Chemistry*, **32**, 271-278. doi:10.1002/jcc.21616

[24] Li, Z.C., Zhou, X.B., Dai, Z. and Zou, X.Y. (2009) Prediction of protein structural classes by Chou's pseudo amino acid composition: Approached using continuous wavelet transform and principal component analysis. *Amino Acids*, **37**, 415-425. doi:10.1007/s00726-008-0170-2

[25] Sahu, S.S. and Panda, G. (2010) A novel feature representation method based on Chou's pseudo amino acid composition for protein structural class prediction. *Computational Biology and Chemistry*, **34**, 320-327.

doi:10.1016/j.compbiolchem.2010.09.002

[26] Zhang, S.W., Chen, W., Yang, F. and Pan, Q. (2008) Using Chou's pseudo amino acid composition to predict protein quaternary structure: A sequence-segmented Pse AAC approach. *Amino Acids*, **35**, 591-598. doi:10.1007/s00726-008-0086-x

[27] Qiu, J.D., Huang, J.H., Shi, S.P. and Liang, R.P. (2010) Using the concept of Chou's pseudo amino acid composition to predict enzyme family classes: An approach with support vector machine based on discrete wavelet transform. *Protein & Peptide Letters*, **17**, 715-722. doi:10.2174/092986610791190372

[28] Zhou, X.B., Chen, C., Li, Z.C. and Zou, X.Y. (2007) Using Chou's amphiphilic pseudo-amino acid composition and support vector machine for prediction of enzyme subfamily classes. *Journal of Theoretical Biology*, **248**, 546-551. doi:10.1016/j.jtbi.2007.06.001

[29] Li, F.M. and Li, Q.Z. (2008) Predicting protein subcellular location using Chou's pseudo amino acid composition and improved hybrid approach. *Protein & Peptide Letters*, **15**, 612-616. doi:10.2174/092986608784966930

[30] Zhang, S.W., Zhang, Y.L., Yang, H.F., Zhao, C.H. and Pan, Q. (2008) Using the concept of Chou's pseudo amino acid composition to predict protein subcellular localization: An approach by incorporating evolutionary information and von Neumann entropies. *Amino Acids*, **34**, 565-572. doi:10.1007/s00726-007-0010-9

[31] Ding, Y.S. and Zhang, T.L. (2008) Using Chou's pseudo amino acid composition to predict subcellular localization of apoptosis proteins: An approach with immune genetic algorithm-based ensemble classifier. *Pattern Recognition Letters*, **29**, 1887-1892. doi:10.1016/j.patrec.2008.06.007

[32] Jiang, X., Wei, R., Zhang, T.L. and Gu, Q. (2008) Using the concept of Chou's pseudo amino acid composition to predict apoptosis proteins subcellular location: An approach by approximate entropy. *Protein & Peptide Letters*, **15**, 392-396. doi:10.2174/092986608784246443

[33] Lin, H., Wang, H., Ding, H., Chen, Y.L. and Li, Q.Z. (2009) Prediction of subcellular localization of apoptosis protein using Chou's pseudo amino acid composition. *Acta Biotheoretica*, **57**, 321-330. doi:10.1007/s10441-008-9067-4

[34] Jiang, X., Wei, R., Zhao, Y. and Zhang, T. (2008) Using Chou's pseudo amino acid composition based on approximate entropy and an ensemble of AdaBoost classifiers to predict protein subnuclear location. *Amino Acids*, **34**, 669-675. doi:10.1007/s00726-008-0034-9

[35] Lin, H. (2008) The modified Mahalanobis discriminant for predicting outer membrane proteins by using Chou's pseudo amino acid composition. *Journal of Theoretical Biology*, **252**, 350-356. doi:10.1016/j.jtbi.2008.02.004

[36] Nanni, L. and Lumini, A. (2008) Genetic programming for creating Chou's pseudo amino acid based features for submitochondria localization. *Amino Acids*, **34**, 653-660. doi:10.1007/s00726-007-0018-1

[37] Zeng, Y.H., Li, M.L., *et al.* (2009) Using the augmented Chou's pseudo amino acid composition for predicting protein submitochondria locations based on auto covariance approach. *Journal of Theoretical Biology*, **259**, 366-372. doi:10.1016/j.jtbi.2009.03.028

[38] Ding, H., Luo, L. and Lin, H. (2009) Prediction of cell wall lytic enzymes using Chou's amphiphilic pseudo amino acid composition. *Protein & Peptide Letters*, **16**, 351-355. doi:10.2174/092986609787848045

[39] Esmaeili, M., Mohabatkar, H. and Mohsenzadeh, S. (2010) Using the concept of Chou's pseudo amino acid composition for risk type prediction of human papillomaviruses. *Journal of Theoretical Biology*, **263**, 203-209. doi:10.1016/j.jtbi.2009.11.016

[40] Fang, Y., Guo, Y., Feng, Y. and Li, M. (2008) Predicting DNA-binding proteins: Approached from Chou's pseudo amino acid composition and other specific sequence features. *Amino Acids*, **34**, 103-109. doi:10.1007/s00726-007-0568-2

[41] Gu, Q., Ding, Y.S. and Zhang, T.L. (2010) Prediction of G-protein-coupled receptor classes in low homology using Chou's pseudo amino acid composition with approximate entropy and hydrophobicity patterns. *Protein & Peptide Letters*, **17**, 559-567. doi:10.2174/092986610791112693

[42] Qiu, J.D., Huang, J.H., Liang, R.P. and Lu, X.Q. (2009) Prediction of G-protein-coupled receptor classes based on the concept of Chou's pseudo amino acid composition: An approach from discrete wavelet transform. *Analytical Biochemistry*, **390**, 68-73. doi:10.1016/j.ab.2009.04.009

[43] Guo, J., Rao, N., Liu, G., Yang, Y. and Wang, G. (2011) Predicting protein folding rates using the concept of Chou's pseudo amino acid composition. *Journal of Computational Chemistry*, **32**, 1612-1617. doi:10.1002/jcc.21740

[44] Lin, H., Ding, H., Guo, F.-B., Zhang, A.Y. and Huang, J. (2008) Predicting subcellular localization of mycobacterial proteins by using Chou's pseudo amino acid composition. *Protein & Peptide Letters*, **15**, 739-744. doi:10.2174/092986608785133681

[45] Mohabatkar, H. (2010) Prediction of cyclin proteins using Chou's pseudo amino acid composition. *Protein & Peptide Letters*, **17**, 1207-1214. doi:10.2174/092986610792231564

[46] Mohabatkar, H., Mohammad Beigi, M. and Esmaeili, A. (2011) Prediction of GABA(A) receptor proteins using the concept of Chou's pseudo-amino acid composition and support vector machine. *Journal of Theoretical Biology*, **281**, 18-23. doi:10.1016/j.jtbi.2011.04.017

[47] Yu, L., Qin, W., *et al.* (2010) SecretP: Identifying bacterial secreted proteins by fusing new features into Chou's pseudo-amino acid composition. *Journal of Theoretical Biology*, **267**, 1-6. doi:10.1016/j.jtbi.2010.08.001

[48] Zhang, G.Y. and Fang, B.S. (2008) Predicting the cofactors of oxidoreductases based on amino acid composition distribution and Chou's amphiphilic pseudo amino acid composition. *Journal of Theoretical Biology*, **253**, 310-315. doi:10.1016/j.jtbi.2008.03.015

[49] Zhang, G.Y., Li, H.C., Gao, J.Q. and Fang, B.S. (2008) Predicting lipase types by improved Chou's pseudo-amino acid composition. *Protein & Peptide Letters*, **15**, 1132-1137. doi:10.2174/092986608786071184

[50] Georgiou, D.N., Karakasidis, T.E., Nieto, J.J. and Torres,

A. (2009) Use of fuzzy clustering technique and matrices to classify amino acids and its impact to Chou's pseudo amino acid composition. *Journal of Theoretical Biology*, **257**, 17-26. doi:10.1016/j.jtbi.2008.11.003

[51] Mohammad Beigi, M., Behjati, M. and Mohabatkar, H. (2011) Prediction of metalloproteinase family based on the concept of Chou's pseudo amino acid composition using a machine learning approach. *Journal of Structural and Functional Genomics*, **12**, 191-197 doi:10.1007/s10969-011-9120-4

[52] Chou, K.C. (2011) Some remarks on protein attribute prediction and pseudo amino acid composition (50th anniversary year review). *Journal of Theoretical Biology*, **273**, 236-247. doi:10.1016/j.jtbi.2010.12.024

[53] Chou, K.C. and Shen, H.B. (2008) Cell-PLoc: A package of Web servers for predicting subcellular localization of proteins in various organisms. *Nature Protocols*, **3**, 153-162. doi:10.1038/nprot.2007.494

[54] Chou, K.C. and Shen, H.B. (2007) MemType-2L: A web server for predicting membrane proteins and their types by incorporating evolution information through Pse-PSSM. *Biochemical and Biophysical Research Communications*, **360**, 339-345. doi:10.1016/j.bbrc.2007.06.027

[55] Shen, H.B. and Chou, K.C. (2007) EzyPred: A top-down approach for predicting enzyme functional classes and subclasses. *Biochemical and Biophysical Research Communications*, **364**, 53-59. doi:10.1016/j.bbrc.2007.09.098

[56] Chou, K.C. and Shen, H.B. (2008) ProtIdent: A web server for identifying proteases and their types by fusing functional domain and sequential evolution information. *Biochemical and Biophysical Research Communications*, **376**, 321-325. doi:10.1016/j.bbrc.2008.08.125

[57] Xiao, X., Wang, P. and Chou, K.C. (2011) GPCR-2L: Predicting G protein-coupled receptors and their types by hybridizing two different modes of pseudo amino acid compositions. *Molecular Biosystems*, **7**, 911-919. doi:10.1039/c0mb00170h

[58] Lin, W.Z., Fang, J.A., Xiao, X. and Chou, K.C. (2011) iDNA-prot: Identification of DNA binding proteins using random forest with grey model. *PLoS One*, **6**, e24756. doi:10.1371/journal.pone.0024756

[59] Wang, P., Xiao, X. and Chou, K.C. (2011) NR-2L: A two-level predictor for identifying nuclear receptor subfamilies based on sequence-derived features. *PLoS One*, **6**, e23505. doi:10.1371/journal.pone.0023505

[60] Chou, K.C., Wu, Z.C. and Xiao, X. (2011) iLoc-Euk: A multi-label classifier for predicting the subcellular localization of singleplex and multiplex eukaryotic proteins. *PLoS One*, **6**, e18258. doi:10.1371/journal.pone.0018258

[61] Chou, K.C., Wu, Z.C. and Xiao, X. (2012) iLoc-Hum: Using accumulation-label scale to predict subcellular locations of human proteins with both single and multiple sites. *Molecular Biosystems*, **8**, 629-641. doi:10.1039/c1mb05420a

[62] Chou, K.C. and Shen, H.B. (2010) Plant-mPLoc: A top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS One*, **5**, e11335.

[63] Chou, K.C. and Shen, H.B. (2007) Signal-CF: A subsite-coupled and window-fusing approach for predicting signal peptides. *Biochemical and Biophysical Research Communications*, **357**, 633-640. doi:10.1016/j.bbrc.2007.03.162

[64] Chou, K.C. and Shen, H.B. (2009) Review: Recent advances in developing web-servers for predicting protein attributes. *Natural Science*, **2**, 63-92.

[65] Chou, K.C. and Chen, N.Y. (1977) The biological functions of low-frequency phonons. *Scientia Sinica*, **20**, 447-457.

[66] Chothia, C. and Janin, J. (1975) Principles of protein-protein recognition. *Nature*, **256**, 705-708. doi:10.1038/256705a0

[67] Painter, P.C. and Mosher, L.E. (1979) The low-frequency Raman spectrum of an antibody molecule: Bovine IgG. *Biopolymers*, **18**, 3121-3123. doi:10.1002/bip.1979.360181217

[68] Painter, P.C., Mosher, L.E. and Rhoads, C. (1982) Low-frequency modes in the Raman spectra of proteins. *Biopolymers*, **21**, 1469-1472. doi:10.1002/bip.360210715

[69] Painter, P.C., Mosher, L.E. and Rhoads, C. (1981) Low-frequency modes in the Raman spectrum of DNA. *Biopolymers*, **20**, 243-247. doi:10.1002/bip.1981.360200119

[70] Urabe, H. and Tominaga, Y. (1982) Low-frequency collective modes of DNA double helix by Raman spectroscopy. *Biopolymers*, **21**, 2477-2481. doi:10.1002/bip.360211212

[71] Urabe, H., Tominaga, Y. and Kubota, K. (1983) Experimental evidence of collective vibrations in DNA double helix Raman spectroscopy. *Journal of Chemical Physics*, **78**, 5937-5939. doi:10.1063/1.444600

[72] Urabe, H., Sugawara, Y., Ataka, M. and Rupprecht, A. (1998) Low-frequency Raman spectra of lysozyme crystals and oriented DNA films: Dynamics of crystal water. *Biophysical Journal*, **74**, 1533-1540. doi:10.1016/S0006-3495(98)77865-8

[73] Martel, P. (1992) Biophysical aspects of neutron scattering from vibrational modes of proteins. *Progress in Biophysics and Molecular Biology*, **57**, 129-179. doi:10.1016/0079-6107(92)90023-Y

[74] Chou, K.C. (1983) Low-frequency vibrations of helical structures in protein molecules. *Biophysical Journal*, **209**, 573-580.

[75] Chou, K.C. (1984) The biological functions of low-frequency phonons: 3. Helical structures and microenvironment. *Biophysical Journal*, **45**, 881-890. doi:10.1016/S0006-3495(84)84234-4

[76] Chou, K.C. (1984) Low-frequency vibration of DNA molecules. *Biophysical Journal*, **221**, 27-31.

[77] Chou, K.C. (1985) Low-frequency motions in protein molecules: Beta-sheet and beta-barrel. *Biophysical Journal*, **48**, 289-297. doi:10.1016/S0006-3495(85)83782-6

[78] Chou, K.C., Maggiora, G.M. and Mao, B. (1989) Quasi-continuum models of twist-like and accordion-like low-frequency motions in DNA. *Biophysical Journal*, **56**, 295-305. doi:10.1016/S0006-3495(89)82676-1

[79] Chou, K.C. (1986) Origin of low-frequency motion in

biological macromolecules: A view of recent progress of quasi-continuity model. *Biophysical Chemistry*, **25**, 105-116. doi:10.1016/0301-4622(86)87001-6

[80] Chou, K.C. (1989) Low-frequency resonance and cooperativity of hemoglobin. *Trends in Biochemical Sciences*, **14**, 212. doi:10.1016/0968-0004(89)90026-1

[81] Chou, K.C., Chen, N.Y. and Forsen, S. (1981) The biological functions of low-frequency phonons: 2. Cooperative effects. *Chemica Scripta*, **18**, 126-132.

[82] Chou, K.C. (1984) The biological functions of low-frequency phonons: 4. Resonance effects and allosteric transition. *Biophysical Chemistry*, **20**, 61-71. doi:10.1016/0301-4622(84)80005-8

[83] Chou, K.C. (1987) The biological functions of low-frequency phonons: 6. A possible dynamic mechanism of allosteric transition in antibody molecules. *Biopolymers*, **26**, 285-295. doi:10.1002/bip.360260209

[84] Chou, K.C. and Maggiora, G.M. (1988) The biological functions of low-frequency phonons: 7. The impetus for DNA to accommodate intercalators. *British Polymer Journal*, **20**, 143-148. doi:10.1002/pi.4980200209

[85] Chou, K.C. and Mao, B. (1988) Collective motion in DNA and its role in drug intercalation. *Biopolymers*, **27**, 1795-1815. doi:10.1002/bip.360271109

[86] Chou, K.C. and Kiang, Y.S. (1985) The biological functions of low-frequency phonons: 5. A phenomenological theory. *Biophysical Chemistry*, **22**, 219-235. doi:10.1016/0301-4622(85)80045-4

[87] Chou, K.C., Zhang, C.T. and Maggiora, G.M. (1994) Solitary wave dynamics as a mechanism for explaining the internal motion during microtubule growth. *Biopolymers*, **34**, 143-153. doi:10.1002/bip.360340114

[88] Sinkala, Z. (2006) Soliton/exciton transport in proteins. *Journal of Theoretical Biology*, **241**, 919-927. doi:10.1016/j.jtbi.2006.01.028

[89] Chou, J.J., Li, S., Klee, C.B. and Bax, A. (2001) Solution structure of Ca^{2+}-calmodulin reveals flexible hand-like properties of its domains. *Nature Structural Biology*, **8**, 990-997. doi:10.1038/nsb1101-990

[90] Gordon, G. (2007) Designed electromagnetic pulsed therapy: Clinical applications. *Journal of Cellular Physiology*, **212**, 579-582. doi:10.1002/jcp.21025

[91] Gordon, G. (2008) Extrinsic electromagnetic fields, low frequency (phonon) vibrations, and control of cell function: A non-linear resonance system. *Journal of Biomedical Science and Engineering*, **1**, 152-156. doi:10.4236/jbise.2008.13025

[92] Madkan, A., Goodman, R., *et al*. (2009) Steps to the clinic with ELF EMF. *Natural Science*, **1**, 157-165.

[93] Chou, K.C. (1988) Review: Low-frequency collective motion in biomacromolecules and its biological functions. *Biophysical Chemistry*, **30**, 3-48. doi:10.1016/0301-4622(88)85002-6

[94] Ebel, J.P., Dirheimer, G., *et al*. (1973) Concept of kinetic specificity for the aminoacylation of tRNAs. *Biochimie*, **55**, 547-557. doi:10.1016/S0300-9084(73)80415-8

[95] Giegé, Lorber, Ebel, Moras and Thierry (1980) First crystallization of tRNA/aaRS complex, that between tRNA Asp and AspRS from yeast. *C R Acad Sci Paris D*2, **291**, 393-396.

[96] Pütz, J., Puglisi, J.D., Florentz, C. and Giegé, R. (1991) Identity elements for yeast tRNAAsp charging. *Science*, **252**, 1696-1699. doi:10.1126/science.2047878

[97] Pütz, J., Puglisi, J.D., Florentz, C. and Giegé, R. (1993) Engineering the structure and function of yeast tRNAAsp. *Nucleic Acids Research*, **21**, 41-49.

[98] Schimmel, P., Giegé, R., Moras, D. and Yokoyama, S. (1993) Operational RNA code for amino acids and possible relationship to genetic code: An HFSP collaboration. *Proceedings of the National Academy of Sciences*, **90**, 8763-8768. doi:10.1073/pnas.90.19.8763

[99] Frugier, M., Söll, D., Giegé, R. and Florentz, C. (1994) Identity switches between tRNAs aminoacylated by class I glutaminyl- and class II aspartyl-tRNA synthetases. *Biochemistry*, **33**, 9912-9921. doi:10.1021/bi00199a013

[100] Giegé, Sissler and Florentz (1998) Universal rules and idiosyncrasies in tRNA identity. *Nucleic Acids Research*, **26**, 5017-5735.

[101] Giegé, Florentz, *et al*. (2011) A rationalization of conserved and variable structural features in the tRNA World, with emphasis to structural plasticity. *Wiley Interdisciplinary Reviews*: *RNA*, in Press.

[102] Florentz, Giegé, *et al*. (1982) Valylation kinetics and organization of the tRNA-like domain of TYMV RNAs. *The EMBO Journal*, **1**, 269-276.

[103] Giegé, R. and Ducruix, A. (1999) An introduction to the crystallogenesis of biological macromolecules. In: Ducruix, A. and Giegé, R., Eds., *Crystallzation of Nucleic Acids and Proteins*: *A Practical Approach*, Oxford University Press, Oxford, 1-16.

[104] Ducruix, A. and Giegé, R. (1999) Crystallzation of nucleic acids and proteins: A practical approach. Oxford University Press, Oxford.

