

# The Significant and Profound Impacts of Protein Subcellular Localization Prediction

Kuo-Chen Chou

Gordon Life Science Institute, Boston, Massachusetts, USA  
Email: kcchou@gordonlifescience.org, kcchou38@gmail.com

**How to cite this paper:** Chou, K.-C. (2021). The Significant and Profound Impacts of Protein Subcellular Localization Prediction. *Voice of the Publisher*, 7, 1-5. <https://doi.org/10.4236/vp.2021.71001>

**Received:** September 16, 2020

**Accepted:** March 1, 2021

**Published:** March 4, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

In this short review paper, the significant and profound impacts of the protein subcellular prediction have been briefly presented with crystal clear convincingness.

## Keywords

Proteins, Subcellular Locations, Multiple Site Samples, Significant Impacts, Profound Impacts

---

The first paper introducing the protein subcellular location prediction was in 2005 [1]. It has stimulated a series of followed-up publications [2]-[10], particularly for those proteins with multiple location sites [11]-[18], as well as the eight master pieces of papers from the then Chairman of Nobel Prize Committee Sture Forsen [19]-[26].

It is indeed very significant by introducing the intriguing approach to deal with those proteins with multiple location sites, and it is indeed very profound and stimulating in this regard (see, e.g., [27]-[46]).

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Shen, H.B. and Chou, K.C. (2005) Predicting Protein Subnuclear Location with Optimized Evidence-Theoretic K-Nearest Classifier and Pseudo Amino Acid Composition. *Biochemical and Biophysical Research Communications*, **337**, 752-756. <https://doi.org/10.1016/j.bbrc.2005.09.117>
- [2] Chou, K.C. and Shen, H.B. (2006) Predicting Protein Subcellular Location by Fusing

- Multiple Classifiers. *Journal of Cellular Biochemistry*, **99**, 517-527. <https://doi.org/10.1002/jcb.20879>
- [3] Chou, K.C. and Shen, H.B. (2006) Predicting Eukaryotic Protein Subcellular Location by Fusing Optimized Evidence-Theoretic K-Nearest Neighbor Classifiers. *Journal of Proteome Research*, **5**, 1888-1897. <https://doi.org/10.1021/pr060167c>
- [4] Chou, K.C. and Shen, H.B. (2006) Large-Scale Predictions of Gram-Negative Bacterial Protein Subcellular Locations. *Journal of Proteome Research*, **5**, 3420-3428. <https://doi.org/10.1021/pr060404b>
- [5] Chou, K.C. and Shen, H.B. (2007) Large-Scale Plant Protein Subcellular Location Prediction. *Journal of Cellular Biochemistry*, **100**, 665-678. <https://doi.org/10.1002/jcb.21096>
- [6] Chou, K.C. and Shen, H.B. (2007) Euk-mPLoc: A Fusion Classifier for Large-Scale Eukaryotic Protein Subcellular Location Prediction by Incorporating Multiple Sites. *Journal of Proteome Research*, **6**, 1728-1734. <https://doi.org/10.1021/pr060635i>
- [7] Chou, K.C. and Shen, H.B. (2007) Recent Progresses in Protein Subcellular Location Prediction. *Analytical Biochemistry*, **370**, 1-16. <https://doi.org/10.1016/j.ab.2007.07.006>
- [8] Shen, H.B. and Chou, K.C. (2007) Hum-mPLoc: An Ensemble Classifier for Large-Scale Human Protein Subcellular Location Prediction by Incorporating Samples with Multiple Sites. *Biochemical and Biophysical Research Communications*, **355**, 1006-1011. <https://doi.org/10.1016/j.bbrc.2007.02.071>
- [9] Shen, H.B., Yang, J. and Chou, K.C. (2007) Euk-PLoc: An Ensemble Classifier for Large-Scale Eukaryotic Protein Subcellular Location Prediction. *Amino Acids*, **33**, 57-67. <https://doi.org/10.1007/s00726-006-0478-8>
- [10] Shen, H.B. and Chou, K.C. (2010) Virus-mPLoc: A Fusion Classifier for Viral Protein Subcellular Location Prediction by Incorporating Multiple Sites. *Journal of Biomolecular Structure and Dynamics*, **28**, 175-186. <https://doi.org/10.1080/07391102.2010.10507351>
- [11] Shen, H.B., Yang, J., Liu, X.J. and Chou, K.C. (2005) Using Supervised Fuzzy Clustering to Predict Protein Structural Classes. *Biochemical and Biophysical Research Communications*, **334**, 577-581. <https://doi.org/10.1016/j.bbrc.2005.06.128>
- [12] Chou, K.C. and Shen, H.B. (2006) Hum-PLoc: A Novel Ensemble Classifier for Predicting Human Protein Subcellular Localization. *Biochemical and Biophysical Research Communications*, **347**, 150-157. <https://doi.org/10.1016/j.bbrc.2006.06.059>
- [13] Chou, K.C. and Shen, H.B. (2006) Addendum to "Hum-PLoc: A Novel Ensemble Classifier for Predicting Human Protein Subcellular Localization" [Biochem. Biophys. Res. Commun. 347 (2006) 150-157]. *Biochemical and Biophysical Research Communications*, **348**, 1479. <https://doi.org/10.1016/j.bbrc.2006.08.030>
- [14] 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. <https://doi.org/10.1016/j.bbrc.2007.03.162>
- [15] 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. <https://doi.org/10.1016/j.bbrc.2007.06.027>
- [16] Shen, H.B. and Chou, K.C. (2007) Signal-3L: A 3-Layer Approach for Predicting Signal Peptide. *Biochemical and Biophysical Research Communications*, **363**, 297-303. <https://doi.org/10.1016/j.bbrc.2007.08.140>
- [17] 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. <https://doi.org/10.1016/j.bbrc.2007.09.098>
- [18] 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. <https://doi.org/10.1016/j.bbrc.2008.08.125>
- [19] Chou, K.C. and Forsen, S. (1980) Diffusion-Controlled Effects in Reversible Enzymatic Fast Reaction System: Critical Spherical Shell and Proximity Rate Constants. *Biophysical Chemistry*, **12**, 255-263. [https://doi.org/10.1016/0301-4622\(80\)80002-0](https://doi.org/10.1016/0301-4622(80)80002-0)
- [20] Chou, K.C. and Forsen, S. (1980) Graphical Rules for Enzyme-Catalyzed Rate Laws. *Biochemical Journal*, **187**, 829-835. <https://doi.org/10.1042/bj1870829>
- [21] Chou, K.C., Forsen, S. and Zhou, G.Q. (1980) Three Schematic Rules for Deriving Apparent Rate Constants. *Chemica Scripta*, **16**, 109-113.
- [22] Chou, K.C., Li, T.T. and Forsen, S. (1980) The Critical Spherical Shell in Enzymatic Fast Reaction Systems. *Biophysical Chemistry*, **12**, 265-269. [https://doi.org/10.1016/0301-4622\(80\)80003-2](https://doi.org/10.1016/0301-4622(80)80003-2)
- [23] Li, T.T., Chou, K.C. and Forsen, S. (1980) The Flow of Substrate Molecules in Fast Enzyme-Catalyzed Reaction Systems. *Chemica Scripta*, **16**, 192-196.
- [24] Chou, K.C., Carter, R.E. and Forsen, S. (1981) A New Graphical Method for Deriving Rate Equations for Complicated mechanisms. *Chemica Scripta*, **18**, 82-86.
- [25] 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.
- [26] Chou, K.C. and Forsen, S. (1981) Graphical Rules of Steady-State Reaction Systems. *Canadian Journal of Chemistry*, **59**, 737-755. <https://doi.org/10.1139/v81-107>
- [27] Cao, J.Z., Liu, W.Q. and Gu, H. (2012) Predicting Viral Protein Subcellular Localization with Chou's Pseudo Amino Acid Composition and Imbalance-Weighted Multi-Label K-Nearest Neighbor Algorithm. *Protein & Peptide Letters*, **19**, 1163-1169. <https://doi.org/10.2174/092986612803216999>
- [28] Li, L.Q., Zhang, Y., Zou, L.Y., Zhou, Y. and Zheng, X.Q. (2012) Prediction of Protein Subcellular Multi-Localization Based on the General Form of Chou's Pseudo Amino Acid Composition. *Protein & Peptide Letters*, **19**, 375-387. <https://doi.org/10.2174/092986612799789369>
- [29] Mei, S. (2012) Multi-Kernel Transfer Learning Based on Chou's PseAAC Formulation for Protein Mitochondria Localization. *Journal of Theoretical Biology*, **293**, 121-130. <https://doi.org/10.1016/j.jtbi.2011.10.015>
- [30] Mei, S. (2012) Predicting Plant Protein Subcellular Multi-Localization by Chou's PseAAC Formulation Based Multi-Label Homolog Knowledge Transfer Learning. *Journal of Theoretical Biology*, **310**, 80-87. <https://doi.org/10.1016/j.jtbi.2012.06.028>
- [31] Zia-ur-Rehman and Khan, A. (2012) 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**, 890-903.
- [32] Huang, C. and Yuan, J. (2013) 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**, 50-57. <https://doi.org/10.1016/j.biosystems.2013.04.005>
- [33] Huang, C. and Yuan, J.Q. (2013) A Multilabel Model Based on Chou's Pseudo Amino Acid Composition for Identifying Membrane Proteins with Both Single and

- Multiple Functional Types. *The Journal of Membrane Biology*, **246**, 327-334. <https://doi.org/10.1007/s00232-013-9536-9>
- [34] Huang, C. and Yuan, J.Q. (2013) 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**, 205-212. <https://doi.org/10.1016/j.jtbi.2013.06.034>
- [35] Pacharawongsakda, E. and Theeramunkong, T. (2013) Predict Subcellular Locations of Singleplex and Multiplex Proteins by Semi-Supervised Learning and Dimension-Reducing General Mode of Chou's PseAAC. *IEEE Transactions on NanoBioscience*, **12**, 311-320. <https://doi.org/10.1109/TNB.2013.2272014>
- [36] Wang, X., Li, G.Z. and Lu, W.C. (2013) 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**, 309-317.
- [37] Mandal, M., Mukhopadhyay, A. and Maulik, U. (2015) Prediction of Protein Subcellular Localization by Incorporating Multiobjective PSO-Based Feature Subset Selection into the General form of Chou's PseAAC. *Medical & Biological Engineering & Computing*, **53**, 331-344. <https://doi.org/10.1007/s11517-014-1238-7>
- [38] Wang, X., Zhang, W., Zhang, Q. and Li, G.Z. (2015) MultiP-SChlo: Multi-Label Protein Subchloroplast Localization Prediction with Chou's Pseudo Amino Acid Composition and a Novel Multi-Label Classifier. *Bioinformatics*, **31**, 2639-2645. <https://doi.org/10.1093/bioinformatics/btv212>
- [39] Zou, H.L. and Xiao, X. (2016) Predicting the Functional Types of Singleplex and Multiplex Eukaryotic Membrane Proteins via Different Models of Chou's Pseudo Amino Acid Compositions. *The Journal of Membrane Biology*, **249**, 23-29. <https://doi.org/10.1007/s00232-015-9830-9>
- [40] Zou, H.L. and Xiao, X. (2016) Classifying Multifunctional Enzymes by Incorporating Three Different Models into Chou's General Pseudo Amino Acid Composition. *The Journal of Membrane Biology*, **249**, 551-557. <https://doi.org/10.1007/s00232-016-9904-3>
- [41] Qiu, W.R., Zheng, Q.S., Sun, B.Q. and Xiao, X. (2017) Multi-iPPseEvo: A Multi-label Classifier for Identifying Human Phosphorylated Proteins by Incorporating Evolutionary Information into Chou's General PseAAC via Grey System Theory. *Molecular Informatics*, **36**, Article ID: 1600085. <https://doi.org/10.1002/minf.201600085>
- [42] Ahmad, J. and Hayat, M. (2018) MFSC: Multi-Voting Based Feature Selection for Classification of Golgi Proteins by Adopting the General form of Chou's PseAAC Components. *Journal of Theoretical Biology*, **463**, 99-109. <https://doi.org/10.1016/j.jtbi.2018.12.017>
- [43] Javed, F. and Hayat, M. (2018) Predicting Subcellular Localizations of Multi-Label Proteins by Incorporating the Sequence Features into Chou's PseAAC. *Genomics*, **17**, 793-821.
- [44] Ahmad, J. and Hayat, M. (2019) MFSC: Multi-Voting Based Feature Selection for Classification of Golgi Proteins by Adopting the General form of Chou's PseAAC Components. *Journal of Theoretical Biology*, **463**, 99-109. <https://doi.org/10.1016/j.jtbi.2018.12.017>
- [45] Du, X., Diao, Y., Liu, H. and Li, S. (2019) MsDBP: Exploring DNA-Binding Proteins by Integrating Multi-Scale Sequence Information via Chou's 5-Steps Rule. *Journal of Proteome Research*, **18**, 3119-3132.

<https://doi.org/10.1021/acs.jproteome.9b00226>

- [46] Du, L., Meng, Q., Jiang, H. and Li, Y. (2020) Using Evolutionary Information and Multi-Label Linear Discriminant Analysis to Predict the Subcellular Location of Multi-Site Bacterial Proteins via Chou's 5-Steps Rule. *IEEE Access*, **8**, 56452-56461. <https://doi.org/10.1109/ACCESS.2020.2982160>