(Mini-review)

Meanings of the Honor Award for Prof Kimito Funatsu

Manabu Sugimoto,a* Kenji Hori,b**, and Shigehiko Kanaya*c***

a Faculty of Advanced Science and Technology, Graduate School of Science and Technology, Kumamoto University, 2-39-1 Kurokami, Chuo-ku, Kumamoto 860-8555, Japan
b Division of Materials Science, Graduate School of Science and Technology for Innovation, Yamaguchi University, Tokiwadai, Ube 755-8611, Japan
c Division of Information Science & Data Science Center, Graduate school of Science and Technology, Nara Institute of Science and Technology (NAIST), Takayama-cho 8916-5, Ikoma, Nara 630-0101, Japan

(Received November 21, 2019; Accepted November 22, 2019)

Prof. Kimito Funatsu received the Honor Award in Division of Chemoinformatics, the Chemical Society of Japan in the 42th Annual Meeting of Chemoinformatics held on Nov. 28th 2019. The awarding recognizes his significant contributions in the development of the cheminformatics discipline in the world as well as in Japan and his efforts propagate multiple domains, i.e. (i) system development including elucidation of chemical structures and prediction of organic reactions, (ii) QSAR: Quantitative structure activity relationship (QSAR), (iii) quantitative structure property relationship (QSPR), and (iv) international collaborations in chemoinformatics. In the present review, we focus on chemoinformatics in the world as well as in Japan based on “Special issue dedicating to Honor Award: Prof. Kimito Funatsu”, which consists of five invited papers by the world-famous distinguished foreign researchers, and six papers from domestic researchers. Taking these papers into consideration, we try to discuss the meanings of the Honor Award dedicating to Prof. Kimito Funatsu.

Key Words: chemoinformatics, material informatics, QSAR, QSPR, international collaboration

1. Introduction

Prof. Kimito Funatsu received Honor Award in Division of Chemoinformatics, the Chemical Society of Japan in the 42th Annual Meeting of Chemoinformatics held on Nov. 28th 2019. The awarding recognizes his significant contributions in the development of the cheminformatics discipline in the world as well as in Japan and his efforts propagate multiple domains, i.e. (i) system development including elucidation of chemical structures and prediction of organic reactions, (ii) QSAR: Quantitative...
structure activity relationship, (iii) QSAR: quantitative structure property relationship, and (iv) international collaboration in chemoinformatics. For more than 35 years, Prof. Funatsu has made pioneering scientific contributions to the developing field of chemical & material informatics. Prof. Funatsu’s seminal contributions include the conceptualization and implementation of algorithms and expert systems for structure elucidation designated by CHEMICS [1-3] and computer-assisted synthetic design and reaction prediction systems AIPHOS [4] and SOPHIA [5]. The structure elucidation and synthesis design systems developed by Prof. Funatsu have been extensively applied in the chemical industry, demonstrating his dedication to first class, problem-solving science. In recent years, he has increasingly focused on inverse QSAR analysis, including de novo structure generation and on the development of the soft sensor methodology for chemical process control [6] which represents another par excellence example of ground-breaking research with immediate practical and industrial application potential. Given his unique research profile, Prof. Funatsu has been able to secure large amounts of funding from the chemical industries and the CREST Program on Big Data Applications, funded by the Japan Science and Technology Agency. These efforts have left their mark on the scientific and industrial landscape of chemoinformatics in Japan. Prof. Funatsu is among the core of leaders of the broadly-defined chemoinformatics field worldwide.

Prof. Funatsu obtained his Doctoral degree (Dr. Sci.) in physical organic chemistry from Kyushu University (1983) and started research concerning chemoinformatics with Prof. Shinichi Sasaki’s group at Toyohashi University of Technology in 1984. In 2004, he moved to the University of Tokyo to continue research in these areas as a full professor, and there, he expanded into material design and soft sensors for monitoring and controlling chemical plants. In addition to his professorship, he is also the research director of the Data Science Center (http://www-dsc.naist.jp/dsc_en/) at the Nara Institute of Science and Technology (NAIST).

We would like to introduce Prof. Shin-ichi Sasaki because he was a pioneering Japanese chemoinformatician. Prof. Sasaki with Japan Electronic Optical Laboratory (JEOL) was already exploiting computers for their utility in chemical structure elucidation in the 1960s, when most chemists were only dimly aware of the possibilities of computers. Finally, at Toyohashi University of Technology, Prof. Sasaki laid the foundations of CHEMICS. In the special issue for papers dedicated by his students and colleagues to Prof. Sasaki to mark his retirement as president of the Toyohashi University of Technology, Prof. Shinichi Sasaki’s group at Toyohashi University of Technology, Prof. Funatsu described in reflection as follows: “Learn while you are young lest you should be an old fool, for youth is such stuff as will not endure. Don’t make light of even a very short time. While you are appreciating spring flowers around a pond, leaves of a Chinese parasol in the garden began to show signs of autumn stealthily” (ancient Chinese poem). “I have not spent much of my time appreciating spring flowers around a pond, but the progress of time has been so fast that the period of nearly 30 years has passed as if in the twinkling of an eye. Now, I am going to retire from the President of TUT, the Toyohashi University of Technology, my last activity, like a leaf of a Chinese parasol falling down.”

“TUT, the Toyohashi University of Technology, my last activity, like a leaf of a Chinese parasol falling down.”

“I am more than a little proud that I have tackled research challenges which other persons did not try to conduct and have complete some of them, even if they were tiny, and that with this research as a momentum I may have worked as one of the pioneers exploiting the new discipline of computers in chemistry.”

Prof. Funatsu has taken over the research challenges by Prof. Sasaki and extended informatics to diverged fields of chemistry based on precise analysis of individual tasks, which reflected to more than 200 peer-reviewed and well-recognized publications, and a plethora of presentations and conference contributions.

The American Chemical Society’s Division of Chemical Information is pleased to announce that Professor Kimito Funatsu has been selected to receive the 2019 Herman Skolnik Award for his contributions to structure elucidation, de novo structure generation, and applications of cheminformatics methods to drug and materials design, and chemical process control.

“Honor ("Eiyo" in Japanese) Award” is different from “Honorary ("Meiyo") Award”, the former dedicates a researcher contributing remarkably to chemoinformatics with the highest research activity, and the latter generally have been used to reward significant achievements. So we would like to respect Prof. Funatsu, and to hope that he contributes further development of chemoinformatics in the world as well as in Japan.

In the present review, we focus on chemoinformatics in the world as well as Japan based on “Special issue dedicating to Honor Award: Prof. Kimito Funatsu”, which consists of five invited papers from around the world (Sp 1-5 in Table 1) and six papers from domestic researchers (Sp 6-11).

2 Chemoinformatics

Gasteiger defined that chemoinformatics is the application of informatics methods to solve chemical problems [7]. Science can be explained by molecules, i.e. the structure of a molecule is a fundamental factor in determining its physical, chemical, and biological properties. Chemoinformatics utilizes techniques such as graph theory for representing chemical structures,
multivariate statistics for characterizing chemical data, and machine learning for systematic understanding and predicting the properties in molecules. In the present paper, we perform to review chemoinformatics focused on this special issue according to contribution of Prof. Funatsu to chemoinformatics, i.e. (i) system development including elucidation of chemical structures and prediction of organic reactions, (ii) QSAR, (iii) QSPR, and (iv) international collaboration in chemoinformatics.

3 System Development for Elucidation of Chemical Structures and Organic Reactions

Kurt Varmuza of the Viena University of Technology introduced discussion between him and Prof. Funatsu on “Similarity, diversity – chemoinformatics” [Sp 1 in Table 1]. German-Japanese Workshop on Spectral Databases was held in 1991 at the Toyohashi University of Technology (TU) where the CHEMICS system [2] had been developed by Prof. Shin-ichi Sakaki and coworkers, with essential contributions by Prof. Funatsu. In the article, Kurt Varmuza mentioned that “CHEMICS is one of only a few approaches for a systematic and exhaustive chemical structure elucidation based on spectroscopic data and the brutto formula of an unknown organic compound. He received from Prof. Funatsu a computer tape with the CHEMICS software, and combined the results of their mass spectra classifiers with the methods implemented in CHEMICS, and realized the improvement by their mass spectra classification was rather limited [8]. Then he could validate that neural nets applied for variable selection improved the performance of mass spectra substructure classifiers considerably [Sp1].

Johann Gasteiger of University of Erlangen-Nuremberg described an article entitled by “Kimito Funatsu – The driving force of chemoinformatics in Japan” [Sp 2]. In this article, he mentioned that “Prof Funatsu completely redesigned the CHEMICS system for Computer-Assisted Structure Elucidation (CASE) based on spectroscopic data [9]. Gasteiger’s research group had also developed systems for synthesis design and reaction prediction: EROS (Elaboration of Reactions for Organic Synthesis) [10] and WODCA (Workbench for the Organization of Data for Chemical Applications) [11]. Thus Prof. Funatsu and Gasteiger became scientific competitors. However, the competition was always friendly and dominated by mutual respect leading to an exchange of many scientific ideas. This led to make a joint publication on the classification of chemical reactions by artificial neural networks [12].

Kenji Hori and his collaborators of Yamaguchi University [Sp 6] pointed out the computational chemistry is applicable for analyzing synthesis routes [13, 14] created by organic chemists or synthesis route design systems such as AIPHOS and EROS and reaction route prediction toward new compounds should also be challenging theme. To attain this purpose, they have developed quantum mechanical calculation result database (QMRDB) including transition state (TS) information of elementary reactions. Those DBs make it possible to confirm the possibility of proposed synthesis routes toward target compounds.

4 QSAR: Quantitative Structure activity Relationship

Francesca Grisoni and Gisbert Schneider of ETH Zurich described application of recurrent networks (RNNs) in the molecular de novo design [Sp 3] and corroborated the potential of transfer learning in small data using long short-term memory (LSTM) networks [15]. In an example of drug design to peroxisome proliferator-activated receptors (PPAR), after pre-training the RNN with LSTM cells on approximately 500,000 bioactive molecules, it
was fine-tuned on 25 fatty acid mimetics with known agonistic activity [16]. From the resulting fine-tuned model, 1,000 SMILES strings were generated by applying fragment growing form from carboxy group and ranked according to the predicted biological target and similarity to known bioactivities. Predicted top-ranking compounds included reasonable activity on PPAR [17-19]. Miyao, Funatsu et al. suggested a fundamental method for exhaustive structure generation by solving inverse QSAR equations in 2010 [20]. Grisoni and Schneider hope that Funatsu and collaborators play a seminal role in the development of new methods for de novo molecular design [Sp 3].

Jürgen Bajorath of Rheinische Friedrich-Wilhelms-Universität explored polypharmacology and molecular promiscuity [Sp 4]. Polypharmacology is the design of pharmaceutical agents that act on multiple targets or disease pathways. Molecular promiscuity means the ability of a molecule to interact with one or more other classes of molecules, in synergistic or antagonistic ways. In the context of polypharmacology, compound promiscuity is often rationalized in different way that promiscuity is understood as the ability of a small molecule to specifically interact with multiple targets [21]. A total of 112,624 inhibitors of human kinases with well-defined activity data have been assembled by Bajorath and these inhibitors were active against 426 kinases, hence providing 82% coverage of kinome, and formed a total of 234,740 unique compound-kinase interactions [22]. 61% of kinase inhibitors were reported to be active against a single kinase, whereas 39% of the inhibitors correspond to be promiscuous. He emphasized that polypharmacology meets chemical biology and it will be interesting to see how the interface between multi-target ligand design and chemical biology will evolve in the future. He also mentioned “In science, we have the fortunate opportunity to cross cultural barriers with ease, a true privilege, which teaches us to be tolerant, humble, and curious. Occasionally, our scientific interactions turn into friendships!”

Manabu Sugimoto and his collaborators examined properties of chemical compounds based on electronic-structure informatics by developing automatic drawing of orbital correlation diagrams [Sp 7]. This provides quantitative prediction of molecular properties concerning to chemical reactivity. They also mentioned that combined with the conventional descriptors in chemoinformatics, various descriptors in next generation will be derived from computer simulations including electronic-structure and molecular dynamics calculations. The Tomoyuki Miyao and Shigehiko Kanaya groups in NAIST examined relationships of polymer properties and predicting polymer properties based on monomer structure information [Sp 10]. PLSR models made it possible to represent density, glass transition temperature, and dissolution parameters by chemical descriptors obtained from the monomer unit structure information. It should be noted that several physicochemical properties in polymers are reflected by the descriptors of the monomer unit.

Shigehiko Kanaya and his collaborators of NAIST also examined how to select the best regression models among machine learning regression models for estimating natural products were classified by DPClus [26] which makes it possible to detect high-density subgraphs with related natural products in terms of 3D structure similarity. Then activity-natural product relationships can be clarified by assessing diverged activities for natural products belonging to those subgroups. The idea of network biology is also highly related with polypharmacology and molecular promiscuity reported by Bajorath et al [Sp 4]. Thus network representation of relationships between biological molecules and candidate compounds can provide a clue of new drug discovery associated with the fact that many effective drugs act via modulation of multiple proteins rather than single targets [27].

5 QSAR: Quantitative Structure Property Relationship

Advanced materials are essential to economic security and human well-being and industries aimed at addressing challenges in clean energy, national security, and human welfare. Accelerating the pace of discovery and deployment of advanced material systems should be crucial to achieving global competitiveness in the 21st century. The Materials Genome Initiative (MGI) was launched in the U.S. to accelerate the discovery, design, development and deployment of new, advanced materials. Thus material informatics is also important topics in chemoinformatics. An examination of quantitative structure-property relationship is the first step in material informatics.

Manabu Sugimoto and his collaborators also examined properties of chemical compounds based on electronic-structure informatics which provide chemical descriptors of QSAR [Sp 7] and pure understanding of chemical compounds [Sp 8]. Half maximal inhibitory concentration (IC_{50}) values of fatty acid synthase inhibitor can be explained by at least 8 electronic descriptors developed by them [Sp 8]. All the selected descriptors play key roles related with redox reactions or chemical reactions accompanied by state changes.

Md. Altaf-Ul-Amin of NAIST is an expert researcher for network biology [23]. He predicted activities of natural products by repetitive clustering of the structural similarity [Sp 9], where natural products were obtained from KNapsACK Core DB [24], and the structure similarity scores between natural product pairs were assessed in 3D level using COMPLIG algorithm [25].
recovery rates of pesticides in fruits and vegetables in mass spectrometry based on chemical descriptors of pesticides [Sp 11]. Here, the recovery rate of pesticides is defined by a ratio of peak area in a crop sample to that in the solvent standard calibration curve. If the ratio is close to 1, the effect of impurity caused by the crop is reduced in the measuring system. In R program language, rdk package is publicly available and can create 174 chemical descriptors derived from SMILES of pesticides. Then regression models for estimating the recovery rate based on those chemical descriptors were created by caret package. Here, 69 ordinary and 20 ensemble learning regression models are available in caret package. In this data set, two machine learning regression methods called SBC and xgbLinear are the best choice in terms of prediction rates and execution times. It should be recognized that currently we can freely utilize a large number of chemical descriptors and machine learning methods, which leads to acceleration of chemoinformatics research and those methodology will lead to systematic understanding of chemical data.

6 International Collaboration in Chemoinformatics

Alexandre Varnex of University of Strasbourg reviewed Japanese-French collaboration in chemoinformatics from 2007 to 2016 and a series of Autumn Schools in Chemoinformatics in Tokyo (2011, 2015), and Nara (2013, 2017) which brought together experienced and young scientists. This was initiated by Funatsu and has been a tremendous success worldwide [Sp 5]. The 8th French-Japanese Workshop on Computational Methods in Chemistry (FJCMC2020) will be held on March 19–20, 2020 at Kumamoto in Japan organized by Alexandre Varnex, Kimito Funatsu, Manabu Sugimoto (http://www.chem.kumamoto-u.ac.jp/~frjp2020/program.html).

7 Conclusive Remarks

Data science can be created by integrating disciplines such as theoretical understanding in selected research fields, informatics (for data systematization, model construction, and prediction), and statistics, for validation. Specifically, chemistry and chemical physics could be overlapped with chemical information (molecular structures), and chemometrics validation. An example is the overlap of deuterium isotope effects in solvolysis reactions [28-30], AIPHOS[4] and computer-aided structure elucidation (CHEMICS [1-3] and soft sensors[31, 32]). Kimito Funatsu sits at the center of overlap of such systems. When we focused molecules, we can integrate bioinformatics and chemoinformatics because both research fields focus on understanding of life systems and ecosystems based on molecules, i.e. macromolecules such as DNA RNA, proteins and small molecules such as metabolites and drugs are targeted in bioinformatics and chemoinformatics. Advanced materials are essential to economic security and human well-being. We believe that the meanings of the Honor Award dedicating to Prof. Kimito Funatsu is to provide us a good opportunity to ask ourselves about how molecular informatics should be advanced in a variety of scientific fields.

Prof. Kimito Funatsu encourages us by saying “It is our time to shine! Let’s make seminal works in the development of chemoinformatics!” We hope that a lot of talented researchers and students would learn from his achievements, and would develop and enjoy their unique, original, and productive cutting-edge chemoinformatics in versatile fields as Prof. Funatsu has been doing in his research carrier.

References and Notes

[1] H. Abe, I. Fujiwara, T. Nishimura, T. Okuyama, T. Kida, S. Sasaki, Comput. Enhanced Spectrosc., 1, 55-62 (1983).
[2] K. Funatsu, N. Miyabayashi, S. Sasaki, J. Chem. Inf. Comput. Sci., 28, 18-28 (1988).
[3] K. Funatsu, Y. Susuta, S. Sasaki, J. Chem. Inf. Comput. Sci., 29, 6-11 (1989).
[4] K. Funatsu, S. Sasaki, Tetrahedron Comput. Method, 1, 27-38 (1988).
[5] H. Satoh, K. Funatsu, J. Chem. Inf. Comput. Sci., 35, 34-44 (1995).
[6] K. Funatsu, Process Control and Soft Sensors in Applied Chemoinformatics: Achievements and Future Opportunities, Editor(s): Thomas Engel Johann Gasteiger (2018).
[7] J. Gasteiger, Molecules, 21, 151 (2016).
[8] K. Varmuz, W. Werther, J. Chem. Inf. Comput. Sci., 36, 323-333 (1996).
[9] K. Funatsu, S. Sakaki, J. Chem. Inf. Comput. Sci., 36, 190-204 (1996).
[10] M. Gasteiger, G. Hunchings, B. Christoph., L. Gann, C. Hiller, P. Löw, M. Marsili, H. Saller, K. Yuki, Design. Topics Curr. Chem., 137, 19-73 (1987).
[11] W.D. Ihlenfeldt, J. Gasteiger, Angew. Chem. Int. Ed. Engl., 34, 2613-2633 (1995).
[12] H. Satoh, O. Sacher, T. Nakata, J. Gasteiger, K. Funatsu, J. Chem. Inf. Comput. Sci., 38, 210-219 (1998).
[13] H. Yamamoto, T. Yamaguchi, K. Yoshimura, M.
Sumimoto, K. Hori, J. Synth. Org. Chem. Jpn., 70, 722 (2012).

[14] K. Hori, T. Sadatomi, K. Okano, A. Miyamoto, H. Kuroda, M. Sumimoto, H. Yamamoto, Molecules, 15, 8289 (2010).

[15] S. Hochreiter, J. Schmidhuber, Neural Comput. 9, 1735-1780 (1997).

[16] A. Gupta, A. T. Müller, B. J. H. Huisman, J. A. Fuchs, P. Schneider, G. Schneider, Mol. Inf. 37, 1700111, (2018).

[17] D. Reker, T. Rodrigues, P. Schneider, G. Schneider, Proc. Natl. Acad. Sci. USA, 111, 4067-4072 (2014).

[18] F. Grisoni, D. Merk, V. Consonni, J.A. Hiss, S. G. Tagliabue, R. Todeschini, G. Schneider, Commun. Chem., 1, 44 (2018).

[19] F. Grisoni, D. Merk, R. Byrne, G. Schneider, Sci. Rep., 8, 16469 (2018).

[20] T. Miyao, M. Arakawa, K. Funatsu, Mol. Inf, 29, 111-125 (2010).

[21] Y. Hu, J. Bajorath, Future Sci. OA, FSO179 (2017).

[22] F. Miljković, J. Bajorath, ACS Omega, 3, 17295-17308 (2018).

[23] Md. Altaf-Ul-Amin, T. Katsuragi, T. Sato, S. Kanaya, BioMed Res. Int., 540297 (2005).

[24] F.M. Afendi, T. Okada, M. Yamazaki, A. Hirai-Morita, Y. Nakamura, K. Nakamura, S. Ikeda, H. Takahashi, Altaf-Ul-Amin, M. Darsman, K. Saito, S. Kanaya, Plant Cell Physiol., 53, e1 (2012).

[25] M. Saito, N. Takemura, T. Shirai, J. Mol. Biol., 424, 379-390 (2012).

[26] Md. Altaf-Ul-Amin, Y. Shinbo, K. Mihara, K. Kurokawa, S. Kanaya, BMC bioinformatics, 7, 207 (2006).

[27] A. L. Hopkins, Nat. Chemical Biol. 4, 682-690 (2008).

[28] K. Funatsu, M. Fujio, Y. Tsuno, Mem. Fac. Sci., Kyushu Univ., Ser. C, 13, 125-34 (1981).

[29] K. Funatsu, M. Fujio, Y. Tsuno, Mem. Fac. Sci., Kyushu Univ., Ser. C, 13, 391-6 (1982).

[30] K. Funatsu, M. Kimura, T. Furukawa, M. Fujio, Y. Tsuno, Mem. Fac. Sci., Kyushu Univ., Ser. C, 14, 343-54 (1984).

[31] K. Funatsu, Process Control and Soft Sensors. In Applied Chemoinformatics; Engel, T., Gasteiger, J., Eds.; Wiley-VCH: Weinheim, Germany, pp 571-584 (2018).

[32] K. Funatsu, Soft Sensors: Chemoinformatic Model for Efficient Control and Operation in Chemical Plants, Mol. Inf. 35, 549-554 (2016).