P1-01

Development status of ABINIT-MP program in 2019

Yuji Mochizuki[1,2] **Kota Sakakura**[3] **Hiromasa Watanabe**[4] **Kouichiro Kato**[5]
fullmoon@rikkyo.ac.jp k-sakakura@ay.jp.nec.com hi-watanabe@hpc.co.jp
kouichiro.kato@mizuho-ir.co.jp
Yoshinobu Akinaga[6] **Kaori Fukuzawa**[2,7] **Koji Okuwaki**[7] **Tatsuya Nakano**[8]
akinaga@vinas.com k-fukuzawa@hoshi.ac.jp k-okuwaki@hoshi.ac.jp nakano@nihs.go.jp

[1] Department of Chemistry & Research Center for Smart Molecules, Faculty of Science, Rikkyo University, 3-34-1 Nishi-ikebukuro, Toshima-ku, Tokyo 171-8501, Japan
[2] Institute of Industrial Science, University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
[3] 1st Government and Public Solutions Division, NEC Corporation, 5-7-1 Shiba, Minato-ku, Tokyo 108-8001, Japan
[4] HPC SYSTEMS Inc., Kaigan 3-9-15, Minato-ku, Tokyo 108-0022, Japan
[5] Mizuho Information & Research Institute Inc., 2-3 Kanda-nishiki-cho, Chiyoda-ku, Tokyo 101-8443, Japan
[6] VINAS Co. Ltd., 2-3-31 Dojima, Kita-ku, Osaka 530-003, Japan
[7] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-Ku, Tokyo 142-8501, Japan
[8] Division of Medicinal Safety Science, National Institute of Health, 3-25-26 Tonomachi Kawasaki-ku, Kawasaki, Kanagawa 210-9501, Japan

Keywords: Fragment molecular orbital, FMO, ABINIT-MP, Interaction energy analysis

The ABINIT-MP program [1] has been widely used in Japan, especially in the pharmaceutical communities, e.g., FMO Drug Design Consortium (FMO-DD) [2]. At the Open Version 1 Revision 10 in Spring, 2018, the decomposition analysis of pair interaction energy (PIEDA) [3] was made usable in ABINIT-MP, as was in GAMESS-US [4]. It has been recently known that PIEDA-based regression analyses of drug activities are frequently better than those with parent IFIE values. As other utilities, a couple of orbital-wise analysis tools, FILM [5] and CAFI [6], have also been available from Revision 10. The latest release of ABINIT-MP is Revision 15 in Spring 2019, providing the fragmentation at the sp2-carbon of peptide bond [7], higher-order correlated calculations (such as CCSD(T)) [8], and non-empirical HF+D type correction. In this presentation, we will brief the status of ABINIT-MP in 2019.

[Acknowledgment] This work was supported by MEXT as a social and scientific priority issue #6 to be tackled by using post-K computer (FS2020) and also as a grant-in-aid (Kaken-hi) No. 16H04635.

[1] Tanaka, S., Mochizuki, Y., Komeiji, Y., Okiyama, Y., Fukuzawa, K., Electron-Correlated Fragment-Molecular-Orbital Calculations for Biomolecular and Nano Systems, *Phys. Chem. Chem. Phys.*, 16:10310-10344, 2014.
[2] Fukuzawa, K., et al., <https://fmodd.jp>.
[3] Fedorov, D. G., Kitaura, K. Pair Interaction Energy Decomposition Analysis, *J. Comp. Chem.* 28: 222-237, 2007.
[4] Fedorov D. G., The fragment molecular orbital method: theoretical development, implementation in GAMESS, and applications, *WIREs, Comp. Mol. Sci.*, e1322, 2017.
[5] Ishikawa, T., Mochizuki, Y., Amari, S., Nakano, T., Tokiwa, H., Tanaka, S., Tanaka, K. Fragment interaction analysis based on local MP2, *Theor. Chem. Acc.*, 118:937-945, 2007.
[6] Mochizuki, Y., Fukuzawa, K., Kato, A., Tanaka, S., Kitaura, K., Nakano, T., A configuration analysis for fragment interaction, *Chem. Phys. Lett.*, 410:247-253, 2005.
[7] Akinaga, Y., Nakano T., Fukuzawa, K., Kato, K., Mochizuki, Y., to be published.
[8] Mochizuki, Y., Yamashita, K., Nakano, T., Okiyama, Y., Fukuzawa, K., Taguchi, N., Tanaka S., Higher-order correlated calculations based on fragment molecular orbital scheme, *Theor. Chem. Acc.*, 130:515-530, 2011.

Prediction of log *P* for cyclodextrin-drug complexes using computational chemistry

Masao Fujisawa[1] fujisawa@waka.kindai.ac.jp
Hirohito Ikeda[2] ikeda@fukuoka-u.ac.jp
Takayoshi Kimura[3] kimura@chem.kindai.ac.jp

[1] Department of Biotechnological Science, Kindai University, 930 Nishimitani, Kinokawa-City, Wakayama, Tokyo 649-6493, Japan
[2] Faculty of Pharmaceutical Science, Fukuoka University, 8-19-1 Nanakuma Fukuoka 814-0180, Japan
[3] Department of Chemistry, Kindai University, Kowakae, Higashi-osaka 577-8502, Japan

Keywords: log *P*, DFT, molecular dynamics, Inclusion complex

It is important to examine the effect of hydrophobicity on the interaction between organic pollutants and water molecules accurately to understand the mechanism of removal of organic pollutants from water. Cyclodextrin (CyD) is used as a carrier to improve the solubility of poorly water-soluble drugs. The accumulation of accurate and quantitative values of the changes in thermodynamic functions on molecular inclusion of guest molecules into CyDs in aqueous solutions have been carried out systematically by microcalorimetry[1-3], in order to clarify the mechanisms of molecular recognition and discrimination in aqueous solutions. However, determining the solubilities of CyD-drug complexes experimentally is difficult. Thus, log P partition coefficients of inclusion complexes obtained from the difference between their solvation Gibbs energy in two-phase octanol/water are utilized to estimate the hydrophobicity of complexes. In this paper, we report the contribution of dispersion forces to the solvation Gibbs energy of isolated molecules in aqueous solution. Solvation Gibbs energies were determined using vibrational frequency calculations with the dispersion-corrected density functional theory (DFT-D). The difference between the solubilities of mefenamic acid and a CyD-mefenamic acid complex in water was estimated using predicted log *P*. Water solubility of 2-hydroxypropyl-beta-CyD improved on the formation of the CyD inclusion complex. Our prediction method reproduced this phenomenon well. Moreover, the slight difference between the solubilities of mefenamic acid and niflumic acid could be reproduced well. In addition, the epinephrine system will be discussed on the day of annual meeting[4].

[1] Kimura T., Yukiyama T., Fujie S., Fujisawa M., Aki H., Kamiyama T., Enthalpy and entropy changes on molecular inclusion of pentane derivertives into α-cyclodextrin cavities in aqueous solutions,, *Journal of Inclusion Phenomena and Macrocyclic Chemistry*, 70(3-4):269-278, 2011.
[2] Kimura T., Yukiyama T., Fujisawa M., Thermodynamic properties of inclusion complexes of α-cyclodextrin + aliphatic nitriles ($H(CH_2)_nCN$: n = 1-8) in aqueous solution, *Journal of Thermal Analysis and Calorimetry*, 108(2):695-704, 2012.
[3] Ikeda H., Fukushige Y., Matsubara T., Inenaga M., M. Kawahara M., Yukawa M., Fujisawa M., Yukawa E., Aki H., Improving water solubility of nateglinide by complexation of β-cyclodextrin, *Journal of Thermal Analysis and Calorimetry*, 123(3):1847-1850, 2016.
[4] Rayenne D., Djameleddine K., Quantum mechanical study of complexation of dopamine and epinephrine with β-cyclodextrin using PM6, ONIOM and NBO analysis, *Journal of Computational and Theoretical Nanoscience*, 9(10):1571-1576, 2012.

P1-03

Binding Energy Calculation of Protein-Peptide Complex Using Unbiased MD Simulations and MSM Analysis

Hiroaki Hata[1]
hiroakihata@bio.titech.ac.jp

Yasutaka Nishihara[2]

Masayoshi Nishiyama[3] **Ikuro Kawagishi**[4] **Akio Kitao**[1]

[1] School of Life Science and Technology, Tokyo Institute of Technology, M6-13, 2-12-1, Ookayama, Meguro-ku, Tokyo 152-8550, Japan
[2] Institute of Molecular and Cellular Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan
[3] Department of Physics, Kindai University, 3-4-1 Kowakae, Higashiosaka, Osaka 577-8502, Japan
[4] Department of Frontier Biosciences, Hosei Univesity, Koganei, Tokyo 184-0002, Japan

Keywords: Molecular dynamics simulation, Protein-peptide binding, Markov state model

Association/dissociation among protein molecules plays central roles in functional regulation of living organisms. It is of great biological and pharmaceutical interest to characterize the dynamics of interactions among proteins. Molecular dynamics (MD) simulations are powerful tool for investigation of the protein-protein interaction in atomic detail. However, direct observation of association/dissociation events of protein molecules using conventional MD simulations are often difficult due to the limitation of the simulation time [1]. Therefore, some widely used MD methods, such as the steered MD, utilize external forces to enhance the event [2], although the artificial force may induce unnatural behavior of molecules.

We developed a MD method called PaCS-MD which efficiently enhances conformational fluctuations not by adding bias forces but by repeating short MD simulations and conformational selections [3]. Recently, we showed that the Markov state model (MSM) analysis can be directly applied to the PaCS-MD trajectory for calculation of binding energy, association/dissociation rate constants, and residence time [4]. In this presentation, we applied the PaCS-MD/MSM method to the protein complex of bacterial flagella motor protein FliM and the chemotaxis signaling protein CheY. The binding of CheY onto the N-terminal segment of FliM has been known to regulate the rotational direction of bacterial flagella motor. We successfully observed multiple dissociation pathways of the complex using PaCS-MD, and calculated the standard binding energy using MSM. We found that the binding energy is significantly dependent on the hydrostatic pressure condition and the calculated pressure dependence agreed well with experimental observations. This result indicated that the binding of CheY-FliM, and consequently the rotation of flagella motor, can be modulated by applying pressure. We will also discuss the mechanism how the environmental pressure modifies the interaction between CheY-FliM.

[1] Pan, A. C. et al. Atomic-level characterization of protein–protein association. Proc. Natl. Acad. Sci. 116, 4244–4249 (2019).
[2] Cuendet, M. A. & Michielin, O. Protein-protein interaction investigated by steered molecular dynamics: The TCR-pMHC complex. Biophys. J. 95, 3575–3590 (2008).
[3] Harada, R. & Kitao, A. Parallel cascade selection molecular dynamics (PaCS-MD) to generate conformational transition pathway. J. Chem. Phys. 139, 035103 (2013).
[4] Tran, D. P. & Kitao, A. Dissociation Process of a MDM2/p53 Complex Investigated by Parallel Cascade Selection Molecular Dynamics and the Markov State Model. J. Phys. Chem. B 123, 2469–2478 (2019).

P1-04

Interaction Analyses between Calcite/Apatite and Peptides by Fragment Molecular Orbital Method

Ryo Hatada[1] **Kouichiro Kato**[2] **Koji Okuwaki**[3] **Kaori Fukuzawa**[3,4] **Yuji Mochizuki**[1,4]
hatada@rikkyo.ac.jp kouichiro.kato@mizuho-ir.co.jp k-okuwaki@hoshi.ac.jp
k-fukuzawa@hoshi.ac.jp fullmoon@rikkyo.ac.jp

[1] Department of Chemistry & Research Center for Smart Molecules, Faculty of Science, Rikkyo University, 3-34-1 Nishi-ikebukuro, Toshima-ku, Tokyo 171-8501, Japan
[2] Mizuho Information & Research Institute Inc., 2-3 Kanda-nishiki-cho, Chiyoda-ku, Tokyo 101-8443, Japan
[3] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-Ku, Tokyo 142-8501, Japan
[4] Institute of Industrial Science, University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

Keywords: Biomineral, Fragment molecular orbital method, Data driven analyses

In the natural world, organisms produce various bio-related minerals (e.g., pearls, shells, bones and exoskeletons), and such actions are called as biomineralizations. Most of biominerals consist of inorganic crystals and biomolecules, providing a coexistence of lightweight and strong properties. Therefore, fundamental understanding for the complex inorganic-bio hybrid structures or the adsorption mechanism of biomolecules on inorganic crystals should be helpful not only to elucidate the mechanism of biomineralization but also to design novel functional materials. In our previous study, we have performed the fragment molecular orbital (FMO) method to analyze the interaction between calcite crystal ($CaCO_3$) and DDGSDD motif having the corresponding specific adsorption ability [1]. A similar FMO-based analysis [2] was carried out for the system of hydroxyapatite ($Ca_5(PO_4)_3(OH)$) and ESQES motif [3]. These two studies revealed that the role in adsorption of each residue is greatly different and also that the C-terminal side plays vital roles in adsorption.

In this study, we have investigated the contribution of each energy component to the interaction between crystal and peptide motif, based on the pair interaction energy decomposition (PIEDA) analysis derived from FMO calculations. In addition, we extracted the interaction energies between residues and close fragments in crystal as well as the distance and arrangement information between residues and crystal. Such data were subjected to the multiple regression analysis by which feature parameters could be identified. As a result, it was found that the nature of interaction between calcite and DDGSDD motif is more ionic than that between hydroxyapatite and ESQES motif and thus that the distance parameters among residues and close fragments in crystal are principal features. In addition, it was suggested that the presence of ionized carboxyl group ($-COO^-$) of D or E is important in both systems. Detailed results will be reported on the presentation day.

[Acknowledgment] This work was supported by MEXT as a social and scientific priority issue #6 to be tackled by using post-K computer (FS2020) and also by a grant-in-aid (Kaken-hi) No. 16H04635.

[1] Gotliv, B., Kessler, N., Sumerel, J. L., Morse, D. E., Tuross, N., Addadi, L., Weiner, S., Asprich: A Novel Aspartic Acid-Rich Protein Family from the Prismatic Shell Matrix of the Bivalve *Atrina rigida*, *ChemBioChem*., 6: 304-314, 2005.
[2] Kato, K., Fukuzawa, K., Mochizuki, Y., Modeling of hydroxyapatite-peptide interaction based on fragment molecular orbital method, *Chem. Phys. Lett*., 629: 58-64, 2015
[3] He, G., Dahl, T., Veis, A., George, A., Nucleation of apatite crystals *in vitro* by self-assembled dentin matrix protein 1, *Nat. Mater*., 2: 552-558, 2003.

P1-05

Fragment Molecular Orbital Method Applied to Factor Xa Inhibitors

Kazufumi Ohkawa
okawa.kf@om.asahi-kasei.co.jp

Laboratory for Medicinal Chemistry, Pharmaceuticals Research Center, Asahi Kasei Pharma Corporation, 632-1 Mifuku, Izunokuni, Shizuoka 410-2321, Japan

Keywords: Fragment Molecular Orbital, Factor Xa, antithrombotic

Factor Xa(FXa) occupies the central position in the blood coagulation cascade and their importance is reflected in the wide clinical use of FXa inhibitors. One of the second generation FXa inhibitors, Rivaroxaban has chlorothiophene as P1 motif instead of benzamidine which is popular in first generation FXa inhibitors. Both benzamidine and chlorothiophene could bind S1 pocket. As a basic P1 group, benzamidine interact with ASP189 by salt bridge formation. For the binding energy of chlorothiophene as neutral P1 group, has been ascribed to (1) direct chlorine atom interaction with TYR228 and (2) polarization of the C-H bond ortho to the chlorine atom, leading to a favorable interaction with ASP189.

An in-depth energetic analysis by Fragment Molecular Orbital method (FMO) suggests that the nature of the interaction energy between chlorothiophene and S1 pockets is due primary to (2) polarization of the C-H bond ortho to the chlorine atom, leading to a favorable interaction with ASP189.

We also applied FMO to several data sets, including a set of reported FXa inhibitors. We found that a strong correlation between binding energy (Total-IFIE) and the activity value on 50% inhibitory concentration (IC50) of the inhibitors.

[1] K Kitaura, E Ikeo, T Asada, T Nakano, and M Uebayasi, Fragment Molecular Orbital Method: An Approximate Computational Method for Large Molecules, *Chem. Phys. Lett*., 313:701– 706, 1999.

[2] S Roehrig., *et al.,* Discovery of the novel antithrombotic agent 5-chloro-N-({(5S)-2-oxo-3-[4-(3-oxomorpholin-4-yl)phenyl]-1,3-oxazolidin-5-yl}methyl)thiophene- 2-carboxamide (BAY 59-7939): an oral, direct factor Xa inhibitor, *J. Med. Chem.,* 48, 19:5900-5908, 2005.

[3] C Chan., *et al.,* Factor Xa Inhibitors: S1 Binding Interactions of a Series of N-{(3S)-1-[(1S)-1-Methyl-2-morpholin-4-yl-2-oxoethyl]-2-oxopyrrolidin-3-yl}sulfonamides, *J. Med. Chem.,* 50, 7:1546-1557, 2007.

[4] D. J. P. Pinto., *et al.,* Factor Xa Inhibitors: Next-Generation Antithrombotic Agents, *J. Med. Chem.,* 53, 17:6243-6274, 2010.

[5] Y. N. Imai, Y. Inoue, I. Nakanishi, and K. Kitaura, Cl–π interactions in protein–ligand complexes, *Protein Sci*., 17, 7:1129–1137, 2008.

P1-06

Influence of phosphorylation on structure and electronic states of tau-protein: MD and *ab initio* fragment MO simulations

Katsumi Suzuki[1] k-suzuki@klab.cs.tut.ac.jp
Riku Sato[1] sato@klab.cs.tut.ac.jp
Pavel Karpov[2] karpov@nas.gov.ua
Sergiy Shulga[2] Shulga5@i.ua
Yaroslav Blume[2] blume@nas.gov.ua
Noriyuki Kurita[1] kurita@klab.cs.tut.ac.jp

[1] Department of Computer Science and Engineering, Toyohashi University of Technology, 1−1 Hibarigaoka, Tempaku−cho, Toyohashi, Aichi, 441−8580, Japan
[2] Institute for Food Biotechnology and Genomics, National Academy of Sciences of Ukraine, 2a. Osypovskogo Str., Kyiv-123, 04123, Ukraine

Keywords: Tau-protein; Aggregate; Phosphorylation; Alzheimer's disease; Molecular simulation; Molecular dynamics; Fragment molecular orbital

Alzheimer's disease (AD) is the most common neurodegenerative disease characterized by neurofibrillary tangles (NFTs) in a brain. It is considered that NFTs are caused by the aggregation of abnormally phosphorylated tau-proteins in neurons. The phosphorylation sites of tau-protein were classified into three categories [1]: ones identified only in the brains of AD patients, ones identified in the brains of both AD patients and healthy people, and ones identified only in the brains of healthy people. However, it is not clear why such a classification occurs and how the phosphorylation of tau-protein affects its structure and aggregation.

In the present study, we investigated the influence of phosphorylation on the structure and electronic states of tau-protein at atomic and electronic levels, using molecular dynamics (MD) and *ab initio* fragment molecular orbital (FMO) calculations in water. As a target structure of tau-protein, we employed the repeat 1 domain among the microtubule binding repeat (MBR) of tau-protein. In the domain, the peptide composed of 89 residues from the 186th to 275th residues in the N-terminal of MBR was selected, because this region has many phosphorylation sites [2]. Based on the PDB structure (PDB ID: 5N5A), we employed protein 3D structure prediction program I-TASSER to make the initial structure of peptide. One serine residue among the twelve ones was phosphorylated to make twelve initial structures of phosphorylated tau-protein. Among them, eight sites are identified only in the brains of AD patients, while the other four sites are identified in the brains of both AD patients and healthy people. We attempted to elucidate the differences in the influence of phosphorylation depending on the site of phosphorylation, using the following molecular simulations.

We first carried out 100 ns MD simulations for the twelve phosphorylated tau-protein as well as the non-phosphorylated tau-protein at 300 K in explicit waters, in order to reveal the influence of phosphorylation on the structure. The snapshot at 100 ns was optimized by MM method, and the electronic states of the optimized structure were investigated by *ab initio* FMO calculations. Based on the results of FMO, we investigated how the interactions between amino acid residues of tau-protein are changed by the phosphorylation of one serine residue in the tau-protein. The details of the results obtained by the MD and FMO calculations will be shown in the poster.

[1] L. Martin, X. Latypova, and F. Terro, "Post-translational modifications of tau protein: Implications for Alzheimer's disease," *Neurochem. Int.* 58, 458–471, 2011.
[2] M. Morishima-Kawashima *et al.*, "Proline-directed and Non-proline-directed Phosphorylation of PHF-tau," *J. Biol. Chem.* 270, 823–829, 1995.

P1-07

Determination of stable protonation states of amino acid residues in metalloproteinase-inhibitor complex: *ab initio* molecular simulations

Takuya Ezawa[1] ezawa@klab.cs.tut.ac.jp
Satoshi Sugiyama[1] sugiyama@klab.cs.tut.ac.jp
Ayami Ara[1] ara@klab.cs.tut.ac.jp
Shirley Weiyan Chen[2] wchen003@citymail.cuny.edu
Ingebrigt Sylte[3] ingebrigt.sylte@uit.no
Noriyuki Kurita[1] kurita@cs.tut.ac.jp

[1] Department of Computer Science and Engineering, Toyohashi University of Technology, Toyohashi, 441-8580, Japan
[2] Department of Biomedical Engineering, The City College of New York, 160 Convent Avenue New York, NY 10031, U.S.
[3] Department of Medical Biology, Faculty of Health Sciences, UiT The Arctic University of Norway, NO-9037 Tromso, Norway

Keywords: Metalloproteinases; Bacterial virulence; Molecular simulation; Fragment molecular orbital; Protein-ligand interactions; Inhibitor; Antimicrobial agent

Pseudolysin (PLN) is a metalloproteinase secreted from bacteria and degrades extracellular proteins to produce bacterial nutrition. It is thus expected that inhibitors against PLN can suppress the growth of bacteria and their pandemic spread. In addition, since these inhibitors do not attack to bacteria directly, there is a reduced risk for producing drug-resistant bacteria. PLN has a Zn ion around its ligand-binding pocket (LBP), so that there is a possibility that the interactions between Zn and some amino acid residues as well as ligand are significantly changed depending on the protonation states of the residues around Zn. Therefore, we here considered some types of the protonation states for the Glu and His residues existing around Zn and elucidated which protonation state is more stable, based on *ab initio* fragment molecular orbital (FMO) calculations. In addition, specific interactions between PLN residues, Zn and its inhibitor were investigated in an electronic level using the *ab initio* FMO method to reveal the following points.

(1) Although Glu141 existing near the LBP of PLN can have a protonated state [1], non-protonated Glu141 is more preferable than the protonated one in the PLN-inhibitor complex [2] from the viewpoint of the PLN structure around Zn.
(2) His223 existing near LBP can have Hid or Hie protonated state depending on the ligand.
(3) In proposing novel ligands for PLN, two types of protonation states for His223 should be considered to elucidate the specific interactions between PLN and ligand.

The FMO calculations are underway now for the PLN complexed with the other inhibitors, in order to make clear the reason for the wide range of binding affinity between PLN and its inhibitors observed by the previous experiments [2].

[1] X. Hu, S. Balaz, and W. H. Shelver, A practical approach to docking of zinc metalloproteinase inhibitors, *J. Mol. Graph. Model.*, 22 (2004) 293–307.
[2] A. M. Kany *et al.*, Tackling *pseudomonas aeruginosa* virulence by a hydroxamic acid-based LasB inhibitor," *ACS Chem. Biol.*, 13 (2018) 2449–2455.

P1-08

Regulation mechanism of agonistic / antagonistic activities of vitamin D receptor analyzed by generalized ensemble method

Takafumi Kudo[1]
w185412d@yokohama-cu.ac.jp

Toru Ekimoto[1]
ekimoto@yokohama-cu.ac.jp

Tsutomu Yamane[1]
tyamane@yokohama-cu.ac.jp

Mitsunori Ikeguchi[1,2]
ike@yokohama-cu.ac.jp

[1]Graduate School of Medical Life Science, Yokohama City University,
1-7-29 Suehirocho, Tsurumi-ku, Yokohama City 230-0045, Japan
[2]RIKEN Medical Sciences Innovation Hub Program,
1-7-29 Suehirocho, Tsurumi-ku, Yokohama City 230-0045, Japan

Keywords: Molecular dynamics simulation, Generalized ensemble method, Vitamin D receptor

Vitamin D receptor (VDR) is one of the nuclear receptors (NR) and an important target for drug discovery. The mechanism of the agonistic/antagonistic activities of VDR is not clear on the basis of X-ray crystallography. According to X-ray crystal structures of other NRs and experimental results, the ligand-binding domain (LBD) undergoes the conformational change upon ligand binding, and a local conformational change around helix 12 is key to regulating agonistic/antagonistic activities. However, since the crystal structures of agonist and antagonist complexes of VDR-LBD were almost identical, it is difficult to discuss the differences of their activities on the basis of crystal structures.

In order to clarify the regulation mechanism of agonistic/antagonistic activities of VDR ligands, molecular dynamics (MD) simulations for VDR-LBD were performed. However, due to the limited time scale of the conventional MD simulations, it was difficult to observe large conformational changes. Therefore, gREST [1], one of the generalized ensemble methods, was employed for the efficient conformational sampling of VDR-LBD. In the gREST simulations, spontaneous conformational changes of VDR were observed, indicating a mechanism underlying the differences of agonistic/antagonistic activities. A typical conformation of the antagonist bound form has a similar feature to that observed by MD-SAXS analysis [2].

[1] Kamiya, M., Sugita Y., *J. Chem. Phys.*, 149, 072304-1-11, 2018
[2] Anami, Y., Shimizu, N., Ekimoto, T., Egawa, D., Itoh, T., Ikeguchi, M., Yamanoto, K., *J. Med. Chem.*, 59, 7888-7900, 2016.

P1-09

The inhibition mechanism of HSP90 function by a medium molecular drug

Lisa Matsukura[1]
liz@miyashita-lab.net

Kazuto Mochizuki[2]
m1743011@edu.cc.uec.ac.jp

Naoyuki Miyashita[1,2,*]
miya@waka.kindai.ac.jp

Masumi Taki[2]
taki@pc.uec.ac.jp

Shinichi Watanabe[2]
shin.watanabe@uec.ac.jp

[1] Department of Biological System Engineering, Graduate school of Biological Oriented Science and Technologies, KINDAI University, 930 Nishimitani, Kinokawa, Wakayama 649−6493, Japan
[2] Department of Engineering Science, Graduate School of Informatics and Engineering, University of Electro-Communications, 1-5-1, Chofugaoka, Chofu, Tokyo 182-8585, Japan

Keywords: Molecular dynamics Simulation, Ligand Docking

HSP90 is known as one of the anticancer drug targets. The previous study has suggested the activity of HSP90 is suppressed by cryptand-aptamer, which consists of 14 residue peptide and cryptand compound. A peptide, which has 14 residue peptide, and the cryptand-aptamer bind to the HSP90. The cryptand-aptamer has a higher binding affinity. However, the inhibition mechanism of HSP90 function has not been understood yet. In this study, we investigated the structural differences between the peptide and cryptand-aptamer using 20μs molecular dynamics simulations; we performed docking simulation of HSP90 and the most reliable structure of cryptand-aptamer to estimate the docking mode; we carried out the molecular dynamics simulation of the complex structure to understand the inhibition mechanism of HSP90 function by the cryptand-aptamer.

Our results have shown that the turn structure in the peptide is more unstable than that in the cryptand-aptamer, and the cryptand-aptamer binds to the beta-strand side of HSP90. The stability of the turn structure of the medium molecular drug affects the binding affinity. The closed conformation of HSP90 usually has the domain swapping around N-terminal beta-strand[1]. Our result suggests that the cryptand-aptamer inhibits the domain swapping in the closed conformation of HSP90.

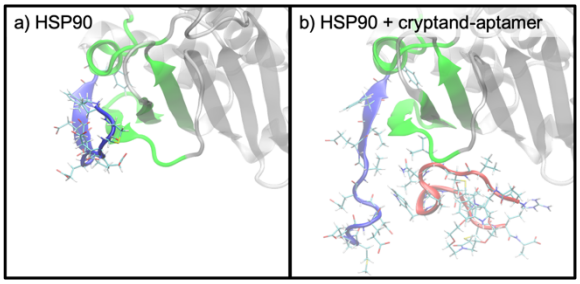


Fig. 1 The interaction between N-terminal region and the cryptand-aptamer.
a) HSP90 without the drug, b) HSP90 with the drug.

[1] Ali M.M., et al., Crystal structure of an HSP90–nucleotide–p23/Sba1 closed chaperone complex , *Nature*, 440:1013-1017, 2006.

Molecular Dynamics Simulations of HIV Tat protein and Amyloid-β peptides

Kazumi Omata[1]
komata@hosp.ncgm.go.jp

Hisashi Okumura[2,3]
hokumura@ims.ac.jp

[1] Department of Data Science, National Center for Global Health and Medicine, 1-21-1 Toyama, Shinjuku, Tokyo 162-0052, Japan
[2] Institute for Molecular Science, 38 Nishigonaka, Myodaiji, Okazaki, Aichi 444-8585, Japan
[3] The Graduate University for Advanced Studies, 38 Nishigonaka, Myodaiji, Okazaki, Aichi 444-8585, Japan

Keywords: HIV Tat protein, Amyloid-β, Ligand docking, Molecular dynamics, neurocognitive dysfunction

Continuing from CBI 2018, we show the results of molecular dynamics (MD) simulations about the docking of an HIV transactivator of transcription (Tat) protein and amyloid-β (Aβ) peptides and identify binding sites of these molecules.

Recently, the experiments by Hategan et al. suggested that HIV Tat binds to the exterior surfaces of Aβ fibrils, increasing β-sheet formation and lateral aggregation into thick multifibrillar structures [1]. As a result, the fibers increase their rigidity and mechanical resistance. They also showed that the aggregate of Tat-Aβ complexes synergistically induced neurotoxicity both *in vitro* and in animal models. Apart from such a study, Okumura and Itoh examined the structure of Aβ fibrils by all-atom MD simulations, and demonstrated that structural fluctuations at the fibril ends play an important role for the fibril formation [2]. Based on the knowledge about the structure of Aβ fibrils derived from their MD simulations, the present study has attempted to find which parts of Aβ peptides are tightly associated with HIV Tat proteins, using MD simulations.

The all-atom MD simulations were carried out using the Generalized-Ensemble Molecular Biophysics (GEMB) program [3]. The AMBER99SB force field was used for the HIV Tat and Aβ, and the TIP3P rigid-body model was used for the water molecules. The electrostatic potential was calculated by the particle mesh Ewald method. Temperature was controlled at 298 K with the Nosé-Hoover thermostat, and pressure was controlled at 0.1 MPa with the Andersen barostat. The symplectic quaternion scheme was employed for the rigid-body water molecules.

We have investigated three sets of initial condition. In the first set, a Tat protein was placed near one of the two beta sheets constituting the Aβ. The Tat moved and was bound to the N-terminal soft region of the Aβ. The second set in which a Tat was placed near the other beta sheet also exhibited the movement and binding of the Tat to the N-terminal region. Considering these results, in the third set, we placed a Tat protein near the loop region of the Aβ which is far from the N-terminal soft region. Again, this set provided a similar movement and binding to the soft region. These results allow us to conclude that the binding of a Tat and Aβ occurs only in the soft region of the Aβ. In addition, we found that this binding may be attributed to hydrophobic interactions among several residues in Tat and Aβ.

[1] Hategan, A., et al., *Nature Structural & Molecular Biology* **24**, 379–386, 2017.
[2] Okumura, H. and Itoh, S. G., *Scientific Reports* 6, Article number:38422, 2016.
[3] Okumura, H., *Proteins*, 80: 2397-2416, 2012.

P1-11

Dynamics of the transmembrane region of ß-secretase in raft environment

Kaori Yanagino (KINDAI Univ.)

Cancelled

Development of the CHARMM force field for Cyclosporine A and application to molecular dynamics simulations using a membrane-water system

Tsutomu Yamane[1] tyamane@yokohama-cu.ac.jp
Ryo Takahashi[2] i140382e@yokohama-cu.ac.jp
Akari Ito[2] i160060f@yokohama-cu.ac.jp
Toru Ekimoto[1] ekimoto@yokohama-cu.ac.jp
Mitsunori Ikeguchi[1,3] ike@yokohama-cu.ac.jp

1 Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
2 Department of Medical Life Science, Faculty of Sciences, International Arts and Science Studies, International College of Arts and Sciences, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
3 RIKEN Medical Sciences Innovation Hub Program, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

Keywords: CHARMM force field, cyclosporin A, cyclic peptide, membrane permeability, molecular dynamics simulation

Cyclosporin A (CsA), which is found in *Tolypocladium inflatum* in 1969 [1], is a cyclic peptide consisting of 11 residues and used as an immune suppresser. CsA is permeable to biological membranes despite its high molecular weight (M. w. ~1200). CsA has two types of conformations with a hydrophobic surface (closed form) and with a hydrophilic surface (open form), and the membrane permeability of CsA is possibly related to structural changes between these conformations during membrane permeation process. CsA contains unnatural residues (N-methyl amino acids), which contribute to the structural change of CsA.

Here, we developed the CHARMM force field [2] of CsA. Specifically, we developed CMAP terms of N-methyl amino acids, which are cross-terms of backbone dihedral angle ϕ and ψ of amino acids, by high-level QM calculations. MD simulations in water and chloroform solvent were performed and these results were compared with crystal structures and NOE from NMR measurements. Also, we performed MD simulations of CsA and its metabolite, cyclosporin E (CsE) in the water-membrane system and analyzed the dynamics in the lipid bilayer.

[1] Svarstad, H. et al., *Biodiversity and Conservation*, 9: 1521–1541, 2000.
[2] Huang, J. et al., *Nature Methods*, 14:71-73, 2016.

Analyzing intramolecular interaction using canonical Kohn-Sham molecular orbital calculation in protein

Toshiyuki Hirano[1] t-hirano@iis.u-tokyo.ac.jp
Fumitoshi Sato[1] satofumi@iis.u-tokyo.ac.jp

[1] Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

Keywords: Canonical Molecular Orbital Calculation, Interaction energy

We have developed a QM application, ProteinDF [1], which is designed for canonical Kohn-Sham molecular orbital (CMO) calculations of large molecular system. By using the ProteinDF, some CMO calculations of protein have been reported [2-5]. The CMO of protein provides useful knowledge, which are hard to observe in experiments, such as molecular orbitals, energy level, charge density and so on. Our previous CMO calculation results showed that the CMO of protein was widely spread throughout the whole molecular region, and suggested that the molecular properties were related to the spread of the CMOs. On the other hand, for phenomenon analysis and molecular design, visualization of characteristic interactions between amino acid residues would assist in understanding. Energy decomposition analysis approaches [6], which are one of the interaction energy calculation methods, have been well-known and needed to evaluate some CMO calculations of complex and its subunits. In the case of protein, however, there are a lot of interaction calculations among subunits so that the computational cost of each subunit is expensive. The purpose of this study is to estimate the interaction energies among amino acid residues and/or substrates from the result of CMO calculation with as little effort as possible.

In this study, the interaction energy was estimated based on the energy density analysis (EDA) method [7], which divided the total energy of molecular system into atomic energies. The total energy of electrons in HF calculation is obtained by the sum of dot product of the density matrix and the sum of core Hamiltonian and Fock matrix. Similar to Mulliken population analysis, the sum of the corresponding matrix elements is estimated to be the interaction energy. One of the advantages of this method is that the interaction energy is given by the sum of one electron term, Coulomb term, and XC term (for DFT). In order to calculate the XC interaction energy, the matrix elements of the energy were needed; however, only the sum of the XC energy is given in the grid method, which is commonly used to estimate XC term in many DFT application. In this study, the grid-free method was applied, and so it was suited for representation of AO based XC energy matrix.

Analyzing intramolecular interactions in some proteins have been performed. The Mulliken population of the amino acid residue could only determine the charge density of itself. However, by the analysis method of this research, it was possible to quantify specific Coulomb interactions between amino acid residues. Not only Coulomb interaction but also interaction of one electron term and exchange correlation term could be observed. We will show the analyzing method and these results.

[1] https://github.com/ProteinDF/ProteinDF
[2] T. Inaba, *et.al.*, *Chem. Phys. Lett.*, **434**, 331-335 (2007).
[3] Y. Tokita, *et. al.*, *Angew. Chem. Int. Ed.*, **50**, 11663 (2011).
[4] K. Chiba, *et.al.*, *Int. J. Quant. Chem.*, **113**, 2345-2354 (2013).
[5] Y. Yagi, *et.al.*, *Trans. Jpn. Soc. Sim Tech.*, **4**, 127 (2012).
[6] M. Phipps, *et.al.*, *Chem. Soc. Rev.*, **44**, 3177 (2015).
[7] H. Nakai, *Chem Phys. Lett.*, **363**, 73-79 (2002).

Homology Modeling of the Protein Structure in the Solution using Deep Auto Encoder and Trajectories of the Long-time Molecular Dynamics Simulation of Template Proteins

Masaya Furue[1]
shiver@miyashita-lab.net

Lisa Matsukura[1]
liz@miyashita-lab.net

Mitsutaka Nemoto[1]
nemoto@waka.kindai.ac.jp

Naoyuki Miyashita[1]
miya@waka.kindai.ac.jp

[1] Department of Biological System Engineering, Graduate school of Biological Oriented Science and Technologies, KINDAI University, 930 Nishimitani, Kinokawa, Wakayama 649−6493, Japan

Keywords: Deep Learning, Protein Structure Prediction, ,Homology Modeling, Molecular Dynamics Simulation

A lot of primary sequences of the proteins have been elucidated by next-generation sequencer and so on, while the tertiary structures of proteins usually have been solved by the NMR spectroscopy, X-ray structure analysis, and Cryo-EM. The total number of known tertiary structures of proteins are less than the total number of known primary sequences. Homology modeling is the methods to predict the tertiary structure of target protein using template protein structures, which have known tertiary structures, and MODELLER[1] and SWISS - MODEL [2] are two of famous homology modeling programs. Recently, a machine learning and deep learning has been widely used as tools for image recognition, self-driving and so on. We have developed a homology modeling program, mDeepHoMe, using deep learning methods (Fig. 1). The trajectories of the molecular dynamics simulations of template structures are used as the learning data in mDeepHoMe. PDB data often have the structure differences between a structure in PDB and the structure in the solution which is after molecular dynamics simulations. In this study, we have tried to predict the equilibrated structure of target protein. We performed the long-time molecular dynamics simulations of template proteins. The trajectories was classified by principal component analysis, and we used it to the learning data of mDeepHoMe.

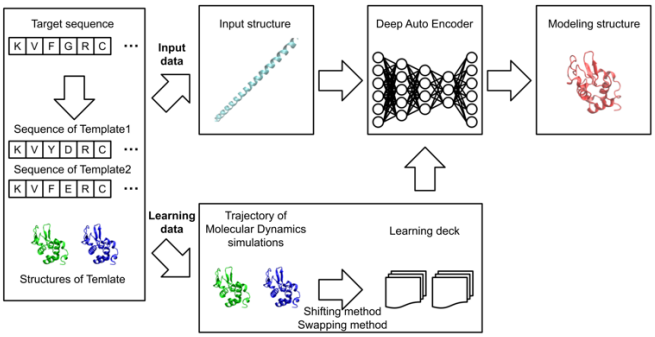


Fig. 1 Schematic figure of mDeepHoMe Program

[1] Fiser A, Sali A, et al, *Methods in Enzymology* **374,** pp 461-491, (2003)

[2] Andrew Waterhouse et al, *Nucleic Acids Research,* **46**, pp.W296–W303, (2018)

[3] Li H et al, *J Proteomics Bioinform* **9**, pp306-313, (2016)

P1-15

Cluster Analysis of Amino Acid by Inter-Fragment Interaction Energy Using Non-Metric Multidimensional Scaling Method

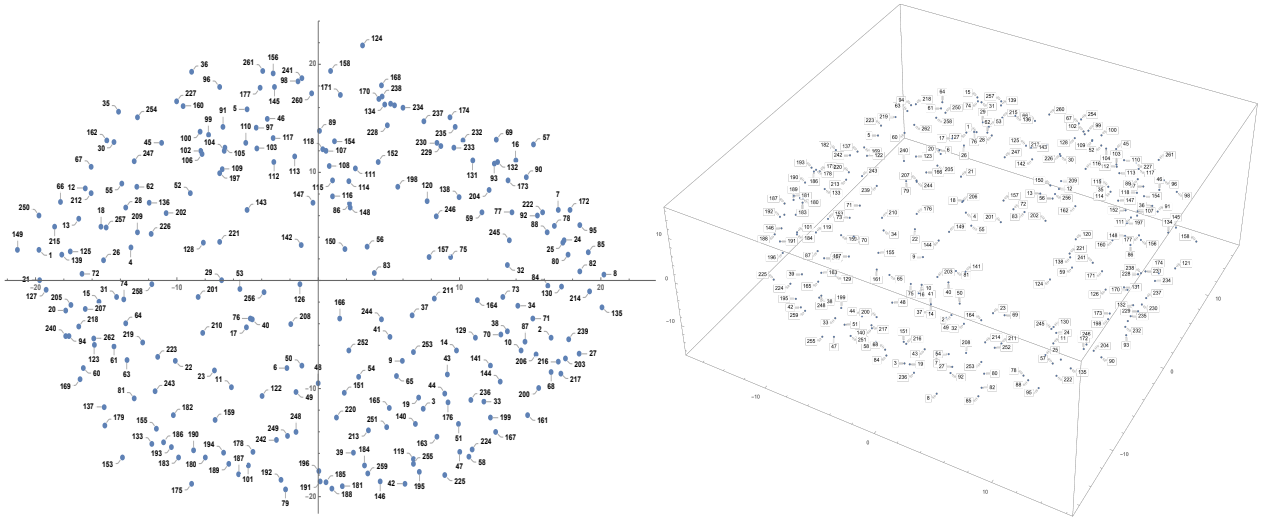
Abe Yuki
csykl6009@g.nihon-u.ac.jp

Yamanaka Masanori
yamanaka.masanori@nihon-u.ac.jp

Department of Physics, College of Science and Technology, Nihon University, 1-8 Kanda-Surugadai, Chiyoda-ku, Tokyo 101-8308, Japan

Keywords: Fragment molecular orbital method, Dimensional scaling, Inter fragment interaction energy

Fragment molecular orbital (FMO) method is a method of calculating the energy and the electron density by dividing a molecular system into fragments, calculating the electronic states of fragments and fragment pairs, and gathering them [1]. Using the method and the Auto-FMO protocol the FMO data base is released [2,3]. In this work, we investigate the inter-fragment interaction energy (IFIE) matrix by the non-metric multidimensional scaling method and perform a cluster analysis for the amino acids and the ligand. The non-metric multidimensional scaling is a method of embedding an object under conditions that preserve the distance relationship between elements in a certain space, which is one of the multivariate analysis, mainly used for the compression of large amounts of information [4]. The method is essentially non-linear analysis, which will be compensate for linear analysis such as the principal component analysis. As an example, we analyzed a kinase-ligand complex, PDB ID: 2X8E, FMODB ID: XYQ2X). The amino acids and the ligand are decomposed into strongly correlated clusters in a non-linear fashion. A part of this research was done in activities of the FMO drug design consortium (FMODD).



The result of embedding the IFIE matrix in two dimensions (left) and three dimensions (right) using non-metric multidimensional scaling. The number from 1 to 261 represent the amino acid which form kinase and the ligand is 262.

[1] K. Kitaura et.al., Chem. Phys. Lett. 313, 701, (1999).
[2] http://drugdesign.riken.jp/FMODB/
[3] C. Watanabe, et.al., Chem-Bio Informatics Journal, 19, 5, (2019)
[4]Modern Multidimensional Scaling, Springer, New York. EEG (2000)

Prediction of Inter-Fragment Interaction Energies in Janus Kinase by Neural Network with Geometrical Information on Ligand and Residues

Shusuke Tokutomi[1] tokutomi@eniac.scitec.kobe-u.ac.jp
Takahiro Hirata[1] thirata@eniac.scitec.kobe-u.ac.jp
Kohei Shimamura[1] shimamura@port.kobe-u.ac.jp
Kaori Fukuzawa[2] k-fukuzawa@hoshi.ac.jp
Shigenori Tanaka[1] tanaka2@kobe-u.ac.jp

[1] Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan
[2] Department of Physical Chemistry, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa, Tokyo 142-8501, Japan

Keywords: Janus kinase; Ligand binding; Fragment Molecular Orbital (FMO) method; Inter-Fragment Interaction Energy (IFIE); Machine learning

Inter-Fragment Interaction Energy (IFIE) obtained by Fragment Molecular Orbital (FMO) method can quantitatively measure the strength of interaction between ligand and a residue in protein, which is therefore useful for drug screening. However, it has not been clarified so far whether IFIE can be predicted using only geometrical (atomic distance) information in biopolymer complexes without resort to explicit FMO calculations.

In this research we have tested a hypothesis that the ligand-residue IFIE can be predicted by using a neural network model with the atomic distances between ligand and the surrounding 50 residues as the feature quantities (descriptors) in Janus kinase (JAK). We have used a JAK1-Tofacitinib complex (PDB code: 3EYG).

If this issue can be elucidated, we can expect the following three gains:

- To clarify what information is important for IFIE between ligand and residue in JAK.
- To obtain IFIE for various structures without FMO calculations.
- To anticipate that it is possible to construct high-precision molecular force field using FMO.

This research was conducted according to the following procedure:

1. Execution of Molecular Dynamics simulation to obtain 150 structures used for data set.
2. Execution of FMO calculation to obtain IFIE which is the target variable in this research.
3. Acquisition from 150 structures of distance information between ligand and residue atoms used as feature values.
4. Construction and evaluation of neural network model with MLP Regressor in Scikit-Learn.

In 3 and 4 above, by repeating trial and error, the optimal parameters were searched. This was performed for each 50 residues. The model was evaluated by dividing the training data set and test data set to 7:3 and using the value of the coefficient of determination (R^2).

The results were categorized into three types: residues that could be learned, those that were overlearned, and those that were not learned. The R^2 value for the residues that could be learned was as high as about 0.8 in the test data set. The fact that there are residues that can be well learned should show an ability of the present approach to improve the results for residues that are overlearned or unlearned.

In the poster presentation, we will discuss these three types and how to improve the accuracy of prediction for overlearned and unlearned cases.

Specific interactions between retinoic acid receptor-related orphan receptor and its ligands: molecular dynamics and *ab initio* fragment molecular orbital calculations

Shusuke Suzuki[1] s_suzuki@klab.cs.tut.ac.jp
Syunya Nakamura[1] nakamura@klab.cs.tut.ac.jp
Ryosuke Saito[1] saito@klab.cs.tut.ac.jp
Kentaro Kawai[2] kentaro.kawai@pharm.setsunan.ac.jp
Midori Takimoto-Kamimura[3] m.kamimura@teijin.co.jp
Noriyuki Kurita[1] kurita@cs.tut.ac.jp

1 Department of Computer Science and Engineering, Toyohashi University of Technology, 1−1 Hibarigaoka, Tempaku−cho, Toyohashi, Aichi, 441−8580, Japan
2 Faculty of Pharmaceutical Sciences, Setsunan University 45-1, Nagaotoge-cho, Hirakata, Osaka, 573-0101, Japan
3 Teijin Institute for Bio-Medical Research, Teijin Pharma Ltd., 4-3-2 Asahigaoka, Hino, Tokyo, 191-8512, Japan

Keywords: Molecular simulation; Molecular dynamics; Fragment molecular orbital; Protein ligand interactions; Conformational change; ROR

Retinoic acid receptor-related orphan receptor gamma (RORγT) is an important transcription factor involved in the differentiation of T helper cell 17 (Th17) and the production of the inflammatory cytokine, interleukin 17 (IL-17). Since IL-17 is considered to be a cause of inflammatory diseases, RORγT has drawn much attentions as an important drug target for the treatments of these diseases [1], and many therapeutic agents have been developed as a potent inhibitor against RORγT. A recent experimental study [2] revealed that minor structural change to tertiary sulfonamide RORγT ligands leads to opposite mechanism of action as agonist or inverse-agonist to RORγT. However, the reason for this significant difference in the actions of two similar RORγT ligands is not clarified yet.

In the present study, to elucidate the effect of ligand binding on the structure and conformation of helix 12 (H12) of RORγT, we performed molecular dynamics (MD) simulations and fragment molecular orbital (FMO) calculations for some RORγT+ligand complexes. As the target ligands, we employed agonist 3SN and inverse-agonist 3SX; 3SN enhances the RORγT activity, while 3SX suppresses the activity by controlling the conformation of H12 existing at the terminal of the ligand binding domain of RORγT.

At first, we optimized the structures of the complexes in explicit waters by molecular mechanics (MM) calculations and investigated the interactions between RORγT residues and the ligand by *ab initio* FMO calculations. The FMO results show that the interactions between RORγT and the ligands are similar to each other. Therefore, the reason for the opposite mechanism of actions of 3SX and 3SN cannot be explained by the molecular simulations based on MM and FMO methods.

To search widely for the stable conformation of H12 depending on the ligand bound to RORγT, we conducted MD simulations for the complexes in explicit waters. In addition, some characteristic structures were picked up from the MD snapshots, and the specific interactions between RORγT residues and the ligands were investigated by *ab initio* FMO calculations. The results of MD and FMO calculations will be shown in our poster.

[1] A. M. Jetten, S. Kurebayashi, and E. Ueda, “The ROR nuclear orphan receptor subfamily: critical regulators of multiple biological processes.,” *Prog. Nucleic Acid Res. Mol. Biol.*, vol. 69, pp. 205–47, 2001.
[2] O. René *et al.*, “Minor Structural Change to Tertiary Sulfonamide RORc Ligands Led to Opposite Mechanisms of Action,” *ACS Med. Chem. Lett.*, vol. 6, no. 3, pp. 276–281, Mar. 2015.

Validation of conformer generation in solution using NMR data.

Paul C. D. Hawkins[1]
phawkins@eyesopen.com

Stanislaw Wlodek[1]
stan@eyesopen.com

Jan Kihlberg[2]
jan.kihlberg@kemi.uu.se

Mate Erdelyi[2]
mate.erdelyi@kemi.uu.se

[1] OpenEye Scientific, 9 Bisbee Court, Santa Fe, NM 87501, USA.
[2] Department of Chemistry – BMC, University of Uppsala, Uppsala, SE-752 37, Sweden.

Keywords: Conformation generation, solution conformations, validation, NMR.

The quality of models for prediction of molecular properties in solution relevant to drug discovery is critically dependent on the quality of the conformations used to build those models. Unfortunately conformation generators have, historically, been validated against conformations found in the solid state and their performance in solution has simply been assumed to be the same as it is for the solid state.

We have therefore built a framework to validate and compare conformation generators using solution NMR data. This framework comprises input NMR conformational data (NOEs and scalar couplings), a conformation ensemble generator and a method to identify those conformations from the ensemble that best match the NMR data as well as their probabilities.

Part of this framework is the new tool NMR_FIT, which is a user-friendly and faster reimplementation of the well-established NAMFIS method.[1] We have validated NMR_FIT against the reference implementation on a variety of molecules, and used it to compare different methods of conformation generation in solution.

[1] Cicero, D. O.; Barbato, G.; Bazzo, R., NMR Analysis of Molecular Flexibility in Solution: A New Method for the Study of Complex Distributions of Rapidly Exchanging Conformations. Application to a 13-Residue Peptide with an 8-Residue Loop. *Journal of the American Chemical Society* **1995**, 117, 1027-1033.

Molecular simulations on aggregation mechanism of microtubule associated Tau proteins

Riku Sato[1] sato@klab.cs.tut.ac.jp
Katsumi Suzuki[1] k-suzuki@klab.cs.tut.ac.jp
Takuya Ezawa[1] ezawa@klab.cs.tut.ac.jp
Shusuke Suzuki[1] s_suzuki@klab.cs.tut.ac.jp
Karpov Pavel[2] karpov@nas.gov.ua
Sergiy Shulga[2] Shulga5@i.ua
Yaroslav Blume[2] blume@nas.gov.ua
Noriyuki Kurita[1] kurita@cs.tut.ac.jp

[1] Department of Computer Science and Engineering, Toyohashi University of Technology, 1−1 Hibarigaoka, Tempaku−cho, Toyohashi, Aichi, 441−8580, Japan
[2] Institute for Food Biotechnology and Genomics, National Academy of Sciences of Ukraine, 2a. Osypovskogo Str., Kyiv-123, 04123, Ukraine

Keywords: Alzheimer's disease, Microtubule associated protein, Tau protein, Paired helical filament, Aggregation, Molecular simulation, Molecular dynamics, Fragment molecular orbital

Alzheimer's disease (AD), a dementia characterized by memory impairment, disorientation and loss of judgment occur, has been a serious problem worldwide. There are various hypotheses about the cause of AD. In the tau hypothesis, the aggregation and accumulation of hyper-phosphorylated microtubule associated protein Tau (MAPT) is considered to be related with the onset of ADs. MAPT originally plays an important role in regulating microtubule stabilization. However, when it is hyper-phosphorylated, it aggregates to form an insoluble fiber called as a paired helical filament (PHF). Therefore, the elucidations of the MAPT aggregation mechanism as well as the aggregation properties of PHFs are expected to be helpful in the development of novel therapeutic agents for inhibiting the onset of ADs.

Since MAPT is an intrinsically disordered protein, it is difficult to determine its 3D structure by X-ray crystallography. However, recently, the 3D structure of PHF was identified by cryo-electron microscopy and registered in the Protein Data Bank (PDB ID: 6HRE)[1]. The structure is a hexamer of MAPT, in which three pairs of C-shaped MAPTs are stacked.

In the present study, we first carried out 100 ns molecular dynamics (MD) simulations at 300 K for the hexamer in an explicit water box, in order to investigate the stability of hexamer in water. In addition, we picked out some characteristic structures of the hexamer to investigate their electronic states using *ab initio* fragment molecular orbital (FMO) calculations. The same analyses were conducted for MAPT monomer, dimer paired in the direction perpendicular to helical axis and trimer stacked in the helical direction.

The MD simulation for the MAPT hexamer indicates that the stacked trimer keeps its structure and conformation, while the MD simulation for the MAPT monomer elucidates that some β sheet structures in MAPT were easily broken. Accordingly, it is suggested that C-shaped MAPT in the structure of PHF cannot be formed in the monomer state. The details of structural changes of monomer, dimer, trimer and hexamer will be shown in the poster. In addition, the change in the interactions between MAPT residues in these structures will be discussed.

[1] B. Falcon, *et al*., *Acta Neuropathologica*, 136: 699-708, 2018.

Selectivity of phosphodiesterase-10A inhibitor for phosphodiesterase family elucidated by free energy perturbation approach

Toru Ekimoto[1]
ekimoto@yokohama-cu.ac.jp

Chisa Yuasa[1]
i130699c@yokohama-cu.ac.jp

Tsutomu Yamane[1]
tyamane@yokohama-cu.ac.jp

Mitsunori Ikeguchi[1,2]
ike@yokohama-cu.ac.jp

[1] Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[2] RIKEN Medical Sciences Innovation Hub Program, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

Keywords: Molecular dynamics simulation, Binding free energy, Free energy perturbation, Phosphodiesterase

Tuning of binding selectivity to isoform proteins is one of the important issues in the drug design. Phosphodiesterase (PDE) hydrolyzes cyclic-AMP and cyclic-GMP, which play important roles in signal transductions as second messengers, and there are 11 isoforms in human. Because each isoform involves specific signal transductions, it is required for the drug design of PDE to improve not only binding affinity to the targeted PDE but also selectivity for other isoforms. In PDE10A, the selectivity pocket [1] is well known as an important area to generate selectivity, and a lot of drug candidates were designed so that the selectivity pocket is filled with a part of the ligand [2]. However, TAK-063 designed by Takeda pharmaceutical Co., Ltd. exhibits high selectivity despite the fact that TAK-063 does not bind to the selectivity pocket [3], suggesting other structural factors except for the selectivity pocket.

To understand structural factors for the selectivity, we calculated the binding-free-energy changes upon single mutations of residues at the binding site using free energy perturbation (FEP) exhaustively. By sequence alignment of 11 isoforms, residues with different sequences among PDE10A and isoforms were detected (28 residues in total). At each residue, FEP calculations of the single mutation were executed by the number of different sequences among PDE10A and other isoforms (85 mutations in total). From the FEP analyses, key residues generating the selectivity were detected successfully, and a combination of residues contributing to the selectivity for each isoform was specified. In addition, the detected residues could be classified in terms of structural features and the interaction manner with TAK-063.

[1] Verhoest, R.P. et al., *J. Med. Chem.*, 52:5188-5196, 2009.
[2] Chappie, A.T., *J. Med. Chem.*, 55:7299-7331, 2012.
[3] Kunimoto, J. et al., *J. Med. Chem.*, 57:9627-9643, 2014.

De Novo Binding Prediction using gREST

Suyong Re[2] suyongre@riken.jp
Ai Niitsu[1] ai.niitsu@riken.jp
Hiraku Oshima[2] hiraku.oshima@riken.jp
Motoshi Kamiya[3] motoshi.kamiya@riken.jp
Yuji Sugita[1,2,3] sugita@riken.jp

[1] Theoretical Molecular Science Laboratory, RIKEN Cluster for Pioneering Research, Hirosawa 2-1, Wako, Saitama 351-0198, Japan
[2] Laboratory for Biomolecular Function Simulation, RIKEN Center for Biosystems Dynamics Research, 6-7-1 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan
[3] Computational Biophysics Research Team, RIKEN Center for Computational Science, 6-7-1 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

S. Re and A. Niitsu are equally contributed to this work.

Keywords: Protein-ligand binding, Molecular dynamics simulation, generalized replica-exchange with solute tempering

Molecular Dynamics (MD) simulations treat molecular interactions and flexibility of a protein-ligand system at atomic resolution, and are increasingly applied to predict protein-ligand recognition. However, *de novo* prediction of protein-ligand recognition presents a significant challenge due to the timescale gap between MD simulations and actual protein-ligand bindings. Conventional simulations can observe only a few binding events even using MD-specialized supercomputers. In this work, we apply the generalized replica-exchange with solute tempering method, gREST [1], combined with a flat-bottom potential to overcome the difficulty. gREST is a generalization of REST [2] and provides a versatile framework for defining the solute. For protein-ligand interactions, the solute region is defined in terms of selected energy components of the ligand and selected protein binding site residues to enhance protein structural flexibility and expedite dynamics relevant to binding. The application of a flat-bottom potential centered on the cavity can further enhances the sampling of binding events by avoiding the irrelevant encounters. We apply the method to the ligand binding with T4 lysozyme Leu99Ala mutant [3]. The method samples bindings in multiple poses and correctly predicts X-ray structures. Based on free-energy profiles, we are able to predict whether the molecule binds to the target protein or not. Furthermore, a coupling of the bindings with protein conformations are clearly illustrated. The protocol is generally applicable to various proteins having buried cavities with limited access for ligands with different shapes, sizes and chemical properties.

[1] Kamiya, M.; Sugita, Y. J. Chem. Phys. 2018, 149, 072304.
[2] Wang, L. et al. J. Phys. Chem. B 2011, 115, 9431–9438. Moors, S. L. C. et al. J. Chem. Theory Comput. 2011, 7, 231–237. Terakawa, T. et al. J. Comput. Chem. 2011, 32, 1228–1234.
[3] Niitsu, A. et al. submitted.

P1-22

A binding free energy calculation method along a modified thermodynamic path which avoids exhaustive enumeration of multiple protein-ligand poses

Yoshitake Sakae[1] sakae@rist.or.jp

Bin W. Zhang[2] bin.w.zhang@temple.edu

Ronald M. Levy[2,3] ronlevy@temple.edu

Nanjie Deng[4] nanjie.deng@gmail.com

[1] Division of Computational Science and Technology, Research Organization for Information Science and Technology, 1-18-16
Hamamatsucho, Minato-ku, Tokyo 105-0013, Japan
[2] Center for Biophysics and Computational Biology, Temple University
Philadelphia, Pennsylvania 19122, United States
[3] Department of Chemistry and Institute for Computational Molecular Science, Temple University
Philadelphia, Pennsylvania 19122, United States
[4] Department of Chemistry and Physical Sciences, Pace University
New York, NY 10038, United States

Keywords: Binding free energy, Ligand docking, Double decoupling, Molecular simulation, Drug design

We propose a free energy calculation method for receptor-ligand binding, which have multiple binding poses that avoid exhaustive enumeration of the poses. For systems with multiple binding poses, the standard procedure is to enumerate orientations of the binding poses, restrain the ligand to each orientation, and then, calculate the binding free energies for each binding pose. In this study, we modify a part of the thermodynamic cycle in order to sample a broader conformational space of the ligand in the binding site. This modification leads to the more accurate free energy calculation without performing separate free energy simulations for each binding pose.

We applied our modification to several host-guest systems as tests. One is β-cyclodextrin and ligands, which contain the carbonyl group. Another is a protein-ligand system, phenol binding to T4 lysozyme. The results showed that the binding free energies obtained from our method without knowing the binding poses were in good agreement with the benchmark results obtained by explicit enumeration of the binding poses.

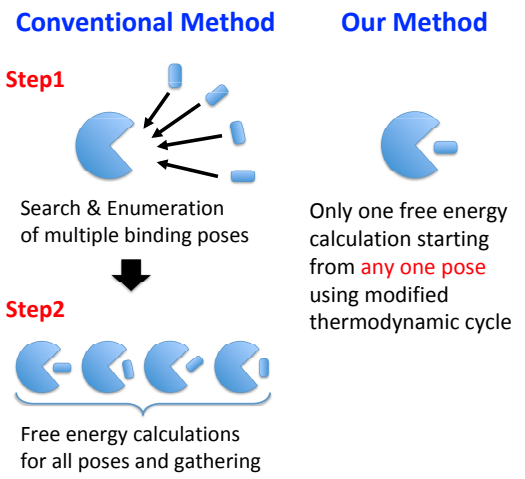


Figure: Procedures of alchemical binding free energy calculation methods of conventional (left) and our method (right).

[1] Sakae, Y., Zhang, B. W., Levy, R. M., and Deng, N., *Journal of Computational Chemistry*, submitted.

P1-23

Improvement of Carbohydrate Force Field for Molecular Dynamics

Makoto Ikejo[1]
ikejo@eniac.scitec.kobe-u.ac.jp

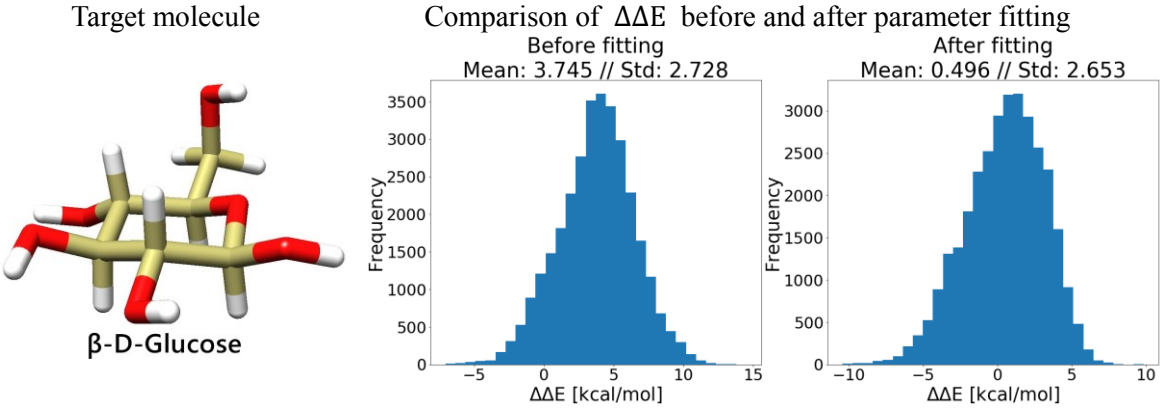
Hirofumi Watanabe[2]
hirofumi.watanabe@port.kobe-u.ac.jp

Shigenori Tanaka[1]
tanaka2@kobe-u.ac.jp

[1] Department of Computational Science, Graduate School of System Informatics, Kobe University, 1-1, Rokkodai, Nada, Kobe 657-8501, Japan
[2] Education Center on Computational Science and Engineering, Kobe University, 7-1-48, Minatojimaminamimachi, Chuo-ku, Kobe 650-0047, Japan

Keywords: Force field, Glycan, Parameter fitting, Machine learning

Although sugar chains have been actively studied in recent years as the “third” biological molecule, details of their structures and functions have not been elucidated yet. Molecular dynamics (MD) simulation is a useful tool to clarify the relationship between structure and function of biomolecules, and force field for carbohydrate called GLYCAM [1] is introduced in MD calculation software AMBER. In GLYCAM force field, force field parameters are determined to reproduce the energy variation of quantum mechanics (QM) calculation. However, when comparing the relative energies of QM and MM calculations, it is observed that deviations of about 4 kcal/mol have occurred. In this study, in order to better reproduce the energy of QM calculation (DFT / 6-31 ++G (2d, 2p)) for β-D-Glucose molecule in vacuum, the atomic charge and dihedral angle parameters were newly fitted using a machine learning library called TensorFlow [2] (see Figure).



[1] Kirschner, K. N., et al., GLYCAM06: a generalizable biomolecular force field. Carbohydrates, *J. Comput. Chem.*, **29**(4):622-55, 2008.

[2] M Abadi, *et al.*, Tensorflow: A system for large-scale machine learning, 12th USENIX Symposium on OSDI’16, 2016.

Systematic construction of the cosolvents sets for cosolvent MD (CMD) with the large-scale computation

Keisuke Yanagisawa[1]
yanagisawa@bi.a.u-tokyo.ac.jp

Yoshitaka Moriwaki[1]
moriwaki@bi.a.u-tokyo.ac.jp

Tohru Terada[2 3]
tterada@iu.a.u-tokyo.ac.jp

Kentaro Shimizu[1 2]
shimizu@bi.a.u-tokyo.ac.jp

[1] Department of Biotechnology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan
[2] Interfaculty Initiative in Information Studies, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[3] Agricultural Bioinformatics Research Unit, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

Keywords: cosolvent MD, probability map, structural similarity, large-scale computation

Cosolvent MD (CMD) is an MD simulation of a protein in explicit water molecules mixed with cosolvent molecules. The simulation has been used for hotspot detection [1,2], binding site identification [3] and binding energy estimation [3,4]. Existing methods utilize small molecules which represent functional groups of compounds, such as isopropanol and benzene. Mahmoud *et al.* demonstrated that the spatial probability distribution map of cosolvent atoms (Pmap) was highly dependent on its chemical context and suggested that a wider variety of cosolvent molecules should be used in CMD [4]; however, the number of cosolvents is practically limited by the computational resources available. For efficient use of the resources, we need a non-redundant set of the cosolvent molecules that covers as wide a range of chemical space as possible.

In the present study, we extracted typical functional groups from FDA approved drugs and generated >100 cosolvent structures. For each cosolvent molecule, we conducted 10-ns CMD simulations 20 times to generate a Pmap. After that, Pmap similarity-based hierarchical clustering was done to analyze the relationship between the cosolvent structure and the Pmap. We found that cosolvents having similar structures tended to be clustered into the same cluster, suggesting that Pmap similarity could be predicted from structural similarities without time-consuming simulations. We here present a method to construct the optimum set of cosolvent molecules for given computational resources using the result of the hierarchical clustering.

[1] Ghanakota P *et al.*, Large-Scale Validation of Mixed-Solvent Simulations to Assess Hotspots at Protein−Protein Interaction Interfaces, *J. Chem. Inf. Model.* 58:784−793, 2018.
[2] Ung PMU *et al.*, Identifying Binding Hot Spots on Protein Surfaces by Mixed-Solvent Molecular Dynamics: HIV-1 Protease as a Test Case, *Biopolymers* 105:21−34, 2016.
[3] Ustach VD *et al.*, Optimization and Evaluation of Site-Identification by Ligand Competitive Saturation (SILCS) as a Tool for Target-Based Ligand Optimization, *J. Chem. Inf. Model.* 59:3018−3035, 2019.
[4] Mahmoud AH *et al.*, Improving Atom-Type Diversity and Sampling in Cosolvent Simulations Using λ-Dynamics, *J. Chem. Theory Comput*. 15:3272−3287, 2019.

A Method for Comparing Structural Ensembles: Applications to Molecular Dynamics Trajectory Data

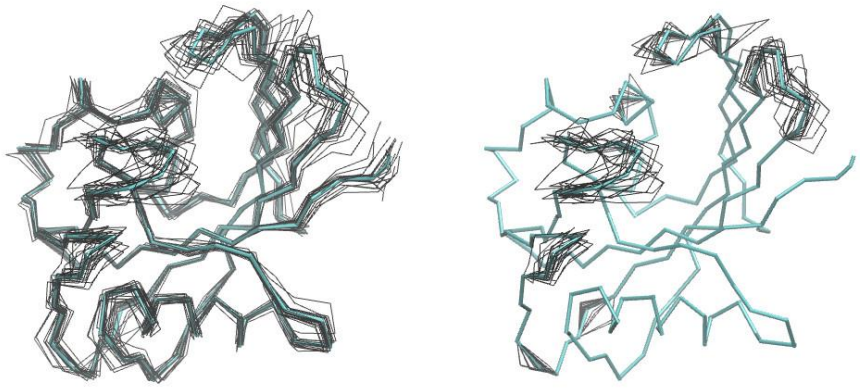
Takashi Amisaki[1]
amisaki@tottori-u.ac.jp

[1] Department of Biological Regulation, Faculty of Medicine, Tottori University, 86 Nishi-cho, Yonago, Tottori 683-8503, Japan

Keywords: Structural superposition, Random effects model, L1 regularization

Molecular dynamics simulations are often used to discover the outcomes of adding some perturbations to a protein molecular system. The perturbations include, for example, amino acid mutations and variations on ligands. The impacts of the perturbations may be investigated by two-stage superposition of conformers in the ensembles (trajectories). Namely, conformers are first superimposed within each ensemble and then the average structures of the ensembles are superimposed. Deviations seen on thus superimposed structures may imply differences due to the perturbations. On the other hand, the method proposed in this work simultaneously superimposes the all conformers using a random effects model (REM) distinguishing between inter- and intra-ensemble variabilities. In addition, distinctions between ensembles can be highlighted by means of L1 regularization on inter-ensemble variability. The geometrical parameters (e.g., rotation matrices) and variance components are jointly estimated/shrunken. In other words, the method seeks portions for intra-ensemble variability while adjusting the orientations of the conformers. If this regularization is turned off and the sizes of ensembles are large enough, then the estimates by the random effects model coincide to those by the two-stage method.

The proposed method was applied to twenty trajectories, each 42ns long, of hMTH1 which is an enzyme that sanitizes nucleotide pools. The trajectories were obtained by twenty simulations under different ligand-binding statuses and/or different protonation states of Asp119 and Asp120. The two aspartate residues have been reported to relate to the recognition of ligand molecules. Two results are shown in the figure. The left panel shows the REM superposition of the twenty ensembles (42 conformer each) without the L1 regularization on inter-ensemble variability. The thin lines are the ensemble averages. The thick line represents the overall average structure. Large inter-ensemble variability can be seen at the loops as well as at the helices. By contrast, in the right panel, which is the result under the regularization with a certain strength, large variability was limited mainly to the loops. The results exemplified the effectiveness of the proposed method.



Fast prediction of binding hotspots on a peptide ligand in complex by Digital Annealer

Yoshiaki Tanida[1]
tanida.yoshiaki@fujitsu.com

Kentaro Takai[2]
takai.kentaro@fujitsu.com

Sakiko Mori[2]
mori.sakiko@fujitsu.com

Hiroyuki Sato[1]

Mutsuyo Wada[2]

Atsushi Tomonaga

Azuma Matsuura[1]

Akihiko Ueda[2]

Seiji Fukui[2]

Nozomu Kamiya[2]

Shunji Matsumoto[2]

Hirotaka Oshima[1]

[1]Fujitsu Laboratories Ltd., Atsugi, Kanagawa 243-0197, Japan
[2]Fujitsu Ltd., 1-2-70 Konan, Minato-ku, Tokyo 108-0075, Japan

Keywords: ligand-receptor complex, cyclic peptide, binding hotspot, combinatorial optimization problem, digital annealer

In middle molecule drug discovery, understanding peptide-protein interactions are of paramount importance for computer-based de novo design. Residue properties in complex are particularly key in binding affinity of a peptide to a target protein. Thus, it is necessary to calculate the interactions between residues and protein to understand their binding features and find out which regions should be modified to improve the affinity. However, calculating the interaction between middle molecules and proteins from the first principles is very time-consuming. To solve this problem, we assume that if the peptide molecule in the complex is bound in its natural form, it gives strong binding [1, 2].

In this report, we propose a simple method to fast search for hotspots in peptide-protein binding based on a packing optimization [3, 4]. Furthermore, our method can be easily accelerated by special purpose hardware, so called “digital annealer”.

[1] M-H Hao, O. Haq and I. Muegge, J. Chem. Inf. Model., 47, 2242 (2007).
[2] K. A. Brameld,†B. Kuhn,‡D. C. Reuter and M. Stahl, J. Chem. Inf. Model, 48, 1 (2008)
[3] C. Lee and S. Subbiah, J. Mol. Biol., 217, 373 (1991).
[4] R. Tanimura, A. Kidera and H. Nakamura, Protein Science, 3, 2358 (1994).

Advanced methods to predict the property of cyclic peptides: reproduction of actual conformations

Sakiko Mori[1] mori.sakiko@fujitsu.com
Kentaro Takai[1] takai.kentaro@fujitsu.com
Yoshiaki Tanida[2] tanida.yoshiaki@fujitsu.com
Atsushi Tomonaga **Hiroyuki Sato**[2] **Azuma Matsuura**[2] **Hirotaka Oshima**[2]
Mutsuyo Wada[1] **Akihiko Ueda**[1] **Seiji Fukui**[1] **Shunji Matsumoto**[1] **Nozomu Kamiya**[1]

[1] Fujitsu Ltd, 1-2-70 Konan, Minato-ku, Tokyo 108-0075, Japan
[2] Fujitsu Laboratories Ltd., Atsugi, Kanagawa 243-0197, Japan

Keywords: conformation search, cyclic peptides, Digital Annealer

Drug discovery in middle sized molecules such as cyclic peptides is an attracting area for their particular properties against small molecules or antibodies. It is important to examine the conformation related to properties of cyclic peptides for the discovery [1]. As peptides have high internal degree of freedom compared to small molecules, the systematic search needs extensive calculations. Therefore, an exhaustive and efficient conformation analysis is required.

Our developed conformation analysis methods consist of three steps.

(i) Generation of cyclic glycine conformers by our “Ring Expansion Method” [2].
(ii) Substitution of an alanine or a proline residue for a glycine residue and addition of the methyl group to –NH of amide bonds in a way corresponding to a target cyclic peptide (called the main chain conformer).
(iii) Sequential addition of side chain conformers to the main chain conformer.

To check the accuracy of our system, we performed the conformation analysis for the 7-mer peptide (DQSEPHP) measured by NMR in water (PDB entry 6BEW), and got stable conformers contained similar one to the NMR structure. The lowest RMSD between PDB and calculated conformers is 1.42 Å under the dielectric constant (DC) set to 80 (the experimental condition).

In the case of the larger peptides, Gramicidin S (PDB entry 1TK2) and Cyclosporin A (PDB entry 1CWA), the lowest RMSD are 1.48 Å and 0.67 Å (Fig. 1), respectively. Here, we set the DC to 4 since these PDB structures were measured in the complex states with proteins.

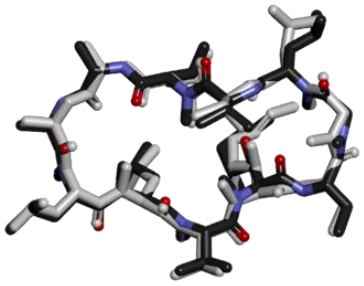


Figure 1: The closest conformer (carbon: black, nitrogen: blue, oxygen: red) to the X-ray structure (white line) by superposing the backbone.

Striving for 15-mer peptides, we are energetically developing our technique applying the special purpose hardware, “Digital Annealer”.

[1] A. Whitty, M. Zhong, L. Viarengo, D. Beglov, D. R. Hall and S. Vajda, *Drug Discovery Today*, 21, 712 (2016).
[2] A. Tomonaga, H. Sugiyama and A. Ueda, WO2018/154789A1.

Advanced methods to predict the property of cyclic peptides: exhaustive and efficient conformation search

Kentaro Takai[1] takai.kentaro@fujitsu.com
Sakiko Mori[1] mori.sakiko@fujitsu.com
Yoshiaki Tanida[2] tanida.yoshiaki@fujitsu.com

Atsushi Tomonaga **Hiroyuki Sato**[2] **Azuma Matsuura**[2] **Hirotaka Oshima**[2]

Mutsuyo Wada[1] **Akihiko Ueda**[1] **Seiji Fukui**[1] **Shunji Matsumoto**[1] **Nozomu Kamiya**[1]

[1] Fujitsu Ltd, 1-2-70 Konan, Minato-ku, Tokyo 108-0075, Japan
[2] Fujitsu Laboratories Ltd., Atsugi, Kanagawa 243-0197, Japan

Keywords: conformation search, cyclic peptides, database, Digital Annealer

Drug discovery in middle sized molecules such as cyclic peptides is an attracting area for their particular properties against small molecules or antibodies. As it is known that the properties of a cyclic peptide depend on its conformation[1], we developed a new "Ring Expansion Method"[2] for the generation of cyclic glycine conformers and a sequential addition method of side chain conformers using their databases (patent pending) to perform the exhaustive but practical conformation analysis.

Our new conformation analysis methods consist of three steps.

(i) Generation of cyclic glycine conformers by "Ring Expansion Method" and creation of their conformer databases. They are used in step (ii) and also reused for new targets.
(ii) Substitution of an alanine or a proline residue for a glycine residue and addition of the methyl group to –NH of amide bonds in a way corresponding to a target cyclic peptide (called the main chain conformer).
(iii) Sequential addition of side chain conformers to the main chain conformer using the databases of side chain conformers calculated in advance.

In step (i), we searched conformers of a cyclic 11-mer glycine to analyze the conformation of Cyclosporin A (a 11-mer peptide). Using our method, the calculation time is drastically reduced from 380 years (estimated) of the existing systematic search to 56 days (actual). While the systematic search is exhaustive for even high-energy conformers not contributing to properties to be measured, our search is exhaustive for only lower-energy ones.

In step (ii) and (iii), we got stable 11,472 conformers of Cyclosporin A in 4.6 days using 400 parallel calculations and the databases of cyclic glycine conformers and side chain ones. The latter databases help us to search stable conformers efficiently; the number of conformers generated by the systematic search is 1.1 billion, but it is reduced to less than 1.4 million for Cyclosporin A.We calculated the RMSD to PDB (1CWA) to confirm the validity of our methods and we found the lowest RMSD is 0.67 Å.

Aiming to accelerate the conformation search and realize the analysis for cyclic 15-mer peptides, we are energetically developing our technique applying the packing optimization method[3,4] and the special purpose hardware, "Digital Annealer".

[1] A. Whitty, M. Zhong, L. Viarengo, D. Beglov, D. R. Hall and S. Vajda, *Drug Discovery Today*, 21, 712 (2016).
[2] A. Tomonaga, H. Sugiyama and A. Ueda, WO2018/154789A1.
[3] C. Lee and S. Subbiah, *J. Mol. Biol.*, 217, 373 (1991).
[4] R. Tanimura, A. Kidera and H. Nakamura, *Protein Science*, 3, 2358 (1994).

Conformational Changes and Interactions of Calcium Ion Signal Transfer Protein Calmodulin and Calmodulin-binding Domain by Multi-scale and Docking Simulation

Hiromitsu Shimoyama [1]
shimoyamah@pharm.kitasato-u.ac.jp

Yasushige Yonezawa [2]
yonezawa-wk@waka.kindai.ac.jp

[1] School of Pharmacy, Kitasato University, Shirokane 5-9-1, Minato-ku, Tokyo 108-8641, Japan
[2] Institute of Advanced Technology, Kindai University, Wakayama 649-6493, Japan

Keywords: signal transduction, conformational change, molecular dynamics, rigid-body docking

Calmodulin (CaM) is a major calcium ion signal transfer protein in many cellular processes. CaM undergoes large conformational change upon calcium ion binding. The conformational change is necessary to form complex with its target protein's binding domain (CaMBD; about 20-residue long peptide). However, the process of conformational change and complex forming are still not well understood.

Recently I analyzed various structures of CaM and clarified interactions driving the conformational change [1]. In this work, mechanism of CaM's conformational change and binding to CaMBD is studied by combined multi-scale molecular dynamics and rigid-body docking. Our method would be able to analyze general protein-protein interactions involved in most cellular signal transductions.

Most cellular signal transductions are protein-protein binding/dissociations. Such phenomena are difficult to treat by molecular dynamics (MD) because binding/dissociations in cellular environment would be repeatedly try-and-error processes.

For such phenomena, a rigid body docking is significantly useful because the docking can obtain complex structures in significantly short time compared with MD. Though the docking is useful, protein's flexibility should be considered. Then, we tried to combine the docking and MD to realize the high performance complex-forming simulation method that considers the flexibility.

First, both protein's conformations are coarsely sampled by coarse-grained model (CGM). After that, CGM structures are reconstructed at atomistic resolution (AAM) by Homology modeling. Then, all-to-all AAM docking of CaM and CaMBD are performed. The flexibility of protein are considered by CGM-MD and translated to AAM, whereas the docking reduces computational cost of binding/dissociation try-and-error.

The structure of the highest docking score is shown in Fig. 1 right (left structure is the native structure PDB ID 4Q5U). As can be seen from the figure, CaMBD is bound in the similar position of native complex. Further MD starting at the right structure would efficiently sample conformations around the native complex and help understanding the binding process.

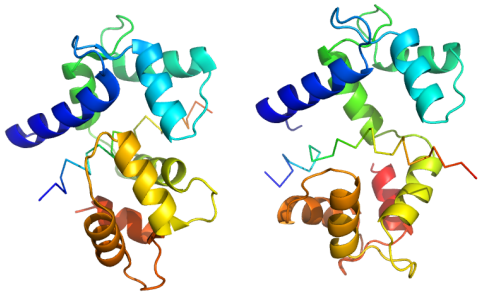


Fig. 1 the native structure (left) and the simulated structure (right)

[1] Shimoyama H., *A structural comparison of 'real' and 'model' calmodulin clarified allosteric interactions regulating domain motion*, Journal of Biomolecular Structure and Dynamics, 2018.

Autoencoder-based Analyses of Dynamic Allostery on Proteins by Regulator Binding

Yuko Tsuchiya[1]
yuko.tsuchiya@aist.go.jp

Kei Taneishi[2]
kei.taneishi@riken.jp

Yasushige Yonezawa[3]
yonezawa-wk@waka.kindai.ac.jp

[1] Artificial Intelligence Research Center, Advanced Industrial Science and Technology, 2-4-7 Aomi, Koto-ku, Tokyo, 135-0064, Japan
[2] Center for Advanced Photonics, RIKEN, 2-1 Hirosawa, Wako, Saitama, 351-0198, Japan
[3] High Pressure Protein Research Center, Institute of Advanced Technology, Kindai University, 930 Nishimitani, Kinokawa, Wakayama, 649-6493, Japan

Keywords: Autoencoder, Molecular dynamics simulation, Dynamic allostery, Molecular mechanisms of disease

Analyses of subtle changes in side-chain dynamic motion, and not large rigid conformational changes, are essential to understand the regulation of protein functions. We propose an autoencoder-based method that can detect dynamic allostery, based on the comparison of protein fluctuations, using distance matrices obtained from the molecular dynamics simulations in regulator-bound and -unbound forms. The method detected that the changes in dynamics by regulator binding led to the reorganization of correlative fluctuations among residue pairs. Other correlative motions were also found as a result of dynamic perturbation. Our method would be usefully applied to the analyses of molecular mechanisms in signal transduction and mutagenesis systems that are involved in protein functions and severe diseases.

[1] Tsuchiya, Y., Tanekshi, K., Yonezawa, Y., Autoencoder-based detection of dynamic allostery triggered by ligand binding based on molecular dynamics, *Journal of Chemical Information and Modeling*, in press.

Molecular dynamics simulation of drug efflux in multidrug ABC transporter

Takumi Someya[1]
w185419d@yokohama-cu.ac.jp

Kei Moritsugu[1]
moritugu@yokohama-cu.ac.jp

Akinori Kidera[1]
kidera@tsurumi.yokohama-cu.ac.jp

Hiroaki Kato[2]
katohiro@pharm.kyoto-u.ac.jp

1 Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
2 Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshida Shimoadachi-cho, Sakyo-ku, Kyoto 606-8501, Japan

Keywords: Multidrug ABC transporter, Drug efflux mechanism, Molecular simulation, Path connecting two crystal structures

Multidrug ATP binding cassette (ABC) transporter, ABCB1, is known as one of the most important molecules regulating the ADME properties of drugs. ABCB1 transports a broad spectrum of lipophilic drugs during a large scale structural change from the inward-open to the outward-open structure, using ATP hydrolysis as the driving force. To investigate the transport mechanism as well as the chemical property determining transport activity, we performed molecular dynamics (MD) simulations tracing the drug transport process accompanied with the structural change, based on the crystal structures of the inward-open and the outward-open forms of CmABCB1 (a eukaryotic homolog of human ABCB1 from *Cyanidioschyzon merolae*), which have recently been solved by Kodan *et al*. [1,2].

The simulations of the transport process were carried out along paths connecting the inward-open to the outward-open structure, using the targeted MD. During the simulation, the protein was forced to move to the target structure, while the drug molecule was translocated solely by the change in protein environment without any influence from the targeting force. Using Rhodamine 6G as the substrate, we examined how the transport behavior was affected by the initial position of the drug and the path of the structural change, and then extended the simulation to the transport of various drugs with different activity. In the presentation, we will show the analysis results on the mechanism how the protein squeezes out a drug molecule, focusing on the detail interactions and solvation, and will compare the transport activities of various drugs.

[1] Kodan, A., Yamaguchi, T., Nakatsu, T., Sakiyama, K., Christopher, J. H., Fujioka, A., Hirokane, R., Ikeguchi, K., Watanabe, B., Hiratake, J., Kimura, Y., Suga, H., Ueda, K., Kato, H., Proc.Natl.Acad.Sci. 111, 4049 (2014).
[2] Kodan, A., Yamaguchi, T., Nakatsu, T., Matsuoka, K., Kimura, Y., Ueda, K., Kato, H., Nature Commun. 10, 1 (2019).

P1-32

Characterization of water-molecule interaction based on fragment molecular orbital method

Hiromichi Tsurui [1]
tsurui@juntendo.ac.jp

[1] Department of Immunological Diagnosis, Juntendo University School of Medicine. 2-1-1, Bunkyo-ku, Tokyo 113-8421, Japan

Keywords: FMO, IFIE, three-body interaction, water-water interaction

Interaction of water molecules are characterized by fragment molecular orbital (FMO) calculation. Water molecules of 2842 were distributed spherically with the diameter of 54 nm. Clusters of 4, 8, 12, 16 and 20 molecules on lattice points at every 60° longitude, 30° latitude and 4.5 nm, 193 in total containing the center, were calculated for their interaction based on FMO2 and FMO3 method. Calculation were performed on ABINIT-MP implemented on the K computer at Riken. There were a little difference of the interaction energies, namely 0.5 ~ 0.8%, between the calculation of FMO2 and FMO3, and it gradually increased as the size of cluster increased. The interaction energies of clusters in absolute value increased almost linearly as the size of cluster increased. Interaction energies of the same cluster size apparently seemed to increase in absolute value as the radius of the distribution shell increased, however, this finding requires more intensive analysis. Effect of protein on the interaction of surrounding water molecules are under analysis, and hopefully can be reported on the meeting.

P1-33

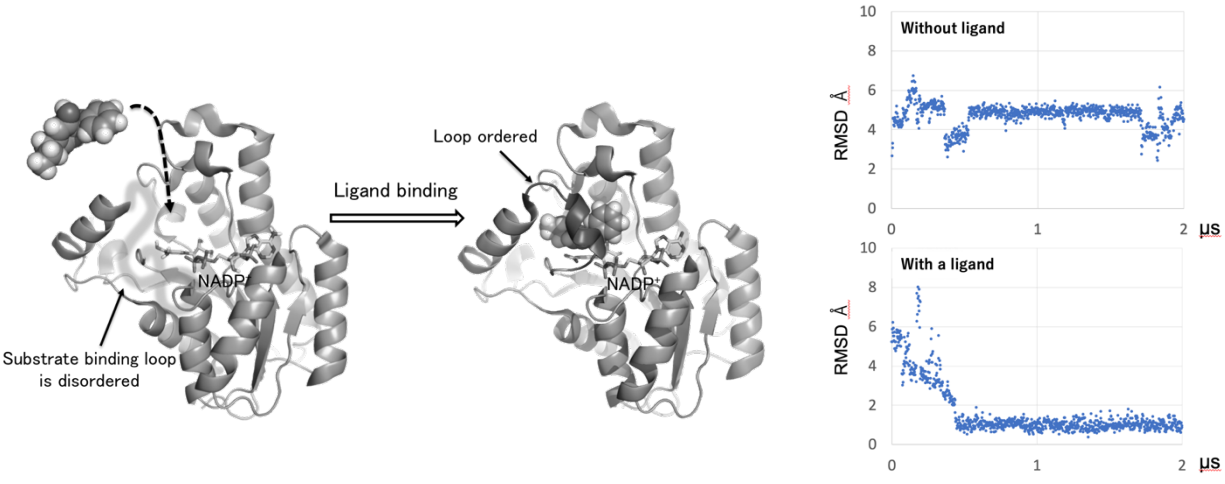
Computational approaches to drug-receptor binding kinetics

Osamu Ichihara[1]
osamu.ichihara@schrodinger.com

[1] Schrödinger K.K., 17F Marunouchi Trust Tower North, 1-8-1 Marunouchi, Chiyoda-ku, Tokyo 100-0005, Japan

Keywords: Binding kinetics, Residence time, MD simulation, Association rate, Dissociation rate

It has been proposed by Pearlstein *et al.* that desolvation and resolvation of the binding sites of proteins can be a key determinant of ligand binding kinetics. They have shown that, for a set of inhibitors of p38 MAP kinase, calculated desolvation and resolvation costs during ligand binding are qualitatively consistent with observed on and off rates respectively. Previously, we extended the application of their hypothesis further to rationalizing the binding kinetics (residence time) of the inhibitors of *Staphylococcus aureus* enoyl-ACP reductase (saFabI), where more complex two-step binding mechanism is suggested to be in operation. It has been proposed that, on ligand binding, saFabI may undergo a directional structural change, i.e. ordering of the substrate-binding loop to form an α-helical structure, and thereby locks the inhibitor into the cavity and increases its residence time. We calculated a set of thermodynamic parameters associated with the solvation properties around the ligands bound to saFabI using molecular dynamics simulations (WaterMap). The observed rate of ligand dissociation (residence time) are shown to be in excellent correlation with the parameters calculated.



Ligand biding loop Cα RMSD to the X-ray structure

In this work, in order to obtain more insight into the role of water molecules in the ligand binding event, we have performed a series of unbiased μ-second MD simulations to investigate the movement of the disordered substrate-binding loop, in the presence or absence of a ligand in the binding site. The MD simulations showed that the disordered loop was folded into the ordered state only in the presence of the ligand in the simulation time frame of a few hundred ns to μs. The RMSDs between the X-ray structure of the ordered loop and the structures observed in the simulations were ~1Å. On the other hand, the disordered loop failed to settle into the well-defined ordered state in the absence of the ligand, being consistent with the reported crystallographic studies. More details of the MD simulations and the role of water molecules will be discussed.

Investigation of stabilization mechanism of amorphous solid dispersion by fragment molecular orbital calculation

Xiaohan Ma[1] ahga4347@chiba-u.jp
Keisuke Ueda[1] keisuke@chiba-u.jp
Kenjirou Higashi[1] ken-h@faculty.chiba-u.jp
Kaori Fukuzawa[2] k-fukuzawa@hoshi.ac.jp
Etsuo Yonemochi[2] e-yonemochi@hoshi.ac.jp
Kunikazu Moribe[1] moribe@faculty.chiba-u.jp

[1] Laboratory of Pharmaceutical Technology, Graduate School of Pharmaceutical Sciences, Chiba University 1-8-1, Inohana, Chuo-ku, Chiba, 260-8675, Japan
[2] Hoshi University, 2-4-41 Ebara, Shinagawa-Ku, Tokyo 142-8501, Japan

Keywords: Amorphous solid dispersion, Molecular dynamics, Fragment molecular orbital calculation

α-Glucosyl rutin (Rutin-G), a new functional food additive, has been proven to improve the dissolution and bioavailability of poorly water-soluble drugs. In our previous study, carbamazepine (CBZ) was used as a model drug. The amorphous solid dispersion (ASD) of CBZ and Rutin-G effectively stabilized the amorphous state of CBZ and increased its solubility. However, the stabilization mechanism was still unknown. The present study aims to investigate the specific interaction between CBZ and Rutin-G which contributes to the stabilization of ASD by fragment molecular orbital (FMO) calculation. An ASD of CBZ and Rutin-G was constructed by molecular dynamics (MD) simulation following the experimentally melt-quenching method. The similar glass transition temperatures (T_gs) obtained from MD simulation and the experiment indicated that the simulated ASD was reasonable. In this ASD, hydrogen bonds were formed at around Tg, followed by the π-π interaction during the cooling process. The interaction energy of the ASD was calculated by FMO calculation. The electrostatic energies dominated the interaction energy between CBZ and sugar groups of Rutin-G, while the electrostatic and dispersion energies were comparable in the interaction energy between CBZ and flavone skeleton of Rutin-G. In general, electrostatic energy is the main components of hydrogen bonding energy, while the CH/π and π-π interaction energies are predominated by dispersion energy. The total molecular interaction energy between CBZ and the flavone skeleton of Rutin-G was slightly higher than that between CBZ and the sugar groups. Therefore, hydrogen bonds as well as π-π interactions mainly existed between CBZ and the flavone skeleton of Rutin-G in ASD, which could contribute to the stabilization of ASD. FMO calculation is considered to be a potential tool to interpret the stabilization mechanism of ASD.

Interaction analysis of HIV protease-inhibitor complex by fragment molecular orbital method

Hirofumi FUJI[1] 1933310128s@kindai.ac.jp
Norihito KAWASITA[1,2] nkawashita@life.kindai.ac.jp

[1] Graduate School of Science and Engineering, Kindai University, 3-4-1 Kowakae, Higashiosaka city, Osaka, 577-8502, Japan
[2] Faculty of Science and Engineering, kindai University, 3-4-1 Kowakae, Higashiosaka city, Osaka, 577-8502, Japan

Keywords: fragment molecular orbital method, inter-fragment interaction energy, enthalpy

Development of protein-protein interaction inhibitors by small molecule inhibitors has become an attractive drug discovery target. However, it is still difficult to predict binding free energy of novel inhibitors having strong activity in docking simulation. In recent years, the fragment molecular orbital (FMO) method[1] has been used to analyze the interaction between proteins and inhibitors and the sum of inter-fragment interaction energies (IFIE) of protein-inhibitor obtained by FMO calculation can evaluate static interactions into consideration the inter-electron interaction. If there is a high correlation between IFIE and ΔH by FMO calculation, a regression expression between ΔH and IFIE can be constructed. Compensation law is established for ΔH and $T\Delta S$[2], so that finally ΔG can be predicted only IFIE. Therefore, it is possible to establish a method that can predict ΔG without considering entropy.

This study examined the correlation between IFIE and ΔH in the HIV protease inhibitor system. First, 14 structures of the HIV protease inhibitor complex in the literature[3-5] were downloaded from the protein databank (PDB), and the structure was corrected by BioStationViewer, DiscoveryStudioViewer, and MOE. In structural optimization, CHARMm was used as a force field. Next, the calculation was carried out by MP2/6-31G, using supercomputer K and FX100 and using ABINIT-MP 6.0+ for software. Finally, using the sum of IFIE obtained by calculation and the experimental value (ΔH) in the literature[3-5], it was examined whether a correlation was established between IFIE and ΔH.

Confirming the correlation between IFIE obtained by FMO calculation and the experimental value, a good correlation of R^2=0.5326 was obtained. Furthermore, classifying the inhibitors charge into 0 and +1 and confirming the correlation, a better correlation value of R^2=0.8071 was obtained in the positively charged inhibitor complex. When confirming between the inhibitor and the amino acid residue, there was a difference in IFIE between inhibitor and GLY49 in each complex. When confirmed the structure, there were complexes with and without water molecules between the inhibitor and GLY49, so we classified the complex with water as group 1 and the complex without water as group 2. After that, a better correlation of R^2=0.9754 for group 1 and R^2=0.7272 for group 2 was obtained. From these results, by calculated the complex at the same conditions, it could be obtained the accurate correlation between IFIE and ΔH needed for the ΔG prediction method.

This research was done in activities of the FMO drug design consortium (FMODD). The results were obtained using the K computer (project ID: hp180147). This work was supported by Grant-in-Aid for Scientific Research (C) from JSPS KAKENHI (No. 18K06615),

[1] Kitaura K. *et al*., *Chem. Phys. Lett*. 312, 319-324, (1999).
[2] John D. C. *et al*., *Annu. Rev. Biophys*. 42, 121–142, (2013).
[3] Nancy M. K. *et al*., *ACS Chem. Biol*., 7, 1536-1546, (2012).
[4] Milan K. *et al*., *J. Mol. Biol*, 374, 1005-1016, (2007).
[5] Milan K. *et al*., *FEBS Journal*, 281, 1834-1847, (2014).

Assessment of a simple pKa estimation scheme for drug molecules by quantum chemical calculations

Takao Otsuka[1] totsuka@riken.jp
Noriaki Okimoto[1] okimoto@riken.jp
Makoto Taiji[1] taiji@riken.jp
Toru Matsui[2] matsui@chem.tsukuba.ac.jp

[1] RIKEN Center for Biosystems Dynamics Research (BDR), Building B, 6-2-4 Furuedai, Suita, Osaka 565-0874, Japan
[2] Department of Chemistry, Graduate School of Pure and Applied Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan

Keywords: acid dissociation constant, pKa, electronic structure calculation,

Computational prediction of acid dissociation constant, pKa is so important for pharmacokinetics of drug molecules and/or rational drug design. There are some theoretical studies for evaluating the pKa, such as empirical parameterized approaches [1], classical or first-principles molecular dynamics simulations [2,3], and electronic structure calculations [4]. Recently, Matsui and co-workers have proposed a simple pKa evaluation scheme based on quantum chemical calculations [5]. Although this scheme needs the experimental values, it can give the semi-quantitative pKa values of specific functional groups. They have also demonstrated the availability for amino acids and/or small proteins by using this scheme.

In this study, we report a study on this pKa evaluation scheme for applying to drug molecules. In order to be able to apply it to drug molecules, we assessed the models for drug molecules and the computational conditions for quantum chemical calculations. We show the results by some molecules systems.

[1] H. Li, A. D. Robertson and J. H. Jensen, Proteins, 2005, 61, 704.
[2] For example, A. H. Juffer and H. J. Vogel, Proteins, 2000, 41, 554.
[3] For example, M. Sprik and G. Ciccotti, J. Chem. Phys., 1998, 109, 7737.
[4] For example, J. Ho and M. L. Coote, Theor. Chem. Acc., 2010, 125, 3.
[5] T. Matsui, T. Baba, K. Kamiya, Y. Shigeta, Phys. Chem. Chem. Phys., 2012, 14, 4181.

Is hydrophobic group in osmolyte hydrophilic? (2): A time series interaction energy decomposition analysis study by means of effective fragment potential molecular dynamics simulation

Tamon Funakura[1]
a14.fp3h@g.chuo-u.ac.jp

Yukina Uchino[2]
g1940628@edu.cc.ocha.ac.jp

Nahoko Kuroki[3,4]
kuroki.91d@g.chuo-u.ac.jp

Hirotoshi Mori[1,5]
qc-forest.19d@g.chuo-u.ac.jp

[1] Department of Applied Chemistry, Faculty of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan
[2] Department of Chemistry and Biochemistry, Graduate School of Humanities and Sciences, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan
[3] Institute of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan
[4] JST ACT-I
[5] JST PRESTO

Keywords: Osmolyte, trimethylamine N-oxide (TMAO), *ab initio* effective fragment potential (EFP), time series interaction energy decomposition analysis, hydrogen bond, hydrophobicity

Molecular architecture of aqueous osmolyte solution is one of the important chemical physics to understand biological mechanism of biological control of osmotic pressure and protein denaturalization. It has been considered that complemental study of spectroscopy and *ab initio* molecular dynamics (AIMD) simulations is powerful tool to clarify the mechanisms. However, as the computational cost of AIMD simulation is very expensive at the moment, there still exist very big wall to observe slight difference in various osmolytes with the above-mentioned schemes.

Recently, in our group, we have focused on effective fragment potential (EFP) [1], which is an *ab initio* polarizable force field defined by a set of quantum chemical wave function of compact constituent molecules. Combining the EFP method with the MD technique, it is possible to perform time series interaction energy decomposition analysis with relatively cheap computational costs. In the last CBI annual meeting 2018 [2], we have reported that methyl group of trimethylamine N-oxide (TMAO), a well-known osmolyte [3], is actually working as hydrophilic group due to strong intramolecular polarization of TMAO molecule. This time, we will focus on the hydration dynamics of TMAO from not only TMAO-water interaction but also water-water interaction around TMAO solutes. Applying originally developed time series interaction energy decomposition analysis based on EFP-MD simulation, we will discuss hydrogen bonding network and hydrophobicity of TMAO in physiological condition. [4]

[1] Kuroki, N., and Mori, H., Applicability of effective fragment potential version 2 – Molecular dynamics (EFP2-MD) simulations for predicting excess properties of mixed solvents, *Chemical Physics Letters*,694:82-85, 2018.
[2] Uchino, Y., Kuroki, N., and Mori, H., Is hydrophobic group in osmolyte hydrophilic? :A study by fragment based molecular theory, *CBI Annual Meeting* 2018, P1-10.
[3] Zou, Q., Bennion, B.J., Daggett, V., and Murphy, K.P., The Molecular Mechanism of Stabilization of Proteins by TMAO and Its Ability to Counteract the Effects of Urea, *Journal of American Chemical Society*, 124 (7):1192–1202, 2002.
[4] Uchino, Y., Kuroki, N., Funakura T., and Mori, H., *in preparation*, 2019.

Artificial ion channels studies by All-atom molecular dynamics simulations

Takahiro Osamura[1] w185408g@yokohama-cu.ac.jp
Toru Ekimoto[1] ekimoto@yokohama-cu.ac.jp
Tsutomu Yamane[1] tyamane@yokohama-cu.ac.jp
Takahiro Muraoka[2] muraoka@go.tuat.ac.jp
Kazushi Kinbara[3] kinbara.k.aa@m.titech.ac.jp
Mitsunori Ikeguchi[1,4] ike@yokohama-cu.ac.jp

[1] Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehirocho, Tsurumi-ku, Yokohama City 230-0045, Japan
[2] Institute of Global Innovation Research, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei-shi, Tokyo 184-8588, Japan
[3] School of Life Science and Technology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550 Japan
[4] RIKEN Medical Sciences Innovation Hub Program, 1-7-29 Suehirocho, Tsurumi-ku, Yokohama City 230-0045, Japan

Keywords: Molecular dynamics simulation, molecular engine, artificial ion channel

Research on artificial molecular machines has been conducted for a long time, and in 2016 the Nobel Prize was awarded for achievements in the basic research on artificial molecular machines. Molecular machines move mechanically by receiving external energies, but it has been a problem that conventional artificial molecular machines have no function. On the other hand, a large number of proteins present in our body are called biomolecular machines, which convert external energies by mechanical movements to support life activities. Under these circumstances, in recent years, the research and development of artificial molecular machines having energy conversion functions, termed molecular engine (ME), have been promoted. Kinbara et al. developed artificial ion channels.[1][2][3] KME that is one of the artificial ion channels developed by Kinbara et al. consists of hydrophobic transmembrane regions and hydrophilic PEG regions, and it has been clarified *in vitro* that its function as a channel is controlled by ligand bindings. However, the detailed mechanism, such as how ligands bind to KME, how KMEs create a pore to the membrane and how many KMEs form to be functionally, has not been clarified. Therefore, we aim to clarify the details of the function of KME by analysis through all-atom molecular dynamics simulations in lipid bilayer membranes. Elucidating the detailed mechanism will lead to more effective KME design. At present, we unravel that KME multimerizes in the presence of a ligand, and it is predicted that cation-π interaction may contribute to ion permeation.[4]

[1] Takahiro M, Tatsuya S, Tsutomu H,Masamune M, Kazushi K, *Chem.Commun*., 2011, **47**, 194-196
[2] Takahiro M, Tatsuya S, Tsutomu H, Masamune M, Mashiro T, Kazuhito VT, Hiroyuki N, Kazushi K, *J. Am. Chem. Soc*., 2012, **134**, 19788-19794
[3] Takahiro M, Takahiro E, Kazuhito VT, Hiroyuki N, Satoru N, Kouhei T, Rui L, Kazushi K, *J. Am. Chem. Soc*., 2014, **136**, 15584-15594
[4] Robert A. Kumpf, Dennis A. Dougherty, *Science*, 1993, **261**, 1708-1710

Thermodynamic, kinetic and computational analyses of the recognition mechanism of a flexible protein antigen by an antibody

Ikuho KANEDA [1] i-kaneda@protein.t.u-tokyo.ac.jp
Satoru NAGATOISHI [1,2] ngtoishi@ims.u-tokyo.ac.jp
Daisuke KURODA[1] d-kuroda@protein.t.u-tokyo.ac.jp
Makoto NAKAKIDO[1] nakakido@protein.t.u-tokyo.ac.jp
Jose M.M. CAAVEIRO[3] jose@phar.kyushu-u.ac.jp
Kouhei TSUMOTO[1,2] tsumoto@bioeng.t.u-tokyo.ac.jp

[1] Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
[2] The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
[3] Department of Global Healthcare, Graduate School of Pharmaceutical Sciences, Kyushu University, 3-1-1 Maidashi, Fukuoka, 812-8582, Japan

Keywords: Flexible epitope, Molecular dynamics simulation, Antibody engineering

Many proteins have both rigid and flexible regions, and change their conformations upon ligand binding. With current methodologies of antibody generation, it is often difficult to rationally obtain specific antibodies that recognize flexible epitopes on antigens due to the lack of our knowledge on molecular recognition between antibodies and flexible antigens. In this study, we focused on the interaction between an antibody and an antigen that binds metal ions, to obtain quantitative relationships about antibody-antigen interactions where the antigen is highly flexible. As a model system, we used the Fab region of a monoclonal antibody 5E1 whose antigen is Sonic hedgehog (Shh) protein, which also binds to Ca^{2+} and Zn^{2+} ions. In a previous study, 5E1 Fab had exhibited high affinity toward Shh regardless of the presence or absence of the metal ions [1]. However, the details of the interactions and physicochemical properties remains unclear.

In this study, to reveal the binding mechanism of 5E1 Fab and Shh, thermodynamic and kinetic parameters of the interactions were experimentally evaluated, and we observed drastic changes of binding enthalpy and entropy depending on the existence of the metal ions. In addition, we analyzed and discussed energetic contributions of some interface residues based on physicochemical measurements and molecular dynamics simulations with the Ala mutants as well as the wild-type. In addition to the interaction analyses, we also experimentally and computationally assessed the thermal stability of Shh with or without the metal ions, showing the dependency of the metal ions on the thermal stability of the antigen.

A co-crystal structure of Shh and 5E1 Fab with the metal ions was already solved [1]. In this study, we also succeeded in obtaining a crystal structure of the complex without the metal ions, enabling the structure-based understanding of the interaction with or without the metal ions. Putting together, our results could be a guideline for rational design of antibodies that recognize flexible antigens.

[1] Henry R. Maun, *et al.*, *J. Biol. Chem.*, 285, 26570-80, 2010

Ligand Binding Mechanism of an Enzyme Studied by Binding Free Energy Analyses for Mutants of the Protein

Yoshiharu Mori moriy@pharm.kitasato-u.ac.jp
Mayuko Takeda-Shitaka shitakam@pharm.kitasato-u.ac.jp

School of Pharmacy, Kitasato University, 5-9-1 Shirokane, Minato-ku, Tokyo 108-8641, Japan

Keywords: Enzyme, Threonyl-tRNA synthetase, Ligand binding, Binding free energy, Molecular dynamics, Sequence alignment

Aminoacyl-tRNA synthetases (AARS) catalyze the aminoacylation of tRNA. Aminoacylation of tRNA is needed before protein synthesis in ribosomes. The function of AARS is therefore important for cells.

In this study, we chose the threonyl-tRNA sythetase (ThrRS) of *E. coli* to study the mechanism of threonine binding to ThrRS and its molecular insights. This enzyme selects a threonine molecule and the threonine molecule is bound to ThrRS. Similar amino acids to threonine such as valine and serine are not bonded to a tRNA corresponding to threonine. We studied the selection mechanism of a threonine molecule in ThrRS using several theoretical and bioinformatics methods.

The catalytic site of ThrRS includes amino acid residues that can have several protonation states. In addition, ThrRS has a zinc ion in the catalytic site. The coordination structure of the catalytic site was modeled and confirmed by molecular dynamics simulations and quantum chemical calculations. We performed test simulations with different protonation states and force fields. We determined an appropriate force field and protonation states for ThrRS. We found that the molecular dynamics simulations can give accurate results, which were confirmed by the QM/MM calculations for the catalytic site including the zinc ion and the threonine molecule.

In the molecular dynamics simulations, we estimated that several amino acid residues in TheRS stabilize the threonine ligand. To identify the conservation of those amino acid residues, we carried out a multiple sequence alignment for all the sequences of ThrRS in the UniProt database.

We performed further molecular dynamics simulations to calculate the free energy profile of amino-acid binding to the catalytic site of ThrRS. To enhance sampling efficiency, we used the umbrella sampling method. Molecular dynamics simulations with umbrella sampling allowed us to obtain the accurate free energy along with the reaction coordinate. The free energy profile of the catalytic site for the ThrRS was obtained.

To compare the results of the molecular dynamics simulations with those of the multiple sequence alignment, we performed molecular dynamics simulations for several mutants of *E. coli* ThrRS and the wildtype of *S. aureus* ThrRS. The binding free energy of these systems was calculated.

From the multiple alignment, we found that several amino acids in ThrRS are highly conserved and that some of them are positioned near to a threonine molecule in the catalytic site of ThrRS. From the binding free energy analyses, we found that amino acid residues that are highly conserved can trap the ligand of ThrRS due to the interactions between the ligand and the sidechain of the highly conserved amino acids. The binding free energy of *E. coli* ThrRS was lowest compared with that of the mutants, where one of the conserved amino acids was replaced by other amino acids.

We studied the selection mechanism of a threonine molecule in ThrRS using molecular dynamics simulations, quantum chemical calculations, and a multiple sequence alignment. We found that several amino acid residues are highly conserved and that these amino acid residues contribute the stability of the ligand binding in threonyl-tRNA synthetase.

Conformation analysis of hydrogen atoms around ligand-binding pocket based on quantum chemical calculation

Chiduru Watanabe[1, 2] chiduru.watanabe@riken.jp
Kaori Fukuzawa[3] k-fukuzawa@hoshi.ac.jp
Yoshio Okiyama[4] okiyama@nihs.go.jp
Motoyasu Adachi[5] adachi.motoyasu@qst.go.jp
Shigenori Tanaka[6] tanaka2@kobe-u.ac.jp
Teruki Honma[1] honma.teruki@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[2] JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan
[3] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-Ku, Tokyo, 142-8501, Japan
[4] National Institute of Health Sciences, 3-25-26 Tonomachi, Kawasaki-ku, Kawasaki City, Kanagawa, 210-9501, Japan
[5] National Institute for Quantum and Radiological Science and Technology, Tokai, Ibaraki 319-1106, Japan
[6] Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan

Keywords: FMO, structure refinement, hydrogen atom,

The conformational state of protein–ligand complex based on structural biology such as X-ray crystal structure analysis is valuable information to realize rational drug design more precisely. On the other hand, to understand biological phenomena with chemical reaction, structural refinement for hydrogen atoms around the ligand-binding or the active sites is one of the important issues because of structural ambiguity in the case of a low-resolution crystal structure. From technical advances of protein crystallography in recent years, it is possible to use the highly reliable conformations including hydrogen atoms obtained from high-resolution X-ray and neutron diffraction analyses. The behavior of molecules regarding chemical reactions that were so far difficult to observe has been becoming to analyze by these high-resolution experimental data. Additionally, since attempts of structure refinement using optimization [1] and electron density analysis based on fragment molecular orbital (FMO) calculations have already begun, the structures commensurate with the accuracy of quantum chemical calculation are becoming available. Therefore, it is expected that the precise structure including these hydrogen atoms will lead to the elucidation of mechanisms of the reaction in an active center and the molecular recognition.

In this research, to evaluate molecular conformation around ligand-binding pocket using the neutron crystal structure including coordinate of hydrogen atoms, we analyze the stability of energy potential surface and the complementary comparison of FMO electron density with X-ray electron density. The findings of this study would be useful to construct fundamental technology of FMO-based structure refinement.

Acknowledgement
This research was done in activities of JST PRESTO Grant Number JPMJPR18GD. This research was conducted as part of the activity of the FMO drug design consortium (FMODD).

[1] T. Tsukamoto et al., *Chem. Phys. Lett.* **535** (2012) 157.

In vitro screening systems for influenza virus RNA polymerase inhibitors

Miho Kobayashi[1]
m_koba@gifu-u.ac.jp

Yuji O. Kamatari[1,2]
kamatari@gifu-u.ac.jp

[1] Life Science Research Center, Gifu University, 1-1 Yanagido, Gifu, 501-1193 Japan
[2] United Graduate School of Drug Discovery and Medical Information Sciences, Gifu University, 1-1 Yanagido, Gifu, 501-1194 Japan

Keywords: in vitro screening system, influenza, RNA polymerase, DSF, NMR

Emergence of resistant viruses against influenza drugs is a major concern and development of anti-influenza compounds with novel mechanisms is required. Because its amino acid sequence of the RNA polymerase was proven to be well-conserved among strains, it is good target for influenza drugs development. It consists of three subunits (PA, PB1, PB2), and compounds that prevent the trimer formation are thought to exert anti-influenza activity. We have constructed in vitro screening systems for influenza virus RNA polymerase inhibitors. First screening system is targeted to the interaction site between PA(257–716) and PB1(1–15) [1]. Recombinant protein, PA(257–716), was expressed using *E. coli* and the thermal denaturation of this protein was observed using differential scanning fluorimetry (DSF). Because DSF offer high-throughput, minimal material consumption, and relatively low false positive rate [2], it is suitable for the in vitro screening system. Addition of a synthetic peptide, PB1(1–15), shifted the transition of PA(257–716), indicating this system, DSF of recombinant PA(257–716), is available to screen compounds to bind PA(257–716). Second system is targeted to the interaction site between PB1(678–757) and PB2(1–37) [3]. Recombinant PB1(678–757) and PB2(1–37) were expressed using *E. coli*. Free PB1(678–757) was folded with relatively large conformational fluctuation, free PB2(1–37) was unfolded, and the complex had fixed structure different from the free forms, indicated by DSF and NMR. There are also available to screen compounds to bind PB1(678–757) and PB2(1–37). Third screening system is targeted to the endnuclase domain, PA(1–198) [4]. Recombinant PA(1–198) was expressed using *E. coli* and the thermal denaturation of this protein was observed using DSF. Addition of a compound, baloxavir acid [5], shifted the transition of PA(1–198), indicating this is also available to screen compounds to bind PA(1–198). Advantages of in such vitro screening systems are, (1) the same strategy can be applied to any other targets, (2) it is applicable to high-risk targets safely, for example, Ebola virus, (3) it is possible to screen compounds to targets that have not yet emerged, and (4) because of the mechanism of action of the compound is clear, its optimization may be straight forward.

[1] Obayashi, E. et al., *Nature*, 454:1127-1131, 2008.
[2] Kabir, A. et al., *J Biochem.* 161:215-222, 2017.
[3] Sugiyama, K. et al., *EMBO J.* 28:1803-1811, 2009.
[4] Omoto, S. et al., *Sci. Rep.* 8:9633, 2018.
[5] Noshi, T. et al., *Antiviral Res.* 160:109-117, 2018.

Can Cryptic Binding Sites Be Characterised by Voids, Pocket Detection, and Molecular Dynamics?

Shinji Iida[1], **Hironori Nakamura**[2], **Tadaaki Mashimo**[1,3]
s-iida@aist.go.jp, nakamura@biomodeling.co.jp, mashimo.tadaaki@imsbio.co.jp

Yoshifumi Fukunishi[4]
y-fukunishi@aist.go.jp

[1] Technology Research Association for Next-Generation Natural Products Chemistry, 2-3-26, Aomi, Koto-ku, Tokyo 135-0064, Japan
[2] Biomodeling Research Co., Ltd., 1-704-2, uedanishi, tenpaku-ku, Nagoya, Aichi, 468-0058, Japan
[3] IMSBIO Co., Ltd., 170-0013, 4-21-1, Higashiikebukuro, Toshima City, Tokyo, Japan
[4] Molecular Profiling Research Center for Drug Discovery, AIST Tokyo Waterfront, 2-3-26 AOMI, KOTO-KU, Tokyo, Japan

Keywords: Molecular dynamics, cryptic binding site, void

For drug discovery, binding sites tell us how to design drugs, although they are often *cryptic* in apo conformation especially in protein-protein interactions. Success of cryptic site detection would impact drug design because the existence of cryptic sites expands druggable targets [1].

Here, we aim to characterise cryptic sites and to propose a simple scheme for the detection. We first attempt to capture characteristics of cryptic sites for a set of PDB data [1] via void generation, pocket detection, and structural comparison. The reason why we focus on voids interior to a protein is that a decrease in void volume could lead to the formation of cryptic sites. We show that about 70 % of the dataset have a void within 10 Angstrom of a ligand-binding site. To examine the possibility that cryptic sites are simply detected from an apo conformation, we apply MolSite [2], which is a software for detection of ligand binding sites, to the dataset, demonstrating that it detects some cryptic sites even though the ratio is not high. Additionally, we investigate several modes of cryptic site opening from the dataset: bending helix, bending beta sheet, a single residue movement, etc.

Furthermore, we conduct atomistic molecular dynamics (MD) simulations with explicit solvent at an ambient condition in order to detect conformers in which cryptic sites is open. Throughout the simulations, we use Bcl-x(L) as a model of the proteins possessing cryptic sites. We indicate that, at an ambient condition, Bcl-x(L) exposes its cryptic sites during 5 micro sec, implying that cryptic site formation may be dominant if it does not require large conformational changes like the case of the cryptic site formation for Calmodulin [1]. The implication is supported by an earlier study of a different protein indicating that about 50 % of opened states of a cryptic site were populated [3].

[1] Peter Cimermancic et al., CryptoSite: Expanding the druggable proteome by characterization and prediction of cryptic binding sites. *J. Mol. Biol*, **2016**, 428(4), 709-719.
[2] Yoshifumi Fukunishi et al., Prediction of ligand-binding sites of proteins by molecular docking calculation for a random ligand library, *Protein Science*, **2011**, 20:95-106.
[3] Gregory R. Bowman et al., Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites, *PNAS*, **2012**, 109(29), 11681-11686.

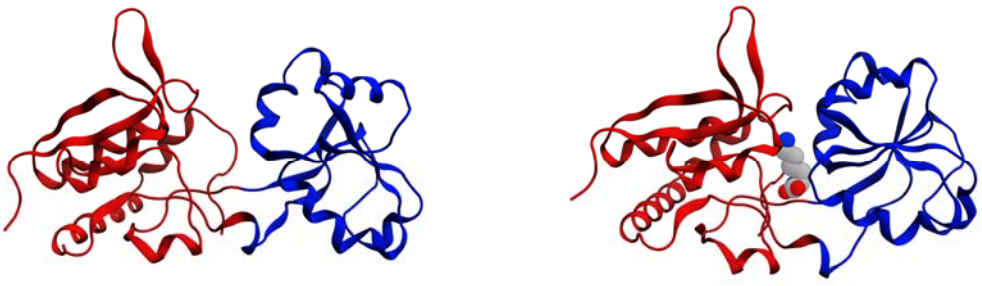
Computational study on conformational transition of protein binding pocket upon ligand binding

Noriaki Okimoto[1] okimoto@riken.jp
Yoshinori Hirano[1] hirano@riken.jp
Makoto Taiji[1] taiji@riken.jp

1 BDR, RIKEN, Building B, 6-2-4 Furuedai, Suita, Osaka 565-0874, Japan, Japan
Keywords: molecular dynamics simulation, binding pocket, ligand binding

The conformation, dynamics and function of proteins are closely linked, and many proteins undergo conformational changes upon ligand binding. The X-ray crystallographic studies have revealed conformational differences in proteins between the liganded and unliganded states. It is important to understand the mechanism of the conformational change on ligand binding at atomic level. Molecular dynamics (MD) simulation is so helpful for understanding the conformational change in depth because this is a powerful tool in providing description of the dynamics and energetics of biomolecules with high spatial and high temporal resolution.

In this study, we have performed several MD simulations of a few proteins (the lysine/arginine/ornithine-binding protein (LAOBP)[1] and so on) to understand the conformational change on ligand binding at atomic level. We have analyzed the domain movement, protein-ligand interactions, and solvation structure. Based on the current and previous results[2], we have investigated the relationship between protein conformational change and ligand binding. We will report the detailed results.



The crystal structures of LAOBP:
The left and right figures show the unliganded (apo) and liganded (holo) structures.

References

1. Oh, B. H.; Pandit, J.; Kang, C. H.; Nikaido, K.; Gokcen, S.; Ames, G. F.; Kim, S. H. J Biol Chem 1993, 268(15), 11348-11355.

2. Kondo, H. X.; Okimoto, N.; Morimoto, G.; Taiji, M. J Phys Chem B 2011, 115(23), 7629-7636.

The computational study of cycloaddition with Rhodium catalyst

Watanabe Kazuki[1] **Kawashima Yusuke**[1]
Watanabe-ka@phs.osaka-u.ac.jp kawashima-y@phs.osaka-u.ac.jp

Yu-Shi TIAN[1] **Kawashita Norihito**[2] **Takagi Tatsuya**[1]
yushi-tian@phs.osaka-u.ac.jp nkawashita@life.kindai.ac.jp ttakagi@phs.osaka-u.ac.jp

[1] Graduate School of Pharmaceutical Sciences, Osaka Universtity, 1-6 Yamadaoka, Suita, Osaka 565-0871, JAPAN
[2] Faculty of Science and Engineering, Kindai University, 3-4-1 Kowakae, Higashi-osaka, Osaka 577-8502, JAPAN

Keywords: Molecular orbital calculation, Rhodium catalyst, AFIR, Metallo-cycleaddition

Although the cycloaddition reaction of VCP with olefins using transition metal catalysts has been considered that the driving force of the reaction is the strain energy of the three-membered ring[1],[2], similar reactions occur when less distorted cyclobutane or cyclopentane analogs having an allenyl group are used[3],[4]. Thus, DFT calculations (B3LYP/6-31G*/SDD+ECP) were applied to reveal the cycloaddition reaction mechanism adopting an allenyl cycloalkane and cyclopentane analogue in the previous research as a ligand. Two pathways, one contains ring cleavage(Fig. 1 [1]) and the other contains the metallocycle as an intermediate(Fig. 1 [2]), are considered. Note that Gaussian 09 was used for electronic structure calculation, and Reacion Plus Pro 1.0.1[5] and GRRM 17[6] were used for reaction path search.

In the allenyl cycloalkane, there was a correlation between strain magnitude of the partial structure cycloalkane and the reaction pathway. Specifically, allenyl cyclopropane having a large three-membered ring distortion was superior to the reaction pathway of [1]. On the other hand, those having a relatively small strained structure such as allenyl cyclobutane and pentane as a partial structure were superior to the reaction pathway of [2], which was particularly remarkable in allenyl cyclopentane.

Therefore, next, in order to verify the superiority of the allenyl structure, a reaction model close to a real reaction system is created with reference to the previous researches, and the AFIR method implemented in the chemical reaction path automatic search program GRRM17 is a more detailed study was conducted using this.

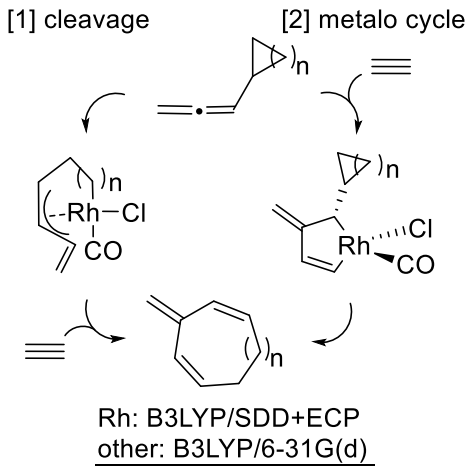


Fig. 1.Modeling of the reaction

[1] Paul A. Wender, H. Takahashi, *J. Am. Chem. Soc*., 117, 4720-4721 (1995).
[2] Zhi-Xiang Yu, Paul A. Wender, K. N. Houk, *J. Am. Chem. Soc*., 126, 9154-9155 (2004).
[3] F. Inagaki, K. Sugikubo, Y. Oura, C. Mukai, *Chem. A Eur. J*., 17, 9062-9065 (2011).
[4] C. Mukai et al., *J. Am. Chem. Soc*., 134, 19580-19583 (2012).
[5] Henkelman G. et al., *J. Chem. Phys*., 113, 9978 (2000).
[6] S. Maeda et al., *J. Comput. Chem.*, 39, 233-250, (2018).

Analysis of target DNA recognition mechanism by Nrf2-small Maf heterodimer using fragment molecular orbital (FMO) method

Toru Sengoku[1] tsengoku@yokohama-cu.ac.jp
Hirotomo Moriwaki[2] hirotomo.moriwaki@riken.jp
Chiduru Watanabe[2] chiduru.watanabe@riken.jp
Teruki Honma[2] honma.teruki@riken.jp
Kazuhiro Ogata[1] ogata@yokohama-cu.ac.jp

1 Department of Biochemistry, School of Medicine, Yokohama City University, 3-9 Fukuura, Kanazawa-Ku, Yokohama City, Kanagawa, 236-0004, Japan
2 RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama City, Kanagawa, 230-0045, Japan

Keywords: FMO, Transcription factor, Neurodegenerative disease, Crystallography

Nrf2, a transcription factor belonging to the CNC-bZIP family, plays a central role in oxidative stress response. Nrf2 forms a heterodimer with other bZIP proteins (small Maf proteins) and specifically recognizes its target sequence (CNC-small Maf binding element: CsMBE) in the vicinity of the target gene to control transcription. The optimal CsMBE is **T**GCTGA(G/C)TCA(T/C), whereas Nrf2-small Maf heterodimers can bind slightly different sequences. Recently, A SNP (rs242561) associated with neurodegenerative diseases such as Parkinson's disease have been found in the CsMBE in the regulatory region of the MAPT gene (encoding Tau protein). rs242561 is located at the first base on the CsMBE (bold position, C or T), and the low-risk T allele increases the binding level of Nrf2-small Maf heterodimer and the expression level of MAPT gene. Thus, transcriptional control of target genes by Nrf2-small Maf heterodimers is quantitatively influenced by slight sequence differences in the CsMBE, which is linked to the risk of neurodegenerative diseases.

To analyze in detail the effects of CsMBE sequence differences on the interaction with Nrf2-small Maf heterodimers, we determined the crystal structure of the Nrf2-MafG (a type of small Maf group factor) heterodimer recognizing the optimal CsMBE. In the structure, the first T-A base pair does not form a hydrogen bond with either Nrf2 or MafG; thus, it was not clear how the heterodimer recognizes this base pair. Therefore, we employed the FMO analysis and found that the T base and Arg57 of MafG form a base-specific interaction mainly driven by dispersion forces. Our study elucidates how the disease-related SNP affects the binding of transcription factors and ultimately disease risk, and shows that the FMO method is a useful tool for quantitative analysis of functions of biomacromolecules.

Interaction Analysis between HEL and HyHEL10 by Fragment Orbital Method

Norihito KAWASHITA[1]
kawashita@life.kindai.ac.jp

Shotaro YAMAGUCHI[1]
1410330081u@kindai.ac.jp

[1] Faculty of Science and Engineering, Kindai University, 3-4-1 Kowakae, Higashiosaka City, Osaka 577-8502, Japan

Keywords: Antibody, Antigen, Fragment molecular orbital (FMO) method, Inter-fragment interaction energy (IFIE)

It is expected that the novel drug can be efficiently developed by the simulation in the computer, while the development cost is suppressed. Fragment molecular orbital(FMO) method which is one of the methods of *in silico* drug discovery, and by shortening the calculation time, the high-precise interaction analysis of the biopolymer has become feasible in short time.

In this study, we focused on egg white lysozyme and its antibody, HEL/HyHEL-10, as an example of antigen-antibody complex in which there are few examples of analysis by FMO method, and examined the evaluation method of the interaction between antigen and antibody, and collected the information which contributes to the molecular design of antibody medicine.

Eight structures of the antigen-antibody complex in the literature [2] was downloaded from the Protein Structure Data Bank (PDB), and the structure was modified by MOE. In the structural optimization, Amber10:EHT was used as the

force field. Modifications were made in two ways: optimized after elimination of solvent water and optimized including solvent water. The computer used the supercomputer K with the ABINIT-MP software and performed the calculation at MP2/6-31G*.

We compared the calculated inter-fragment interaction energies (IFIE) with the experimental values in the literature, and analyzed them on the basis of structural changes in amino acid residues and changes in IFIE in response to changes in antibodies. Comparing the calculated IFIE with the experimental values, it was found that the optimized ones including solvent water had a better correlation with the experimental values. Therefore, it seems to be better to carry out the structural modification without removing solvent water. Next, we compared the structures of 2dqf and WT (wild-type) which showed large differences in IFIE by MP2 correlation. Aromatic amino acids such as tyrosine, which are commonly found in antigen-antibody binding sites in WT, had π-π interactions with other residues, but no such interactions were found at the mutation site in 2dqf, in which tyrosine was mutated to alanine. And, the interaction by tyrosine residue showed the good correlation, when the IFIE by the MP2 correlation was indicated that the assessment using MP2 correlation was important in the antigen-antibody complex.

This research was done in activities of the FMO drug design consortium (FMODD). The results were obtained using the K computer (project ID: hp170183 and hp180147). PIEDA calculation was done by using MIZUHO/BioStation software package.

[1] Lin *et al.*, *Science*, 290: 150-153, 2008.
[2] Congreve *et al.*, *J. Med. Chem.* 50: 1124-1132, 2007.
[3] Chemical Computing Group, MOE. 2015.
[4] Nakano *et al.*, *Chem. Phys. Lett.* 351: 475-480, 2002.

Combined computational and experimental study on the binding mechanism of RNA aptamer to human Immunoglobulin G

Keisuke Masukawa[1] **Hisae Yoshida** [1,2] **Seiichiro Ishii**[1] **Yuuki Yatabe**[3]
ceke18021@g.nihon-u.ac.jp

Takeshi Ishikawa[4] **Taiichi Sakamoto**[3] **Kenji Yamagishi**[1]

[1] College of Engineering, Nihon University, 1 Nakagawara, Tokusada, Tamura, Koriyama, Fukushima 963-8642, Japan
[2] JSPS Research Fellow, Kojimachi Business Center Building, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo 102-0083, Japan
[3] Faculty of Advanced Engineering, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino, Chiba 275-0016, Japan
[4] Faculty of Engineering, Kagoshima University, 1-21-24 Koorimoto, Kagoshima, Kagoshima, 890-8580, Japan

Keywords: RNA aptamer, Molecular interaction, Dynamical feature

RNA aptamers are short, single-stranded oligonucleotides that bind to specific target molecules such as proteins, nucleic acids, and small molecules. The single strand of RNA aptamers can fold into various secondary structure motifs such as stems, hairpins, bulges, and pseudoknots, and combination of these secondary structure motifs creates unique tertiary structures that enable the RNA aptamer to interact with high affinity and specificity for its target molecule. RNA aptamers, therefore many potential applications in medicine and technology.

Recently, an optimized 23-nucleotide aptamer was designed, which was shown to bind with high affinity to the Fc domain of human Immunoglobulin G (IgG) [1]. The crystal structure of the IgG complexed with the RNA aptamer has been determined [2]. However, the structural basis of the binding mechanism of the RNA aptamer with IgG is poorly understood.

In this study, to clarify the binding mechanism of RNA aptamer, we have utilized computational simulations and experimental approaches. At first, we have performed the *ab initio* fragment molecular orbital calculation and experimental isothermal titration calorimetry analysis for RNA aptamer/IgG complex to elucidate the molecular interaction between RNA aptamer and IgG. Furthermore, we performed the molecular dynamics calculations and experimental nuclear magnetic resonance analysis for the RNA aptamer - IgG complex in order to elucidate the conformational behaviors and the dynamical features of aptamers.

[1] S. Miyakawa, et al., *RNA*, **14**, 1154 (2008).
[2] Y. Nomura, et al., *Nucleic Acids Research*, **38**, 7822 (2010).

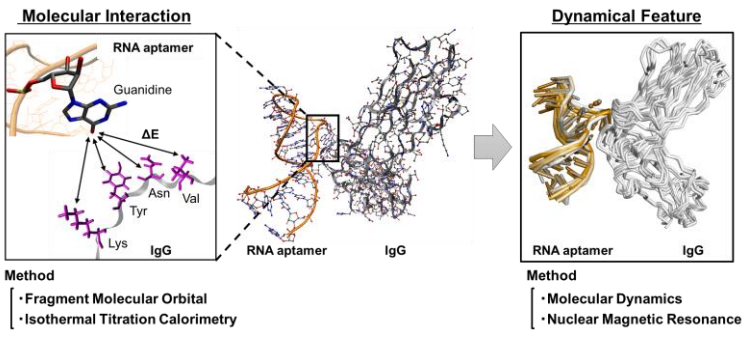


Figure. Combined computational and experimental study

Investigation of drug resistance acquisition mechanism of influenza cap-dependent endonuclease against Baloxavir marboxil by molecular dynamics simulation

Ryunosuke Yoshino[1,2]
yoshino.r.aa@md.tsukuba.ac.jp

Nobuaki Yasuo[3]
yasuo@cbi.c.titech.ac.jp

Masakazu Sekijima[3]
sekijima@c.titech.ac.jp

[1] Transborder Medical Research Center, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan
[2] Center for Computational Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan
[3] Advanced Computational Drug Discovery Unit, Tokyo Institute of Technology, J3-23-4259 Nagatsutacho, Midori-ku, Yokohama 226-8501, Japan

Keywords: Baloxavir marboxil, Influenza virus, Cap-dependent endonuclease, Molecular dynamics

Baloxavir marboxil (BXM), with the trade name Xofluza, has been approved by the FDA in October 2018 as antiviral drug for influenza virus [1, 2]. This drug developed by Shionogi Inc and inhibits RNA replication of influenza virus by binding to RNA replication cap-dependent endonucleases (CEN) [2]. However, in clinical trials, drug resistant viruses were detected with I38 mutation [2, 3].

In order to clearly the reduction of drug sensitivity by the I38 mutant variants, we conducted molecular dynamics (MD) simulation to compare the interaction mode between wild type and mutant type. MD simulation results suggest that the methyl groups of I38 side chain interacts with aromatic ring of BXM through a CH-pi interaction. In contrast, the MD simulation results for the mutation models indicated that the I38T mutation results in a loss of interaction. The side chain methyl group of isoleucine is connected to Cα atom via three single bonds. In the case of threonine, 2 single bonds connect the methyl group with Cα atom. Therefore, MD simulation suggest that the methyl group of Thr38 is at a distance too far to interact with aromatic ring of BXM.

[1] Food and Drug Administration. FDA approves new drug to treat influenza. 2018. https://www.fda.gov/news-events/press-announcements/fda-approves-new-drug-treat-influenza
[2] Omoto, S., *et al.*, Characterization of influenza virus variants induced by treatment with the endonuclease inhibitor baloxavir marboxil, *Scientific Reports*. 8:9633, doi: 10.1038/s41598-018-27890-4 2018.
[3] Hayden, F.G., *et al.*, Baloxavir marboxil for uncomplicated influenza in adults and adolescents. *N. Engl. J. Med.*, 379:913–923, 2018.

P2-01

Suitable chemical library for academic researchers in Japan

Hirotatsu Kojima
kojima@mol.f.u-tokyo.ac.jp

Drug Discovery Initiative, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Keywords: Academic drug discovery, Library design, Chemical screening

Drug Discovery Initiative (DDI) of the University of Tokyo was launched to promote academic drug discovery. We are maintaining the largest public chemical library in Japan, which consists of about 280,000 compounds including non-commercial unique samples from universities and deposited samples from pharmaceutical companies. The samples of the chemical library have been chosen mainly based on druggability and structural diversity, and added every year since 2007. We have been making improvements of the chemical library such as exclusion of promiscuous hitters based on reported data.

The chemical samples are provided to any researchers in Japan who are willing to confidentially disclose their research goals and report the assay results to us. The general samples themselves are free of charge, but payment for microplates and shipping is needed. We do not claim any right to the results of screening in the absence of intellectual collaboration. We have provided more than 25 million samples to more than 600 users so far.

It is difficult for many academic researchers to carry out random screening of more than 10,000 samples. In such cases, we can provide "core library," which is a diverse 9,600-sample set. While taking into consideration the users' circumstances, "advanced core library" has been recently prepared in order to encourage larger-scale assays, which is another diverse 22,400-sample set.

Short training courses on chemical screening are held several times a year to help beginners. The researchers may also access HTS facilities in DDI. Our organization founded the Lead Exploration Unit in 2016, which has started a hit-to-lead synthesis service, as well as ADMET support, upon request.

P2-02

Classification QSAR with Vanishing Kernels and a single parameter

Francois Berenger[1] **Yoshihiro Yamanishi**[1]
beren314@bio.kyutech.ac.jp yamani@bio.kyutech.ac.jp

[1] Department of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka, 820-8502, Japan.

Keywords: Ligand-Based Virtual Screening, Classification, QSAR, Kernel Density Estimate.

In this poster, a technique to rank-order molecules is shown. Being able to rank-order molecules allows to perform classification QSAR. The technique described here needs a single parameter (called kernel bandwidth) to be optimized during model training.
In essence, Kernel Density Estimates (KDE) are revisited and used to rank-order molecules. In order to make KDE compute-efficient, two modifications are proposed: i) using ``vanishing kernels''; i.e. kernel functions with a bounded support and ii) using the Tanimoto distance between chemical fingerprints as a radial basis function in order to work in one dimension. We call this modified construction "Vanishing Ranking Kernels"[1].
Equipped with this construction and using two real-world datasets, one from toxicology and one from High Throughput Screening (HTS) experiments, we show that Ranking Kernels can compete in performance with a state of the art deep-learning implementation for molecules.
Ranking Kernels are conceptually simple. They only require a single parameter to be optimized and hence are fast to train. Once trained, they also define a Boolean Applicability Domain[2], for free. In our experiments, this AD allows to screen candidate molecules at least 69% faster.

[1] Berenger, F., & Yamanishi Y. Ranking Molecules With Vanishing Kernels. Manuscript to be submitted, 2019.
[2] Berenger, F., & Yamanishi, Y. (2018). A Distance-Based Boolean Applicability Domain for Classification of High Throughput Screening Data. *Journal of chemical information and modeling*, *59*(1), 463-476.

P2-03

Automated Assessment of Binding Affinity via Free Energy Perturbation

Giovanna Tedesco[1]
giovanna@cresset-group.com

Maximilian Kuhn[1,2]
max@cresset-group.com

Paolo Tosco[1]
paolo@cresset-group.com

Antonia Mey[2]
antonia.mey@ed.ac.uk

Mark Mackey[1]
mark@cresset-group.com

Julien Michel[2]
julien.michel@ed.ac.uk

[1] Cresset, New Cambridge House, Cambridgeshire, UK
[2] University of Edinburgh, Department of Chemistry, Edinburgh EH9 3FJ, UK

Keywords: Free Energy Perturbation, Relative Binding Affinity

Free Energy Perturbation (FEP) calculations allow for prediction of relative binding affinity changes within a congeneric ligand series.[1] In these calculations a molecule is gradually converted into a structurally closely related ligand via a non-physical ('alchemical') pathway. By assessing the free energy difference (ΔΔG) between the end states of such transformations, accuracies of about 1 kcal/mol compared to experimental values can be achieved for large datasets.[2,3]

However, routine application of FEP calculations is still hampered by a lack of automation. Though several open-source applications strive to make FEP more accessible (e.g., the Sire molecular simulation package[4,5]), installing and using these relatively complex tools requires expertise.

To make this method accessible and user-friendly, a fully automated workflow for performing FEP calculations has been implemented in Flare, Cresset's structure-based design platform.[6]

The workflow was validated using different datasets, including the FEP+ dataset as a benchmark.[2,3] Results obtained with this method were broadly comparable to published reports, considering the overall reduced simulation time.

Full control over the simulation parameters is available within the Flare FEP project GUI and through the Flare Python API, to enable power users to explore and identify the ideal conditions for a given set of ligands and their target protein.

[1] Cournia, Z. *et al., Journal of Chemical Information and Modeling*, 57: 2911-2937, 2017.
[2] Wang, L. *et al.*, *Journal of the American Chemical Society*, 137:2695-2703, 2015.
[3] Song, L.F. *et al.*, *Journal of Chemical Information and Modeling*, 59:3128-3135, 2019.
[4] Woods, C. *et al.*, Sire molecular simulations framework, http://siremol.org, 2016.
[5] Mey, A.S.J.S. *et al.*, *Journal of Computer-Aided Molecular Design* 32:199-210, 2018
[6] http://www.cresset-group.com/flare/

P2-04

Discovery of novel selective CSF-1R Inhibitors with *de novo* drug design of FBDD and MD simulation.

Mutsuyo Wada[1] mutsu@fujitsu.com
Atsushi Tomonaga a.tomonaga1215@jcom.zaq.ne.jp
Shunji Matsumoto[1] shun@fujitsu.com
Yuichiro Tabunoki[2] y-tabuno@kowa.co.jp
Ayumu Okuda[2] a-okuda@kowa.co.jp
Tatsuhiko Kodama[3] kodama@lsbm.org

[1] Fujitsu Ltd., 1-2-70 Konan, Minato-ku, Tokyo 108-0075, Japan
[2] Kowa Co., Ltd., Tokyo New Drug Research Laboratories, 2-17-43, Noguchi-cho, Higashimurayama-sh, Tokyo 22, 189-0022, Japan
[3] Laboratory for Systems Biology and Medicine, Research Center for Advanced Science and Technology, The University of Tokyo, Building 4, Komaba Research Campus, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan

Keywords: *de novo* drug design, FBDD, in silico screening, MD simulation

The conventional low-molecular-weight drug discovery technology search for the compounds shows a certain threshold of reference inhibitory activity from the market or original chemical libraries. However, the hit compound has the problem with a novelty, diversity and selectivity to the target protein. We report that novel CSF-1R inhibitors were identified with in silico drug design technology based on FBDD (named OPMF [1]) and molecular dynamics (MD) simulation.

In this study, we designed the various compounds with OPMF using different types of DFG-out (PDB: 3KRJ, 3LCO) structures of CSF-1R, an anti-cancer target protein [2, 3], synthesized 8 compounds and obtained 3 novel hit compounds (IC50: < approximately 5 μM, hit rate 38%) measuring the inhibition of phosphorylation in CSF-1R enzyme assay. The modified compounds of the highest activity of hit compounds were assessed with the hydrogen bond analysis of MD simulation. We identified the highly potent inhibitors that increased formation time of hydrogen bond with D796 and E633 of CSF-1R. The compounds showed the kinase-selective inhibition and the suppression of CSF-1 dependent cell growth.

We showed that *de novo* drug design with OPMF and ligand optimization and selection with MD simulation were useful for screening and optimizing the novel and selective compounds for lead discovery.

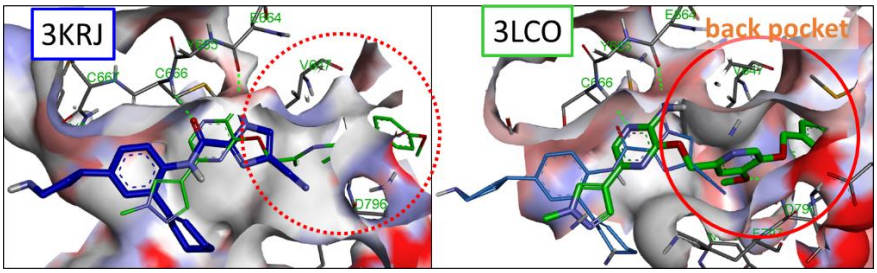


Fig. Two X-ray co-crystal structures of CSF-1R

[1] T. Yamashita, *et al.*, *Chem Pharm Bull (Tokyo)*. 2014; 62(7): 661-7
[2] C. R. Illig, *et al.*, *J Med Chem*. 2011 Nov 24; 54(22): 7860-83.
[3] M. J. Meyers, *et al. Bioorg Med Chem Lett*. 2010 Mar 1; 20(5): 1543-7.

P2-05

Predicting pKa Using a Combination of Quantum and Machine Learning Methods

Peter Hunt[1] peter@optibrium.com
Sumie Tajima[2] tajima@hulinks.co.jp
Layla Hosseini-Gerami[1] info@optibrium.com
Tomas Chrien[1] tomas@optibrium.com
Matthew Segall[1] matt@optibrium.com

[1] Optibrium Ltd., F5-6 Blenheim House, Cambridge Innovation Park, Denny End Road Cambridge, CB25 9PB, UK
[2] Hulinks Inc., 5-14 Nihonbashi Hakozaki-cho, Chuo-ku, Tokyo 103-0015, Japan

Keywords: Dissociation of a proton, pKa prediction, Radial Basis Function method

The dissociation of a proton from a heteroatom has a significant impact on the charge distribution and interactions of a molecule. These influence many important molecular properties, including binding to target and off-target proteins, absorption, distribution, metabolism and excretion (ADME) and pharmacokinetic (PK) properties such as solubility, tissue or cellular distribution and permeability. Therefore, the ability to predict the propensity of a molecule to lose or gain a proton in water is crucial for the development of new chemical entities with desirable PK, ADME and binding properties. We describe a method for prediction of the acid dissociation constant (pK_a) of a heteroatom that combines quantum-mechanical (QM) descriptors, calculated using the semi-empirical AM1 method, with machine learning to generate an accurate quantitative structure-activity relationship (QSAR) model. The QM descriptors capture the geometric and electronic properties of the environment of site of (de)protonation, in the context of the whole molecule. This provides greater discrimination between potentially acidic or basic sites and greater transferability than simple fragment descriptors. The resulting model achieves a coefficient of determination of greater than 0.9 and a root-mean-square error less than one log unit on an external test set containing both mono- and multi-protic compounds. We also present results and comparisons with other methods for published benchmarking sets.

P2-06

An in silico approach for integrating phenotypic and target-based approaches in drug discovery

Hiroaki Iwata[1]
Iwata.hiroaki.3r@kyoto-u.ac.jp

Ryosuke Kojima[1]
kojima.ryosuke.8e@kyoto-u.ac.jp

Yasushi Okuno[1]
okuno.yasushi.4c@kyoto-u.ac.jp

[1] Graduate School of Medicine, Kyoto University, 53 Kawahara-cho, Syogoin, Sakyo-ku, Kyoto 606-8507, Japan

Keywords: target deconvolution, polypharmacology, phenotypic approach, target-based approach, machine learning, drug discovery

Phenotypic and target-based approaches have been widely used for drug discovery [1]. Phenotypic approach is an experimental approach to evaluate the phenotypic response [2]. Target-based approach is a rational approach to screen drug candidates targeting a biomolecule that causes disease [3, 4]. These approaches are useful experimental methods in drug discovery. However, target deconvolution and polypharmacology have remained in these conventional experimental approaches [3, 5, 6]. To overcome these two serious problems, we developed a new in-silico method with a Bayesian probabilistic framework. Our method is based on a machine learning method, which integrates the data of the compound-target protein interactions obtained from target-based approach and the data of the compound-phenotype associations obtained from phenotypic approach. Our method can computationally examine target deconvolution considering polypharmacology, and then could give us keys for understanding pathway and mechanism from compound to phenotype, leading to promotion of drug discovery.

[1] Heilker, R., Lessel, U., and Bischoff, D., The power of combining phenotypic and target-focused drug discovery, *Drug discovery today*, 2018.
[2] Feng, Y., Mitchison, T. J., Bender, A., Young, D. W., and Tallarico, J. A., Multi-parameter phenotypic profiling: using cellular effects to characterize small-molecule compounds, *Nature reviews. Drug discovery*, 8:567-78, 2009.
[3] Croston, G. E., The utility of target-based discovery, *Expert opinion on drug discovery*, 12:427-429, 2017.
[4] Samsdodd, F., Target-based drug discovery: is something wrong? *Drug discovery today*, 10:139-147, 2005.
[5] Moffat, J. G., Vincent, F., Lee, J. A., Eder, J., and Prunotto, M., Opportunities and challenges in phenotypic drug discovery: an industry perspective, *Nature reviews. Drug discovery*, 16:531-543, 2017.
[6] Peters, J. U., Polypharmacology - foe or friend? *Journal of Medical Chemistry*, 56:8955-71, 2013.

P2-07

Analysis of subtype selectivity in estrogen-like compounds

Yuya Seki[1] m1814@hoshi.ac.jp
Tsukasa Kato[1] s151080@hoshi.ac.jp
Satoshi Anzaki[2] anzaki@eniac.scitec.kobe-u.ac.jp
Shigenori Tanaka[2] tanaka2@kobe-u.ac.jp
Koji Okuwaki[1] k-okuwaki@hoshi.ac.jp
Takayuki Furuishi[1] t-furuishi@hoshi.ac.jp
Kaori Fukuzawa[1] k-fukuzawa@hoshi.ac.jp
Etsuo Yonemochi[1] e-yonemochi@hoshi.ac.jp

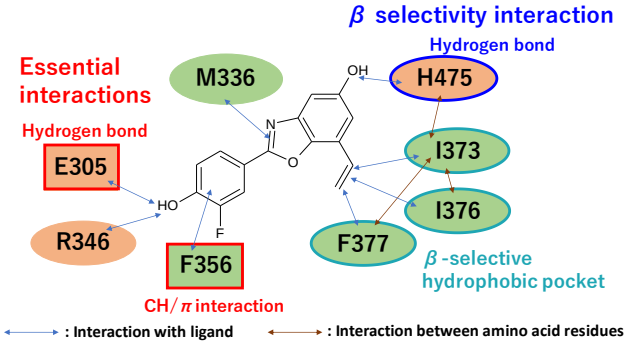
[1] Hoshi University, 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan
[2] Kobe University Graduate school, 1-1, Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501, Japan

Keywords: Estrogen Receptor β, Subtype Selectivity, Protein-Ligand Interaction, Fragment Molecular Orbital (FMO) Method, Pair Interaction Energy Decomposition Analysis (PIEDA)

Estrogen receptor (ER) has two subtypes (α, β) and it is known that phytoestrogen is selective for ERβ while 17β-Estradiol binds strongly to both subtypes. Complex structures between ER and various ligands have been revealed by X-ray crystallography. The difference between these two subtypes in the ligand binding pocket is that two hydrophobic amino acid residues (M336L, I373M) are mutated. However, ligand binding properties and the mechanism of their subtype selectivity have not been well understood. In this study, we exhaustively analyzed the PDB complex structures of ERα/β and ligand using the fragment molecular orbital (FMO) method. We examined changes in the shape of the ligand binding pocket for each subtype in terms of the interaction with the ligand and surrounding amino acid residues.

We calculated 34 PDB structures of the ER-ligand complex at the FMO2-MP2/6-31G* level. We analyzed the inter-fragment interaction energies (IFIE) and their energy component (PIEDA) between the ligand and surrounding amino acid residues. We also performed a clustering of ligands by similarity of interactions with receptor using the Visualized Cluster Analysis (VISCANA).

As a result of the cluster analysis based on the binding modes of ERβ- ligand complex, several findings were obtained: the hydrogen bond with E305 and the CH/π interaction with F356 were essential for ligand binding. The β-selective ligand interacted with H475 via a strong hydrogen bond. The amino acid residue (I373M) mutated between subtypes formed CH/π interaction with H475. The orientation of the side chain of I376 was different among the subtypes. Such hydrophobic amino acid residues, I373, I376, and F377 of ERβ formed a β-selective hydrophobic pocket. This pocket of ERβ was found to be smaller than that of ERα, and it is suggested that a highly planar structure of ligand with π electrons can be fitted to this hydrophobic pocket of ERβ. Therefore, a functional group with π electrons that can form an interaction network in the hydrophobic pocket is a characteristic of β-selective ligands.



This research was done in activities of the FMO drug design consortium (FMODD). The results were obtained using the K computer (project ID: hp190119). PIEDA calculation was done by using MIZUHO/BioStation software package.

P2-08

FMO analysis of binding property of estrogen receptor and β selective ligands

Tsukasa Kato s151080@hoshi.ac.jp
Yuya Seki m1814@hoshi.ac.jp
Takayuki Furuishi t-furuishi@hoshi.ac.jp
Kaori Fukuzawa k-fukuzawa@hoshi.ac.jp
Etsuo Yonemochi e-yonemochi@hoshi.ac.jp

School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-Ku, Tokyo, 142-8501, Japan

Keywords: estrogen receptor β, β-selective ligand, phytoestrogens, protein-ligand interaction, fragment molecular orbital (FMO) method, Molecular docking

The estrogen receptor (ER) has two subtypes, ERα and ERβ. It is known that 17β-estradiol, which is an endogenous estrogen, has no selectivity, but phytoestrogens are several tens of times and prinaberel, a synthetic ligand, is about 200 times more selective for ERβ. In this study, fragment molecular orbital (FMO) method is used to evaluate the interaction between ERβ and its selective ligand to clarify the relationship between them and β selectivity.

Docking calculations with ten ligands were performed using the ERβ-prinaberel complex structure (PDBID: 1X7B) as a template in the Molecular Operating Environment (MOE). Amber10:EHT was used for the molecular force field in the doking calculations, and FMO calculations were performed at the FMO2-MP2/6-31G* level. The interaction energies between the ligand and the amino acid residues of ERβ were calculated, and the docking scores were re-evaluated. ABINIT-MP was used for the FMO calculation program, and K supercomputer was used for calculations.

Comparing the docking score by the molecular force field and that by the FMO method, the best position (pose1) of both methods were the same pose as the crystal structure. In pose1, stable hydrogen bonds were found between the hydroxyl group of the phenol group and Glu305, and the hydroxyl group of the benzoxazole group and His475. Other poses were found lacking these hydrogen bonds. In the FMO calculation, the interaction energy between Glu305 and the ligand with hydrogen bond was about -40 kcal/mol, whereas when the hydrogen bond was not formed, it changed dramatically to +4 kcal/mol. On the other hand, the force field showed no significant change in any pose, indicating that the FMO score can properly evaluate hydrogen bonding. Similarly, docking calculations were performed on 10 compounds with various β-selectivity (mainly flavonoids and steroids), and the interactions of β-selective ligands were evaluated. Correlation was seen between calculated binding energy and IC_{50} values. In particular, it was shown that there is a strong correlation between His475 and β selectivity.

This research was done in activities of the FMO drug design consortium (FMODD). The results were obtained using the K computer (project ID: hp190119).

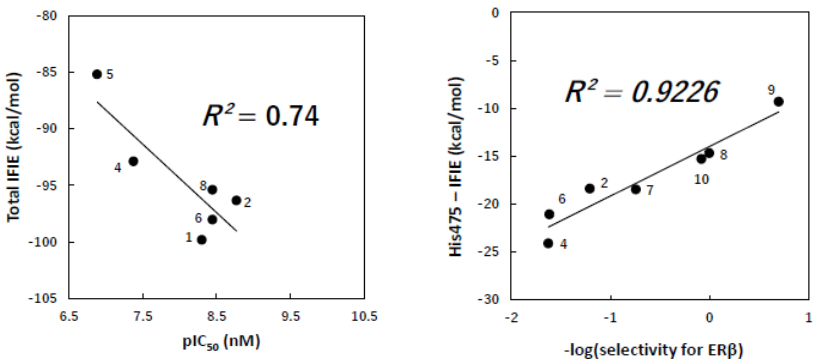


Figure 1. Correlation between calculated and experimental

P2-09

High-speed geometry optimization for protein active site with the fragment molecular orbital method

Koji Okuwaki[1] k-okuwaki@hoshi.ac.jp
Naoki Watanabe[2] naoki.watanabe@mizuho-ir.co.jp
Tatsuya Nakano[3] nakano@nihs.go.jp
Yoshio Okiyama[3] okiyama@nihs.go.jp
Yuji Mochizuki[4, 5] fullmoon@rikkyo.ac.jp
Chiduru Watanabe[6] chiduru.watanabe@riken.jp
Teruki Homma[6] honma.teruki@riken.jp
Kaori Fukuzawa[1, 5] k-fukuzawa@hoshi.ac.jp

[1] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-Ku, Tokyo, 142-8501, Japan
[2] Mizuho Information and Research Institute Inc., 2-3 Kanda Nishiki-cho, Chiyoda-ku, Tokyo, 101-8442, Japan
[3] National Institute of Health Sciences, 3-25-26 Tonomachi, Kawasaki-ku, Kawasaki City, Kanagawa, 210-9501, Japan
[4] Department of Chemistry, Rikkyo University, 3-34-1 Nishi-Ikebukuro, Toshima-ku, Tokyo, 171-8501, Japan
[5] Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan
[6] Center for Biosystems Dynamics Research, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan

Keywords: Fragment Molecular Orbital Method, Frozen Domain, Geometry Optimization, Protein–Ligand Complex

In the analysis of protein–ligand interactions, preparing highly accurate structures are crucial for precise drug design. Although high-resolution X-ray structures have been obtained in recent years, several structures still have an ambiguous near the active site. Therefore, refinement of X-ray structure by using quantum chemical calculation is highly expected. In our ABINIT-MP program for the fragment molecular orbital (FMO) calculation, partial optimization (POpt) has been possible for selected residues of the active site. However, because of the high computational cost, POpt has not been used for a large structure. In this study, frozen domain (FD) and FD dimer (FDD) methods [1] were introduced to realize high-speed geometry optimization.

In this study, we optimized several protein–ligand complexes consist of 89 to 441 residues at the HF/6-31G* level; Estrogen receptor, Neuraminidase, and G protein-coupled receptor. The optimized region contained the ligand and surrounding 4 or 5 residues. The computational time of one optimization cycle is shown in Fig. 1. Note that 2 nodes (56 cores) of TSUBAME3.0 were used for the calculation. As the size of the entire system increased, the calculation time for POpt increased, but the calculation costs for FD and FDD were not significantly affected. In the case of a system with 441 residues, the Performance of FDD was increased 4.2 times compared to its POpt.

This research was done in activities of the FMO drug design consortium (FMODD, project ID: hp190119) and was supported by AMED-BINDS under Grant Number JP19am0101113j.

[1] D. G. Fedorov et al., *J. Phys. Chem. Lett.*, **2** (2011) 282.

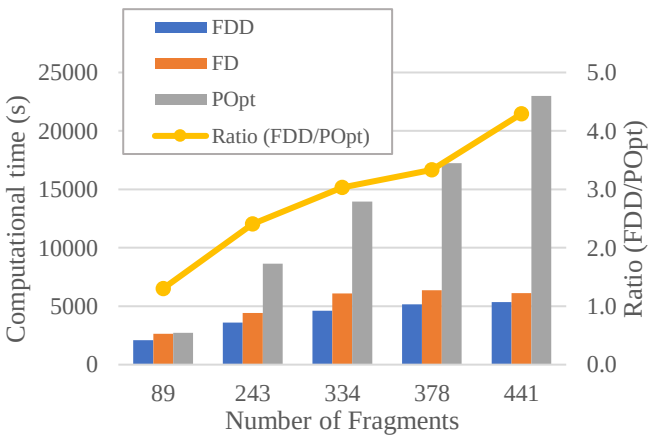


Fig. 1. Comparison between POpt and FD method

Development of FMODB for analyzing protein-ligand interactions in 2019

Daisuke Takaya[1]
daisuke.takaya@riken.jp

Chiduru Watanabe[1,2]
chiduru.watanabe@riken.jp

Kikuko Kamisaka[1]
kikuko.kamisaka@riken.jp

Teruki Honma[1]
honma.teruki@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[2] JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan

Keywords: FMO, database, IFIE, PIEDA, protein-ligand interaction.

To aid in structure-based drug design, it is important to understand protein-ligand interactions such as hydrogen bonds, electrostatic interactions, and the van der Waals interactions. Our group has been focusing on not only molecular mechanics force fields but also quantum mechanics (QM) which calculate the effects of donating and withdrawing electrons and can appropriately deal with the CH-π, π-π, and cation-π interactions. Fragment molecular orbital (FMO) method [1] enables us to efficiently perform *ab initio* QM calculations for large-biomolecules, and its fragmentation technique is capable of analyzing an inter-fragment interaction energy (IFIE).

The IFIE data provides useful information for analyzing protein-ligand interactions. To analyze IFIE data by a statistical approach, the accumulation of a large-scale of FMO calculations is essential. Therefore, we have developed FMODB [2], which is a database for providing results of FMO calculations and user-friendly web interfaces to access IFIE/PIEDA data. Currently, the interface is constructed for analyzing a ligand-binding interaction; however, the interface will be improved for analyze various molecular interaction in our plan. In this poster, we introduce new features of the FMO database and recent applications of the FMODB data.

Acknowledgement

This research was done in activities of the FMO drug design consortium, http://eniac.scitec.kobe-u.ac.jp/fmodd/index.html. The results of FMO calculations were obtained using the K computer (project ID: hp190119). PIEDA calculations were done by using MIZUHO/BioStation software package. This research was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP19am0101113. Finally, DT and CW acknowledge JSPS KAKENHI Grant Number 18K06619.

1. K. Kitaura *et al.*, *Chem. Phys. Lett.* **1999**, *313*, 701-706.
2. FMODB URL: https://drugdesign.riken.jp/FMODB

Analysis of FMO-based intramolecular interaction energies for structural stability of apo structures

Kikuko Kamisaka[1] kikuko.kamisaka@riken.jp
Shunpei Nagase[1] shunpei.nagase@riken.jp
Chiduru Watanabe[1] chiduru.watanabe@riken.jp
Daisuke Takaya[1] daisuke.takaya@riken.jp
Hitomi Yuki[1] hitomi.yuki@riken.jp
Teruki Honma[1] honma.teruki@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Japan

Keywords: FMO, IFIE, SBDD

In structure-based drug design (SBDD), such as *in silico* drug discovery, it is expected that understanding the structural stability and the function of the target protein using the three-dimensional structural simulations leads to rational drug design. However, if there is not enough structural data of the protein-ligand complexes, it is desirable to understand accurate structural information around the binding site based on available structures such as apo proteins, before the evaluation of the ligand-binding by docking or molecular dynamics simulations.

The fragment molecular orbital (FMO) method is an approach of dividing a large biomolecule into small fragments in the unit of an amino acid residue and a ligand in a protein-ligand system or a protein-protein system. The FMO method enables to calculate an electronic state of the whole system based on quantum mechanics (QM). Inter-fragment interaction energy (IFIE) based on FMO calculation can accurately estimate not only the intermolecular interaction between the target protein and the ligand molecule but also the intramolecular interaction of the biomolecule in units of an amino acid residue. Identifying which sites are energetically stable and which interactions may be involved in a ligand binding helps us to understand the capability of the target protein structure in SBDD.

In recent years, with the improvement of computer performance, analyses of interaction energies using the FMO method in a protein-ligand system have increased. On the other hand, there are only a few reports [1][2] of intramolecular interaction analysis regarding structural stability, which is important for constructing secondary and tertiary structures for a protein. In this research, we performed FMO calculations and analyzed how the structural characters such as secondary structures and the type of an amino acid residue affect the inter- and intramolecular interaction energies for structural stability. We mainly used apo-protein structures extracted from the protein data bank. The calculation was performed using the Auto FMO protocol [3]. The FMO calculation results are registered in the FMODB [4]. This research was done in activities of the FMO drug design consortium, http://eniac.scitec.kobe-u.ac.jp/fmodd/index.html.

[1] “Visualization analysis of inter-fragment interaction energies of CRP–cAMP–DNA complex based on the fragment molecular orbital method”, I. Kurisaki, et al., *Biophysical Chemistry*, **2007** 130, 1-9.
[2] “Fragment molecular orbital (FMO) study on stabilization mechanism of neuro-oncological ventral antigen (NOVA)–RNA complex system”, I. Kurisaki, et al., *Molecular Structure (THEOCHEM)*, **2010**, 962, 45-55.
[3] “Development of an automated fragment molecular orbital (FMO) calculation protocol toward construction of quantum mechanical calculation database for large biomolecules”, C. Watanabe, et al., *Chem-Bio Informatics Journal*, **2019**, 19, 5-18.
[4] https://drugdesign.riken.jp/FMODB/

P2-12

Transformation-driven Molecule Generation as Another Benchmarking Model

Ryuichiro Hara hara@lifematics.co.jp
Makoto Oku oku@lifematics.co.jp

Lifematics Inc., Sanshin Hatchobori Bldg. 5F, 2-25-10 Hatchobori, Chuo-ku, Tokyo 104-0032, Japan

Keywords: Virtual screening library, Chemical Space, Metrics of Generative Model

Sane molecule sampling from whole the not-yet-seen galactic drug chemical space has long been sought by computational and medicinal chemists for efficiency demand of their drug discovery tasks. The space is anticipated to be of well over 10^{63} sub-Lipinski molecules in size which makes it totally impractical for scientists to collect all the possible compounds as a library in hand. On the other end, modern software technology in conjunction with ever-growing compute resource has promisingly allowed us to screen through billions of molecules in ligand- and structure-based manner, however, such an ironically minuscule set of molecules is required to consort both the diversity and moderate concentration of prospective molecules in a limited reasonable size for success of screening tasks.

Transformations in a substantial set of molecules are collectively and efficiently extracted by certain graph analysis of molecules, such as matched molecular pair (MMP)[1,2] or finite differential graph analysis, which in turn applied to self set of molecules to generate molecules on-the-fly. This methodology in molecule generation is proved to be practically serviceable with reasonable retrosynthetic convenience for medicinal chemistry study.[3] Compared to such like more challenging recurrent generative network[4] on molecule string representations, the transformation-derived *de novo* molecules inherit sensible motifs (often of chemical importance) from their parents. When applied complicated cuts for retrosynthetic fragmentations, visually observed are balanced core replacement and side chain growth without overly unwanted divergence off the Lipinski-rule if appropriately filtered. Besides this legitimate expectation with extremely fast process of generating molecules, the question still remains how far the emerging molecules satisfactorily fill the chemical space to persuade medicinal chemists to be soaked on trust. Along this context, we herein seek for an effective assessment methodology based on Fréchet ChemNet (FCN) distance[5] for various set of generated compounds on Guacamol benchmarking system[6] to measure the diversity in chemical space.

[1] Hussain, J.; Rea, C.; Computationally Efficient Algorithm to Identify Matched Molecular Pairs (MMPs) in Large Data Sets., *J. Chem. Inf. Model*, **50** (3), 339, 2018.
[2] Dalke, A.; Hert, J.; Kramer, C.; mmpdb: An Open-Source Matched Molecular Pair Platform for Large Multiproperty Data Sets, *J. Chem. Inf. Model*, **58** (5), 902, 2018.
[3] León, A. de la V.; Bajorath, J.; Matched molecular pairs derived by retrosynthetic fragmentation, *Med. Chem. Commun.*, **5**, 64, 2014.
[4] Gupta, A.; Müller, A. T.; Huisman, B. J. H.; Fuchs, J. A.; Schneider, P.; Schneider G; Generative Recurrent Networks for De Novo Drug Design, *Mol Inf.*, **37**, 1700111, 2018.
[5] Preuer, K.; Renz, P.; Unterthiner, T.; Hochreiter, S.; Klambauer, G.; Fréchet ChemNet Distance: A Metric for Generative Models for Molecules in Drug, *J. Chem. Inf. Model*, **58** (9), 1736, 2018.
[6] Brown, N.; Fiscato, M.; Seglar, M. H. S.; Vauchar, A. C.; GuacaMol: Benchmarking Models for de Novo Molecular Design, *J. Chem. Inf. Model*, **59** (3), 1096, 2019.

P2-13

Inter-lobe motion of EGFR kinase: Determinants of structural variation in the crystal structures

Kei Moritsugu moritugu@yokohama-cu.ac.jp
Yoshihiko Nishino nisshy155416@yokohama-cu.ac.jp

Akinori Kidera kidera@yokohama-cu.ac.jp

Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

Keywords: Protein kinase, EGFR, KLIFS database, Structural bioinformatics, Molecular dynamics simulation, MSES

Protein kinases play important roles in cellular signaling processes, and have been one of the best studied drug targets. The structural information of protein kinases can be retrieved from a database based on protein data bank, *KLIFS*, compiling comprehensive data on protein-drug interactions for almost 300 kinase species that bind more than 3,000 distinct compounds. A whole variety of crystal structures in the database gives us a marked impression that protein kinase is remarkably flexible or intrinsically dynamic, and leads us to a naïve question: What are the determinants of these complex structural variation? Drug interactions, mutations, crystal environment, or some other causes?

In this study, we try to answer this question by focusing on the kinase domain of epidermal growth factor receptor (EGFR) as the target (a set of crystal structures consisting of 148 PDB entries and 206 domain structures are deposited in KLIFS), using database analyses and molecular dynamics simulations. EGFR kinase is a receptor tyrosine kinase that has been extensively studied for understanding the mechanism of activation occurring during the dimer formation and for developing inhibitors for cancer treatments.

We found that the inter-lobe motion between N-lobe and C-lobe constitutes the most significant structural variation, which can definitely be described by opening/closing and twisting motions of the beta-sheet in N-lobe against C-lobe. Hinge/linker moves collectively with these lobe motions, affecting the binding pose of ATP competitive inhibitors. The descriptors of the lobe motion allow us to classify the 206 kinase structures rigorously into nine groups. Molecular dynamic simulation using the enhanced sampling method (MSES) revealed that a monomer of EGFR kinase stays in the inactive state due to the interaction between a short helix formed in A-loop and N-lobe, whereas αC-helix largely fluctuates between the αC-in and αC-out states. Asymmetric dimer, regarded as the active state, maintains large structural variation in the 121 crystal structures, due to strong crystal packing that forms various head-to-tail chains with three-fold symmetry (active) – four-fold symmetry (inactive) in the crystal. Catalytic activation is understood to be regulated by exchanging the salt-bridge partner of K721 from D831 (inactive) to E738 (active). DFG-out structures occur in the mutants of L834R/V924R (or MIG6 binding) to keep the interaction between A-loop and N-lobe. These findings were firstly derived from the present survey of the whole EGFR crystal structures, and yield quite different view from the conventional studies focusing only on the local structures close to the drug binding sites.

P2-14

AUTOMATED PREDICTIVE MODELING OF BIOACTIVES

Toshikazu DEWA[1] t.dewa@elsevier.com
Matthew CLARK[2] m.clark@elsevier.com

[1] Elsevier Japan KK. 4F, 1-9-15 Higashi-Azabu, Minato-ku, Tokyo 106-0044, JAPAN
[2] Elsevier Inc., 1600 John F Kennedy Boulevard, Suite 1600, Philadelphia, PA 19103, USA

Keywords: Drug Discovery, Predictive Model, QSAR, Automation

We built a KNIME workflow to generate prediction models using bioactivity data stored in Elsevier Reaxys Medicinal Chemistry. By combining Reaxys API and KNIME, one can automatically generate predictive models with various search criteria and descriptors for any target molecules. In this presentation, we will present use case of the prediction models.

P2-15

Insights into the Mechanism of NA-I117V-Mediated Oseltamivir Resistance in H5N1 Avian Influenza Virus

Mohini Yadav[1] **Manabu Igarashi**[2]
s1879602EC@s.chibakoudai.jp igarashi@czc.hokudai.ac.jp

Norifumi Yamamoto[1]
norifumi.yamamoto@p.chibakoudai.jp

[1] Department of Engineering, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino, Chiba 275-0016, Japan
[2] Research Center for Zoonosis Control, Hokkaido University, Kita-20, Nishi-10, Kita-ku, Sapporo, Hokkaido 001-0020, Japan

Keywords: Influenza, Neuraminidase, H5N1, NA-I117V, Drug Resistance, Oseltamivir Resistance

Influenza is a contagious respiratory illness that infects the nose, throat, and lungs, causing mild to severe illness. Glycoproteins, hemagglutinin (HA) and neuraminidase (NA), of the influenza virus play essential roles in its replication process. HA binds to sialic acid (SA) at the end of a sugar chain present on the surface of a host cell, and then the virus is taken into the cell. NA on the virus surface plays a role of scissors to cleave the bond between SA and the sugar chain, which allows the daughter virus to liberate to begin infecting the surrounding cells.

Anti-flu drugs that inhibit the interaction between NA and SA have been developed. Currently, four types of NA inhibitors (oseltamivir, zanamivir, peramivir, laninamivir) have been used for the treatment of influenza in Japan. But resistance against available anti-flu drugs is emerging rapidly. Hence, it's crucial to find out the detailed mechanism of drug resistance to develop anti-flu drugs which are less prone to resistance.

Most of the receptor mutations that lead to drug resistance also affect its enzymatic activity to a significant extent. But an Ile-to-Val substitution at position 117 in influenza A virus subtype H5N1 neuraminidase (NA-I117V) causes a reduction in susceptibility to oseltamivir carboxylate while maintaining its enzymatic activity.[1]

In this study, to clarify the binding affinities of SA to the NA-I117V, we computed the corresponding binding free-energies using the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method. To clarify the influence of the I117V mutation to the enzymatic reactivity of NA, we analyzed the minimum energy path (MEP) along with the enzymatic reaction using the QM/MM hybrid method combining quantum mechanics (QM) and molecular mechanics (MM).

[1] Takano, R., Kiso, M., Igarashi, M., Le Q.M., Sekijima M., Ito, K., Takada, A., Kawaoka, Y., Molecular Mechanisms Underlying Oseltamivir Resistance Mediated by an I117V Substitution in the Neuraminidase of Subtype H5N1 Avian Influenza A viruses, *The Journal of Infectious Diseases,* 207:89-97, 2013.

Biotherapeutics modeling and developability assessment by Bio-MOE

Kentaro Kamiya[1], **Takashi Ikegami**[1], **Ryoichi Kataoka**[1], **Nels Thorsteinson**[2], **Will Long**[2]
kamiya.kentaro@molsis.co.jp

[1] Life Science Department, MOLSIS Inc., 3-19-9 Hacchobori, Chuo-ku, Tokyo 104-0032, Japan
[2] Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal QC, Canada H3A 2R7

Keywords: Antibody modeling, Protein property calculation, Liability detection

In recent years, the development of various biotherapeutics including antibody drugs has been significantly increased. It is important to estimate molecular interactions and physicochemical properties with the 3D structure at the early stage of drug discovery in order to clarify problems and lead to rational drug design. Especially, it is crucial for developing biotherapeutics to control physicochemical properties relevant to aggregation, solubility and viscosity. To design a better biotherapeutics, it is necessary to predict sites where the chemical modifications such as degradation and/or isomerization should be introduced.

The antibody modeling application implemented in Molecular Operating Environment[1] (MOE) automatically assigns the appropriate templates to each CDR and framework of the target antibody sequence and builds the homology models (Fig. 1) from multiple antibody sequences in a database. To investigate the homology models, MOE detects the protein surface patches (Fig. 2) closely related to aggregation and solubility and calculates over 30 kinds of protein properties.

Bio-MOE, a custom application package of MOE, contains a series of applications useful for biologics modeling. It can detect the protein surface patches under user defined pH values, calculate protein properties considering flexible conformations of CDR loops, detect liabilities such as chemical modification sites, search humanized antibody sequences from the human germline sequences and so on.

In this study, we present various *in silico* methods for antibody modeling and developability assessment with MOE and Bio-MOE.

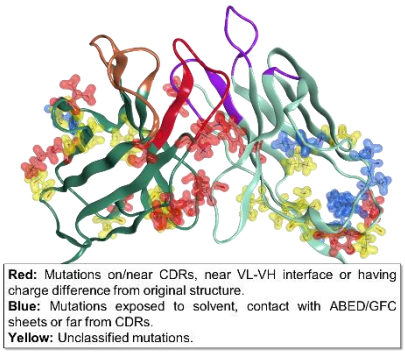


Figure 1. Humanized antibody model and visualized mutation sites.

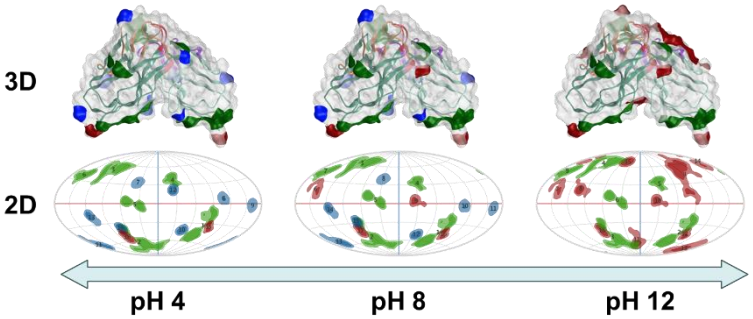


Figure 2. 3D and 2D diagrams of protein patches at each pH calculated from modeled antibody structure.

[1] *Molecular Operating Environment (MOE)*, 2018.01; Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, **2019**.

Design and Implementation of an Easy-to-Use High Performance Computing Service System for Large-Scale Molecular Dynamics Simulations

Yoshio Nakao[1] ynakao@fujitsu.com
Tetsuya Ito[1] itou.tetsuya@fujitsu.com
Toyokuni Sumita[2] sumita.toyokuni@fujitsu.com
Yusuke Sugihara[2] ysugihara@fujitsu.com

[1] Fujitsu Limited, 4-1-1 Kamikodanaka, Nakahara-ku, Kawasaki, Kanagawa 211-8588, Japan
[2] Fujitsu Kyushu Systems Limited, 5-13 Higashihie 1-Chome, Hakata-ku, Fukuoka 812-0007, Japan

Keywords: Molecular Dynamics Simulation, Absolute Binding Free Energy Calculation, High Performance Computing

In recent years, computational power has been increasing rapidly by such a rate of ten times faster every three years, and also large-scale computational resources are available as of spreading cloud computing services. In this environment, we have developed a high performance computing (HPC) service system for pharmaceutical researchers to conduct large-scale simulations easily.

The system is designed so that users unfamiliar with HPC can easily use an absolute binding free energy calculation solution[1-2], called MAPLECAFEE V2. The absolute binding energy calculation based on molecular dynamics simulation[1-4] is a promising tool for in silico drug design, however it requires sophisticated knowledge and skills on application settings and on parallel computing tuning. For example, it is required to execute efficiently several tens of molecular dynamics simulation processes corresponding to different thermodynamic states by tuning execution parameters concerning parallel computing while each of that process needs to appropriate simulation parameter settings by each steps of different functions (e.g., energy minimization, temperature and pressure control, etc.).

We then use a workflow framework[6] to describe simulation flows in an automatically executable form, and have developed a calculation process management function on the server. With provided workflow data, the user can execute calculation with the KNME workbench on his/her desktop PC while the calculation processes are executed and maintained by the HPC system.

In the poster, we introduce the system and its features in details with several usage examples.

[1] Fujitani, H., et al., Direct calculation of the binding free energies of FKBP ligands. Journal of Chemical Physics, AIP, 123:084108, 2005.
[2] Fujitani, H., et al., Massively Parallel Computation of Absolute Binding Free Energy with Well-Equilibrated States. Physical Review E, APS, 78:021914, 2009.
[3] Boresch, S., et al., Absolute binding free energies: a quantitative approach for their calculation. Journal of Physical Chemistry B, ACS, 107:9535-9551, 2003.
[4] Wang, J., et al., Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. Biophysical Journal, BPS, 91:2798-2814, 2006.
[5] Fujitani, H., et al., High-level ab initio calculations to improve protein backbone dihedral parameters. Journal of Chemical Theory and Computation, ACS, 5:1155-1165, 2009.
[6] https://www.knime.com/

Development of a Library Containing Billions of Virtual Compounds

Aki Hasegawa[1] ahasegawa@riken.jp
Yasuhiro Fujihara[2] fujihara@rist.or.jp
Noriaki Okimoto[1] okimoto@riken.jp
Yoshinori Hirano[1] hirano@riken.jp
Gentaro Morimoto[1] gentaro.morimoto@riken.jp
Makoto Taiji[1] taiji@riken.jp
Kimito Funatsu[3] funatsu@chemsys.t.u-tokyo.ac.jp

[1] Laboratory for Computational Molecular Design, RIKEN Center for Biosystems Dynamics Research (BDR), Bldg. B, 6-2-4 Furuedai, Suita, Osaka 565-0874, Japan
[2] Research Organization for Information Science and Technology (RIST) Tokai, 2-4, Shirakata, Tokai-mura, Ibaraki 319-1106, Japan
[3] Department of Chemical System Engineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Keywords: Virtual library, Virtual screening, Big data

Virtual compound libraries of various scales have been developed in private, public, commercial or non-commercial form so far [1,2,3]. In virtual screening from these libraries, it is important how to find promising lead compounds against drug target proteins at low cost and efficiency. However, there is a problem that the hit rate of lead compounds is still low, so it is necessary to increase this problem. Recently, it has been reported that it is possible to find compounds comparable to highly optimized drugs by increasing the number of compounds contained in the virtual library while including many more diverse structures [4,5]. One solution is considered to be a significant increase in the number of compounds. In addition, since the hit compound needs to be actually available, at least enough information to evaluate the synthesis route of the compound must be obtained. Based on the above, we have developed a library containing billions of virtual compounds with hints on synthesis routes. We calculated predictive assay values for all compounds in the library using a program [6] that employs deep learning techniques. We also implemented a function to export synthetic route information to a high-speed transition state evaluation system [7]. We will introduce a virtual library as part of a platform that links the flow from virtual screening to synthesis route evaluation.

[1] Ruddigkeit, L.; van Deursen, R.; Blum, L. C.; Reymond, J.-L. Enumeration of 166 Billion Organic Small Molecules in the Chemical Universe Database GDB-17. J. *Chem. Inf. Model.* 2012, 52 (11), 2864–2875.
[2] Peng, Z. Very Large Virtual Compound Spaces: Construction, Storage and Utility in Drug Discovery. *Drug Discovery Today: Technologies.* September 2013, pp e387-94.
[3] Hoffmann, T.; Gastreich, M. The next Level in Chemical Space Navigation: Going Far beyond Enumerable Compound Libraries. *Drug Discov. Today* 2019, 24 (5), 1148–1156.
[4] Gloriam, D. E. Bigger Is Better in Virtual Drug Screens. *Nature* 2019, 566 (7743), 193–194.
[5] Lyu, J.; Wang, S.; Balius, T. E.; Singh, I.; Levit, A.; Moroz, Y. S.; et al. Ultra-Large Library Docking for Discovering New Chemotypes. *Nature* 2019, 566 (7743), 224–229.
[6] Hamanaka, M.; Taneishi, K.; Iwata, H.; Ye, J.; Pei, J.; Hou, J.; Okuno, Y. CGBVS-DNN: Prediction of Compound-Protein Interactions Based on Deep Learning. *Mol. Inform.* 2016.
[7] Hori, K.; Sumimoto, M.; Murafuji, T. Quantum Chemistry-Assisted Synthesis Route Development. *AIP Conference Proceedings*; AIP Publishing LLC, 2015; Vol. 1702, p 090019.

This work was supported by JST CREST Grant Number JPMJCR1311 including AIP challenge program, and AWS Research Grants.

P3-01

IsdH: mechanism of action and novel antibacterial strategies

Sandra Valenciano Bellido[1] s-valenciano@protein.t.u-tokyo.ac.jp

Jose Manuel Martínez Caaveiro[2] jose@phar.kyushu-u.ac.jp

Kouhei Tsumoto[1] tsumoto@bioeng.t.u-tokyo.ac.jp

[1] Faculty of Bioengineering, University of Tokyo, Japan
[2] Kyushu University, Fukuoka, Japan

Keywords: *Staphylococcus aureus*, IsdH linker-NEAT3, protein structure, heme extraction mechanism, heme rotation, antibody, VHH, inhibition.

The bacterium *Staphylococcus aureus* is the major cause of deadly hospital-acquired infections worldwide[1]. The Isd (Iron surface determinant) system is a group of proteins that obtain iron from the host organism[2,3], helping the bacterium to proliferate, and therefore a promising antibacterial target. The protein IsdH binds to hemoglobin and acquires the heme[4,5]. IsdH comprises three NEAT (Near-iron transporter) domains connected by linkers of unknown function[6,7]. The objectives of this study are deciphering the molecular mechanism of heme extraction, to explain the role of the linker region and to obtain an antibody that recognizes IsdH specifically, inhibiting the heme binding. The first evidence of the structure of IsdH linker-NEAT3 bound to heme was reported in this research. The heme extraction mechanism of IsdH has been deciphered and the role of the linker as key for heme extraction has been explained through structure analysis. A novel VHH that binds to IsdH linker-NEAT3 specifically and that inhibits the heme binding was obtained. The complex crystal structure of the VHH bound to linker-NEAT3 was resolved and it was stablished that the obtained antibody specifically recognizes the heme binding pocket.

[1] Mody, L. *et al.* Multidrug-resistant Organisms in Hospitals: What Is on Patient Hands and in Their Rooms? *Clin. Infect. Dis.* 1–8 (2019). doi:10.1093/cid/ciz092
[2] Smith, A. D. & Wilks, A. Extracellular heme uptake and the challenges of bacterial cell membranes. *Curr Top Membr.* 1–28 (2013). doi:10.1016/B978-0-12-394390-3.00013-6. Extracellular
[3] Hazmanian, S. K. *et al.* Passage of heme-iron across the envelope of Staphylococcus aureus. *Science*.. **299**, 906–909 (2003).
[4] Vu, N. T. *et al.* Selective binding of antimicrobial porphyrins to the heme-receptor IsdH-NEAT3 of Staphylococcus aureus. *Protein Soc.* **22**, 942–953 (2013).
[5] Dickson, C. F. *et al.* Structure of the Hemoglobin-IsdH Complex Reveals the Molecular Basis of Iron Capture by Staphylococcus aureus *. *J. Biol. Chem.* **289**, 6728–6738 (2014).
[6] Sjodt, M. *et al.* Energetics underlying hemin extraction from human hemoglobin by Staphylococcus aureus. *J. Biol. Chem.* **293**, 6942–6957 (2018).
[7] Villareal, V. A. *et al.* Transient Weak Protein À Protein Complexes Transfer Heme Across the Cell Wall of Staphylococcus aureus. *J. Am. Chem. Soc.* 14176–14179 (2011).

P3-02

Effect of binning size of XFEL Diffraction Patterns on the resolution of reconstructed 3D-molecular structure

Miki Nakano[1]
miki.nakano@riken.jp

Osamu Miyashita[1]
osamu.miyashita@riken.jp

Florence Tama[1,2,3]
florence.tama@riken.jp

[1] RIKEN Center for Computational Science, 6-7-1, Minatojima-minamimachi, Chuo-ku, Kobe, 650-0047, Japan
[2] Department of Physics, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8602, Japan
[3] Institute of Transformative Bio-Molecules, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan

Keywords: Single particle analysis, coherent X-ray diffraction imaging, phase retrieval

Single-particle analysis (SPA) by X-ray free electron laser (XFEL) is a novel method that can observe biomolecules and living tissue difficult to crystallize in a state close to nature. To reconstruct three-dimensional (3D) molecular structure from two-dimensional (2D) XFEL diffraction patterns, we have to retrieve phase information of the 3D-diffraction intensity distribution assembled from 2D-diffraction patterns. Usually, experimental diffraction patterns are binned before analysis to increase signal and noise ratio. However, binning process would make the loss of detailed information within the binned pixels. In this study, we examined the optimal binning size to keep the quality of restored structures. Interpolation parameters used to assemble 2D-diffraction patterns into 3D-diffraction intensity distribution also should be adjusted. We found that resolution of restored molecular structure is sensitive to the interpolation parameters but insensitive to binning size. Using the optimal parameter set, a linear oversampling ratio of around four is found to be sufficient.

P3-03

Mail-in X-ray Diffraction Data Collection for High-throughput Crystal Structure Determination

Hideaki Niwa[1] hideaki.niwa@riken.jp
Shin Sato[1] shin.sato@riken.jp
Kentaro Ihara[1] kentaro.ihara@riken.jp
Takashi Umehara[1] takashi.umehara@riken.jp
Mikako Shirouzu[1] mikako.shirouzu@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan

Keywords: X-ray crystallography, Drug discovery, Protein-ligand complex, Inhibitors

Structure-based drug design needs high-throughput structure determination of the target protein complexed with a variety of drug candidate molecules. Since 2010, RIKEN has been running a program of Drug Discovery and Medical Technology Platforms (DMP) in order to bridge academia-based researches and pharmaceutical industry for drug development. As the Drug Discovery Structural Biology Platform Unit of RIKEN DMP based in Yokohama, we are equipped with a pipeline for protein expression, crystallization and structure analysis. For crystal structure determination, we have been collecting X-ray diffraction data from dozens of crystals at a time, mainly using BL26B2 of SPring-8 in a mail-in system with the help of the beamline staff at RIKEN SPring-8 center. In this system, we remotely examine X-ray diffraction images from diffraction tests and input data collection conditions before starting automated overnight data collection. Over the last eight years, we have sent about 7,800 crystals and collected about 2,300 data sets. In this presentation, we will introduce the mail-in data-collection system, as well as our protein production and crystallization pipeline. Furthermore, we report representative studies of structure-based drug design targeting epigenetics-related proteins.

P3-04

Structural basis for the binding of histo-blood group antigens to the norovirus capsid protein

Tomomi Kimura-Someya[1] tomomi.someya@riken.jp
Yuichi Someya[2] someya@niid.go.jp
Miyuki Kato-Murayama[1] miyuki.kato@riken.jp
Kazushige Katsura[1] kazushige.katsura@riken.jp
Kazutaka Murayama[1,3] kazutaka.murayama.d4@tohoku.ac.jp
Mikako Shirouzu[1] mikako.shirouzu@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[2] Department of Virology II, National Institute of Infectious Diseases, 4-7-1 Gakuen, Musashi-Murayama, Tokyo 208-0011, Japan
[3] Tohoku Univ. School of Biomedical Engineering, 6-6-12 Aramaki Aza Aoba, Aoba-ku, Sendai, Miyagi 980-8579, Japan

Keywords: norovirus, capsid, X-ray crystallography, histo-blood group antigen

Noroviruses are major causative agents of acute non-bacterial gastroenteritis in humans worldwide. There are neither antiviral therapeutic agents nor vaccines for the disease. Several genogroups (GI to GVII) are currently identified, among which GI and GII noroviruses are major pathogens for humans. The Vancouver isolate is the representative strain of GI.9, which is the genotype with the shortest history within GI. It is well known that histo-blood group antigens (HBGAs) play a critical role in norovirus infection and that the HBGA recognition profile varies by genotype. It was found by ELISA-based assay using the synthetic HBGAs and virus-like particles (VLPs) that the GI.9 norovirus preferentially bound to Lewis antigens, but not to any of ABH antigens.

The X-ray crystal structure of the protruding (P) domain dimers derived from the Vancouver strain was determined in the presence or absence of Lewis b tetrasaccharide or Lewis x trisaccharide. As suggested by the difference in amino acid sequences, it was apparent that the P domain structure was different from the corresponding domains from other GI strains, indicating that the P2 subdomain in the P domain is likely responsible for antigenicity of each norovirus strain. Nevertheless, the HBGA binding site was located at a position equivalent to those for other GI strains, which was formed on the P2 subdomain. The crystal structures clearly explained why the Vancouver strain favored Lewis antigens. Accumulating the information on the molecular structure of VLPs [1,2] and/or P domains [3,4] from various strains will help us design antiviral agents and vaccines which are hopefully effective against a wide variety of norovirus genotypes.

[1] Prasad BV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK. X-ray crystallographic structure of the Norwalk virus capsid. *Science* 286:287–290, 1999.
[2] Jung J, Grant T, Thomas DR, Diehnelt CW, Grigorieff N, Joshua-Tor L. 2019. High-resolution cryo-EM structures of outbreak strain human norovirus shells reveal size variations. *Proc Natl Acad Sci USA* 116:12828-12832, 2019.
[3] Choi JM, Hutson AM, Estes MK, Prasad BV. Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. *Proc Natl Acad Sci USA* 105:9175-9180, 2008.
[4] Kubota T, Kumagai A, Ito H, Furukawa S, Someya Y, Takeda N, Ishii K, Wakita T, Narimatsu H, Shirato H. Structural Basis for the Recognition of Lewis Antigens by Genogroup I Norovirus. *J Virol* 86:11138–11150, 2012.

P3-05

Crystal structures of GAS41 YEATS domain in complex with acylated histone peptides

Masaki Kikuchi[1]
masaki.kikuchi@riken.jp

Takashi Umehara[1]
Takashi.Umehara@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research,
1-7-22 Suehiro-cho
Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Keywords: YEATS, acylated lysine, GAS41

The YEATS domain, named after five proteins (Yaf9, ENL, AF9, Taf14 and Sas5) where it was originally found, is evolutionarily conserved from yeast to human. The human YEATS domain-containing proteins (GAS41, AF9, ENL and YEATS2) are found in chromatin-modifying and transcription regulatory complexes, this domain is proposed to function in histone and chromatin interactions. GAS41 is frequently amplified in various types of cancer and is required for cancer cell proliferation, survival, and transformation. Recently, the YEATS domains of GAS41 were reported to recognize acyl-lysine, including Kac, Kcr and Ksuc, modifications in histone H3. The GAS41 in complex with Tip60 or SRCAP colocalizes with acetylated histone H3 on the promoters of actively transcribed genes. It was suggested that GAS41 recruits chromatin remodeling factor to gene promoters and promotes histone H2A.Z deposition through binding acylated lysines. Moreover, GAS41 functions as an oncogene also plays a role in repressing the p53 tumor suppressor pathway and activating Wnt/β-catenin signaling pathways in the nucleus through binding acylated lysines. However, the acyl-lysine substrate preference of GAS41 and the structural information of substrate-recognition sites are yet to fully understood. Here, we report the binding of GAS41-YEATS to modified histone peptides with different acetyl groups using a modified histone peptide array, as well as the crystal structure of GAS41 YEATS domain in complex with actylated histone peptides.

P3-06

Crystal structure of the Hepatitis B virus core protein complexed with a novel drug candidate

Wakana Iwasaki[1,2] wiwasaki@riken.jp
Kazushige Katsura[1,2] kazushige.katsura@riken.jp
Kenji Ogawa[3] kkogawa@riken.jp
Soichi Kojima[2,4] skojima@riken.jp
Hideyuki Matsunami[2] hideyuki.matsunami@riken.jp
Tomomi Uchikubo[1,2] tomomi.kamo@riken.jp
Takeshi Yokoyama[1,2] takeshi.yokoyama@riken.jp
Hideki Shigematsu[2,5] hideki.shigematsu@riken.jp
Yuri Tomabechi[1,2] yuri.tomabechi@riken.jp
Mikako Shirouzu[1,2] mikako.shirouzu@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research, Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[2] RIKEN Center for Life Science Technologies, Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[3] RIKEN Center for Sustainable Resource Science, 2-1, Hirosawa, Wako, Saitama 351-0198, Japan
[4] RIKEN Center for Integrative Medical Sciences, 2-1, Hirosawa, Wako, Saitama 351-0198, Japan
[5] RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5148 Japan

Keywords: X-ray crystallography, electron microscopy, virus, capsid

Core protein of the hepatitis B virus (HBc) forms a dimer, and 120 HBc homodimers assemble into an icosahedral capsid that packages the viral pregenomic RNA. HBc not only constitutes capsids but plays multiple essential roles in HBV lifecycle. Therefore, HBc is a promising drug target. HBc can be divided into the N-terminal assembly domain (residues 1-149) and the highly basic C-terminal domain (residues 150-183) that is responsible for interaction with RNA. The mutant consisting of the N-terminal assembly domain only (Cp149) is known to be sufficient for self-assembly into capsid particles. In recent years, antiviral small molecules targeting HBc, called core protein allosteric modulators (CpAMs) have been developed. CpAMs enhance HBc (or Cp149) assembly, forming empty capsids or large aberrant structures.

Ogawa *et al* established a novel high-throughput method for screening drug candidates which affect Cp149-Cp149 interactions (unpublished). Some of the hit compounds of the method significantly inhibited HBV DNA replication. To elucidate the antiviral mechanism of these compounds, we examined the effect of these compounds on Cp149-Cp149 interaction using electron microscopy and dynamic light scattering. After incubating Cp149 in proper salt concentration, capsids were observed by negative staining electron microscopy. On the other hand, Cp149 dimer in buffer of low salt concentration did not assemble into structured polymer. However, after incubating Cp149 dimer with the hit compounds in low-salt buffer, morphologically “normal” capsids appeared. The compound strongly promotes assembly of Cp149, and it can even induce polymerization of the capsid assembling deficient mutant Cp149 (F23A).

Furthermore, we determined the crystal structure of Cp149 in complex with the compound. The compound binds in the hydrophobic pocket of interdimer interface of Cp149. This binding site and the above-mentioned effects on Cp149 assembly indicate that the hit compounds satisfy the features of the CpAMs that promote forming empty capsids and prevents encapsidation of pregenomic RNA.

P3-07

Structural insights into cyclic peptides in complex with target proteins

Satoshi Sogabe
satoshi.sogabe@axcelead.com

Axcelead Drug Discovery Partners, Inc., Molecular Screening,
26-1, Muraoka-Higashi 2-chome, Fujisawa, Kanagawa, 251-0012, Japan

Keywords: cyclic peptide, X-ray crystallography, crystal structure

Recently, cyclic peptides have emerged as one of new therapeutic modalities for the modulation of tough targets in drug discovery. Some attractive drug targets for therapeutic intervention have been considered undruggable because protein-protein interaction (PPI) interfaces are typically broad and/or flat binding pockets for enzyme or receptor targets. As a result, drugging these intractable targets has been deemed to be challenging because of a limited understanding of structural insights of protein-ligand interactions. In general, cyclic peptides adopt flexible conformations, resulting that it is difficult to predict active conformations by computational approach. X-ray crystallography facilitates the identification of cyclic peptide conformation and leads to motivate medicinal chemists for the structural-based design of small molecule inhibitors. We determined crystal structures of several drug targets to contribute to lead generation and optimization.

The phospholipid hydroperoxidase glutathione peroxidase 4 (GPX4) is an enzyme that reduces lipid hydroperoxides in lipid membranes and has been investigated as a target molecule that induces iron-dependent cell death (ferroptosis) selectively in cancer cells. We successfully identified one cyclic peptide that binds to near its catalytic site and two cyclic peptides that bind to another site on GPX4.[1] This is the first example reporting GPX4 inhibitory peptides and structural information of the protein surface.

The Ras proteins play important roles in cell differentiation, proliferation, and survival. We performed structural determination of K-Ras(G12D) in complex with GDP and a cyclic peptide, KRpep-2d.[2] The structure revealed that the peptide binds near Switch II and allosterically blocks protein–protein interactions with the guanine nucleotide exchange factor. This structural analysis identified a unique allosteric binding site for the future design of small-molecule inhibitors.

These structures revealed that cyclic peptides exhibit diverse binding conformations and are adjusted for a variety of molecular surfaces of targets, taking advantage of their flexibility. Cyclic peptides are considered to be a potential modality, although cell permeability and bioavailability are key limitations. Our results provide valuable structural insight to understand molecular properties of cyclic peptides and could expand the knowledge for small molecule design principles with innovative strategies.

[1] Sakamoto, K., Sogabe, S., Kamada, Y., Matsumoto, S., Kadotani, A., Sakamoto, J., and Tani, A. Discovery of GPX4 inhibitory peptides from random peptide T7 phage display and subsequent structural analysis. *Biochem. Biophys. Res. Commun.* 482:195−201, 2017.

[2] Sogabe, S., Kamada, Y., Miwa, M., Niida, A., Sameshima, T., Kamaura, M., Yonemori, K., Sasaki, S., Sakamoto, J., and Sakamoto, K. Crystal structure of a human K-Ras G12D mutant in complex with GDP and the cyclic inhibitory peptide KRpep-2d. *ACS Med. Chem. Lett.* 8:732−736, 2017.

P3-08

Prospects for applying in-silico crystal structure prediction to drug development

Okimasa Okada[1], **Hiroomi Nagata**[2], **Aaron J. Nessler**[3], **Michael J. Schnieders** [3]

[1] Mitsubishi Tanabe Pharma Corporation, Production Technology & Supply Chain Management Division,
[2] Mitsubishi Tanabe Pharma Corporation, Discovery Technology Laboratories
[3] University of Iowa, Biomedical Engineering

Keywords: Crystal structure prediction, Molecular simulation, Polarizable force field, Polymorphs

An organic compound forms different crystals (polymorphs), which influences physical properties such as solubility and stability. It is known that 80 % of drug compounds have polymorphs. So, it is meaningful to study polymorphism in drug development and quality control of medicine.

Elaborate polymorph screening experiments are performed to select the most suitable crystal form for each drug. However, screening with all of crystallization conditions is difficult. For this reason, pharmaceutical companies sometimes encounter a sudden emergence of a new stable form to cause serious damage in business.

Recently, in-silico crystal structure prediction (CSP) is getting attention to support the experiments. Due to enhancement of computational power, interatomic interactions can be evaluated accurately, which made CSP promising in practical use [1-3]. In the methods for CSP, stable packing structures with low lattice energy and high density are extracted from a large number of candidates produced by static methods (ex. Monte Carlo). However, free energy including an entropy term would be preferable for the evaluation of stability.

In this study, we predicted packing structures of a drug compound by molecular dynamics simulations with polarizable force field [4]. Experimental and obtained crystal structures are compared using lattice energy, density, and packing structures. Sublimation free energy is also evaluated. Based on the results, we would like to discuss applicability of the in-silico CSP and issues to be solved to support experimental polymorph screening.

[1] https://www.ccdc.cam.ac.uk/Community/initiatives/CSPBlindTests/
[2] M. Reilly et al. Acta Cryst. B72, 439-459, 2016
[3] Elizabeth Gibney, Nature, 527, 20-21, 2015
[4] https://ffx.biochem.uiowa.edu/

P3-09

Influence Analysis of Amino Acid Residues on Protein Functions using Attention-based Neural Networks

Takato Kameyama[1] takato.kameyama@akane.waseda.jp
Risa Igarashi[1] risa.igarashi@toki.waseda.jp
Masaki Aota[1] masaki.aota@ruri.waseda.jp
Shigeharu Mukai[1] shigeharu.mukai@moegi.waseda.jp
Hideaki Kanehara[1] hideaki.kanehara@toki.waseda.jp
Daisuke Kiga[1] kiga@waseda.jp
Noboru Murata[1] noboru.murata@eb.waseda.ac.jp

[1] Waseda University, Shinjuku-ku, Tokyo, 169-8555, Japan.

Keywords: neural network, attention mechanism, green fluorescent protein

Analyzing the influence of amino acid residues on the protein function is often a repeated process of trial and error, heavily relying on experimenters' knowledge. However, those experiments require a considerable amount of time and cost, therefore it is expected to make good use of relevant experimental data to solve the target task. In this study, we propose a method to analyze the influence of amino acid residues on the function using attention-based neural networks utilizing an accumulated dataset.

We assume that the function of the protein is determined by a set of combinations between residues of amino acid sequences of the protein, particularly some combinations have a strong influence. To distill important relations, we first extract the features of the residues with convolutional neural networks, which can take into account the short-range relations and can handle sequences in different length [1, 2]. Then, we utilize attention-based regression analysis on pair-wise correlations of the features to compute the long-range relations among residues.

We analyze the fluorescence intensity of green fluorescent protein (GFP) based on a dataset by Sarkisyan et al [3]. The dataset contains a lot of different amino acid sequences of GFP variants with their individual fluorescence intensity. Our proposed model suggests the change of fluorescence intensity highly depends on the number of amino acid replacements, and the fluorescence intensity of GFP is strongly affected by its chromophore and the combinations of neighboring residues.

[1] Sheng Wang, Siqi Sun, Zhen Li, Renyu Zhang, and Jinbo Xu. Accurate de novo prediction of protein contact map by ultra-deep learning model. PLoS computational biology, Vol. 13, No. 1, p. e1005324, 2017.

[2] Zhen Li and Yizhou Yu. Protein secondary structure prediction using cascaded convolutional and recurrent neural networks. arXiv preprint arXiv:1604.07176, 2016.

[3] Karen S Sarkisyan, Dmitry A Bolotin, Margarita V Meer, Dinara R Usmanova, Alexander S Mishin, George V Sharonov, Dmitry N Ivankov, Nina G Bozhanova, Mikhail S Baranov, Onuralp Soylemez, et al. Local fitness landscape of the green fluorescent protein. Nature, Vol. 533, No. 7603, p. 397, 2016.

A Novel Inhibitor Stabilizes the Inactive Conformation of MAPK-interacting Kinase 1

Yumi Matsui [1]
matsui.yumi.gk@rdn.daiichisankyo.co.jp

Isao Yasumatsu [1]
yasumatsu.isao.bw@rdn.daiichisankyo.co.jp

Ken-ichi Yoshida [2] **Shin Iimura** [2] **Yutaka Ikeno** [1] **Takako Nawano** [1]

Hajime Fukano [1] **Osamu Ubukata** [1] **Hiroyuki Hanzawa** [1] **Fumie Tanzawa** [2]

Takeshi Isoyama [2]
isoyama.takeshi.y5@daiichisankyo.co.jp

[1] Daiichi Sankyo RD Novare Co., Ltd., 1-16-13 Kitakasai, Edogawa-ku, Tokyo 134-8630, Japan
[2] Daiichi Sankyo Co., Ltd., 1-2-58 Hiromachi, Shinagawa-ku, Tokyo 140-8710, Japan

Keywords: X-ray crystal structure analysis, Structure-based drug design (SBDD), Protein Data Bank, Protein kinase, Enzyme inhibitor, Drug discovery, Tumor therapy

Mitogen-activated protein kinase (MAPK)-interacting kinases 1 (Mnk1) and 2 (Mnk2) are protein serine/threonine kinases that are activated by extracellular signal-regulated kinase (Erk) or p38 MAPK [1, 2]. Either activated Mnk1 or Mnk2 phosphorylates Ser209 of eukaryotic translation initiation factor 4E (eIF4E) [3], which promotes tumorigenesis [4]. However, Mnk1 and Mnk2 are dispensable in normal cells [5], suggesting that the inhibition of Mnk1 and Mnk2 could be effective in cancer therapy.

To provide a structural basis for Mnk1 inhibition, a novel Mnk1 inhibitor was discovered and the crystal structure of Mnk1 in complex with this inhibitor was determined [6]. The crystal structure revealed that the inhibitor binds to the autoinhibited state of Mnk1, stabilizing the Mnk-specific DFD motif in the DFD-out conformation, thus preventing Mnk1 from switching to the active conformation and thereby inhibiting the kinase activity. These results provide a valuable platform for the structure-guided design of Mnk1 inhibitors.

[1] Waskiewicz A.J. *et al.*, Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2, *EMBO J.*, 16(8):1909-1920, 1997.
[2] Fukunaga R. and Hunter T., MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates, *EMBO J.*, 16(8):1921-1933, 1997.
[3] Waskiewicz A.J. *et al.*, Phosphorylation of the cap-binding protein eukaryotic translation initiation factor 4E by protein kinase Mnk1 in vivo, *Mol. Cell. Biol.*, 19(3):1871-1880, 1999.
[4] Scheper G.C. and Proud C.G., Does phosphorylation of the cap-binding protein eIF4E play a role in translation initiation? *Eur. J. Biochem.*, 269(22):5350-5359, 2002.
[5] Ueda T. *et al.*, Mnk2 and Mnk1 are essential for constitutive and inducible phosphorylation of eukaryotic initiation factor 4E but not for cell growth or development, *Mol. Cell. Biol.*, 24(15):6539-6549, 2004.
[6] Matsui Y. *et al.*, A novel inhibitor stabilizes the inactive conformation of MAPK-interacting kinase 1, *Acta Cryst.*, F74(Pt 3):156-160, 2018.

Structural changes in Cl^- pump rhodopsin by time-resolved serial femtosecond crystallography

Toshiaki Hosaka[1] toshiaki.hosaka@riken.jp
Eriko Nango[2] e.nango@mfour.med.kyoto-u.ac.jp
Minoru Kubo[3] minoru@sci.u-hyogo.ac.jp
Takashi Nomura[3] takashi.nomura@riken.jp
Takanori Nakane[4] tnakane@mrc-lmb.cam.ac.uk
Tomomi Kimura-Someya[1] tomomi.someya@riken.jp
So Iwata[2] s.iwata@mfour.med.kyoto-u.ac.jp
Mikako Shirouzu[1] mikako.shirouzu@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[2] Department of Cell Biology, Graduate School of Medicine, Kyoto University, Yoshidakonoe-cho, Sakyo-ku, Kyoto, 606-8501, Japan
[3] Graduate School of Life Science, University of Hyogo, 3-2-1 Koto, Kamigori-cho, Ako-gun, Hyogo 678-1297, Japan
[4] Structural Studies Division, MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Cambridge, UK

Keywords: chloride ion pump rhodopsin, time resolved serial femtosecond crystallography

The X-ray free electron laser (XFEL) pulse is characterized by its high brightness and its short pulse duration[1-3]. We have applied time-resolved serial femtosecond crystallography (TR-SFX) using XFEL to observe structural change of proteins at room temperature without radiation damage. These TR-SFX experiments have revealed the dynamic structural changes of various proteins. In particular, the study of bacteriorhodopsin (BR) was successful in tracking photochemical reactions within femtosecond to millisecond after photoexcitation, and interesting knowledge on the correlation between structural changes and proton transport was obtained [4].

Nonlabens marinus rhodopsin-3 (NM-R3) is a retinal containing protein which functions as a light-driven chloride ion-pump. NM-R3 and halorhodopsin (HR), a well known light-driven chloride ion-pumping rhodopsin, have distinct motif sequences that are important for chloride ion binding and transport. We have previously determined the high-resolution crystal structures of NM-R3 and clarified the difference with HR [5], although detailed chloride ion transporting mechanisms remain unknown. In this study, we observed the conformational changes during ion pumping in NM-R3, from the initial changes to 1 msec following photoactivation by TR-SFX. These results provide important insights for understanding the photochemical reaction of other retinal proteins in comparison with BR.

[1] Amann, J., Berg, W., Blank, V., and Decker, F.J., Demonstration of self-seeding in a Hard-X-ray free-electron laser. *Nature Photonics*, 6:693–698, 2012
[2] Ishikawa, T., Aoyagi, H., Asaka, T., and Asano, Y.A., Compact X-ray free-electron laser emitting in the sub-angstrom region. *Nature Photonics* 6:540-544, 2012
[3] Emma, P., Akre, R., Arthur, J., and Bionta, R., First lasing and operation of an angstrom wavelength free-electron laser. *Nature Photonics* 4:641-647, 2010
[4] Nango, E., *et al*., A three-dimensional movie of structural changes in bacteriorhodopsin. *Science*, 354:1552-1557, 2016
[5] Hosaka, T., *et al*., Structural Mechanism for Light-driven Transport by a New Type of Chloride Ion Pump, Nonlabens marinus Rhodopsin-3. *J. Biol. Chem.* 291:17488-17495, 2016

Crystal Structure of EGFR T790M/C797S in complex with Brigatinib

Mutsuko Kukimoto-Niino[1] kukimoto@riken.jp
Yuri Tomabechi[1] yuri.tomabechi@riken.jp
Naoki Sakai[1] naoki.sakai@riken.jp
Ryohei Katayama[2] ryohei.katayama@jfcr.or.jp
Mikako Shirouzu[1] mikako.shirouzu@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[2] Cancer Chemotherapy Center, Japanese Foundation for Cancer Research, 3-8-31, Ariake, Koto-ku, Tokyo 135-8550, Japan

Keywords: X-ray crystallography, Tyrosine kinase inhibitor, Drug resistance

Activating mutations in the epidermal growth factor (EGFR) kinase domain are a common cause of non-small cell lung cancer (NSCLC). Osimertinib is a tyrosine kinase inhibitor (TKI) that selectively inhibits the EGFR activating mutation including T790M acquired mutation displaying resistance to first generation EGFR TKIs such as gefitinib. However, the additional C797S mutation of EGFR confers resistance to osimertinib, by impairing covalent linking of the drug with Cys797.

We have recently shown that brigatinib, an ALK tyrosine kinase inhibitor, is effective against the EGFR-T790M/C797S/activating triple mutation, and that its effect has been markedly enhanced by the combination of anti-EGFR antibody [1]. Our original computational simulation demonstrated that brigatinib fits into the ATP-binding pocket of the EGFR-T790M/C797S/activating triple mutant.

Here, we report the crystal structure of the EGFR-T790M/C797S double mutant in complex with brigatinib. The structure reveals why brigatinib prefers to bind the T790M mutant without relying on Cys797. The structure will be very powerful information to develop new TKIs to overcome drug resistance due to the C797S mutation.

[1] Uchibori, K., Inase, N., Araki, M. Kamada, M., Sato, S., Okuno, Y., Fujita, N., Katayama, R., Brigatinib combined with anti-EGFR antibody overcomes osimertinib resistance in EGFR-mutated non-small-cell lung cancer, *Nat Commun.*, 8:14768. doi: 10.1038/ncomms14768, 2017.

P3-13

Analyses based on statistical thermodynamics for large difference between thermophilic rhodopsin and xanthorhodopsin in terms of thermostability

Satoshi Yasuda[1,2,3] **Tomohiko Hayashi**[3] **Yuta Kajiwara**[4]
S.yasud@chiba-u.jp thayashi@iae.kyoto-u.ac.jp y-kajiwara@aist.go.jp
Takeshi Murata[1,2] **Masahiro Kinoshita**[3]
t.murata@faculty.chiba-u.jp kinoshit@iae.kyoto-u.ac.jp

[1]Graduate School of Science, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan
[2]Molecular Chirality Research Center, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan
[3]Institute of Advanced Energy, Kyoto University, Uji, Kyoto 611-0011, Japan
[4]Graduate School of Energy Science, Kyoto University, Uji, Kyoto 611-0011, Japan

Keywords: protein folding, solvation entropy, hydrogen bonding, integral equation theory

Although the two membrane proteins, thermophilic rhodopsin (TR) and xanthorhodopsin (XR), share a high similarity in amino-acid sequence and an almost indistinguishable three-dimensional structure, TR is much more thermostable than XR [1]. This is counterintuitive also because TR possesses only a smaller number of intramolecular hydrogen bonds (HBs) than XR. Here we investigate physical origins of the remarkable difference between XR and TR in the stability. Our free-energy function (FEF) is improved so that not only the portion within the transmembrane (TM) region but also the extracellular and intracellular portions within the water-immersed (WI) regions can be considered in assessing the stability. The assessment is performed on the basis of the FEF change upon protein folding, which is calculated for the crystal structure of XR or TR. Since the energetics within the TM region is substantially different from that within the WI regions, we determine the TM and WI portions of XR or TR by analyzing the distribution of water molecules using all-atom molecular dynamics simulations. The energetic component of the FEF change consists of a decrease in energy arising from the formation of intramolecular HBs and an increase in energy caused by the break of protein-water HBs referred to as “energetic dehydration penalty.” The entropic component is a gain of the translational, configurational entropies of hydrocarbon groups within the lipid bilayer and of water molecules. The entropic component is calculated using the integral equation theory combined with our morphometric approach. The energetic one is estimated by a simple but physically reasonable method. We show that TR is much more stable than XR for the following reasons: The decrease in energy within the TM region is larger, and the energetic dehydration penalty within the WI regions is smaller, leading to higher energetic stabilization, and tighter packing of side chains accompanying the association of seven helices confers higher entropic stabilization on TR [2].

[1] T. Tsukamoto, K. Mizutani, T. Hasegawa, M. Takahashi, N. Honda, N. Hashimoto, K. Shimono, K. Yamashita, M. Yamamoto, S. Miyauchi, S. Takagi, S. Hayashi, T. Murata, and Y. Sudo, J. Biol. Chem. **291**, 12223 (2016).
[2] S. Yasuda, T. Hayashi, Y. Kajiwara, T. Murata, M. Kinoshita, J. Chem. Phys. **150**, 055101 (2019).

Analysis by an Accelerated Quantum Chemical Molecular Dynamics Method for the 8-oxoG added DNA Structure

Ai Suzuki[1]
ai@aki.niche.tohoku.ac.jp

Akira Sassa[2]
a-sassa@chiba-u.jp

Manabu Yasui[3]
m-yasui@nihs.go.jp

Masamitsu Homma[3]
honma@nihs.go.jp

[1] New Industry Creation Hatchery Center, Tohoku University, 6-6-10 Aoba Aramaki, Aoba-ku, Sendai, Miyagi 980-8579, Japan
[2] Laboratory of Chromatin Metabolism and Epigenetics, Graduate school of Science, Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba, 263-8522, Japan
[3] Division of Genetics and Mutagenesis, The National Institute of Health Sciences, 3-25-26, Tono-machi, Kawasaki-ku, Kawasaki, Kanagawa 210-9501, Japan

Keywords: 8-oxoguanine, DNA, water

The oxidative DNA lesion 7, 8-dihydro-8-oxoguanine (8oxoG) is particularly removed by the human repair enzyme 8-oxoguanine glycosylase (hOGG1). The hOGG1 has catalytic role for the rotation or cleavage of N-glycosydic linkage between the 8oxoG and ribose [1]. Besides carrying numbers, clustered or isolated of 8oxoG somehow changed the degree of the reactivity of hOGG1 [2], the mechanism of recognition has not been elucidated especially for the several 8oxoG adducted cases. The hOGG1 recognizes the single 8oxoG legion, but does not recognize for the two base pair remote tandem 8oxoG likewise the healthy DNA [3]. Hence, we comprehensively compared the electronic configuration of water organization around the 8oxoG of the DNA and its influence on the backbone conformation by an accelerated quantum chemical molecular dynamics method which guarantees computational accuracy. The binding energies for 8oxoG, Guanine, Adenine, Thymine, and Cytosine obtained by an accelerated quantum chemical molecular dynamic method agree well with those of the density functional methods as shown in Table 1. Backbone dihedral angles and water orientation to the bases were analyzed for healthy, single and tandem 8oxoG damaged DNA.

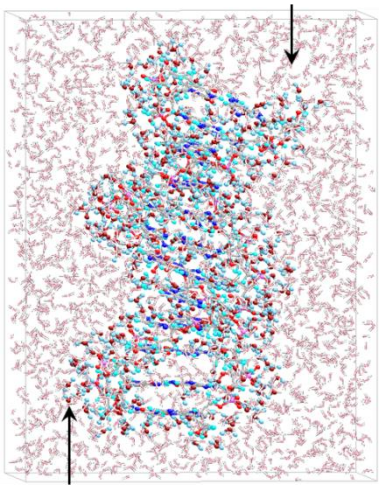


Fig.1 Stereo views of a double-stranded native DNA with water molecules.

Table 1. Comparison of the binding energies of 8oxoG and bases

Binding Energy	8oxoG	Guanine	Adenine	Thymine	Cytosine
DFT (Dmol3-P91)	-2064	-1925.1	-1797	-1768	-1508.2
DFT (Dmol3-BLYP)	-1990	-1860	-1741	-1715	-1464
Accelerated QC	-2046	-1863	-1688	-1700	-1401

[1] David, S. S., *Nature*, Structural biology: DNA search and rescue. 434:569–570, 2005.
[2] H. Terato, R. Tanaka, Y. Nakaarai, T. Nohara, Y. Doi , S. Iwai, R. Hirayama, Y. Furusawa, H. Ide, *J. Radiat. Res, Quantitative Analysis of Isolated and Clustered DNA Damage Induced by Gamma-rays, Carbon Ion Beams, and Iron Ion Beams*, 49:133–46, 2008.
[3] Sassa, A., Kamoshita, N., Kanemaru, Y., Honma, M., Yasui, M., *PLOS ONE*, Xeroderma Pigmentosum Group A Supresses Mutagenesis Caused by Clustered Oxidative DNA Adducts in the Human Genome, 1/15-15/15, 2015.

P3-15

Validation of protein structure from low resolution density map using Deep Learning

Miwa Sato[1] sato-miwa@mki.co.jp
Hiroyuki Nakagawa[1] nakagawa-hiroyuki@mki.co.jp
Yuichi Kokabu[1] kokabu-yuichi@mki.co.jp
Ikuko Miyaguchi[2] miyaguchi.ikuko@mv.mt-pharma.co.jp
Akiko Kashima[2] kashiima.akiko@mt-pharma.co.jp
Biao Ma[3] baio.ma@riken.jp
Sigeyuki Matsumoto[4] sigeyuki.matsumoto@riken.jp
Atsushi Tokuhisa[4] tokuhisa@riken.jp
Masateru Ohta[5] masateru.ota@riken.jp
Mitsunori Ikeguchi[5,6] ike@tsurumi.yokohama-cu.ac.jp

[1] Mitsui Knowledge Industry Co., LTD, 2-5-1 Atago, Minato-ku, Tokyo 105-6215, Japan
[2] Mitsubishi Tanabe Pharma Co., LTD, 1000 Kamoshida, Aoba-ku, Yokohama, Kanagawa, Japan
[3] Foundation for Biomedical Research and Innovation at Kobe, 2-2 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
[4] Riken Kobe Campus, 6-1-26 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 6500047, Japan
[5] Riken Yokohama Campus,1-7-22 Suehirocyo, Tsurumi-ku, Yokohama, Kanagawa, Japan
[6] Yokohama City University, 1-7-29 Suehirocyo, Tsurumi-ku, Yokohama, Kanagawa 230-0034, Japan

Keywords: Protein structure prediction, Deep learning, X-ray

Protein functions are regulated by the interaction between the protein and its specific substrate. The shape complementary of the protein-substrate complex can help us to elucidate the regulation mechanism of protein function and to design medicinal candidates. In drug discovery, it is required the detailed analysis of the interaction between target protein and its ligands with atomic resolution when designing medical candidates based on the information of protein structures. Therefore, the structural resolution of target protein is desired within 2 Å. At present, X-ray crystal structure analysis, one of the methods to determine protein structures, is dominantly used, and determined 90% of protein structures in the Protein Data Bank (PDB). However, about more than half of the protein structures which have been registered in the PDB show less than 2 Å resolution. Moreover, at the scene of drug discovery, low resolution of protein structures negatively may affect all over the structure-based analysis.

On the other hand, rapid progress in deep learning has accelerated its application to predict protein structure and function.[1][2] It became possible to take as input the comprehensive 3D representation of protein - ligand interaction. [3][4] In LINC (Life Intelligence Consortium), so as to convert low resolution X-ray crystallographic data into high resolution data, we are developing a validation system of protein structure using 3D convolutional neural network that take protein structure and electron density map as input. This system will evaluate coordinates of model structure with low resolution.

*Part of this research was carried out as activities of Life Intelligence Consortium (LINC).

[1] Maxwell L.B., et al., Using Deep Learning to annotate the protein universe., bioRxiv, May. 3, 2019
[2] Wang S., Sun S., Li Z., Zhang R., Xu J., Accurate de novo prediction of protein cont act map by Ultra-Deep Learning model., PLoS Comput. Biol., 13(1),2017
[3] Jimez J., Skalic M., Martinez-Rosell G., De Fabritiis G., KDEEP:Protein-liagand absolut e binding affinity predicton via 3D-Convolutional Neural Network., J. Chem. Inf. Model. 2018, 58, 287-296
[4]Jimenez J., Doerr S., Martinez-Rosell G., Rose AS., DeFabritiis G., DeepSite:Protein-bin ding site predictor using 3D-convolutional neural networks., Bioinformatics, 2017 Oct 1;3 3(19):3036-3042

P3-16

Automated X-ray crystallographical inhibitor screening against an insect ecdysteroidogenic enzyme, Noppera-bo

Kotaro Koiwai[1] koiwaiko@post.kek.jp
Kazue Inaba[2] kze1221@gmail.com
Fumiaki Yumoto[1] fumiaki.yumoto1@gmail.com
Yusuke Yamada[1] yusuke.yamada@kek.jp
Ryusuke Niwa[1,3] ryusuke-niwa@tara.tsukuba.ac.jp
Toshiya Senda[1,4] toshiya.senda@kek.jp

[1] Structural Biology Research Center, Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization, 1-1 Oho, Tsukuba, Ibaraki 305-0801, Japan
[2] Graduate School of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
[3] Life Science Center for Survival Dynamics, Tsukuba Advanced Research Alliance (TARA), University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8577, Japan
[4] School of High Energy Accelerator Science, SOKENDAI University, Tsukuba, Ibaraki 305-0801, Japan

Keywords: X-ray protein crystallography, protein-ligand complex

Compounds which inhibit protein function are called as “inhibitors”. Inhibitors can be utilized as probes to reveal *in vivo* function of target proteins. Structures of protein-inhibitor complex provide beneficial information not only for molecular mechanism of interaction between a target protein and an inhibitor, but also for that of biological function of the protein. In addition, the experimentally determined complex structures also provide higher reliability and precision for *in silico* calculation or *in vitro* assays than free-form structures. For these application, complex structures determined at atomic resolution are necessary, and X-ray crystallography has been acknowledged as the most powerful method to determine such structures.

X-ray crystallography has required a reliable and high-performance system due to huge efforts and costs at steps of crystallization, data collection, and structure determination. Currently, high-throughput X-ray crystallography for inhibitor screening has become to be realized by development of highly automated systems at the steps. Toward the high-throughput X-ray crystallography, we have proceeded automation of structure determination followed by data collection, and developed a comprehensive MR-native SAD pipeline for X-ray crystallography, *COMPASS*. In *COMPASS*, an MR and electron-density map visualization module, *PEINTS*, is implemented for structure analysis of protein and small compound inhibitor complex, which collectively processes X-ray diffraction data and results in high-performance inhibitor screening.

Here, we report structural analysis of an insect-specific development and differentiation regulator, Noppera-bo (Nobo), and small compound inhibitor complexes as one of the case studies of application of *PEINTS*. Nobo has been shown to be involved in biosynthesis of an insect steroidogenic compound, ecdysteroid. However, its endogenous substrate and regulatory mechanism of ecdysteroid biosynthesis by Nobo have remained to be elucidated. To reveal them, structural analysis of Nobo and 13 small compound inhibitor complexes was performed. Among them, *D. melanogaster* Nobo (DmNobo)-17β-estradiol (EST) complex demonstrated EST was captured in a ligand-recognition pocket of DmNobo. Therefore, we utilized EST as a probe to reveal biological function of Nobo. Structural analysis revealed that Asp113 of DmNobo interacted with EST and the Asp113 was crucial for *in vivo* function of Nobo. This study showed effectiveness of *COMPASS* toward realization of high-throughput X-ray crystallography.

Development of the screening system to create GPCR mutants stabilized in active states

Ryosuke Nakano[1]
nryosu1995@chiba-u.jp

Nanao Suzuki[1]
nanao@chiba-u.jp

Simon Hikiri[2]
simon.hkr3i@iae.kyoto-u.ac.jp

Satoshi Yasuda[1,2]
s.yasud@chiba-u.jp

Mitsunori Shiroishi[3]
shiroishi@rs.tus.ac.jp

Masahiro Kinoshita[2]
kinoshit@iae.kyoto-u.ac.jp

Takeshi Murata[1]
t.murata@faculty.chiba-u.jp

[1] Graduate school of Science, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan
[2] Institute of Advanced Energy, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan
[3] Graduate school of Engineering Science, Tokyo University of Science, 6-3-1 Niijuku, Katsushika-ku, Tokyo 125-8585, Japan

Keywords: GPCR, active state, theoretical prediction method, *S. cerevisiae*

G protein-coupled receptors(GPCRs) are the largest family of membrane proteins in the human genome, and these proteins exist in equilibrium between inactive and active states. Upon binding extracellular ligands, GPCRs undergo conformational changes that activate heterotrimeric G proteins and transmit the signals to the cell interior. Thus, GPCRs are located upstream of signal transduction and involved in various diseases. Structure determination of GPCRs in active states is important to elucidate the precise mechanism of signal transduction and to facilitate optimal drug design. However, their instability is an obstacle for structural studies and stabilization of GPCRs is essential.

Here, we developed the screening system to create GPCR mutants stabilized in active states. We applied our theoretical prediction method[1] to adenosine A_{2a} receptor in the agonist-bound state and selected mutants from the prediction result. The mutants were expressed in *S. cerevisiae* strain YB14, which have been engineered so that heterologously expressed GPCRs activate the yeast pheromone signaling pathway and enable growth on selective media. We evaluated the mutants by using agonist-dependent yeast growth assay[2].

In conclusion, we succeed in screening a mutant which has higher affinity for agonists than the wild type. The screening system using the theoretical prediction method and the yeast growth assay will be useful for creating mutants of other GPCRs stabilized in active states.

[1] Satoshi Yasuda, Yuta Kajiwara, Yuuki Takamuku, Nanao Suzuki, Takeshi Murata, Masahiro Kinoshita, Identification of thermostabilizing mutations for membrane proteins: rapid method based on statistical thermodynamics., *The Journal of Physical Chemistry B,* 120, 3833-3843, 2016.
[2] Dong S., Rogan SC., Roth BL., Directed molecular evolution of DREADDs: a generic approach to creating next-generation RASSLs., *Nature Protocols,* 5:561-573, 2010.

Identification of a conserved allosteric site in Heme-copper oxygen reductase

Yuya Nishida[1,2]
nishida.yuya@ncvc.go.jp

Seiji Takashima[2]
takasima@cardiology.med.osaka-u.ac.jp

Hitomi Yuki[3]
hitomi.yuki@riken.jp

Teruki Honma[3]
honma.teruki@riken.jp

Yasunori Shintani[1,2]
shintani.yasunori@ncvc.go.jp

[1] Department of Molecular Pharmacology, National Cerebral and Cardiovascular Center Research Institute,
6-1 Kishibeshinmachi, Suita, Osaka 564-8565, Japan
[2] Medical Biochemistry, Osaka University Graduate School of Medicine,
2-2 Yamadaoka, Suita-shi, Osaka, 565-0871, Japan
[3] RIKEN Center for Life Science Technologies,
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan

Keywords: Heme-copper oxygen reductase, cytochrome c oxidase, ubiquinol oxidase allosteric regulation, ligand-based virtual screening

Heme-copper oxygen reductase (HCO) works as a terminal component of the respiratory electron transport chain. HCO couples the oxygen reducing reaction to the proton pumping. This reaction forms a proton gradient across the mitochondrial inner membrane, which is used by ATP synthase to produce ATP. Here we describe two HCOs: mitochondrial cytochrome c oxidase (mtCcO) and bacterial ubiquinol oxidase (UqO). Recently we identified novel allosteric inhibitors of mitochondrial mtCcO and determined mtCcO structure in complex with some of them. These structures showed two inhibitors bound to the same site of mtCcO, which is distant from the ligand-binding sites. There is similarity between the inhibitor binding site of mtCcO and the homologous position in bacterial UqO, although mtCcO has additional subunits that are found only in mammals. These led us to hypothesize that the novel allosteric site is conserved and has a potential to be a novel antibiotic target.

To this end, we established a custom library by using the ligand-based virtual screening system, LAILAPS, with the mtCcO inhibitors as queries from commercially available compounds. As results of enzyme assay with mtCcO, the most improved compound had IC50 of 100 nM, which is one-order better than that of the query inhibitors. In addition, we identified multiple inhibitors targeting UqO, and importantly some of them were specific to UqO. Next, to confirm the inhibitor-binding site in UqO, we performed the enzyme assay with mutant bacterial UqOs which had an amino acid substitution in the corresponding site with the inhibitor-binding site of mtCcO. As a result, the effect of the inhibitors was significantly changed in several mutants. These results strongly suggest that the inhibitors-binding site is conserved between mtCcO and bacterial UqO. Finally, to verify the target specificity of the novel UqO inhibitor *in vivo*, antibacterial activity of these inhibitors was analyzed. The UqO inhibitor prevented the growth of UqO-dependent *E. coli* strain, whereas the effect was abrogated in UqO-independent strain.

From these results, we conclude that the allosteric site is conserved among HCOs. Based on structural difference between mammalian and bacterial oxidase, it is possible to develop novel antibiotics with specificity.

P3-19

The influence of cosolvent on thermal stability of membrane proteins

Kazuki kazama[1] **Satoshi Yasuda**[1,2,3]
kazkaz10nogodc@chiba-u.jp s.yasud@chiba-u.jp

Tomoki Akiyama[1] **Masahiro Kinoshita**[3] **Takeshi Murata**[1,2]
tomotomoaki01310131@yahoo.co.jp kinoshit@iae.kyoto-u.ac.jp t.murata@faculty.chiba-u.jp

[1] Graduate School of Science, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan
[2] Molecular Chirality Research Center, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan
[3] Institute of Advanced Energy, Kyoto University, Uji, Kyoto 611-0011, Japan

Keywords: Membrane proteins, Solvation entropy, Protein stabilization, Clear-native PAGE

Membrane proteins play an important role in basic processes such as signal transduction through biological membranes to maintain homeostasis, substance transport, and energy production. More than 50% of currently marketed pharmaceutical products target membranes proteins. However, many membrane proteins have low thermal stability in detergents and are easily denatured in the purification process. Therefore, it is difficult to elucidate the structures and functions of these unstable membrane proteins. In general, it is known that the addition of a cosolvent such as sugar and alcohol affects the thermal stability of water-soluble proteins. Recently, we have proposed that the solvent entropy is the key factor of the cosolvent effect. Upon protein folding, the excluded volume (EV) (i.e., the volume of the space which the centers of water molecules cannot enter) decreases to a large extent, which is followed by a corresponding increase in the total volume available to the translational displacement of water molecules in the system and a reduction in the water crowding. The folding thus leads to a large gain of the water entropy. When the water crowding is made more serious by the cosolvent addition, the solvent-entropy gain upon protein folding is magnified, leading to the enhanced thermal stability. In this study, we focused on membrane proteins (MPs). We have shown that the entropic effect of hydrocarbon groups constituting the lipid bilayer is essential for stabilization of the MPs. The lipid bilayer is immersed in water, and it is expected that if water crowding becomes more serious, crowding of hydrocarbon groups also become higher. We can regard hydrocarbon groups and water as one solvent. This interpretation is also true for the MP inside the micelle formed by detergents. Based on this unique view, the effects of cosolvent addition for the two MPs, adenosine A_{2a} receptor ($A_{2a}R$) and thermophilic rhodopsin (TR), are experimentally examined. Six cosolvents (the order of the stabilization effect reported previously is sucrose > glucose ~ mannitol > erythritol > glycerol > 2-propanol) are tested. Cosolvent addition can stabilize both of the MPs, and the order of stabilization is completely consistent with our previous analytical results for the water-soluble protein. By sucrose addition, T_m increases by 2.6℃ and 5.5℃ for TR and A2aR, respectively. Thus, the cosolvent addition according to the thermotical prediction is a useful method to stabilize purified MPs.

[1] Shota Murakami, and Masahiro Kinoshita. The Journal of Chemical Physics 144, 125105(2016)
[2] Hiraku Oshima, and Masahiro Kinoshita. The Journal of Chemical Physics 138, 245101(2013)
[3] Nanao Suzuki et al. Analytical Biochemistry 548 (2018) 7-14

P4-01

Identification of a biomarker for disease progression in heart failure using single-cell RNA sequencing data

Momoko Hamano[1] haman929@bio.kyutech.ac.jp
Seitaro Nomura[2,3] senomura-cib@umin.ac.jp
Hiroyuki Aburatani[3] haburata-tky@umin.ac.jp
Issei Komuro[2] komuro-tky@umin.ac.jp
Yoshihiro Yamanishi[1] yamani@bio.kyutech.ac.jp

[1] Department of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan,
[2] Department of Cardiovascular Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan.
[3] Genome Science Division, Research Center for Advanced Science and Technologies, The University of Tokyo, Tokyo 153-0041, Japan.

Keywords: Heart failure, Single-cell RNA sequencing, Biomarker

Heart failure is one of the most serious problems in cardiovascular diseases, as reflected by an increase in the number of patients. Heart failure has multiple risk factors and various pathological conditions, which makes it difficult to understand the molecular pathological mechanisms. We previously reported that metabolic and shape remodeling was induced by activating p53 in hypertrophic cardiomyocytes, which was critical for progression of heart failure [1]. In this study, we analyzed gene expression patterns of individual cardiomyocytes using single-cell RNA sequencing data in order to explore biomarkers characteristic of each cardiomyocyte and clarify the molecular pathological mechanism of heart failure deterioration. First, we found that the expression of natriuretic peptide A (*Nppa*) was significantly increased after stress loading by transverse aortic constriction (TAC), and then, characterized each cardiomyocyte by the expression level of *Nppa*. The tSNE analysis of gene expression data showed that *Nppa* high-expressing cardiomyocytes were located farthest away from the TAC non-loading cardiomyocytes. Also, the clustering analysis showed different gene expression patterns between *Nppa* high-expressing and *Nppa* low-expressing cardiomyocytes. Furthermore, multiple NADH ubiquinone complex family proteins associated with the mitochondrial electron transport system were negatively correlated with *Nppa* expression at the early stage of cardiac hypertrophy, implying that mitochondrial dysfunction was induced in *Nppa* high-expressing cardiomyocytes. These results suggest that the expression level of *Nppa* is a useful marker for understanding the molecular pathogenesis of heart failure.

[1] Nomura S., Satoh M., Fujita T., Higo T., Sumida T., Ko T., Yamaguchi T., Tobita T., Naito AT., Ito M., Fujita K., Harada M., Toko H., Kobayashi Y., Ito K., Takimoto E., Akazawa H., Morita H., Aburatani H., and Komuro I.. Cardiomyocyte gene programs encoding morphological and functional signatures in cardiac hypertrophy and failure, *Nat Commun*. (2018) 9(1):4435.

P4-02

An importance of polyamine metabolism regulated by cancer stem cells highlighted by our trans-omics method

Jun Koseki[1]
jkoseki@cfs.med.osaka-u.ac.jp

Masamitsu Konno[2]
mkonno@cfs.med.osaka-u.ac.jp

Ayumu Asai[1,2]
aasai@cfs.med.osaka-u.ac.jp

Hideshi Ishii[1]
aasai@cfs.med.osaka-u.ac.jp

[1] Department of Medical Data Science, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita-shi, Osaka 565-0871, Japan
[2] Department of Frontier Science for Cancer and Chemotherapy, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita-shi, Osaka 565-0871, Japan

Keywords: Trans-omics analyses, reaction flow modeling, polyamine metabolism

In cancer treatment, generally, chemotherapy and radiation therapy are performed after surgery. These treatments might result in recurrence of cancer, even though it appears to be eradicated. Previous studies have shown that the subpopulations of tumor-initiating cells, or cancer stem cells (CSCs), are responsible for the heterogeneity of tumors. It has been thought that this heterogeneity is involved in their refractoriness to chemotherapy and radiation therapy as well as their subsequent relapse. Therefore, it is very important to understand the properties of CSCs for the cancer therapy. In order to understand the differences of biological characteristic between CSCs and non-CSCs, gene expression, amount of metabolite, etc., have been widely analyzed. It is very important to observe their behavior changes over time, not snapshots at each such observation for understanding the response of cells to external stimuli such as antitumor drug. However, there were no method for analyzing to observe biological changes over time at that point. Therefore, we have developed a novel trans-omics method to comprehend temporal changes in metabolism [1].
Using our trans-omics method, the difference of metabolic reaction flow changes between CSCs and non-CSCs were analyzed. In this analyses, we have focused on the ornithine metabolic pathway. The reason is that, it is well known that the proteasome activity in CSCs is lower than in non-CSCs, in addition, it was reported that the relationship between proteasome activity and ornithine decarboxylase in esophageal, cervical, colorectal and bone cancer was investigated [2, 3]. Our analyses show that CSCs exposed to antitumor drugs rapidly promoted the reverse reaction to ornithine. On the other hand, in non-CSCs, the reaction from ornithine to putrescine was promoted and the after the reaction tended to be stopped. This result means that in non-CSCs, the change of reaction flow make the total amount of polyamine increasing. By the way, previous reports have suggested that polyamine catabolic enzymes such as SAT1, SAT2, and PAO share high similarity in overall structure with the histone demethylase LSD1. In particular, the amino acid sequence of PAO shares over 60% similarity with that of LSD1. So, these results suggest that high level polyamines inhibit LSD1 enzyme activity in non-CSCs and also changes gene expression levels to influence cell survival. It may result in the death of cancer cells.

[1] Koseki, J., et al., A Trans-omics Mathematical Analysis Reveals Novel Functions of the Ornithine Metabolic Pathway in Cancer Stem Cells. *Scientific Reports*, 6:20726, 2016.
[2] Tamari, K., et al., Identification of chemoradiation-resistant osteosarcoma stem cells using an imaging system for proteasome activity. *Int. J. Oncol*., 45:2349, 2014.
[3] Hayashi, K. et al., Visualization and characterization of cancer stem-like cells in cervical cancer. *Int. J. Oncol.*, 45:2468, 2014.

P4-03

Interpreting Japanese GWAS Results on Multi-Omics Drug Target Validation Platform

Wirawit Chaochaisit
wirawit@genesis-healthcare.jp

He Huang
huang@genesis-healthcare.jp

Youssef Darzi
youssef_darzi@genesis-healthcare.jp

Yosvany Lopez
lopez@genesis-healthcare.jp

Kosuke Negishi
kosuke_negishi@genesis-healthcare.jp

Genesis Healthcare Co., Yebisu Garden Place Tower 15F/26F, 4-20-3
Ebisu, Shibuya-ku, Tokyo 150-6015, Japan

Keywords: GWAS, Multi-Omics Data Integration, Drug Target Validation, Drug Discovery

The interpretation of GWAS results for drug target validation requires an exhaustive evaluation based on a multitude of research evidences from heterogeneous data sources. Open Targets Platform (OTP) [1] has integrated over 20 multi-omics data sources to provide a comprehensive compendium of more than 7 million evidences. It offers wide array of analysis and visualization pipelines, allowing a more thorough validation, comparison, and prioritization of drug targets for diseases of interest. We processed our GWAS results for gout, hyperlipidemia, myopia, and asthma through OTP's data processing pipelines to map significant SNPs (P-value $\leq 10^{-5}$) to genes using Ensembl's Variant Effect Predictor [2] or by considering their distance (within 500kb) to the 5' end of the closest gene for intergenic variant. Then, the target-disease association score is calculated using p-value, sample size, and variant's functional consequence coefficient [3][4]. As a result, we were able to cross-validate and confirm 10 target-disease associations, five of which having our GWAS as the only primary genetic association evidence. We also discovered five potential novel SNP's loci for associated targets. Integrating with the platform, we were able to conduct exploratory evaluation of drug targets identified by our GWASs in comparison with other omics evidences, giving an invaluable insight into further research on drug target selection.

[1] Carvalho-Silva, D., Pierleoni, A., Pignatelli, M., Ong, C., Fumis, L., Karamanis, N., ... & Miranda, A. (2018). Open Targets Platform: new developments and updates two years on. Nucleic acids research, 47(D1), D1056-D1065.
[2] McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., ... & Cunningham, F. (2016). The ensembl variant effect predictor. Genome biology, 17(1), 122.
[3] Koscielny, G., An, P., Carvalho-Silva, D., Cham, J. A., Fumis, L., Gasparyan, R., ... & Pierleoni, A. (2016). Open Targets: a platform for therapeutic target identification and validation. Nucleic acids research, 45(D1), D985-D994.
[4] Association score. (n.d.). Retrieved July 31, 2019, from https://docs.targetvalidation.org/getting-started/scoring

P4-04

Development of Integrated Database “dbTMM” for stratification of cohort participant toward drug development

Satoshi Nagaie[1] nagaie@megabank.tohoku.ac.jp
Yukie Konno[1] y.konno@megabank.tohoku.ac.jp
Naoko Kasahara[1] n.kasahara@megabank.tohoku.ac.jp
Satoshi Mizuno[1] samizuno@med.tohoku.ac.jp
Hiroshi Tanaka[2] tanaka@megabank.tohoku.ac.jp
Soichi Ogishima[1] ogishima@megabank.tohoku.ac.jp

[1] Department of Informatics for Genomic Medicine, Tohoku Medical Megabank Organization, Tohoku University, Seiryo 2-1, Aoba-ku, Sendai Miyagi 980-8573 Japan
[2] Department of Bioclinical Informatics, Tohoku Medical Megabank Organization, Tohoku University, Seiryo 2-1, Aoba-ku, Sendai Miyagi 980-8573 Japan

Keywords: Biobank, Cohort, Stratification, Drug Discovery

Tohoku Medical Megabank project was started to reconstruct from the Great East Japan Earthquake. We started cohort studies in 2013, and successfully recruited more than 150,000 participants for baseline assessment by 2017. Though these cohorts, biospecimens including serum, plasma, mononuclear cells, urine, and the answers to the questionnaires of their lifestyles were collected. A part of participants took detailed health survey in assessment centers and satellites.

We developed the integrated database of Tohoku Medical Megabank project (dbTMM). The dbTMM stores large-scale data of 84,000 cohort participants including whole genome (n=3,500) and SNP genotyping data (n=23,000) (genetic factors), questionnaire data (environmental factors), metabolome and proteome data (molecular phenotypic factors), and clinical data (phenotypic factors). By using dbTMM, researchers can stratify cohort participants having a certain phenotype by genetic and environmental factors. For example, researchers can search dbTMM under the following conditions: chromosome 8, position 41519462 (rs515071) = “TT” & health status (laboratory test data): HbA1c > “6.2” & lifestyle (questionnaire data): alcohol drinking = “Yes” & disease history (questionnaire data): Type II diabetes = “Yes”. As for phenotypic data, we have developed the pipeline for retrieval and structuring clinical data from medical information SS-MIX2 storage. By using structured clinical data, we will conduct deep phenotyping, and will store phenotypic data into dbTMM. We believe our integrated database dbTMM will contribute to stratification of cohort participant toward drug development for common disease which is caused by complex interplay between genetic and environmental factors.

[1] Kuriyama, S. *et al.* The Tohoku Medical Megabank Project: design and mission. *J. Epidemiol.* 26, 493–511 (2016).

P4-05

Development of phenotyping algorithms for hypertensive disorders of pregnancy (HDP) for precise stratification toward drug development

Satoshi Mizuno[1] samizuno@med.tohoku.ac.jp
Maiko Wagata[2] wagata@med.tohoku.ac.jp
Satoshi Nagaie[1] satoshi.nagaie@gmail.com
Gen Tamiya[3] gtamiya@genetix-h.com
Shinichi Kuriyama[4] kuriyama@med.tohoku.ac.jp
Hiroshi Tanaka[5] tanaka@cim.tmd.ac.jp
Nobuo Yaegashi[6] nobuo.yaegashi@gmail.com
Junichi Sugawara[2,6] jsugawara@med.tohoku.ac.jp
Soichi Ogishima[1] ogishima@megabank.tohoku.ac.jp

[1] Department of Informatics for Genomic Medicine, Group of Integrated Database Systems, Tohoku Medical Megabank Organization, Tohoku University, Miyagi, Japan
[2] Department of Feto-Maternal Medical Science, Group of Community Medical Supports, Tohoku Medical Megabank Organization, Tohoku University, Miyagi, Japan
[3] Department of Statistical Genetics and Genomics, Group of Disease Risk Prediction, Tohoku Medical Megabank Organization, Tohoku University, Miyagi, Japan
[4] Department of Molecular Epidemiology, Group of the Birth and Three-Generation Cohort Study, Tohoku Medical Megabank Organization, Tohoku University, Miyagi, Japan
[5] Department of Bioclinical Informatics, Group of Integrated Database Systems, Tohoku Medical Megabank Organization, Tohoku University, Miyagi, Japan
[6] Department of Gynecology and Obstetrics, Tohoku University Graduate School of Medicine, Tohoku University, Miyagi, Japan

Keywords: Phenotyping, hypertensive disorders of pregnancy, Stratification

In recent years, many phenotyping algorithms have been developed for scalable and high-throughput cohort identification from electronic health records (EHRs) [1]. Phenotyping is expected to stratify patients into subgroups that reflect clinical diversity of disease. Precise phenotyping is key for stratified medicine and drug development, because stratification of patients expected to improve lack of efficacy of drugs caused by heterogeneity of patients.

In this study, we developed rule-based phenotyping algorithms of hypertensive disorders of pregnancy (HDP) according to international guideline. The developed algorithm includes blood pressure, proteinuria and unstructured medical data. The developed phenotyping algorithms were applied to phenotype 22,256 pregnancy women in the BirThree cohort study [2]. We also compared the phenotyped subclasses with diagnoses to evaluate the performance of our algorithms against the diagnosis as the gold standard.

Of the 22,256 research subjects, 2,038 (9.16%) were phenotyped as having HDP by the phenotyping algorithm. The performance of our phenotyping algorithms shows high PPV (0.94) and modest NPV (0.65). Our phenotyping algorithm enables us to conduct precise stratification which reflect clinical heterogeneity of the subgroups of HDP toward stratified medicine and drug development.

[1] Wei WQ and Denny JC, *Extracting research-quality phenotypes from electronic health records to support precision medicine*, Genome Med, 30; 7(1):41, 2015.
[2] S Kuriyama *et al*, *The Tohoku Medical Megabank Project: Design and Misson*, J Epidemiol, 26(9):493-511, 2016.

P5-01

Development of a Liver Toxicity Informatics System (AMED project)

Hiroshi Yamada[1] h-yamada@nibiohn.go.jp
Yoshinobu Igarashi[1] y-igarashi@nibiohn.go.jp
Noriyuki Nakatsu[1] nnakatsu@nibiohn.go.jp
Yuki Yamagata[1] y-yamagata@nibiohn.go.jp
Hiroshi Akimaru[2] h-akimaru8@nibiohn.go.jp
Erika Akimaru[2] eakimaru@nibiohn.go.jp
Katsuhisa Horimoto[3] k.horimoto@aist.go.jp
Kazuhiko Fukui[3] k-fukui@aist.go.jp
Takumi Era[4] tera@kumamoto-u.ac.jp
Tadahiro Numakawa[4] numakawa@kumamoto-u.ac.jp
Shirou Matsumoto[5] s-pediat@gpo.kumamoto-u.ac.jp
Yoshihiro Uesawa[6] uesawa@my-pharm.ac.jp
Junko Nagai[6] nagai-j@my-pharm.ac.jp
Tamotsu Noguchi[7] noguchit@my-pharm.ac.jp

1 Toxicogenomics Informatics Project, National Institutes of Biomedical Innovation, Health and Nutrition(NIBIOHN), 7-6-8 Asagi Saito Ibaraki-City Osaka 567-0085, Japan
2 Center for Rare Disease Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Asagi Saito Ibaraki-City Osaka 567-0085, Japan
3 Molecular Profiling Research Center for Drug Discovery, National Institute of Advanced Industrial Science and Technology(AIST), 2-4-7 Aomi,Koto-ku,Tokyo 135-0064, Japan
4 Department of Cell Modulation, Institute of Molecular Embryology and Genetics, Kumamoto University, 2-2-1 Honjo,Chuo-ku,Kumamoto 860-0811, Japan
5 Department of Pediatrics, Graduate School of Medical Sciences, Kumamoto University, 1-1-1 Honjo,Chuo-ku,Kumamoto 860-8556, Japan
6 Department of Medical Molecular Informatics, Meiji Pharmaceutical University, 2-522-1 Noshio, Kiyose, Tokyo 204-8588, Japan
7 Laboratory of Computational Biology, Meiji Pharmaceutical University, 2-522-1 Noshio, Kiyose, Tokyo 204-8588, Japan

Keywords: Drug induced liver injury (DILI), Liver toxicity, Database, Prediction system

In order to contribute to the improvement of the ability related to safety evaluation in an academic drug discovery, the five-yearly program for “Development of a Drug Discovery Informatics System” supported by Japan Agency for Medical Research and Development (AMED) was started in 2015. In this program, a collaborative project of NIBIOHN, AIST, Kumamoto university and Meiji pharmaceutical university for “Liver Toxicity Informatics System Development” is organized. The project has established 3 databases related to liver toxicity (DILI-cSEARCH, TOXPILOT, toxBridge) and 4 toxicity prediction systems (DILI-PANEL, toxRANK, LIVER/MIE-QSAR, LUNG/MIE-QSAR). Some databases and prediction systems are already open and can be accessed via portal site of the project (DILI-TOOLBOX, https://dili-toolbox.nibiohn.go.jp).
In this poster, we will introduce an overview of our databases and prediction systems.

Acknowledgment
This research is supported by Japan Agency for Medical Research and Development, AMED under Grant Number 19nk0101103h0005.

P5-02

DILI-cSEARCH: a DILI database for drug safety assessment

Yoshinobu Igarashi[1] y-igarashi@nibiohn.go.jp
Noriyuki Nakatsu[1] nnakatsu@nibiohn.go.jp
Yuki Yamagata[1] y-yamagata@nibiohn.go.jp
Hiroshi Akimaru[2] h-akimaru8@nibiohn.go.jp
Erika Akimaru[2] eakimaru@nibiohn.go.jp
Shirou Matsumoto[3] s-pediat@gpo.kumamoto-u.ac.jp
Tadahiro Numakawa[4] numakawa@kumamoto-u.ac.jp
Takumi Era[4] tera@kumamoto-u.ac.jp
Hiroshi Yamada[1] h-yamada@nibiohn.go.jp

[1] Toxicogenomics Informatics Project, National Institutes of Biomedical Innovation, Health and Nutrition (NIBIOHN), 7-6-8 Asagi Saito Ibaraki-City Osaka 567-0085, Japan
[2] Center for Rare Disease Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Asagi Saito Ibaraki-City Osaka 567-0085, Japan
[3] Department of Pediatrics, Graduate School of Medical Sciences, Kumamoto University, 1-1-1 Honjo, Chuo-ku, Kumamoto 860-8556, Japan
[4] Department of Cell Modulation, Institute of Molecular Embryology and Genetics, Kumamoto University, 2-2-1 Honjo, Chuo-ku, Kumamoto 860-0811, Japan

Keywords: DILI, Transcriptomics, Human primary hepatocyte, Database

Among drug-induced toxicity, liver toxicity is one of the leading causes of discontinuation of drug development and market withdrawal after-sales. For this reason, it is desired to create a new method which is capable of finding and evaluating potential hepatotoxic compounds in the early stage of drug development using in vitro or in silico approaches, which are faster and less expensive than animal experiments. In this study, we constructed the DILI (drug-induced liver injury) database as an information infrastructure in "Development of an informatics system for predicting drug-induced liver injury," which plays a part of AMED's "the Drug Discovery Support Promotion Project." We assembled and organized data on hepatotoxicity from the four independent databases (FDA / LTKB, NIH / LINCS, NTP / Drug Matrix, NIBIOHN / Open TG-GATEs) to a new integrated database. This new database contains more than 800 compound information (mainly on pharmaceutical drugs), gene expression information, biochemistry, and hematology information. We also implemented a flexible search interface that enables to explore in complex data structures. This database called "DILI information cross-search system" is available at https://dili-csearch.nibiohn.go.jp.

Acknowledgment
This research is supported by Japan Agency for Medical Research and Development, AMED under Grant Number 19nk0101103h0005.

Ontology-based Toxic Process Interpretable Knowledge System for Drug-Induced Liver Injury

Yuki Yamagata[1] y-yamagata@nibiohn.go.jp
Yoshinobu Igarashi[1] y-igarashi@nibiohn.go.jp
Noriyuki Nakatsu[1] nnakatsu@nibiohn.go.jp
Katsuhisa Horimoto[2] k.horimoto@aist.go.jp
Kazuhiko Fukui[2] k-fukui@aist.go.jp
Yoshihiro Uesawa[3] uesawa@my-pharm.ac.jp
Hiroshi Yamada[1] h-yamada@nibiohn.go.jp

[1] Toxicogenomics, Informatics Project, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Asagi Saito Ibaraki-City Osaka 567-0085, Japan
[2] Molecular Profiling Research Center for Drug Discovery, National Institute of Advanced Industrial Science and Technology (AIST), 2-4-7 Aomi, Koto-ku, Tokyo, 135-0064, Japan
[3] Department of Clinical Pharmaceutics, Meiji Pharmaceutical University, 2-522-1 Noshio, Kiyose, Tokyo 204-8588, Japan

Keywords: Ontology, Drug induced liver injury, Hepatotoxicity, Toxic course, Toxic process

Hepatotoxic mechanisms are complex and require multidisciplinary expertise. Therefore, new technology is necessary for the understanding of mechanisms, knowledge sharing, and knowledge reusing across domains. We have constructed an ontology, TXPO, related to the hepatotoxicity process [1]. In this research, we report on the toxic process interpretable support knowledge system (TOXPILOT) as application work. TOXPILOT is a web application system that aims to provide necessary information for interpretation of toxicity mechanism. By utilizing Semantic Web technology, it has functions such as visualization of hepatotoxic course, display of functioning processes in the organism, and provides various related information based on ontology. Furthermore, it is also possible to search for routes that focus on the specific process. This system is expected to be a knowledge sharing tool for users involved in safety assessment because it can grasp toxic mechanisms cross-disciplinary. The prototype of TOXPILOT is open at the following site: https://toxpilot.nibiohn.go.jp. Users can also access this knowledge system via our portal site (https://dili-toolbox.nibiohn.go.jp).

Acknowledgment
This research is supported by Japan Agency for Medical Research and Development, AMED under Grant Number 19nk0101103h0005.

[1] Yamagata Y., Igarashi Y., Nakatsu N., Horimoto K., Fukui K., Uesawa Y., and Yamada H., Development of Liver Toxicity Ontology for Drug Safety Evaluation and its Application, *Transactions of the Japanese Society for Artificial Intelligence* 34(2), D-I81_1-18, 2019 [in Japanese].

P5-04

Development of an informatics system for predicting cardiotoxicity: 6. Update of the AMED cardiotoxicity database and the hERG prediction model with additional assays for the compounds selected by active learning.

Tomohiro Sato **Hitomi Yuki** **Teruki Honma**

tomohiro.sato@riken.jp hitomi.yuki@riken.jp honma.teruki@riken.jp

Center for Biosystems Dynamics Research, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama City, Kanagawa, 230-0045, Japan

Keywords: hERG, QT prolongation, database

The inhibition of hERG potassium channel is closely related to the prolonged QT interval [1], and to assess the risk could greatly contribute to the development of safer therapeutic compounds. In the optimization stage of drug development, quantitative prediction of hERG inhibitory activity is crucial to design drug candidates without cardiotoxicity risk.

We previously reported the construction of AMED cardiotoxicity database which integrated the information of small molecule inhibitors of hERG, Nav1.5, Kv1.5, and Cav1.2 from various public data sources, and the prediction models based on the integrated data. [2, 3] Despite the successful development of the largest dataset for the cardiac ion channels, the applicability of the prediction models was still greatly affected by the chemical space coverage of the dataset. In this study, we performed additional hERG inhibition assays employing both patch clamp and binding assays to improve the hERG prediction model. Table 1 shows the number of hERG activity information in the public databases and newly measured data in this study. The compound selection was performed based on 1. known hERG inhibitors to obtain IC_{50} values determined by a single assay protocol, and 2. active learning to add novel compounds for which the prediction model based on public data cannot predict the potency. The effect by the additional compounds to the prediction performance and the applicability domain of the hERG prediction model is to be discussed.

Table 1. The number of hERG inhibitors and inactive compounds in each database

Database	hERG inhibitors (IC50≦10uM)	Inactive compounds (IC50>10uM)	Compounds (total)
ChEMBL (v24)	5,092	5,660	10,752
GOSTAR	3,043	3,124	6,167
PubChem (AID588834)	382	2,964	3,346
hERG Central	4,321	274,536	278,857
public databases	10,269	282,611	292,880
RIKEN hERG assay	924	1,725	2,649
AMED cardiotoxicity database	10,940	283,865	294,805

[1] S. Picard, S. Goineau, P. Guillaume, J. Henry, J.L. Hanouz, R. Rouet, Supplemental studies for cardiovascular risk assessment in safety pharmacology: a critical overview, *Cardiovasc. Toxicol.*11, 285–307, 2011

[2] T. Sato, H. Yuki, T. Honma, Construction of an integrated database for hERG blocking small molecules. *PLOS ONE* 13(7), e0199348.

[3] K. Ogura, T. Sato, H. Yuki, T. Honma, Support Vector Machine model for hERG inhibitory activities based on the integrated hERG database using descriptor selection by NSGA-II. *Scientific Reports* (in press)

P5-05

Development of an informatics system for predicting cardiotoxicity: 7. hERG prediction model based on docking simulation and interaction descriptors with hERG residues

Hitomi Yuki hitomi.yuki@riken.jp
Tomohiro Sato tomohiro.sato@riken.jp
Teruki Honma honma.teruki@riken.jp

RIKEN Center for Biosystems Dynamics Research, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Keywords: hERG, Ligand docking, Machine learning, Interaction fingerprint

hERG (human ether-à-go-go-related potassium channel) is a member of potassium ion channels and involved in cardiac repolarization. Drug-induced hERG blocking causes long QT syndrome and it is a major cause of cardiotoxicity. Drug design for avoiding the hERG inhibition is quite important through the all stages of drug discovery. Although 3D models of hERG would provide worthy compound design ideas for reducing hERG inhibition, there was no experimentally validated 3D model for several decades. Using fingerprints or 2D descriptors of compounds have been a common way to predict their hERG inhibitory. Finally the molecular structure of hERG was determined by cryo-electron microscopy in 2017.[1] However, the drug-binding site seemed to be too narrow for known inhibitors such as dofelitide to bind.

To utilize the hERG 3D structures and design compounds for avoiding the hERG inhibition, we developed prediction models based on docking simulation and interaction descriptors with hERG residues. We performed MD simulation of the hERG cryo-electron microscopy structure and subsequent docking simulations of known representative hERG inhibitors and similar compounds measured their hERG binding affinity at RIKEN. Then we made and validated classification models for hERG binding inhibitory based on interaction descriptors between a compound and each residue/atom of hERG. We would like to show and discuss the analysis of these results and how to use hERG 3D structures.

This work was done as a part of “Construction of Drug Discovery Informatics System” supported by Japan Agency for Medical Research and Development (AMED). The computing resources were provided by the HOKUSAI-GreatWave system at RIKEN.

[1] Wang, W., and MacKinnon, R. Cryo-EM Structure of the Open Human Ether-à-go-go-Related K+ Channel hERG. *Cell* 169:422–430.e10, 2017

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 13. Development of DruMAP, Drug Metabolism and pharmacokinetics Analysis Platform

Hitoshi Kawashima[1] hkawashi@nibiohn.go.jp
Reiko Watanabe[1] reiko-watanabe@nibiohn.go.jp
Tsuyoshi Esaki[1,2] tsuyoshi-esaki@biwako.shiga-u.ac.jp
Chioko Nagao[3] chio@nibiohn.go.jp
Yayoi Natsume-Kitatani[1] natsume@nibiohn.go.jp
Rikiya Ohashi[1,4] rikiya-ohashi@nibiohn.go.jp
Hiroshi Komura[5] h-komura@ado.osaka-cu.ac.jp
Kenji Mizuguchi[1,3] kenji@nibiohn.go.jp

[1] Laboratory of Bioinformatics, AI Center for Health and Biomedical Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8, Saito Asagi, Ibaraki, Osaka 567-0085, Japan
[2] The Center for Data Science Education and Research, Shiga University, 1-1-1, Banba, Hikone, Shiga 522-8522, Japan
[3] Laboratory of In-silico Drug Design, Center of Drug Design Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8, Saito Asagi, Ibaraki, Osaka 567-0085, Japan
[4] Discovery Technology Laboratories, Mitsubishi Tanabe Pharma Corporation, 2-2-50, Kawagishi, Toda, Saitama 335-8505, Japan
[5] Research Center for Industry Innovation and Osaka City University Incubator, Osaka City University, 1-2-7, Asahimachi, Abeno, Osaka, 545-0051, Japan

Keywords: Pharmacokinetics, Drug metabolism, Database, Prediction models, Web application

We started the initiative “Development of a Drug Discovery Informatics System” four years ago with support from the Japan Agency for Medical Research and Development (AMED). The main aim of this initiative is to develop accurate prediction systems for DMPK (Drug Metabolism and PharmacoKinetics), primarily targeting academic scientists.

We collected pharmacokinetic and physicochemical parameters from ChEMBL or other public sources. However, since ChEMBL compiles data obtained in different experimental conditions, we selected the data measured in compatible conditions and reformatted the results as appropriate for our prediction system. In addition to the public data, we have acquired *in vitro* and *in vivo* experimental data under unified conditions.

We stored these data to a database named DruMAP (Drug Metabolism and pharmacokinetics Analysis Platform). We have constructed prediction models for several DMPK parameters using these data, and additional model construction is currently ongoing. DruMAP contains experimental and predictive values for those parameters for about 18,000 compounds and the users can search the database by key words, compound IDs, structural descriptions such as InChI, InChI Key or SMILES, and by parameter values. DruMAP also provides the function to predict DMPK parameters for the user input compounds using our prediction models. DruMAP can be used for early DMPK studies and for candidate compound selection to accelerate novel drug development. DruMAP can be accessed at: https://drumap.nibiohn.go.jp/.

This work was conducted as a part of “Development of a Drug Discovery Informatics System” supported by Japan Agency for Medical Research and Development (AMED).

P5-07

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 14. *In silico* three-class predictor of human intestinal absorption with Caco-2 permeability and dried-DMSO solubility

Tsuyoshi Esaki[1,2]
tsuyoshi-esaki@biwako.shiga-u.ac.jp

Reiko Watanabe[2]
reiko-watanabe@nibiohn.go.jp

Rikiya Ohashi[2,3]
rikiya-ohashi@nibiohn.go.jp

Yayoi Natsume-Kitatani[2]
natsume@nibiohn.go.jp

Hitoshi Kawashima[2]
hkawashima@nibiohn.go.jp

Chioko Nagao[4]
chio@nibiohn.go.jp

Hiroshi Komura[5]
h-komura@ado.osaka-cu.ac.jp

Kenji Mizuguchi[2,4]
kenji@nibiohn.go.jp

[1] The Center for Data Science Education and Research, Shiga University, 1-1-1, Banba, Hikone, Shiga 522-8522, Japan
[2] Laboratory of Bioinformatics, AI Center of Drug Design Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8, Saito Asagi, Ibaraki, Osaka 567-0085, Japan
[3] Discovery Technology Laboratories, Mitsubishi Tanabe Pharma Corporation, 2-2-50, Kawagishi, Toda, Saitama 335-8505, Japan
[4] Laboratory of In-silico Drug Design, Center of Drug Design Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Saito Asagi, Ibaraki, Osaka 567-0085, Japan
[5] University Research Administrator Center, Osaka City University, 1-2-7 Asahi, Abeno-ku, Osaka, 545-0051, Japan

Keywords: Human intestinal absorption, Membrane permeability, Solubility, Machine learning

Oral drug administration is one of the most preferred delivery routes and is currently the goal of new drug development owing to the ease of administration and patient compliance. In the drug discovery stage, lead compounds with high intestinal absorption have been found to be very important because intestinal absorption is the first step after dosing. Orally administered drugs reaching the intestine pass through the intestinal epithelial cell membrane to the circulating blood and interact with targets such as receptors, channels, and enzymes in several tissues. The administration of compounds with poor human intestinal absorption (HIA) may cause substantial individual variation in pharmacokinetics and may affect the desired pharmacological effects.

Therefore, we developed a freely available model classifying compounds using a three-level prediction capacity for the fraction of absorption in human (Fa) with membrane permeability measured using Caco-2 cell line (Papp) and solubility measured using dried-DMSO method (D-Sol); only the chemical structure of compounds was employed [1]. To support Fa prediction, binary classification models of Papp and D-Sol were constructed. To construct the Fa prediction model with high performance, we compared several machine learning methods and descriptors such as physicochemical, substructure, and pharmacokinetics (Papp and D-Sol) parameters and found that adding both predicted Papp and D-Sol was effective in classifying Fa (Accuracy: 0.836 and Kappa: 0.560).

In this presentation, we would like to present the results and to introduce an open web application (https://drumap.nibiohn.go.jp/fa) for three parameters (Fa, Papp and D-Sol) prediction.

This work was conducted as part of “Development of a Drug Discovery Informatics System” supported by Japan Agency for Medical Research and Development (AMED).

Reference: [1] Esaki T., et al., *J. Pharma. Sci.*, 2019, DOI: 10.1016/j.xphs.2019.07.014

P5-08

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 15. Development of an *in silico* prediction system of human renal excretion and clearance from chemical structure information

Reiko Watanabe[1] reiko-watanabe@nibiohn.go.jp
Rikiya Ohashi[1,2] rikiya-ohashi@nibiohn.go.jp
Tsuyoshi Esaki[1,3] tsuyoshi-esaki@biwako.shiga-u.ac.jp
Hitoshi Kawashima[1] hkawashi@nibiohn.go.jp
Yayoi Natsume-Kitatani[1] natsume@nibiohn.go.jp
Chioko Nagao[1,4] chio@nibiohn.go.jp
Kenji Mizuguchi[1,4] kenji@nibiohn.go.jp

[1] Laboratory of Bioinformatics, AI Center for Health and Biomedical Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8, Saito Asagi, Ibaraki, Osaka 567-0085, Japan
[2] Discovery Technology Laboratories, Mitsubishi Tanabe Pharma Corporation, 2-2-50, Kawagishi, Toda, Saitama 335-8505, Japan
[3] The Center for Data Science Education and Research, Shiga University, 1-1-1, Banba, Hikone, Shiga 522-8522, Japan
[4] Laboratory of In-silico Drug Design, Center of Drug Design Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8, Saito Asagi, Ibaraki, Osaka 567-0085, Japan

Keywords: Renal clearance, protein unbound in plasma, machine learning

In the early stages of drug development, predicting the pharmacokinetic profiles of new chemical entities is essential to minimize the risks of potential withdrawals. The excretion of unchanged compounds by the kidney constitutes a major route in drug elimination and plays an important role in pharmacokinetics. Excretion in the urine is the consequence of a complex mechanism that involves three main processes: glomerular filtration, tubular secretion, and reabsorption. We had previously developed *in silico* models to predict the values of fraction unbound in plasma ($f_{u,p}$), which is an important determinant of drug efficacy in pharmacokinetic and pharmacodynamic studies. In this study, we have extended our approach and built predictive models for two important pharmacological indicators of drugs in renal excretion: the fraction of drug excreted unchanged in urine (f_e) and renal clearance (CL_r). Experimental values of f_e and CL_r were extracted from the literature and the ChEMBL database. The binary classification model for f_e demonstrated a balanced accuracy of 0.74. The two-step prediction system for CL_r was generated by using the combination of a classification model to predict excretion-type compounds and regression models to predict the value of CL_r in each excretion type; moreover, the accuracies of the regression models were increased by adding observed and predicted $f_{u,p}$ values as a descriptor. Our predictive models can be integrated into other pharmacokinetic modeling systems, which would be highly useful in academic drug discovery.

[1] Bohnert T, Gan L. S, Plasma protein binding: from discovery to development. J Pharm Sci 2013, 102 (9), 2953-94.
[2] Varma M, Obach Bo R, Troutman M, Physicochemical determinants of human renal clearance. J Med Chem 2009, 52, 4844-52.
[3] Ito S, Ando H, Ose A, Kitamura Y, Ando T, Kusuhara H, Sugiyama Y, Relationship between the urinary excretion mechanisms of drugs and their physicochemical properties. J Pharm Sci 2013, 102(9), 3294-301

This work was conducted as part of “Development of a Drug Discovery Informatics System” supported by Japan Agency for Medical Research and Development (AMED).

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 16. Prediction of sites of metabolism of drug by CYP2C9 by molecular simulation

Hiroaki Saito, hiroaki.saito@riken.jp, **Takao Otsuka,** totsuka@riken.jp, **Noriaki Okimoto,** okimoto@riken.jp, **Makoto Taiji** taiji@riken.jp

Laboratory of computational molecular design computational biology, RIKEN Center for Biosystems Dynamics Research (BDR)

Keywords: Molecular dynamics, Molecular docking, Site of metabolism, Cytochrome P450 (CYP)

Cytochrome P450s (CYPs), a superfamily of hem-containing enzymes, are the major enzymes involved in drug metabolism. In humans, it has been estimated that CYPs metabolize approximately 75% of all marketed drugs, 95% of which are metabolized by CYP3A4, 2D6, 2C9 and 1A2 [1]. In the drug metabolism by the CYPs, identification of sites of metabolism (SOMs) on molecules and the structure of their metabolites can be decisive for the design of molecules with favorable metabolic properties. However, experimental techniques to determine SOMs and structures of metabolites are still highly resource-demanding and challenging [2]. Thus, developing fast and accurate computational methods to predict the SOMs/products of compounds metabolized by the CYPs is one of the important tasks for the optimization of ADME and toxicity properties.

In this study, we present a new computational method (score function) to predict SOMs of compounds metabolized by CYP2D6. The new score function is composed of accessibility and reactivity scores. The accessibility scores for the sites (atoms) of compounds are estimated by ensemble docking simulation, while the activation energies of atoms estimated from the SMARTCyp software[3] are used as the reactivity scores in the method. We carried out a molecular dynamics (MD) simulation of apo type of CYP2D6 for 10 micro second to sample the receptor (pocket) structures for the ensemble docking. For the molecular docking, myPresto program with grid score, which are optimized for CYP system [4], was used. We prepared 24 test set compounds metabolized by the CYP2D5 and 100,000 receptor snapshots (which correspond to the 10 micro second MD simulation) for the ensemble docking simulation. We found the success rate of the predicted SOMs was 88 %, showing better result than those by the ensemble docking alone.

[1] Wienkers LC, heath TG., Nature Rev. Drug. Discov. 4 (2005), 825-833.
[2] Johannes Kirchmair et al., J. Chem. Inf. Model. 52 (2012), 617−648.
[3] P. Rydberg et al., ACS Med. Chem. Lett., 1 (2010), 96-100
[4] Kiumars Shahrokh, Anita Orendt, Garold S. Yost, Thomas E. Cheatham, III, J. Comp. Chem., 33 (2012), 119-33.

This work is conducted as a part of “Construction of Drug Discovery Informatics System” supported by Japan Agency for Medical Research and Development (AMED).

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 17. Accuracy and performance of the MDGRAPE-4A system

Gentaro Morimoto gentaro.morimoto@riken.jp
Yousuke Ohno ohno@riken.jp
Teruhisa Komatsu teruhisa.komatsu@riken.jp
Hao Zhang hao.zhang@riken.jp
Itta Ohmura ohmura@riken.jp
Keigo Nishida keigo.nishida.jg@riken.jp
Yohei M. Koyama ym.koyama@riken.jp
Hiroshi Koyama hkoyama@riken.jp
Makoto Taiji taiji@riken.jp

Laboratory for Computational Molecular Design, RIKEN Center for Biosystems Dynamics Research (BDR), 6-2-4 Furuedai, Suita, Osaka 565-0874, Japan

Keywords: Molecular dynamics simulations, Ligand binding, Site of metabolism, Drug design

We have been developing a series of special purpose computers for molecular dynamics simulations. The target performance is to make it possible to simulate typical protein-ligand complex surrounded by water for microseconds per day, which is two orders of magnitude faster than commercially available systems. One of the applications of long-term MD simulations of proteins is to predict the site of metabolism. In 2019 we have completed the MDGRAPE-4A, an improved version of MDGRAPE-4[1]. The MDGRAPE-4 hardware development was completed in 2014, and the stabilization of operation and software development process succeeded for two years. However, the final performance did not reach the expected one. In the process of the development, we realized that not only accelerating the computationally demanding part but also more specialized hardware for the type of computation is required to improve the performance.

The MDGRAPE-4 and 4A systems both consist of 512 custom System on Chip (SoC) LSI chips in 3D torus network. The biomolecular systems are spatially divided into small cells and each LSI communicate with each other to update the allocated cells. To improve the performance, we revised the SoC LSI to enhance the functionalities as follows:

1) support of management and migration of atoms by the memory hardware
2) more sophisticated treatment of exclusion and reduced non-bonded interaction pairs in the special purpose pipelines
3) implementation SIMD functions and special instructions required in the MD simulations in the general-purpose cores
4) reduction of the latency and more flexible routing patterns in the network interface
5) support of a long-range electrostatic interaction calculation method that is suitable for the parallel systems with torus network topology

Currently the hardware installation has been completed. Software is still under development to improve the accuracy and the performance and to implement more functionalities. In this presentation, we would like to show some results of the evaluation of the accuracy and the performance of the MD simulations by using the MDGRAPE-4A system.

[1] Ohmura, I., *et al.*, MDGRAPE-4: a special-purpose computer system for molecular dynamics simulations, *Phil. Trans. R. Soc. A*, 372:20130387, 2014.

Some of this work is conducted as a part of “Construction of Drug Discovery Informatics System” supported by Japan Agency for Medical Research and Development (AMED).

QSAR model to predict $K_{p,uu,brain}$ with small-scale dataset -incorporating predicted values of related parameters-

Yuki Umemori[1]
y.umemori@teijin.co.jp

Koichi Handa[1]
ko.handa@teijin.co.jp

Seishiro Sakamoto[1]
se.sakamoto@teijin.co.jp

Shin Umeda[1]
s.umeda@teijin.co.jp

[1] DMPK Research Department, Teijin Pharma Limited. 4-3-2, Asahigaoka, Hino, Tokyo, 191-8512, Japan

Keywords: $K_{p,uu,brain}$, CNS, QSAR model

[Objective] Unbound brain-to-blood ratio ($K_{p,uu,brain}$) is widely used as an indicator of brain penetration (BP). However, it is hard to obtain the experimental value of $K_{p,uu,brain}$ for many compounds in drug discovery stage, because it requires many experiments. So, some pharmaceutical companies targeting central nervous system (CNS) are actively performing in silico prediction of $K_{p,uu,brain}$ using their in-house QSAR model[1]. Their models were constructed using not only small-scale published dataset but also large scale in-house dataset. For other pharmaceutical companies which does not have enough in-house dataset, it is difficult to develop such QSAR model[2]. Recently, we can use some software which provide each predicted value for many PK parameters. Although they cannot predict $K_{p,uu,brain}$ itself, they can predict some BP-related parameters (BPrPs). In this study, with the aim of constructing an accurate $K_{p,uu,brain}$ prediction model with only a small-scale published dataset, we examined an effect of combination of structure descriptors and the predicted values of some BPrPs as explanatory variables.

[Method] We used 88 compounds which experimental $K_{p,uu,brain}$ were available as our dataset. Structure descriptors were calculated using ADMET Predictor® 9.0 (AP) and Schrödinger 2016. The predicted values of several BPrPs were also calculated using AP. Random Forest [software: R3.4.4 (caret)] was used as a machine learning model. RMSE and R^2 calculated by 5-fold cross validation were used as the indices of prediction accuracy. Firstly, we developed prediction models using structure descriptors. Then, we verified a predictive accuracy of each model which were incorporated predicted value(s) of BPrPs in various combinations as explanatory variables. In order to confirm the practicability of our model, $K_{p,uu,brain}$ of in-house compounds were predicted, and compared with the experimental values.

[Result and Discussion] Compared to basic model, most models which were incorporated the predicted value(s) of BPrPs showed much higher prediction accuracy. Among BPrPs, brain-to-blood ratio (LogBB) contributed to much improve predictability. Our final model constructed by only small-scale published dataset showed good predictability for in-house compounds. In conclusion, we could build a highly accurate $K_{p,uu,brain}$ prediction model by incorporating BPrP for a small-scale dataset. Based on this finding, it was shown that it is possible to build a highly accurate prediction model by using relevant information even with a small-scale dataset.

[1] Liu, H. et al. (2018) Prediction of brain: Blood unbound concentration ratios in CNS drug discovery employing in silico and in vitro model systems. Drug. Discov. Today. (7): 1357-1372
[2] Zhang, YY et al. (2016) Integrate in silico and in vitro approaches to predict drug accessibility to the central nervous system. Mol. Pharm. 13, 1540–1550

P5-12

Cluster Gauss-Newton method for efficiently estimating multiple sets of parameters: Application to Physiologically-Based PharmacoKinetic models

Ken Hayami[1,2] hayami@nii.ac.jp
Yasunori Aoki[1,3] yaoki@uwaterloo.ca
Kota Toshimoto[3] kota.toshimoto@riken.jp
Yuichi Sugiyama[3] yu-one.sugiyama@nifty.com

[1] National Institute of Informatics
[2] Department of Informatics, School of Multidisciplinary Sciences, SOKENDAI (The Graduate University for Advanced Studies)
[3] Sugiyama Laboratory, RIKEN Baton Zone Program, RIKEN

Keywords: parameter estimation, nonlinear least squares problem, multiple global minimizers PKPK model, pharmacokinetics

Parameter estimation problems of mathematical models can often be formulated as nonlinear least squares problems. Typically, these problems are solved numerically using iterative methods. The local minimizer obtained using these iterative methods usually depends on the choice of the initial iterate. Especially, when there is no unique global minimum to the nonlinear least squares problem, the algorithm finds one of the solutions near the initial iterate. Hence, the estimated parameter and subsequent analyses using the estimated parameter depends on the choice of the initial iterate. One way to reduce the analysis bias due to the choice of the initial iterate is to repeat the algorithm from multiple initial iterates. However, the procedure can be computationally intensive and is not often implemented in practice. In pharmaceutical science, there are many mathematical models where the parameters are not uniquely identifiable from the observation, and this issue of non-unique global minimizers appears often. For example, the model parameters of the physiologically based pharmacokinetics models are usually not uniquely identifiable from the plasma concentration.

To overcome this problem, we propose the Cluster Gauss-Newton (CGN) method [1], an efficient algorithm for finding multiple global minimizers of nonlinear least squares problems. CGN simultaneously solves the nonlinear least squares problem from multiple initial iterates. Then, CGN iteratively improves the solutions from these initial iterates similarly to the Gauss-Newton method. However, it uses a global linear approximation instead of the gradient. The global linear approximations are computed collectively among all the initial iterates to minimize the computational cost and increase the robustness against convergence to local minima.

We use three PBPK model parameter analysis examples to demonstrate the use of CGN. We show that the proposed algorithm is computationally more efficient and robust against local minima compared to many of the conventional algorithms. Besides the algorithm development and the numerical experiments, we have implemented a Graphical User Interface based software to disseminate our algorithm to pharmaceutical scientists.

[1] http://arxiv.org/abs/1801.06714v2

Prediction of Health Effects of Food Peptides and Elucidation of The Mode-of-action Using Multi-task Graph Convolutional Neural Networks

Itsuki Fukunaga[1]
q238056i@mail.kyutech.jp

Ryusuke Sawada[1]
sawad330@bio.kyutech.ac.jp

Tomokazu Shibata[1]
shiba535@bio.kyutech.ac.jp

Kazuma Kaitoh[1]
kaito168@bio.kyutech.ac.jp

Yukie Sakai[1]
sakai541@bio.kyutech.ac.

Yoshihiro Yamanishi[1]
yamani@bio.kyutech.ac.jp

[1]Department of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan

Keywords: foods, peptides, health effects, peptide-protein interaction, graph convolutional neural network, multi-task learning, deep learning

Food proteins work not only as nutrients but also modulators for physiological functions of human body, and they are at the forefront of scientific and industrial attentions. These physiological functions are basically regulated by some peptides encrypted in the food protein sequences, which we call “food peptides”. These bioactive food peptides can have beneficial health effects and can be a lead compound for the development of functional and nutraceutical foods. In fact, there are a number of bioactive food peptides that are beneficial for human health, but the mode-of-action of bioactive food peptides has not been well understood, and their potential health effects remain to be identified. In this study, we propose a novel deep learning-based method to elucidate the mode-of-action of functional food peptides and to predict new health effects of the food peptides, which is made possible by an extension of the concept of drug repositioning from drugs to food peptides. In the algorithm, we represent each peptide structure by a chemical graph, estimate its potential target proteins by multi-task graph convolutional neural networks (GCNNs), and predict new its health effects using the information on therapeutic targets for diseases [1]. We construct predictive models based on 23,933 peptide-protein interactions involving 10,951 peptides and 2533 proteins, and apply the models to analyzing available food peptides including foods for specified health use (referred to TOKUHO in Japan). We also compare the performance of the GCNN model with that of the recurrent neural network (RNN) model in which each peptide structure is represented by a string of amino acids. We show some examples of the mode-of-action analysis and newly predicted health effects of food peptides. The proposed methods are expected to be useful for practical applications in the development of functional foods, drugs, and cosmetics.

[1] Sawada, R., Iwata, H., Mizutani, S., and Yamanishi, Y., Target-based drug repositioning using large-scale chemical-protein interactome data, *Journal of Chemical Information and Modeling*, 55(12), 2717–2730, 2015.

P5-14

Case study of machine learning for drug metabolism and pharmacokinetics properties

Kazuyoshi Yoshii[1]
kazuyoshi-yoshii@zeria.co.jp

Yusuke Hoshino[1]

Ryoko Toda[1]

Yoshihiro Kawabata[1]

[1] Toxicology&Pharmacokinetics Research, Central Research Laboratories,
Zeria Pharmaceutical Co., Ltd, 2512-1, Numagami, Oshikiri, Kumagaya-shi, Japan

Keywords: drug metabolism, cytochrome P450, machine learning

Machine learning is used in natural language processing, analysis of stock market, customer consumption etc. As data in these fields are complex and diverse, a gold standard method hasn't been built for machine learning in these fields. Machine learning in drug metabolism and pharmacokinetics (DMPK) is basically based on compound structure information, so a common strategy in feature engineering and classifier selection might be developed. To explore the common strategy for machine learning in DMPK, the effects of feature engineering and classifier selection on prediction were evaluated using open source data of cytochrome P450 (CYP) inhibition.

Features of compounds were generated using open source software. Number of features, conversion to log-scale and deletion of noise were evaluated. One to eight classifiers were used. The effects of voting and stacking were evaluated.

Accuracy of prediction using untransformed features depended on classifier in the range of 0.8 to 0.85. Deletion of noise in features highly improved accuracy of prediction to approximately 0.9. The classifier using gradient boosting accurately predicted positive and negative about CYP inhibition. The voting and stacking using some classifiers slightly improved accuracy of prediction.

These results suggested that feature engineering like deletion of noise was most effective in this study. If cases like this study were accumulated in machine learning of DMPK, common strategy would be found in the near future.

Free Energy Landscapes of Cyclic Hexapeptide Diastereomers by Multicanonical Molecular Dynamics Simulations

Satoshi Ono[1] **Matthew R. Naylor**[2] **Chad E. Townsend**[2]
Ono.Satoshi@mg.mt-pharma.co.jp

Chieko Okumura[1] **Okimasa Okada**[1] **R. Scott Lokey**[2]

[1] Discovery Technology Laboratories, Mitsubishi Tanabe Pharma Corporation, 1000, Kamosida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan
[2] Department of Chemistry and Biochemistry, University of California Santa Cruz, 1156 High Street, Santa Cruz, California 95064, United States

Keywords: Cell permeability, Cyclic peptide, Molecular dynamics simulations

Efficient conformational sampling of cyclic peptides is an emerging interest for assigning physical properties, such as cell permeability. Conventionally, conformational sampling in chloroform has been used to mimic the membrane environment. Recently, experimental hydrocarbon-water distribution coefficients, $\log D_{hc/w}$, showed excellent correlation with cell permeability [1], suggesting that hydrocarbons, such as cyclohexane, can be alternative solvents for mimicking membrane environments.

Multicanonical molecular dynamics (McMD) simulations were performed to obtain conformational ensembles of eight cyclic hexapeptide diastereomers in explicit cyclohexane, chloroform, and water. Free-energy landscapes (FELs) for each compound and solvent were obtained from principal component analysis at T = 300 K. From the FELs, we introduced a new flexibility index and an overlap between FELs.

Detailed analysis of ensembles at T = 300 K revealed that the average solvent-accessible surface area (SASA) in cyclohexane showed excellent correlation with cell permeability (R^2=0.872), whereas this correlation was weaker in chloroform. We also found that single stereoisomeric change affects the conformational patterns, resulting in permeability differences. A possible strategy for designing permeable cyclic peptides from FELs by McMD simulations is proposed.

[1] Naylor, M. R., et al., *J. Med. Chem.*, 61:11169-11182, 2018.
[2] Ono, S., et al., *J. Chem. Info. Model.*, 59:2952-2963, 2019.

Interpretable Reaction Prediction using Graph Convolutional Networks

Shoichi Ishida[1] ishida.shouichi.57a@st.kyoto-u.ac.jp
Kei Terayama[2,3,4] kei.terayama@riken.jp
Ryosuke Kojima[4] kojima.ryosuke.8e@kyoto-u.ac.jp
Kiyosei Takasu[1] takasu.kiyosei.6r@kyoto-u.ac.jp
Yasushi Okuno[3,4,5] okuno.yasushi.4c@kyoto-u.ac.jp

[1] Graduate School of Pharmaceutical Sciences, Kyoto University, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan
[2] RIKEN Center for Advanced Intelligence Project, Chuo-ku, Tokyo 103-0027, Japan
[3] Medical Sciences Innovation Hub Program, RIKEN Cluster for Science, Technology and Innovation Hub, Tsurumi-ku, Kanagawa 230-0045, Japan
[4] Graduate School of Medicine, Kyoto University, Shogoin-kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan
[5] Foundation for Biomedical Research and Innovation at Kobe, Center for Cluster Development and Coordination, Chuo-ku, Kobe, Hyogo 650-0047, Japan

Keywords: Reaction prediction, Graph convolutional networks, Integrated gradients, Visualization, Interpretability

Improving reaction prediction and its interpretability are essential issues for spreading data-driven retrosynthesis approaches to chemists. Various research groups have been addressing data-driven approaches for reaction prediction and retrosynthetic analysis, and the approaches have shown excellent performances[1-3]. Although the performances have progressed due to recent advances of machine learning and deep learning techniques, problems such as improving capability of reaction prediction and the black-box problem of neural networks persist for practical use by chemists. To spread data-driven approaches to chemists, we focused on two challenges: improvement of reaction prediction and interpretability of the prediction. Here, we propose an interpretable prediction framework using Graph Convolutional Networks (GCN)[4] for reaction prediction and Integrated Gradients (IGs)[5] for visualization of contributions to the prediction to address these challenges[6]. As a result, from the viewpoint of average accuracies, our model showed better performances than the existing model based on Extended-Connectivity Fingerprint (ECFP), which was used in Segler's Alphachem[2]. Furthermore, IGs based visualization of the GCN prediction successfully visualized reaction-related atoms.

[1] Segler, M. H. S., Waller, M. P., Neural-Symbolic Machine Learning for Retrosynthesis and Reaction Prediction, *Chemistry - A European Journal*, 23:5966-5971, 2017
[2] Segler, M. H. S., Preuss, M., and Waller, M. P., Planning chemical syntheses with deep neural networks and symbolic AI, *Nature*, 555:604-610, 2018.
[3] Coley, C. W., Green, W. H., and Jensen, K. F., Machine Learning in Computer-Aided Synthesis Planning, *Accounts of Chemical Research*, 51:1281-1289, 2018.
[4] Kipf, T. N., and Welling, M. S., Semi-Supervised Classification with Graph Convolutional Networks, *International Conference on Learning Representations (ICLR)*, 2017.
[5] Sundararajan, M., Taly, A., and Yan, Q., Axiomatic Attribution for Deep Networks, *International Conference on Machine Learning (ICML)*, 2017.
[6] Ishida, S., Terayama, K., Kojima, R., Takasu, K., and Okuno, Y., Prediction and Interpretable Visualization of Synthetic Reactions using Graph Convolutional Networks, *ChemRxiv*, 10.26434/chemrxiv.8343995, 2019.

P6-02

Applicable Machine Learning Method to Predict Site of Metabolism ~Comparison of Various Methods Using In-house Compounds ~

Katsunori Sasahara Sasahara.Katsunori@otsuka.jp
Tomoki Suzuki Suzuki.Tomoki@otsuka.jp
Hiroyuki Sasabe Sasabe.Hiroyuki@otsuka.jp
Masakazu Shibata Shibata.Masakazu@otsuka.jp
Takahiro Shimada Shimada.Takahiro@otsuka.jp
Kenji Takeuchi Takeuchi.Kenji@otsuka.jp
Ken Umehara Umehara.Ken@otsuka.jp

Nonclinical Research Center, Tokushima Research Institute, Otsuka Pharmaceutical Co., Ltd.
463-10, Kagasuno, Kawauchi-cho Tokushima 771-0192, Japan

Keywords: Machine Learning, ADME Prediction, Drug design, Artificial intelligence (AI)

Pharmaceutical companies have a large amount of high-quality data obtained by high throughput screening (HTS) under the same experimental conditions, which are expected to utilize effectively as part of a big data analysis. We previously focused on machine learning in many *in silico* methods including molecular dynamics, quantum mechanics[1] and have succeeded in building a highly accurate ADME prediction models for membrane permeability, protein binding, metabolic stability, CYP inhibition, and some transporter recognitions using HTS data[2].

The established ADME prediction models, which turn compound structures into each ADME profile, are important to understand ADME profile of candidates. Moreover, considering that these predictions are used for design of new compound, determination of structurally active sites (atoms or parts affected to study results) is more important. If structurally active sites are shown by *in silico* methods, we can change the corresponding structure before synthesis, which enables us to accelerate design process.

Machine learning methods are generally considered as black boxes which just only turn input into output and inclined to believe that only prediction results can be obtained. However some methods can keep track of the internal decision-making process. This means structurally active sites are understandable by analysis of the machine learning process.

In this study, we developed the some active atom visualization models which prioritized and highlighted metabolically active sites (atoms or parts), using a data set of human liver microsomal metabolic activity of approximately 30000 in-house compounds with various methods (Soft used: PipelinePilot, Chainer Chemistry, MOE etc). Proposed sites of metabolism on the structure were evaluated by a comparison with the actual site which are revealed in the microsomal metabolic experiments of some in-house compounds using LC-MS/MS. The active atom visualization models, which can show actual sites of metabolism, were successfully developed.

The procedure of active atom visualization models are considered to be applicable not only to the prediction of metabolic site but also to other studies related to efficacy, toxicology and other DMPK issues. The molecular design at a discovery stage which is required simultaneous optimization of various issues is expected to become more efficient by showing structurally active sites in many experiments.

[1] Sasahara K, Mashima A, Yoshida T, Chuman H (2015) Molecular dynamics and density functional studies on the metabolic selectivity of antipsychotic thioridazine by cytochrome P450 2D6: Connection with crystallographic and metabolic results. Bioorg Med Chem 23:5459-5465
[2] Sasahara K (2018) 薬物動態予測システムを用いた 0 次スクリーニング, 第 41 回薬物動態談話会年会, invited lecture など

P6-03

In Silico Modeling to Predict Drug-induced Phospholipidosis Based on Machine Learning Approach

Yui Migura[1] **Katsunori Sasahara**[2]
Migura.Yui@otsuka.jp

Chifumi Nagano[1] **Yasuhiro Iwata**[1] **Kaori Abe**[1]

[1] Department of Investigative Toxicology, Nonclinical Research Center, Tokushima Research Institute, Otsuka Pharmaceutical Co., Ltd., 463-10 Kagasuno, Kawauchi-cho Tokushima 771-0192, Japan
[2] Department of Drug Metabolism and Pharmacokinetics, Nonclinical Research Center, Tokushima Research Institute, Otsuka Pharmaceutical Co., Ltd., 463-10 Kagasuno, Kawauchi-cho Tokushima 771-0192, Japan

Keywords: Machine Learning, QSAR, Phospholipidosis, Toxicity, Drug Design

The onset of drug-induced phospholipidosis leads to a delay in the development of new drugs, thus early detection of it is desirable. We have previously developed an in vitro screening system to detect the potential of the onset of phospholipidosis.

In this study, we created an in silico predictive model of phospholipidosis by machine learning using the in vitro data and a public database. This model was applied to even earlier detection and de novo design of compounds.

We evaluated various predictive models using several data sets of descriptors for 717 in-house candidate compounds and 793 known compounds in ChEMBL by accuracy and F-value. 2D physicochemical descriptors were calculated by MOE, and structure descriptors such as ECFP were calculated by Discovery Studio.

As a result, better predictions were obtained when using 2D physicochemical descriptors of internal and external compounds as training-data set. And, compared with known models (Ploemen[1], Pelletier[2], Tomizawa[3] and DEREK), our in silico model obtained better results with accuracy of 78%, F-value of 78% and area under an ROC curve of 0.86.

In addition, we created a system that proposes structures avoiding phospholipidosis by evaluating new compounds generated by substituent change automatically.

[1] Plomen JP., Use of physicochemical calculation of pKa and CLogP to predict phospholipidosis-inducing potential: a case study with structurally related piperazines, *Exp. Toxicol. Pathol.*, 2004, 55, 347-55
[2] Pelletier D., Evaluation of a Published in Silico Model and Construction of a Novel Bayesian Model for Predicting Phospholipidosis Inducing Potential, *J. Chem. Inf. Model.*, 2007, 47, 1196-1205.
[3] Tomizawa, K., Physicochemical and cell-based approach for early screening of phospholipidosis-inducing potential, *J. Toxicol. Sci.*, 2006, 31, 315– 324

P6-04

In silico models for predicting hepatotoxicity and renal toxicity based on HESS database

Tatsuya Ochibe
C182711@ed.nagoya-cu.ac.jp

Kaori Ambe
ambek@phar.nagoya-cu.ac.jp

Masahiro Tohkin
tohkin@phar.nagoya-cu.ac.jp

Graduate School of Pharmaceutical Sciences, Nagoya City University, 3-1 Tanabe-dori, Mizuho-ku, Nagoya 467-8603, Japan

Keywords: Machine learning method, Random forest, Hepatotoxicity, Renal toxicity

It is desirable to develop alternative methods for animal experiments to reduce the number of experimental animals and improve efficiency of the risk assessment of chemical substances. *In silico* methods for predicting chemical toxicity are considered to be useful for alternative methods. A high-quality toxicity testing database and appropriate prediction methods are necessary for the *in silico* toxicity evaluation. Hazard Evaluation Support System Integrated Platform (HESS) [1] by Japanese National Institute of Technology and Evaluation, which is open toxicity knowledge database, contains information on repeated dose toxicity (RDT) of rat. In this study, we developed *in silico* models focused on the hepatotoxicity and renal toxicity which are often observed in RTD studies by using HESS and machine learning methods, and tried to examine the toxicity endpoints that could be predicted high performance.

At first, we investigated hepatic and renal toxicological findings reported in HESS. A wide variety of toxicological findings related to hepatotoxicity or renal toxicity were reported in HESS. For the predicting hepatotoxicity and renal toxicity, we developed 4 prediction models, respectively. In hepatotoxicity, we predicted (1) the increase of liver weight, (2) hepatic histopathological findings, (3) the combination of the increase of liver weight and hepatic histopathological findings and (4) alanine aminotransferase (ALT) and aspartate aminotransferase (AST) level rise. In renal toxicity, we predicted (1) the increase of kidney weight, (2) renal histopathological findings, (3) the combination of the increase of kidney weight and renal histopathological findings and (4) creatinine and blood urea nitrogen (BUN) level rise. We developed prediction models by the random forest using descriptors calculated by dragon 7 [2]. The performance of each model was evaluated in terms of matthews correlation coefficient (MCC) [3] using 5-fold cross-validation. All our prediction models achieved MCC of above 0.45. For predicting both hepatotoxicity and renal toxicity, the model showed the best prediction performance (MCC: hepatotoxicity 0.558, renal toxicity 0.631) when predicting the combination of the increase of organ weight and the histopathological findings. These results suggested that our *in silico* models are robust and promising method for hepatotoxicity and renal toxicity prediction.

[1] http://www.nite.go.jp/en/chem/qsar/hess-e.html
[2] https://www.affinity-science.com/dragon.html
[3] Brown JB., et al., Classifiers and their Metrics Quantified, Molecular Informatics, 37(1-2): 1700127, 2018

P6-05

Prediction of pharmacological activities from chemical structures with graph convolutional neural network

Miyuki Sakai[1,2] sakai.miyuki.28w@st.kyoto-u.ac.jp
Kazuki Nagayasu[1] nagayasu@pharm.kyoto-u.ac.jp
Shuji Kaneko[1] skaneko@pharm.kyoto-u.ac.jp

[1] Department of Molecular Pharmacology, Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshida-Shimoadachi-cho, Sakyo-ku, Kyoto, 606-8501, Japan
[2] Medical Database Ltd., 2-5-5 Sumitomoshibadaimon building, Shibadaimon, Minato-ku, Tokyo, 105-0012, Japan

Keywords: Graph convolutional neural network, ChEMBL

The first step of pharmacological activity is the binding of a chemical compound to its target protein, whose efficacy depends on the affinity of ligand to the target. Accordingly, enormous experimental efforts have successfully determined the affinity of molecule-target interactions. Even though precise docking simulation may help estimate the receptor affinity when the three-dimensional (3D) structure of the receptor is given, it is impracticable in the case of large transmembrane proteins or ion channel complexes, since most of the structures still remains to be uncovered, and many of them have allosteric binding sites. In this study, we collected the experimentally-confirmed affinity data (for example, IC50 values) and SMILES data of the ligands from ChEMBL25 database to build regression models by graph convolutional neural networks, and optimized the prediction performance by tuning hyperparameters. A wide range of 18 targets are selected from kinases with known 3D structures to transporters and ion channels with unknown or limited structures. Overall, better prediction performance of the model was obtained in parallel with the increase of the number of raw data. In a case of dopamine transporter with a large number of 5908 records, the coefficient of determination was 0.74. Even in the case of TRPA1 channel with a small number of 621 records, a good value of 0.72 was obtained. We will present more examples and discuss the significance of this approach in the presentation.

Design of Selective GPCR Antagonists with a desired Pharmacophore using Deep Reinforcement Learning

Chisato Kanai[1]
kanai@intage.com

Enzo Kawasaki[1]
kawasaki@intage.com

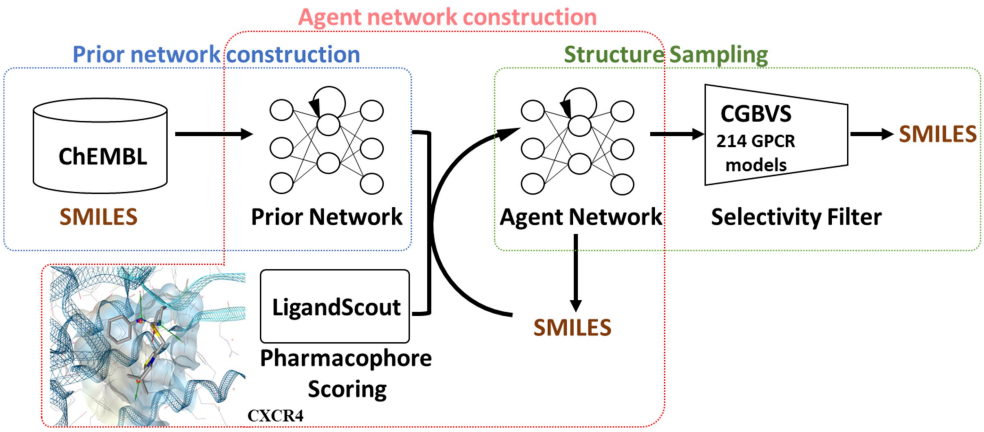
Atsushi Yoshimori[2]
yoshimori@itmol.com

Tomohiko Tasaka[3]
tasaka@affinity-science.com

[1] INTAGE Healthcare, Inc.; 79 Kankoboko-cho, Shimogyo-ku, Kyoto, 600-8009, Japan
[2] Institute for Theoretical Medicine, Inc., 26-1 Muraoka-Higashi 2-chome, Fujisawa, Kanagawa 251-0012, Japan.
[3] Affinity Science Corporation, Aios Gotanda-ekimae, 1-11-1 Nishi-Gotanda, Shinagawa, Tokyo 141-0031, Japan.

Keywords: *De novo* design, Deep reinforcement learning, Pharmacophore model, Chemical genomics-based virtual screening, Selective GPCR antagonist

The objective of drug design is to find chemical structures that fit a pocket of a target protein in both shape and pharmacophore features. In recent years, the use of deep generative models using reinforcement learning is getting a lot of attention as an effective method of *de novo* design [1]. We have previously developed a strategy to generate selective kinase inhibitors with a desired pharmacophore using deep reinforcement learning and chemical genomics-based virtual screening (CGBVS) method [2].



G protein-coupled receptors (GPCRs) are the most intensively focused class of target proteins, because GPCRs are some of the most successful therapeutic targets for serious diseases.

Our strategy can be adopted into general use for generating selective GPCR antagonists or agonists with a desired pharmacophore. In this work, we have employed this strategy to generate molecular structures of CXCR4 antagonists and evaluated their selectivity by predicting interactions with 214 GPCRs using CGBVS. A detailed analysis and application will be presented in the meeting.

[1] Olivecrona, M., Blaschke, T., Engkvist, O., Chen, H., J. Molecular de-novo design through deep reinforcement learning, *Journal of Cheminformatics*, 9:1, 2017.
[2] Yoshimori, A., Kawasaki, E., Kanai, C., Tasaka, T., Strategies for design of molecular structures with a desired pharmacophore using deep reinforcement learning, *submitted for publication.*

P6-07

High-performance predication model utilizing a novel deep learning-based QSAR analysis using Deep Snap and the Tox 21 10k library

Yasunari Matsuzaka **Yoshihiro Uesawa**

matsuzys@my-pharm.ac.jp uesawa@my-pharm.ac.jp

Department of Medical Molecular Informatics, Meiji Pharmaceutical University, Tokyo,

Keywords: Deep Snap, Tox21, Deep learning, QSAR, Tox21 10K library

The Toxicology in the 21st Century (Tox21) project has profiled a library of more than 10,000 chemical compounds including commercial chemicals, pesticides, food additives, and drugs, referred to as the Tox21 10K library [1]. To this data, most of the toxicological assessment methods for novel chemical compounds displaying agonist or antagonist activities rely on animal models. Such approaches can, however, be expensive, time-consuming, and most importantly difficult to translate for humans. Hence, alternative approaches to traditional animal models must be investigated [2]. On the other hand, *in silico* approaches require the feature selection to avoid the combinatorial optimization problems. Furthermore, finding the suitable combination and feature learning to reduce overfitting by numerous molecular descriptors requires complex technical skills and is time-consuming [3]. Recently, a novel technique for capturing molecular features from molecular structure input data has been utilized [4]. In this study, three-dimensionally (3D) optimized molecular structures were rotated at an arbitrary angle on the x-, y-, and z-axes and photographed. In the obtained images, the arrangement of each atomic symbol was depicted as a ball-and-stick model, with different colors denoting different atoms. Automatically inputting as much structural information as possible into the deep learning (DL) models was referred to as Deep Snap [4]. Previously, the performance of the DL prediction model using the image input data produced by Deep Snap showed that the receiver operatic characteristic (ROC-AUC) value in the Deep Snap-DL approach outperformed that of the previously reported state-of –the -art methods using random forest (RF) and DL with 3D or extended-connectivity fingerprint descriptors [4]. Furthermore, optimization of some parameters in the Deep Snap technique has revealed the existence of optimal thresholds to attain the best performance of the prediction models [5]. In the current study, our aim was to assess the input datasets and parameters in the Deep Snap-DL method and investigate how they affect the prediction performance. Excellent performance of the prediction model was observed and calculation cost was reduced using the Deep Snap approach. The Deep Snap software is developed by grant in Long-Range Research Initiative (LRI), Japan Chemical Industry Association (JCIA) and AI-Substances Hazardous Integrated Prediction System (AI-SHIPS) project.

[1] Sipes, N.S. et al., An Intuitive Approach for Predicting Potential Human Health Risk with the Tox21 10k Library, Environ Sci Technol, 51(18):10786 -10796, 2017.

[2] Freires, I.A., Sardi, J.C., de Castro, R.D., Rosalen, P.L. Alternative Animal and Non-Animal Models for Drug Discovery and Development: Bonus or Burden ?, Pharm Res, 34(4), 681 -686, 2017.

[3] Ponzoni, I., et al., Hybridizing Feature Selection and Feature Learning Approaches in QSAR Modeling for Drug Discovery, Sci Rep, 7(1):2403, 2017.

[4] Uesawa, Y. Quantitative structure-activity relationship analysis using deep learning based on a novel molecular image input technique, Bioorg Med Chem Lett, 28(20), 3400 -3403, 2018.

[5] Matsuzaka, Y., Uesawa, Y. Optimization of a Deep-Learning Method Based on the Classification of Images Generated by Parameterized Deep Snap a Novel Molecular-Image-Input Technique for Quantitative Structure-Activity Relationship (QSAR) Analysis, Front Bioeng Biotechnol, 7:65, 2019.

P6-08

Development of AI-aided hit compound finding/profiling system for imaging-based high content screening

Hiroki Terauchi[1] h-terauchi@hhc.eisai.co.jp
Makiyo Iida[1] m3-iida@hhc.eisai.co.jp
Naoko Tsukahara[1] n-tsukahara@hhc.eisai.co.jp
Kaoru Mitsuhashi[1] k2-mitsuhashi@hhc.eisai.co.jp
Hajime Shimizu[1] h4-shimizu@hhc.eisai.co.jp
Sachie Taniguchi[1] s3-taniguchi@hhc.eisai.co.jp
Shuichi Suzuki[1] s3-suzuki@hhc.eisai.co.jp
Katsuyuki Fukuda[1] k4-fukuda@hhc.eisai.co.jp

[1]Tsukuba Research Laboratories, Eisai Co., Ltd., Tokodai 5-1-3, Tsukuba-shi, Ibaraki 300-2635, Japan

Keywords: Artificial Intelligence (AI), deep neural network, cell imaging, hit profiling, Deep Metric Learning

Background:

Recently Artificial Intelligence (AI) especially deep neural network is attracting attention in drug discovery.

We have been using high content cell imaging in early drug discovery. However, the image analysis has high computer complexity, and requires a lot of time and effort to identify true hit samples [1]. To address this issue, we tried to develop imaging analyzing system with AI.

Results:

We established an AI-aided imaging hit finding system using deep neural network. AI learned the difference between negative and positive control images. AI identified compounds with the similar phenotype as the positive controls with high accuracy. Furthermore, by applying the reference compound library as a training set, we were able to predict the mode of action (MOA) of various compounds.

In addition, we recently found a brand-new learning method, Deep Metric Learning [2], also performed possible similar profiling without reference compound library. We would like to discuss the usefulness of our methodology.

[1] Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O. & Golland, P. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome biology, 7(10), R100.

[2] Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). Arcface: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4690-4699).

Computational drug target prediction using PU learning approach.

Kazuto Nakata[1]
k-nakata@ay.jp.nec.com

Hiroaki Iwata[2]
iwata.hiroaki.3r@kyoto-u.ac.jp

Takeshi Fujiwara[3]
takeshi.fujiwara@nibiohn.go.jp

Mari Ito[3]
mari@nibiohn.go.jp

[1] 1st Government and Public Solutions Division, NEC Corporation, 5-7-1, Shiba, Minato-Ku, Tokyo, 108-8001, Japan
[2] Graduate School of Medicine, Kyoto University, 53 Kawahara-cho, Syogoin, Sakyo-ku, Kyoto 606-8507, Japan
[3] Bioinformatics project, National Institutes of Biomedical Innovation, Health and Nutrition 7-6-8 Saito-Asagi, Ibaraki City, Osaka 567-0085, Japan

Keywords: Machine Learning, PU Learning, Drug Target Prediction

Drug target discovery is one of the most important processes in drug development. Computational drug target prediction reduces times and costs to evaluate drug target. Recently, some drug target prediction models were proposed by using machine learning approaches [1][2]. However, these proposed models use several data sources which include few positive samples. To solve this problem, we employed positive-unlabeled learning (PU learning) approach [3] using publicly available data sources. By compared with known models, PU learning model achieved comparable results. These results indicated that our model using publicly available data sources is an effective tool for exhibiting proof of concept to predict drug targets with few positive data sources. This research was carried out as one of the activities of Life Intelligence Consortium (LINC).

[1] Ferrero E, Dunham I, Sanseau P., In silico prediction of novel therapeutic targets using gene-disease association data. *J Transl Med. 2017 Aug 29;15(1):182.*
[2] Yao J, Hurle MR, Nelson MR, Agarwal P., Predicting clinically promising therapeutic hypotheses using tensor factorization. *BMC Bioinformatics. 2019 Feb 8;20(1):69.*
[3] Elkan C, Noto K., Learning Classifiers from Only Positive and Unlabeled Data, *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining Pages 213-220.*

Meta-modeling for Optimization in QSAR Modeling Processes and Application to Estrogen Receptor Agonist Activity Prediction

Kota Kurosaki[1]
d196955@std.my-pharm.ac.jp

Yosihiro Uesawa[1]
ueswa@my-pharm.ac.jp

[1] Department of Medical Molecular Informatics, Meiji Pharmaceutical University, Tokyo

Keywords: Machine learning, Virtual screening, Tox21

Virtual screening is a powerful tool for the rapid and comprehensive screening of compound properties in safety assessment and drug discovery. Although many empirical studies have compared the various algorithms involved in model construction, there are still no definite conclusions about the cause of a specific algorithm's ability or inability to work properly in a specific domain. Therefore, it is difficult to artificially optimize selection decisions in the modeling process from specific domains for the changing environment of each project.

We aimed to build a model that could use machine learning for selecting the best option in each step of the modeling process. First, the optimization, which is known as the Bayesian optimization (BO) introducing the Gaussian process, of the hyperparameters of the decision-tree-based models (Random Forest (RF), Light GBM (LGB), and XG Boost (XGB)), was performed using machine learning. This optimization method is a technique that is able to explore an optimal solution more efficiently [1]. Second, under the assumption that the entire QSAR modeling process itself was a function (meta-model), we extended the BO from determining the dataset fraction to setting the classifier hyperparameters in a simple modeling process. These models were implemented using Python for the prediction of ERlbd signaling pathway active compounds stored in the Tox 21 10 K compound library.

For the first approach, BO was superior to the conventional optimization method in terms of both time efficiency and trial efficiency. Also, the classifier built by BO yielded better ROC-AUC in the test set as compared to the non-optimized model's one. In addition, RF and LGB were able to automatically construct a high-performance model comparable to the Tox 21 data challenge world ranking. However, for the second approach, neither further improvement in ROC-AUC was observed nor was the dataset split rate's convergence to a certain value confirmed.

Previous studies have shown that human optimization can build high-performance prediction models for ERlbd active compounds [2,3] but is very time- and labor-intensive. On the other hand, the optimization with BO enabled us to construct a highly efficient and high-performance toxicity prediction model. Also, no performance improvement, including the data set partitioning process, was confirmed by meta-modeling, suggesting that hyperparameter regulation was the most influential process on the model's performance for the prediction model developed at this time. Thus, we expect that constructing the meta-model would allow us to comprehensively and efficiently determine the process or compound representation that would be influential (or insignificant) and appropriate for the final evaluation.

[1] J. Snoek, H. Larochelle, and R.P. Adams, IN *NIPS*, 2012
[2] Y. Uesawa, *Front. Environ. Sci.*, 4, 1–6, 2016
[3] Y. Aasako, Y. Uesawa, *Molecules*, 22, 1–10, 2017

P6-11

Predicting drug-induced transcriptome responses of a wide range of human cell lines by a novel tensor-train decomposition algorithm

Michio Iwata[1] iwata121@bio.kyutech.ac.jp
Longhao Yuan[2,3] longhao.yuan@riken.jp
Qibin Zhao[3,4] qibin.zhao@riken.jp
Yasuo Tabei[3] yasuo.tabei@riken.jp
Francois Berenger[1] beren314@bio.kyutech.ac.jp
Ryusuke Sawada[1] sawad330@bio.kyutech.ac.jp
Sayaka Akiyoshi[5] akiyoshi-s@bioreg.kyushu-u.ac.jp
Momoko Hamano[1] haman929@bio.kyutech.ac.jp
Yoshihiro Yamanishi[1] yamani@bio.kyutech.ac.jp

[1] Department of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
[2] Graduate School of Engineering, Saitama Institute of Technology, 1690 Fusaiji, Fukaya, Saitama 369-0293, Japan
[3] RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan
[4] School of Automation, Guangdong University of Technology, Panyu, Guangzhou, Guangdong, China
[5] Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, Fukuoka 812-8582, Japan

Keywords: imputation, tensor decomposition, gene expression, human cell line, drug discovery

Genome-wide identification of the transcriptomic responses of human cell lines to drug treatments is a challenging issue in medical and pharmaceutical research. However, drug-induced gene expression profiles are largely unknown and unobserved for all combinations of drugs and human cell lines, which is a serious obstacle in practical applications. In this study [1], we developed a novel computational method to predict unknown parts of drug-induced gene expression profiles for various human cell lines and predict new drug therapeutic indications for a wide range of diseases. We proposed a tensor-train weighted optimization (TT-WOPT) algorithm to predict the potential values for unknown parts in tensor-structured gene expression data. Our results revealed that the proposed TT-WOPT algorithm can accurately reconstruct drug-induced gene expression data for a range of human cell lines in the Library of Integrated Network-based Cellular Signatures. The results also revealed that in comparison with the use of original gene expression profiles, the use of imputed gene expression profiles improved the accuracy of drug repositioning. We also performed a comprehensive prediction of drug indications for diseases with gene expression profiles, which suggested many potential drug indications that were not predicted by previous approaches.

[1] Iwata *et al.*, Predicting drug-induced transcriptome responses of a wide range of human cell lines by a novel tensor-train decomposition algorithm, *Bioinformatics*, **35** i191–i199, 2019.

3D-RISM-AI: A machine learning approach to predict protein-ligand binding affinity using 3D-RISM

Kazu Osaki[1]
W195405a@yokohama-cu.ac.jp

Tsutomu Yamane[1]
tyamane@yokohama-cu.ac.jp

Toru Ekimoto[1]
ekimoto@yokohama-cu.ac.jp

Mitsunori Ikeguchi[1, 2]
ike@yokohama-cu.ac.jp

[1] Graduate School of Medical Life Science, Yokohama City University,
1-7-29 Suehirocho, Tsurumi-ku, Yokohama City 230-0045, Japan
[2] RIKEN Medical Science Innovation Hub Program,
1-7-29 Suehirocho, Tsurumi-ku, Yokohama City 230-0045, Japan

Keywords: Machine learning, 3D-RISM, Binding affinity

Accurately predicting protein-ligand binding affinities is important in drug discovery, because the protein-ligand affinity is a primary criteria to evaluate stabilities for abundant protein-ligand complexes in virtual screening and/or docking simulation. The protein-ligand affinity is given by the (relative) binding free energy, and many computational methods have been developed using empirical scores and molecular dynamics simulations etc. Recently, the prediction of the affinity has been evaluated by machine learning in which intermolecular interactions between protein and ligand, structural features, and amino acid sequences were taken into account as descriptors [1].

In the binding free energy calculation, interactions of protein/ligand with solvent molecules are considered as an important factor to improve accuracy of the prediction. Therefore, we propose a machine learning approach taking into account the interactions with solvent molecules (3D-RISM-AI). The interactions with solvent molecules were evaluated by the hydration free energy (HFE), and the HFE was calculated by the three-dimensional reference interaction model (3D-RISM) [2, 3]. In 3D-RISM, a thermally averaged distribution of solvent molecules around the solute is theoretically obtained, and the HFE can be calculated through an integral equation using the distribution function. We calculated ~3800 of the HFE for protein-ligand complexes in the PDBbind database (v.2018) and used the data in the learning process. Whereas the binding free energies solely evaluated using 3D-RISM are not correlated with the experimental data, the binding free energies predicted using 3D-RISM-AI show a good correlation with the experimental values.

[1] Ain, Q. U. et al. Wiley Interdiscip. Rev.: Comput. Mol. Sci. 2015, 5, 405−424
[2] D. Beglov and B. Roux, J. Chem. Phys. 103, 360, 1995.
[3] D. Beglov and B. Roux, J. Chem. Phys. 104, 8678, 1996.

P6-13

Deep Learning-aided Label-free, Real-time and Time-lapse Cell Visualization System that Enables Live/Dead Cell Discrimination and Counting

Tamio Mizukami[1, 2], **Mie Tsuchida**[1, 2], **Miho Hosoi**[1, 2], **Makoto Hasegawa**[1], **Ryuzo Sasaki**[1, 2], **Shigeru Nemoto**[3], **Katsumi Kishimoto**[2]

mizukami@nagahama-i-bio.ac.jp

[1] Nagahama Institute of Bio-Science and Technology, 1266 Tamura, Nagahama, Shiga, 526-0829, Japan
[2] Frontier Pharma Inc., 1281-8 Tamura, Nagahama, Shiga, 526-0829, Japan
[3] Broad Band Tower Inc., 2-1-6 Uchisaiwaicho, Chiyoda-ku, Tokyo 100-0011, Japan

Keywords: Deep learning, Imaging, Cytotoxicity

At present, as a method for analysis of cell characteristics such as life and death and differentiation state, a fluorescence observation method in which a specific molecule of a cell is labeled with an antibody or a fluorescence reagent to observe fluorescence is generally used. However, in this method, cells are invaded by fluorescence labeling and excitation light irradiation, and continuous observation as live cells can not be performed. Therefore, we constructed a model in which the relationship between the fluorescently labeled teacher cell image and the corresponding bright field cell image was finely learned by deep learning. By utilizing the model, we developed a technology that enables prediction of cell characteristics from a normal microscope bright field cell image and counting of cells and nuclei.

The technology is based on learning models of pseudo-fluorescent image generation and cell nuclear localization as follows.

(1) First step (visualization of cells): A model is constructed in which the relationship between a fluorescence microscope image of a cell treated with a fluorescent reagent for staining live and dead cells and a corresponding bright field image is learned by deep learning. The input of the unknown bright field cell image to the constructed model makes it possible to generate a pseudo-fluorescent cell image showing the characteristics of the cell without experimental work such as fluorescence labeling.

(2) Second step (cell counting): The position of the cell nucleus detected on the generated pseudo-fluorescent cell image is determined. For that purpose, a second model (position determination model) is constructed in which a teacher image manually hit at the center position of the cell nucleus on the fluorescence cell image is created and the relationship with the corresponding fluorescence cell image is learned by deep learning. The input of the pseudo-fluorescent cell image to the present position determination model can generate an image in which the center position of the cell nucleus on the fluorescent cell image is hit. A separately created program can automatically count the number of detected representative points and count the number of cells.

In the validation test of this technology, HepG2, a human hepatoma cell line widely used in hepatotoxicity evaluation and MG132, a proteasome inhibitor, were used. As a result, clear discrimination and counting of live and dead cells were possible.

This technology is capable of noninvasive, real time and time-lapse observation, and we hope that it will be widely used in the fields of life sciences, medicine and drug discovery in the future.

Deep-learning of cancer stem cell morphology

Tomoyasu Sugiyama[1]
tsugiyama@stf.teu.ac.jp

Saori Aida[2]
aidasor@stf.teu.ac.jp

Tomonari Kasai [1]
kasaitnr@stf.teu.ac.jp

Atsushi Sato [1]
atsato@stf.teu.ac.jp

Hiroyuki Kameda[2]
kameda@stf.teu.ac.jp

[1] School of Bioscience and Biotechnology, Tokyo University of Technology, 1404-1 Katakura, Hachioji, Tokyo 192-0982, Japan
[2] School of Computer Science, Tokyo University of Technology, 1404-1 Katakura, Hachioji, Tokyo 192-0982, Japan

Keywords: Artificial intelligence, Cancer stem cell, miPS-LLCcm cells

Cancer stem cells (CSCs) are thought to have an important role in tumour development. In cancer therapies it has been concerned that tumour might possess a capability of regrowth after suppression of large amount of tumour tissue after the therapy. In fact, many clinical results have suggested the presence of cells which have not affected by those therapies. It is believed that CSCs in tumour tissue could serve cancer cells which lacked the stem cell characteristics.

An experimental model showed that CSCs developed malignant tumour tissues (PLoS ONE 2012, 7: e33544). It is conceivable that mouse iPS-LLCcm (miPS-LLCcm) cell, a model of CSC, may differentiate to cancer cell while the cells may maintain self-renewal in mouse. It is useful to know whether a drag affects the characteristics of stem cell such as marker gene expression and cell morphology. Accordingly, in vitro culture of miPS-LLCcm cells makes many kinds of cell-based assays available.

In recent years, image recognition technologies are making remarkable advances by using artificial intelligence (AI). An image-to-image translation of any photographs can be processed by a single algorithm without any specific settings to each photograph (arXiv 2016, 1611.07004v1). Those examples of image translation were significant in accuracy and creativity. It is interesting whether the code of conditional Generative Adversarial Networks (CGAN) can recognize cell morphology. Accordingly, we showed the application of CGAN for the identification of CSCs from phase contrast miPS-LLCcm cell image at the last CBI annual meeting 2018.

Here, we will report the progress of the research.

Methods: miPS-LLCcm cell harboring *Nanog-GFP* reporter gene was used to monitor the pluripotency of CSC. To evaluate cell morphology, miPS-LLCcm cells were cultured in 96-well plate. CSC images of phase contrast and fluorescence pair were obtained using fluorescence microscopy BZ-X810. Each image was processed to make small size images. Then pairs of images were subjected to deep-leaning using CGAN. CGAN was trained to learn translation from phase contrast image to fluorescence image.

Results and Discussion: In the previous study, we obtained cell images using manually-operated conventional fluorescence microscopy equipped with a color CCD camera. Almost 10,000 pairs were used for the deep-learning, however, the trained-AI was hard to output appropriate fluorescence image from phase contrast image. In contrast, current system enables us to obtain CSC images under same condition even if many images are required to take. In addition, monochrome CCD has potential to take more detailed image. We will discuss the output of AI trained under the system which is supposed to use in drug screening.

P6-15

Study on adverse outcome pathways related to drug-induced rhabdomyolysis using machine learning

Shunichi Sasaki y161150@std.my-pharm.ac.jp

Yoshihiro Uesawa uesawa@my-pharm.ac.jp

Department of Medical Molecular Informatics, Meiji Pharmaceutical University, Tokyo, Japan

Keywords: JADER, Tox21, machine learning, QSAR, Rhabdomyolysis

Introduction: Drug-induced rhabdomyolysis (DIR) caused by various medicines including dyslipidemia drugs is a serious adverse effect in pharmacotherapies. Destruction of the cell membrane due to reduction of cholesterol is considered as one of the factor for onset of DIR mechanism [1]. However, information on the adverse outcome pathway (AOP) related to DIR is limited. Therefore, in this study, the relationship between DIR and molecular initiating events (MIE), such as activation of nuclear receptors and stress response pathways in the AOP was investigated using the Tox21-MIE database and the Japanese Adverse Drug Event Report database (JADER) published by the Pharmaceuticals and Medical Devices Agency (PMDA). Furthermore, we constructed a DIR discrimination model based on quantitative structure-activity relationship (QSAR) analysis.

Methods: Initially, the odds ratio and *P*-value in Fisher's exact test for rhabdomyolysis were comprehensively calculated from JADER, and drugs inducing rhabdomyolysis were classified based on both statistics. After that, based on the Tox21-MIE database, QSAR discrimination models for 59 types of MIE related activities were constructed using LightGBM [2]. The QSAR model was used to predict the MIE related activities of medicines listed in JADER. In addition, the prediction results of the medicines were applied to the DIR discrimination model.

Results and Discussion: As a result, it was suggested that estrogen and thyroid-stimulating hormones may be related to DIR. A previous study [3] indicated that rhabdomyolysis occurred as a complication of hypothyroidism; however, the exact involvement of estrogen and thyroid-stimulating hormones to trigger rhabdomyolysis has not yet been investigated.

[1] Dirk Moßhammer, et al., Mechanisms and Assessment of Statin-Related Muscular Adverse Effects, Br J Clin Pharmacol. 78(3): 454–466, Sep 2014.
[2] Guolin Ke, et al., LightGBM: A Highly Efficient Gradient Boosting Decision Tree, NIPS, 2017.
[3] Nooshin Salehi, et al., Rhabdomyolysis in a Patient with Severe Hypothyroidism, Am J Case Rep. 18: 912–918, 2017.

A New Scoring Function for Protein Structure Assessment Based on the Hydration Structure Information

Yasuomi Kiyota [1]
kiyotay@pharm.kitasato-u.ac.jp

Mayuko Takeda-Shitaka [1]
shitakam@pharm.kitasato-u.ac.jp

[1] School of Pharmacy, Kitasato University, 5-9-1 Shirokane Minato-ku, Tokyo 108-8641, Japan

Keywords: Structure bioinformatics, machine learning, 3D-RISM theory

In the field of protein complex study, to understand the hydration structure is becoming an increasingly important issue. A famous international contest for the prediction of protein complex structure, the critical assessment of predicted interactions (CAPRI), also considered the importance of water prediction inside the protein interface (Round 34 in 2015, http://www.ebi.ac.uk/msd-srv/capri/round34/). This challenge was two complexes of pyocin, which is one of the bacteriocin, and the inhibited immunity protein. For solving this session, we focused on the position of conserved water molecules around the interface and reflected the information in the alignments of the amino acid sequence. We got good results in this session, however, we noticed that predicted models which had been much higher accuracy would not be evaluated “good”. We consider that this reason is caused by selection scores for protein complex and we aim to develop a new structure evaluated score for the predicted structure. In the process of development, we focused on the hydration structure again. Therefore, a new structure evaluated score is based on the protein hydration structure information. We calculated the 3D-RISM theory[1], which is strict statistical mechanics theory, and performed the machine learning on the solvation energy obtained from this theory.

At first, we analyzed the structure dataset SCOP2 (Structural classification of proteins)[2], which have a detailed and comprehensive description of the structural and evolutionary relationships, and selected 670 chains which could be carried out the 3D-RISM theory. After the 3D-RISM calculations, the obtained solvation energies are decomposed by each residue. These decomposed energies are accumulated with some structural indices for the amino-acid residues (e.g. accessible surface area, secondary structure, and surrounding residues). We built the dataset for the solvation energy and structural index, and performed the machine learning approach based on Deep Learning with Keras Python library and Theano backend. After the reproducing solvation energies by structure indices, we tried to evaluate the predicted models which are submitted in CASP12. As a results of this verification, we reached that developed score should be used with ProQ3[3], which is also structure evaluated score for protein, for highly precision and stability. The developed structure evaluated score is faster and more accurate than the 3D-RISM theory.

[1] Kovalenko, A., and Hirata, F., Self-consistent description of a metal–water interface by the Kohn–Sham density functional theory and the three-dimensional reference interaction site model, *J. Chem. Phys.*, 110:10095−10112, 1999.
[2] Andreeva, A., Howorth, D., Chothia, C., Kulesha, E. and Murzin, A. G., SCOP2 prototype: a new approach to protein structure mining, *Nucleic Acid Res.*, 42:D310-D314, 2014.
[3] Uziela, K., Shu, N.m, Wallner, B. and Elofsson, A., ProQ3: Improved model quality assessments using Rosetta energy terms. *Sci. Rep.*, 6:35509, 2016.

Development of *in silico* prediction model for skin sensitization using the alternative tests dataset

Masaharu Suzuki[1] c162518@ed.nagoya-cu.ac.jp
Kaori Ambe[1] ambek@phar.nagoya-cu.ac.jp
Masahiro Tohkin[1] tohkin@phar.nagoya-cu.ac.jp
Takashi Yamada[2] t-yamada@nihs.go.jp
Takao Ashikaga[2] takao.ashikaga@nihs.go.jp

[1] Graduate School of Pharmaceutical Sciences, Nagoya City University, 3-1 Tanabe-dori, Mizuho-ku, Nagoya 467-8603, Japan
[2] National Institute of Health Sciences, 25-26, Tonomachi, Kawasaki-ku, Kawasaki 210-9501, Japan

Keywords: Direct Peptide Reactivity Assay, human Cell Line Activation Test, KeratinoSens™, Local Lymph Node Assay, skin sensitization, Support Vector Regression, prediction model

Skin sensitization is a key endpoint for safety assessment of chemicals especially in cosmetics. Although the animal tests such as Local Lymph Node Assay (LLNA) had been widely used for the skin sensitization, the European Union (EU) has banned animal tests in safety testing of cosmetic ingredients since 2013. Therefore, the hazard identification method of skin sensitization is now widely used based on three *in vitro* tests listed in Organization for Economic Co-operation and Development (OECD) test guideline [1]: Direct Peptide Reactivity Assay (DPRA) [2], KeratinoSens™ [3], and human Cell Line Activation Test (h-CLAT) [4]. However, this hazard identification method is not quantitative. In this study, we tried to develop *in silico* quantitative prediction model of LLNA EC3 (Estimated Concentration needed to produce a stimulation index of 3) values which indicate skin sensitization potency [5] using data of those three *in vitro* tests. We used a set of 120 substances from the Cosmetics Europe database [6] to construct the *in silico* prediction model using Support Vector Regression. The data was randomly divided the training dataset for the model construction (90 substances) and the test dataset for the evaluation (30 substances) of constructed models. As a result of the evaluation of SVR model, the contribution rates (r^2values) and root mean squared error (RMSE) were 0.433 and 0.753, respectively, which suggested the satisfied performance prediction model. In conclusion, *in silico* approach using three *in vitro* tests dataset is useful for prediction LLNA EC3 values.

[1] M. Hirota, T. Ashikaga, H. Kouzuki., Development of an artificial neural network model for risk assessment of skin sensitization using human cell line activation test, direct peptide reactivity assay, KeratinoSens™ and in silico structure alert parameter, *J Appl Toxicol*, 38 (4): 1-13, 2017
[2] https://ntp.niehs.nih.gov/iccvam/suppdocs/feddocs/oecd/oecd-tg442c-508.pdf
[3] https://ntp.niehs.nih.gov/iccvam/suppdocs/feddocs/oecd/oecd-tg442d-508.pdf
[4] http://www.asas.or.jp/jsaae/eng/pdf/info_20170208.pdf
[5] https://ftp.cdc.gov/pub/Documents/OEL/06.%20Dotson/References/OECD_2002.pdf
[6] Nicole C. Kleinstreuer et al., Non-animal methods to predict skin sensitization (II): an assessment of defined approaches, Crit Rev Toxicol, 48 (5): 359-374, 2018

Focused Library Generative Model for GPCR Family

Gen Li[1]
li.g.af@m.titech.ac.jp

Nobuaki Yasuo[2]
yasuo@cbi.c.titech.ac.jp

Masakazu Sekijima[1,2]
sekijima@c.titech.ac.jp

[1] Department of Computer Science, Tokyo Institute of Technology, 4259-J3-23, Nagatsuta-cho, Midori-ku, Yokohama, Japan
[2] Advanced Computational Drug Discovery Unit, Tokyo Institute of Technology, 4259-J3-23, Nagatsuta-cho, Midori-ku, Yokohama, Japan

Keywords: drug discovery, GPCR family, selectivity, generative model

Developing novel drug for protein target in GPCR family is a difficult task, where the drug candidates are required for high off-target selectivity [1]. Some ligand based and structure based computational approaches have been reported. However, scanning chemical space is a time-consuming work and candidate molecules are lack of novelty.

In this work, we report a machine learning based model to generate novel drug-like molecules for protein targets in GPCR family with high selectivity. Model training consists of two steps. Doc2vec [2] model, which is used in natural language processing, is used for vectorizing protein sequence to fixed-length vector which is the input condition of generative model. Subsequently, protein sequence vectors and SMILES representations of corresponding Molecules are applied to conditional variational autoencoder [3] for model training. Novel molecules generated by this model represent high distributional similarity to training molecules which have high selectivity and shown that this model can potentially expand ligand library for any specific protein family.

[1] Alexander S. Hauser, Misty M. Attwood, Mathias Rask-Andersen, Helgi B. Schiöth & David E. Gloriam, Trends in GPCR drug discovery: new agents, targets and indications, *Nature Reviews Drug Discovery*, 16:829–842, 2017.
[2] Le, Q. V., and Mikolov, T., Distributed representations of sentences and documents, *ICML*, 14:1188–1196, 2014.
[3] K. Sohn, H. Lee, and X. Yan, Learning structured output representation using deep conditional generative models, *Advances in Neural Information Processing Systems*, 2015.

Prediction of G4MP2-level Molecular Properties from DFT-level Structures Using Deep Tensor Neural Network

Wan Mingda[1]
wan.m.aa@m.titech.ac.jp

Nobuaki Yasuo[2]
yasuo@cbi.c.titech.ac.jp

Masakazu Sekijima[1,2]
sekijima@c.titech.ac.jp

[1] Department of Computer Science, Tokyo Institute of Technology, J3-23-4259 Nagatsutacho, Midori-ku, Yokohama 226-8501, Japan
[2] Advanced Computational Drug Discovery Unit, Tokyo Institute of Technology, J3-23-4259 Nagatsutacho, Midori-ku, Yokohama 226-8501, Japan

Keywords: Property prediction, Deep learning, Computational chemistry, Machine learning

It is important to predict molecular properties with high throughput and accuracy for drug discovery and materials science. *ab initio* quantum mechanical methods, such as density functional theory (DFT) have widely been used for molecular property calculations [1]. When the required accuracy is as high as experimental error (<1kcal/mol), higher level calculations such as G4MP2 are used [2]. However, it is difficult to screen large number of molecules by high-level method, because the computational cost of these methods is much larger than that of DFT. Machine learning is a potential method to predict the molecular property accurately without using high-level calculations. Recently, the Δ-machine learning approach, which learns the difference between rough calculations and accurate calculations, is reported to reduce the calculation cost of property prediction [3]. However, the total prediction accuracy of previous works could not be as high as chemical experiments because the dataset was based on DFT, whose error is larger than experimental error. In this study, we 1) constructed a G4MP2-level dataset for molecular property prediction using TSUBAME3 supercomputer, and 2) performed Δ-machine learning-based prediction by using deep tensor neural network. Our database named QM9-G4MP2, which is the extension of QM9 database [4], contains 133,885 molecules with DFT-level input structure and G4MP2-level properties. The prediction results showed that deep tensor neural network successfully predicted G4MP2-level molecular properties from DFT-level structures.

[1] Hohenberg, P. and Walter K., Inhomogeneous electron gas. *Physical review*, 136.3B: B864, 1964.
[2] Curtiss, L. A., Paul C. R., Krishnan R., Gaussian-4 theory. *The Journal of chemical physics*, 126.8: 084108, 2007.
[3] Ramakrishnan, R., Dral, P. O., Rupp, M., Lilienfeld, O. A., Big data meets quantum chemistry approximations: The Δ-machine learning approach. *Journal of chemical theory and computation* 11.5: 2087-2096, 2015.
[4] Ramakrishnan, R., Dral, P. O., Rupp, M., Lilienfeld, O. A., Quantum chemistry structures and properties of 134 kilo molecules, *Scientific data* 1: 140022, 2014.

P7-01

Data set selection in deep learning based CYP3A4 binding mode prediction

Atsuko Sato[1] sato.a.al@m.titech.ac.jp
Naoki Tanimura[2] naoki.tanimura@mizuho-ir.co.jp
Teruki Honma[1,3,4] honma.teruki@riken.jp
Akihiko Konagaya[1] kona@c.titech.ac.jp

[1] School of Computing, Department of Computer Science, Tokyo Institute of Technology; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa, 226-8503, Japan
[2] Science Solutions Division, Mizuho Information & Research Institute, Inc.; 2-3, Kanda-Nishikicho, Chiyoda-ku, Tokyo, 101-8443, Japan
[3] Center for Biosystems Dynamics Research, RIKEN; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
[4] Medical Sciences Innovation Hub Program, RIKEN; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan.

Keywords: Computational drug design, Convolutional neural network, CYP3A4, binding mode prediction, data selection

For rational drug design, it is essential to predict the binding mode of protein–ligand complexes. Although various machine learning-based models have been reported that use convolutional neural networks to predict binding modes from three-dimensional structures, there are few detailed reports on how best to construct and use datasets. Here, we examined how different datasets affected the prediction of the binding mode of CYP3A4 by a three-dimensional neural network when the number of crystal structures for the target protein was limited. We used four different training datasets: one large, general dataset containing various protein complexes and three smaller, more specific datasets containing complexes with CYP3A4-like pockets, complexes with CYP3A4-binding ligands, and complexes with CYP protein family members. We then trained models with different combinations of datasets with or without subsequent fine-tuning and evaluated the binding mode prediction performance of each model. The ROC AUC model with respect to area under the receiver operating characteristic curve was obtained by training with a combination of the general protein and CYP family datasets. However, the ROC AUC – recall balanced model was obtained by training with this combination of datasets followed by fine-tuning with the CYP3A4-binding ligands dataset. Our results suggest that datasets that balance protein functionality and data size are important for optimizing binding mode prediction performance. In addition, datasets with large median binding pocket sizes may be important for the binding mode prediction specifically of CYP3A4.

[1] Sato, A.; Tanimura, N.; Honma, T. and Konagaya, A. A role of data selection in deep learning based CYP3A4 binding mode prediction. (submitted)

P7-02

An Approach to Investigate Disease Mechanisms by FAERS Data Mining

Yoshito Ohya[1] yoshito.ohya.yf@kyowakirin.com
Yoko Ono[1] yoko.ono.jg@kyowakirin.com
Katsuya Kobayashi[2] katsuya.kobayashi.xe@kyowakirin.com
Kensuke Kojima[1] kensuke.kojima.th@kyowakirin.com
Takashi Yoneya[1] takashi.Yoneya.sk@kyowakirin.com

[1] Biometrics Dept., R&D Division, Kyowa Kirin Co., Ltd., Otemachi Financial City Grand Cube, 1-9-2 Otemachi, Chiyoda-ku, Tokyo 100-0004, Japan
[2] Immunology & Allergy lab., R&D Division, Kyowa Kirin Co., Ltd., 3-6-6 Asahi-machi, Machida-shi, Tokyo 194-8533, Japan

Keywords: Real World Data, RWD, FAERS, Drug discovery, Reporting Odds Ratio

Recently, decline of success rate and cost escalation become notable in drug discovery process [1], and more efficient approach is required. In this decade, the significant progress of technology allows us to collect and analyze the clinical big data from real world not only for epidemiological study but for drug discovery and development [2]. FDA Adverse Event Reporting System (FAERS) is one of the largest real world data which consists of adverse drug events collected all over the world [3], and the number of researches which succeeded to show the possibility of drug repositioning has been increased [4].

Here, we explored the FAERS database to identify effective drugs to lower the occurrence of systemic lupus erythematosus (SLE). We defined the set of adverse events similar to the clinical findings of SLE by some PT codes of Medical Dictionary for Regulatory Activities (MedDRA) [5], followed by calculation of Reporting Odds Ratio (ROR) with its 95% confidence interval for signal detection. According to this ROR, we selected the drugs that could suppress the SLE-like adverse events and evaluated the effects of a confounding factor. After that, we listed common biological functions of these drugs and made assumptions about their molecular mechanisms.

We would like to discuss this scheme, e.g., selection of effective drugs, dealing with confounding factors, identification of the common factors of the selected drugs and the construction of hypotheses.

[1] https://www.mhlw.go.jp/content/10801000/000398096.pdf
[2] Alemayehu, D and Berger, M. L., Big Data: transforming drug development and health policy decision making, *Health Services and Outcomes Research Methodology*, Health Services and Outcomes Research Methodology, 16:92-102, 2016
[3]https://www.fda.gov/drugs/drug-approvals-and-databases/fda-adverse-event-reporting-system-faers
[4] Nagashima,T., Shirakawa, H., Nakagawa, T., Kaneko, S., Prevention of antipsychotic-induced hyperglycaemia by vitamin D: a data mining prediction followed by experimental exploration of the molecular mechanism, *Scientific Reports*, 6: 26375, 2016
[5] https://www.meddra.org/

P7-03

Predicting drug indications and therapeutic target modules based on disease similarity by interpretable machine learning models

Ryusuke Sawada[1]
sawad330@bio.kyutech.ac.jp

Yasuo Tabei[2]
yasuo.tabei@riken.jp

Yoshihiro Yamanishi[1]
yamani@bio.kyutech.ac.jp

[1] Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan.
[2] RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan.

Keywords: drug indication, drug repositioning, drug targets, disease similarity, protein modules, machine learning, interpretability

The identification of drug indications and understanding of the mode-of-action are important issues in drug discovery. Machine learning plays a key role in computational prediction, but predictive models of most machine learning methods are black-box and unexplainable. In this study, we propose a novel machine learning method to predict new indications of drugs from chemical-protein interactome data and disease similarity. Disease-specific predictive models are constructed based on their molecular features such as diagnostic markers, disease-related genes/pathways, and environmental factors. We also estimate informative sets of proteins and biological pathways that contribute to the drug indication prediction by exploring disease-protein modules using a module search algorithm. In the results, we show the superior performance of the proposed method over the other methods in terms of accuracy and interpretability. Herein, we comprehensively predicted drug-disease association networks for 1,124 drugs and 365 human diseases, and validated some of these predictions using independent resources. The proposed method is expected to facilitate drug discovery and repositioning.

A Method for Systematic Analog Searching Using the Mega SAR Matrix Database

Atsushi Yoshimori[1]
yoshimori@itmol.com

Yuichi Horita[2]
horita.yuichi@infogram.co.jp

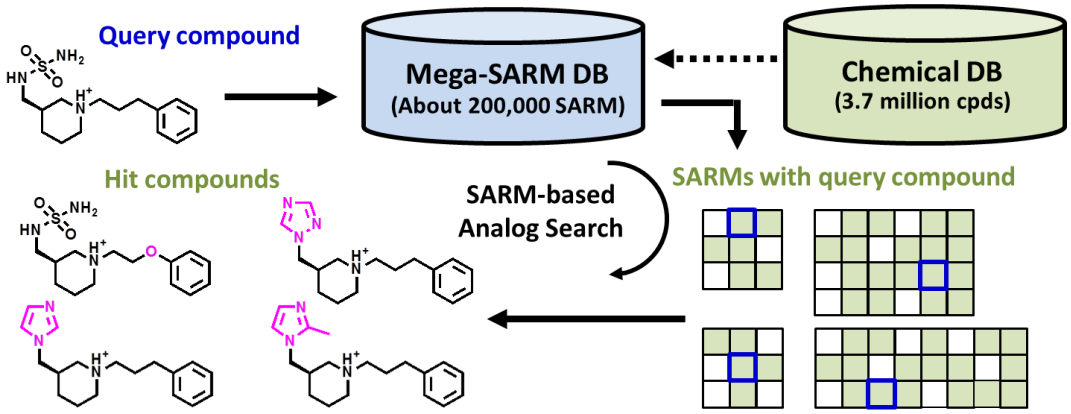
Toru Tanoue[2]
tanoue.toru@infogram.co.jp

Jürgen Bajorath[3]
bajorath@bit.uni-bonn.de

[1] Institute for Theoretical Medicine, Inc., 26-1 Muraoka-Higashi 2-chome, Fujisawa, Kanagawa 251-0012, Japan.
[2] Infogram, Inc., 2-17-19 Hakataekimae, Hakata-ku, Fukuoka 812-0011, Japan
[3] Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universität, Endenicher Allee 19c, D-53115 Bonn, Germany

Keywords: Structure-activity relationships, Matched molecular pair, Matching molecular series, SAR Matrix (SARM), SARM-based analog search

Analog searching is an important part of hit expansion, hit-to-lead, and lead optimization efforts during the early stage of drug development. A new computational methodology is introduced to search for real and virtual structural analogs of active compounds in large chemical databases [1]. The approach is based on the SAR Matrix (SARM) data structure that was originally designed to elucidate SAR information associated with groups of structurally related active compounds [2].



The SARM-based analog search method extends the capacity of current substructure-based methods by (i) considering real and virtual analogs of query compounds , (ii) permitting not only R-group replacements but also well-defined chemical modifications in core structures, (iii) automatically extracting all possible analogs form a large chemical database, and (iv) using an index tables to perform a computationally efficient structure search.

In this study, we introduce several examples of the SARM-based analog search to get useful knowledge from analog compounds linked to ChEMBL database.

[1] Yoshimori, A., Horita, Y., Tanoue, T., Bajorath, J. Method for Systematic Analog Search Using the Mega SAR Matrix Database, *submitted for publication*.
[2] Wassermann, A.M., Haebel, P., Weskamp, N., and Bajorath, J., SAR matrices: automated extraction of information-rich SAR tables from large compound data sets, *Journal of Chemical Information and Modeling*, 52:1769-1776, 2012.

P7-05

The Multiple Representation of Protein Sequence Motifs Using Sequence Binary Decision Diagrams

Hiroaki Kato [1]
kato @ hiroshima-cmt.ac.jp

Kohei Yamato [2]
k143377 @ edu.tut.ac.jp

[1] Department of Distribution and Information Engineering, National Institute of Technology, Hiroshima College, 4272-1 Higashino, OsakiKamijima-cho, Hiroshima 725-0231 Japan
[2] Department of Computer Science and Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi, Aichi 441-8580 Japan

Keywords: sequence binary decision diagram, Aho-Corasick algorithm, sequence motif, matrix metalloproteases, olfactory receptors

A protein is a biological macromolecule consisting of one or more chains of amino acid residues. It is well known that there are characteristic patterns called sequence motifs among proteins which have a common function. Since sequence motifs closely related to its structure and function, they have been applied to the functional analysis of proteins. Usually, it has employed sequence motif representations (motif models) such as regular expressions and weight matrices. In the present work, the authors have proposed a novel representation for a set of sequence motifs using SeqBDDs (sequence binary decision diagrams).

A SeqBDD is a directed acyclic graph, which represents a set of character strings, and is capable of using various set operations [1]. A SeqBDD contain comprehensive patterns associated with a particular function, and also several types of sequence motifs, i.e., it can represent a sequence motif group. We have also developed a string search algorithm and a homology search algorithm using SeqBDDs. The information of a sequence motif group is compressed and stored in a SeqBDD. The string search algorithm is based on the Aho-Corasick algorithm [2], and the homology search algorism based on local and global sequence alignments using dynamic programming.

At first, the sequence motifs on matrix metalloproteases (MMPs) were searched from all amino acid sequences in UniProtKB/Swiss-Prot using SeqBDDs and PROSITE patterns. The results were compared with their precision, recall and f-measure. As a result, the proposed motif model can properly search the sequences that were in false-negative when using regular expressions. Then, we verified the time efficiency of the homology search algorithm using artificial data. The homology search algorithm is faster than existing method. Finally, similar sequences were obtained with the homology search on sequence motif group on human olfactory receptors (ORs), and used to update the motif model on human ORs.

References:

[1] E. Loekito, J. Bailey and J. Pei, A binary decision diagram based approach for mining frequent subsequences, *Knowledge and Information Systems*, **24**, 235-268 (2010).
[2] A. V. Aho and M. J. Corasick, Efficient string matching: an aid to bibliographic search, *Communications of the ACM*, **18**, 333-340 (1975).

P7-06

Development of preventive drugs against oxaliplatin-induced peripheral neuropathy using large-scale medical database

Takahiro Niimura[1]
c201756030@tokushima-u.ac.jp

Yoshito Zamami[1, 2]
zamami@tokushima-u.ac.jp

Yutaro Naitoh[1]
c401831018@tokushima-u.ac.jp

Takehiro Kawashiri[3]
tkawa@med.kyushu-u.ac.jp

Mitsuhiro Goda[2]
mgoda@tokushima-u.ac.jp

Kenshi Takechi[4]
takechi.kenshi@tokushima-u.ac.jp

Masayuki Chuma[4]
chuma.masayuki@tokushima-u.ac.jp

Yuya Horinouchi[5]
horinouchi@tokushima-u.ac.jp

Yuki Izawa-Ishizawa[6]
ishizawa.yuki@tokushima-u.ac.jp

Yasumasa Ikeda[5]
yasuike@tokushima-u.ac.jp

Keisuke Ishizawa[1, 2]
ishizawa@tokushima-u.ac.jp

[1] Department of Clinical Pharmacology and Therapeutics, Tokushima University Graduate School of Biomedical Sciences, 2-50-1 Kuramoto-cho, Tokushima 770-8503, Japan.
[2] Department of Pharmacy, Tokushima University Hospital, 2-50-1 Kuramoto, Tokushima, 770-8503, Japan.
[3] Department of Clinical Pharmacy and Pharmaceutical Care, Graduate School of Pharmaceutical Sciences, Kyushu University, 3–1–1 Maidashi, Higashi-ku, Fukuoka 812–8582, Japan.
[4] Clinical Trial Center for Developmental Therapeutics, Tokushima University Hospital, 2-50-1 Kuramoto-cho, Tokushima, 770-8503, Japan.
[5] Department of Pharmacology, Institute of Biomedical Sciences, Graduate School, Tokushima University, Tokushima, 770-8503, Japan.
[6] AWA Support Center, Tokushima University, 3-18-15 Kuramoto-cho, Tokushima, 770-8503, Japan.

Keywords: oxaliplatin, peripheral neuropathy, adverse event, medical database

Oxaliplatin, a platinum-based anticancer drug, causes oxaliplatin-induced peripheral neuropathy (OIPN). Because OIPN not only reduces patients' quality of life (QOL) but also leads to dose reduction of oxaliplatin, the development of preventive agents against OIPN is desired. In this study, we searched for preventive drugs against OIPN using large-scale medical database, and experimentally validated the findings with neuron-like cells and OIPN rat models.

First, we searched for approved drugs that cancel the gene expression change caused by oxaliplatin, using the drug discovery tool LINCS, and found 23 approved drugs that met the requirements. Candidate drugs were then evaluated for their mitigating effects on OIPN, using the FDA Adverse Event Reporting System (FAERS) database. Based on the FAERS analysis, Drug X was found to significantly reduce the risk of patients developing OIPN. Using PC12 cells, we observed significant improvement in oxaliplatin-associated axonal damage with Drug X treatment. In addition, in vivo experiments showed that Drug X significantly reduced the expression of oxaliplatin-induced neuropathy in OIPN rat models.

Present study suggested that Drug X could be used as a preventive agent for OIPN.

P7-07

Kampo Drug Repositioning and Compound Mixture Analyses using Multi-task Graph Convolutional Neural Networks

Akihiro Douke[1]
akihiro.douke605@mail.kyutech.jp

Ryusuke Sawada[1]
sawad330@bio.kyutech.ac.jp

Michio Iwata[1]
iwata121@bio.kyutech.ac.jp

Yukie Sakai[1]
sakai541@bio.kyutech.ac.jp

Makoto Kadowaki[2]
makotok@inm.u-toyama.ac.jp

Yoshihiro Yamanishi[1]
yamani@bio.kyutech.ac.jp

[1] Department of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan.
[2] Division of Gastrointestinal Pathophysiology, Institute of Natural Medicine, University of Toyama, Toyama 930-0194, Japan.

Keywords: Kampo medicines, mode-of-action, target prediction, repositioning, graph convolutional neural network, multi-task learning, deep learning

Kampo medicine is useful to treat multifactorial and chronic diseases. Today in Japan, more than 80% of medical doctors prescribe Kampo medicines in addition to western medicine. The mechanisms of Kampo medicines are different from those of western medicines. The efficacies of Kampo medicines stem from multiple compound–multiple target interactions. Therefore, it is indispensable to establish fundamental technologies to comprehensively analyze the underlying mechanisms of every pharmacological action of multi-component Kampo medicines in the human body as a complex system. In this study, we propose a novel computational method to elucidate the mode-of-action of Kampo medicines and to predict new indications of Kampo medicines, which we call Kampo drug repositioning. This is an extension of our previous work [1]. In the algorithm, we estimate potential target proteins of each of constituent compound by multi-task graph convolutional neural networks (GCNNs), and predict its new indications by a mixture analysis of indications at the compound level and those at the crude drug level taking into account the ratio of crude drugs in each of Kampo medicines. We also estimate target pathways of Kampo medicines by an enrichment analysis of target proteins of constituent compounds. We construct predictive models based on 1,588,885 compound-protein interactions involving 1,111,415 compounds and 4,393 proteins, and predict new indications of 194 Kampo medicines involving 99 crude drugs and 2791 constituent compounds. The proposed methods are expected to be useful for health science and for clinical applications.

[1] Sawada, R., Iwata, M., Umezaki, M., Usui, Y., Kobayashi, T., Kubono, T., Hayashi, S., Kadowaki, M. & Yamanishi, Y., KampoDB, database of predicted targets and functional annotations of natural medicines, *Scientific Reports*, 8:11216, 2018.

P7-08

Prediction of Compound-Protein Interactions and Visualization Based on Graph Convolutional Networks

Marie Ikeguchi[1]
ikeguchi.marie.44n@st.kyoto-u.ac.jp

Shoichi Ishida[2]
ishida.shoichi.57a@st.kyoto-u.ac.jp

Ryosuke Kojima[1]
kojima.ryosuke.8e@kyoto-u.ac.jp

Hiroaki Iwata[1]
iwata.hiroaki.3r@kyoto-u.ac.jp

Yasushi Okuno[1,3,4]
okuno.yasushi.4c@kyoto-u.ac.jp

[1] Graduate School of Medicine, Kyoto University, Shogoin-Kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan
[2] Graduate School of Pharmaceutical Science, Kyoto University, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan
[3] Medical Sciences Innovation Hub Program, RIKEN Cluster for Science, Technology and Innovation Hub, Tsurumi-ku, Kanagawa 230-0045, Japan
[4] Foundation for Biomedical Research and Innovation at Kobe, Center for Cluster Development and Coordination, Chuo-ku, Kobe, Hyogo 650-0047, Japan

Keywords: Compound-protein interactions, Graph convolutional networks, Multimodal, Visualization, Integrated gradients

Predicting compound-protein interactions (CPIs) has played an important role in drug discovery [1], CPIs prediction methods using deep learning have achieved excellent performances [2,3]. However, effective methods for visualizing the interactions between each atom of a compound and each residue of a protein sequence have not been established. In this study, we developed a new prediction method together with the visualization of the predicted CPIs. Our method uses a multimodal neural network that combining graph convolutional networks [4] for chemical structure and convolutional neural networks for a protein sequence. Also, our method can visualize the contribution of the prediction results by using integrated gradients [5]. To evaluate the prediction performance of our methods, we applied this method to CPI datasets collected from the ChEMBL database. As a result, we indicated that a prediction performance of our model is higher than 0.9 AUC score. Furthermore, using our visualization method, we compared the visualized interactions with crystal structures. We want to give details of the results in this presentation.

[1] Michael, J. K., Vincent, S., Predicting new molecular targets for known drugs, *Nature*, 462:175-181, 2009.
[2] Tsubaki, M., Tomii, K., Sese, J., Compound-Protein interaction prediction with end-to-end learning of neural networks for graphs and sequences, *Bioinformatics*, 35:309-318, 2018.
[3] Hamanaka, M., Taneishi, K., Iwata, H., Ye, J., Pei, J., Hou, J., and Okuno, Y., CGBVS-DNN: Prediction of Compound‐protein Interactions Based on Deep Learning, *Molecular Informatics*, 2016.
[4] Kipf, T. N., and Welling, M. S., Semi-Supervised Classification with Graph Convolutional Networks, *International Conference on Learning Representations (ICLR)*, 2017.
[5] Sundararajan, M., Taly, A., and Yan, Q., Axiomatic Attribution for Deep Networks, *International Conference on Machine Learning (ICML)*, 2017.

P7-09

Development of data curation and integration protocol for chemical library in early drug discovery

Yugo Shimizu shimizu-yg@pha.keio.ac.jp
Masanori Osawa osawa-ms@pha.keio.ac.jp

Kazuyoshi Ikeda
ikeda-kz@pha.keio.ac.jp

Keio University, Faculty of Pharmacy, Division of Physics for Life Functions, 1-5-30 Shibakoen, Minato-ku, Tokyo 105-8512, Japan

Keywords: Chemical data integration, Curation, Library design

It is known that there are a huge number of chemical structures that can be synthesized in drug development. Therefore, it is important to develop data science technology in chemical information, especially chemical data curation and integration protocols related to existing pharmaceutical compounds.

We developed a new library-design method of data curation and integration for high-quality compound-library using open drug discovery data. We aim to design a core-library consisting of drug-like compounds for satisfying diversity in physical properties and chemical characteristics. Furthermore, we considered the bioactivity diversity for known and unknown drug targets. We created ligand-based target models by using several machine learning (ML) methods. We extracted compounds and analysis their bioactivity information from public databases such as ChEMBL and PubChem to train and test the models. By using our method, we mapped and annotated large-scale chemical data extracted from commercial libraries. In this presentation, we will show how to develop our chemical data science technologies for analyzing known and unknown drug developing compounds, and also discuss how to apply it for a core library in an early stage of drug discovery.

This work is conducted as a part of “Construction of Drug Discovery Informatics System” supported by Japan Agency for Medical Research and Development (AMED).

Clustering therapeutic drugs based on similarities of indications and side effects reported in public database

Ryutaro Shinkawa[1], **Yuya Suzumra**[1], **Taichi Yamada**[1], **Takafumi Moriguchi**[1],
319057@m.mie-u.ac.jp

Yuhei Nishimura[1]
yuhei@doc.medic.mie-u.ac.jp

[1]Department of Integrative Pharmacology, Mie University School of Medicine,
2-174 Edobashi, Tsu, Mie 514-8507, Japan

Keywords: Real world data, Public database, Cosine similarity, Multidimensional clustering

Therapeutic drugs have been classified based on pharmacodynamics and disease indications. However, it has gradually been revealed that profiling of side effects can be used to classify therapeutic drugs and to find novel disease indications of drugs. In this study, we generated multidimensional vectors for each therapeutic drug based on the cosine similarity of indications or side effects in the US FDA Adverse Event Reporting System (FAERS) published for 2018. Using the spatial density, the multidimensional vectors were clustered into 54 and 52 groups based on the indications and side effects, respectively, reported in FAERS. By comparing the 54 and 52 groups, we were able to identify 45 sets of therapeutic drugs that were common in the two clusterings. Several sets were comprising of therapeutic drugs with different pharmacodynamics and different disease indications. These findings suggest that clustering therapeutic drugs based on similarities of indications and side effects reported in public database can be useful to find novel similarity between therapeutic drugs and to reveal novel disease indications for these drugs.

Drug Discovery Raid Battle 2018: an open challenge to discover PD-1/PD-L1 small-molecule inhibitors

Kazuki Z Yamamoto[1, 2]
kayamamoto-tky@umin.net

Masahiro Mochizuki[3]
f09l001a@alumni.niigata-u.ac.jp

Hiroyuki Yamasaki[3]
yamasaki.phone@gmail.com

Kentaro Rikimaru[3]
kristall19@gmail.com

[1] Isotope Science Center, University of Tokyo, 2-11-16, Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan
[2] SHaLX Inc., 5-23-11 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[3] Drug Discovery Raid Battle 2018 Contributive Committee, 5-23-11 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Keywords: Social drug discovery, Virtual screening, Screening assay

PD-1/PD-L1 immune checkpoint inhibitors, which have recently been established to be effective in the treatment of cancer, have been put to practical use as antibody drugs. However, they have disadvantages such as cost and inability to be administered orally. Therefore, if substitution of those biologics by small molecules is possible, the resulting practical merits are great [1]. Here, in order to discover small molecules that inhibit PD-1/PD-L1 immune checkpoint signaling, we held a drug discovery event in which anyone could participate. We provided a digital library of screening compounds [2] and the participants individually screened the candidate compounds and submitted the candidates to us. And we assay them and confirm the inhibitory activity [3], whose results are supposed to be available by the fall this year. If promising hit compounds are obtained, we aim for development of a medicine with the compound proposer.

[1] Shaabani, S., Huizinga, H. P., Butera, R., Kouchi, A., Guzik, K., Magiera-Mularz, K., ... & Dömling, A. A patent review on PD-1/PD-L1 antagonists: small molecules, peptides, and macrocycles (2015-2018). *Expert opinion on therapeutic patents*, 28(9), 665-678, 2018.
[2] Enamine Screening Collection
https://enamine.net/hit-finding/compound-collections/screening-collection
[3] Cisbio, Human PD1/PD-L1 biochemical binding assay
https://www.cisbio.jp/human-pd1-pd-l1-biochemical-binding-assay-40578

Automatic Reading of Tables and Figures in Scientific Papers

Hiroyuki Shindo[1]
shindo@is.naist.jp

[1] Division of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan

Keywords: Natural Language Processing, Machine Learning, Computer Vision

Automatic reading of tables and figures is a critical component for data mining and information extraction from scientific papers since many experimental results in scientific research are presented in the form of tables and figures. In computer science, previous work [1, 2, 3] has shown that tables and figures can be successfully detected and parsed with the help of machine learning techniques.

In this work, I propose an automated end-to-end pipeline for information extraction from scientific papers. Given a PDF of scientific papers, the body texts, tables and figures are firstly detected with deep neural networks, then the extracted objects are parsed and structured with natural language processing and image processing. For figure parsing, this work focuses on data extraction from line charts and scatter plots. Experimental results show that my proposed approach accurately extracts data from tables and figures.

[1] Mathieu Cliche, David Rosenberg, Dhruv Madeka and Connie Yee, Scatteract: Automated extraction of data from scatter plots, *Proc. of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 2017.

[2] Noah Siegel, Nicholas Lourie, Russell Power and Waleed Ammar, Extracting Scientific Figures with Distantly Supervised Neural Networks, *Proc. of Joint Conference on Digital Library*, 2018.

[3] Minghao Li, Lei Cui, Shaohan Huang, Furu Wei, Ming Zhou, Zhoujun Li, TableBank: Table Benchmark for Image-based Table Detection and Recognition, *arXiv*, 2019.

P8-01

Functionalization of liposomes using artificially evolved peptides against lipid membranes

Takuya Terai[1] **Shota Kobayashi**[1] **Yuki Yoshikawa**[1]
tterai@mail.saitama-u.ac.jp

Yasuhito Utsugi[1] **Naoto Nemoto**[1]
nemoto@mail.saitama-u.ac.jp

[1] Graduate School of Science and Engineering, Saitama University, 255 Shimo-okubo Sakura-ku, Saitama, 338-8570, Japan

Keywords: Molecular Robotics, Liposome, In vitro selection, Peptide

A molecular robot is a bottom-up artificial molecular system comprised of sensors, actuators, and computing machineries, using organic (bio)molecules such as (modified) nucleic acids, proteins, and lipids. They are potentially useful in a number of applications such as drug delivery, information storage, in vivo diagnosis, and so on [1]. In most cases, the outer layer, or scaffold, of molecular robots is giant unilamellar vesicles (GUVs), or liposomes. Hence, to promote molecular robotics, it is important to develop a facile methodology to stably attach molecules of interest on the surface of lipid membranes.

In the last two decades, our group has established a methodology of in vitro evolution of polypeptides using a puromycin-modified DNA as a linker to covalently attach a polypeptide to its corresponding cDNA (= cDNA display) [2][3]. In this study, we used cDNA display to discover novel peptide sequences that had affinity to lipid membranes but did not disrupt liposome structures. We prepared an arginine-rich peptide library over 10^{12} members, and performed seven rounds of *in vitro* selection against DOPC GUVs. Liposome fractions were collected and attached cDNA were amplified. After cloning of selected cDNA display molecules, a peptide called LB-1 was identified as a novel peptide that bound to lipid membranes [4]. LB-1 consisted of amphiphilic region and cationic region, and both parts played important roles for membrane anchoring. Notably, fusion of nucleic acids and proteins (including an antibody) at the C-terminus of LB-1 did not hamper its specific localization on membrane. It is expected that the peptide will be a useful toolkit to attach sensors and targeting molecules to liposome-based molecular robots.

Furthermore, when we investigated several LB-1 derivatives, it was found that one of the peptides showed permeability to liposome membrane. Although further study is necessary, it seems that the composition of lipid membrane is important for this phenomenon. More details will be presented at the conference venue.

[1] Murata, S., Konagaya, A., Kobayashi, S., Saito, H., Hagiya, M., Molecular Robotics: A New Paradigm for Artifacts, *New Generation Computing*, 31:27-45, 2013.
[2] Yamaguchi, J., Naimuddin, M., Biyani, M., Sasaki, T., Machida, M., Kubo, T., Funatsu, T., Husimi, Y., Nemoto, N., cDNA display: a novel screening method for functional disulfide-rich peptides by solid-phase synthesis and stabilization of mRNA–protein fusions, *Nucleic Acids Research*, 37:e108, 2009.
[3] Terai, T., Anzai, H., Nemoto, N., Selection of Peptides that Associate with Dye-Conjugated Solid Surfaces in a pH-Dependent Manner Using cDNA Display, *ACS Omega*, 4:7378-7384, 2019.
[4] Kobayashi, S., Terai, T., Yoshikawa, Y., Ohkawa, R., Ebihara, M., Hayashi, M., Takiguchi, K., Nemoto, N., In vitro selection of random peptides against artificial lipid bilayers: a potential tool to immobilize molecules on membranes, *Chemical Communications*, 53:3458-3461, 2017.

Experimental Investigation of DNA Generation Circuits toward Molecular Robot Control

Ken Komiya[1] **Chizuru Noda**[1]
komiya@c.titech.ac.jp

Teruya Enomoto[1] **Masayuki Yamamura**[1]
my@c.titech.ac.jp

[1] School of Computing, Tokyo Institute of Technology, 4259, Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa, 226-8503, Japan

Keywords: Single-stranded DNA, DNA amplification, Physiological temperature, Molecular robot

We had developed a modular reaction cascade that generates pre-designed, single-stranded DNA (ssDNA) molecules as signal transducers in a programmable way [1]. With the use of tandemly aligned template sequences of signal DNA [2], we later applied our reaction cascade to exponential DNA amplification at a physiological temperature [3]. This isothermal ssDNA amplification reaction, termed low-temperature amplification (L-TEAM) reaction, allows a high amplification rate and concurrent hybridization of the amplified ssDNA under low-temperature conditions during amplification. Therefore, the L-TEAM reaction seems promising for versatile applications including automated nucleic acid test and a cell-like molecular robot that requires generation of a large amount of ssDNA signals to control a swarm of DNA-directed nanodevices in response to a tiny amount of nucleic acid stimuli [4, 5]. In the present study, we experimentally investigated the performance of DNA generation cascades as elementary molecular circuits. Leak-free DNA amplification and temporal modulation of DNA generation would be discussed from the viewpoint of molecular robots, where our reaction cascade is integrated as the controller.

[1] Komiya, K. and Yamamura, M., Cascading DNA generation reaction for sequential control of DNA nanomachines at a physiological temperature, *New Generation Computing*, 33:213-229, 2015.
[2] Van Ness, J., Van Ness, L. K. and Galas, D. J., Isothermal reactions for the amplification of oligonucleotides, *Proc. Natl. Acad. Sci. USA*, 100(8):4504-4509, 2003.
[3] Komiya, K., Komori, M., Noda, C., Kobayashi, S., Yoshimura, T. and Yamamura, M., Leak-free million-fold DNA amplification with locked nucleic acid and targeted hybridization in one pot, *Organic and Biomolecular Chemistry*, 17(23):5708-5713, 2019.
[4] Komori, M., Komiya, K., Shirakawa, T., Morikawa, T. J. and Yoshimura, T., Measurement of microRNA with isothermal DNA amplification on fully automated immunoassay analyzers, *Analytical and Bioanalytical Chemistry*, 411(17):3789-3800, 2019.
[5] Sato, Y., Komiya, K., Kawamata, I, Murata, S., Nomura, S. M., Isothermal amplification of specific DNA molecules inside giant unilamellar vesicles, *Chemical Communications, in press*, 2019, DOI: 10.1039/c9cc03277k

P8-03

A database to store design information of DNA nanostructure

Ibuki Kawamata[1] kawamata@molbot.mech.tohoku.ac.jp
Sho Aradachi[1] aradachi@molbot.mech.tohoku.ac.jp
Satoshi Murata[1] murata@molbot.mech.tohoku.ac.jp

[1] Department of Robotics, School of Engineering, Tohoku University, Sendai, Miyagi, Japan

Keywords: Molecular Robotics, DNA origami, DNA nanotechnology, database, caDNAno file

In the field of structural DNA nanotechnology, DNA duplex is utilized as build blocks to assemble nano-scale structures with a complex geometry [1]. In the past few decades, a number of DNA nanostructures have been demonstrated with a potential application in molecular-scale robotics, smart drug delivery systems, and so on [2]. To rationally design such DNA nanostructures, several softwares have been developed such as caDNAno [3]. Although those softwares can be downloaded, raw design information (e.g. caDNAno file) is not usually provided together with a paper publication. The situation causes an underlying problem that one has to pay a great effort to reuse/recycle the DNA nanostructures of the others. Getting rid of the obstacle enables us to engineer novel DNA systems on top of well-qualified DNA nanostructures.
Here, we propose a web-based database that stores the design information of DNA nanostructures, especially targeting the caDNAno format (Fig. 1). In the currently developed trial version, the input to the database is a caDNAno file, which is then converted to several useful formats such as PDB file, rendered images, DNA sequence table etc. It is also possible to store additional information such as digital object identifier (DOI) to identify the publication about the structure. Moreover, we prepare some standard objects such as rectangles and cuboids, which are available from the database. From the web-based interface, users can upload a caDNAno file and download the all of the data for the purpose of lab experiment and/or computer simulations. At the presentation, we would like to discuss further requirements and get some feedbacks of the database to improve the user experience of the service. In the future, the database may serve as a catalog of DNA objects, which will accelerate the engineering process of complex DNA nanostructures.

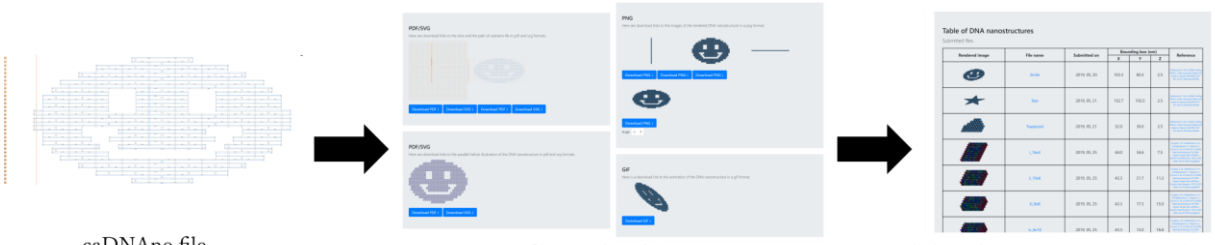


Figure 1: The schematic figure of the DNA nanostructure database. When the caDNAno file is uploaded, it is converted to several file formats, which will be listed in the database.

[1] N. C. Seeman, Molecular Biotechnology. 2007, 37, 246—257.
[2] N. C. Seeman and H, F. Sleiman, Nature Reviews Materials. 2017, 3, 17068.
[3] S. M. Douglas, A. H. Marblestone, S. Teerapittayanon, A. Vazquez, G. M. Church, and W. M. Shih, Nucleic Acids Research. 2009, 37, 5001–5006.

Haptic Interactive Virtual Reality Simulation on Biomolecules

Arif Pramudwiatmoko[1,2]
Pramudwiatmoko.a.aa@m.titech.ac.jp

Satoru Tsutoh[3]
tsutoh.satoru@fujixerox.co.jp

Gregory Gutmann[1]
gutmann.g.aa@m.titech.ac.jp

Yutaka Ueno[4]
yutaka.ueno@aist.go.jp

Akihiko Konagaya[1]
kona@c.titech.ac.jp

[1] Tokyo Institute of Technology, J3-18, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan
[2] Universitas Teknologi Yogyakarta, Jl Siliwangi, Jombor, Sleman, D.I. Yogyakarta 55285, Indonesia
[3] Fuji Xerox Co., Ltd., 430 Sakai, Nakai-machi, Ashigarakami-gun, Kanagawa 259-0157, Japan
[4] National Institute of Advanced Industrial Science and Technology, 2-4-7, Aomi, Koto-ku, Tokyo 135-0064, Japan

Keywords: Haptic rendering, hand tracking, molecular modelling

In order to provide an intuitive way to observe molecular objects, virtual reality (VR) system for biomolecules have been implemented with haptic technology and hand tracking module. An immersive VR simulation requires a specific framerate to perform and so does the interaction of hands with the biomolecules. Anchors with spring forces are implemented to propagate hand interaction force to biomolecule objects.

[1] Pramudwiatmoko, A., Tsutoh, S., Gutmann, G., Ueno, Y., Konagaya, A., Haptic rendering applied to hand tracking 3D user interface for a molecular modeling environment, *24th International Symposium on Artificial Life and Robotics*, International Society of Artificial Life and Robotics, 109–114, 2019.
[2] Gutmann, G., Inoue, D., Kakugo, A., Konagaya, A., Parallel Interaction Detection Algorithms for a Particle-based Live Controlled Real-time Microtubule Gliding Simulation System Accelerated by GPGPU, *New Gener Comput*, 35:157–180, 2017.
[3] Löwe, J., Li, H., Downing, K.H., Nogales, E., Refined structure of $\alpha\beta$-tubulin at 3.5 Å resolution, *J Mol Biol*, 313:1045–1057, 2001.

Molecular dynamics study on breakage of photoresponsive DNA

Ryuzo Azuma azuma.r.ac@m.titech.ac.jp
Greg Gutmann gutmann.g.aa@m.titech.ac.jp
Akihiko Konagaya kona@c.titech.ac.jp

Tokyo Institute of Technology, J3-25, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan

Keywords: Molecular robotics, azobenzene attached photoresponsive DNA, molecular dynamics

Mechanical stretching of DNA has been one of hot topics in nanotechnology and nano-science research fields for nearly 30 decades. Whereas conventional studies have mostly focused on characterization of mechanisms underlying force-stretching curves obtained from experiments, engineering novel functional devices by utilizing these properties has also become popular since the development of technologies such as DNA linkers, switches, etc. In particular, photoresponsive DNA (*p*-DNA) has been incorporated into tubulines, the protein dimer subunit of microtubles, as conjugate and lead to successful observation of switching on and off the collective motion of microbutules (referred to as 'swarms') in response to visible and UV light irradiation [1].

Here we present our molecular dynamics studies on breakage of *p*-DNA under a pulling force applied to the 5' ends. The *p*-DNA consists of an 8 repeat of the same sequence, CAA, each of which (except for the last one) is followed by one azobenzene. Azobenzene is known to take two different forms depending on wave lengths of irradiation: under visible or UV light they may take trans- or cis-forms, respectively [2]. Our strategy is to use atomic structures of *p*-DNA prepared with those 2 different forms of azobenzenes and carry out MD simulations so that we can distinctively analyze the dynamics of breakage. In addition, as a control, we carried out same MD analysis on native duplex with a sequence $(CAAT)_{11}CAA$. The pulling force is due to ideal elastic strings attached on both ends of the *p*-DNA, and terminals of those strings are moving apart at a constant speed. In our measurements, the force constants of the elastic string are ranged within 0.144-14.4 [10^{-2}pN/Å], while the pulling speeds are ranged within 5-100 [10^{-3}Å/ps]. The system initially undergoes an NVT MD run up to 36ps for equilibration, then goes through a production NPT run under a pulling force for 8-120ns. The simulation results exhibit common behavior wherein breakage involves structural transitions. The latter are shown to proceed as follows: the B→ S structural change continues until the pulling force reaches the maximum (as found experimentally [3]), which is followed by gradual sliding and rapid decrease in the pulling force, and then separation of the strands involving rapid shrinkage takes place after breakage of all the h-bonds. The insertion of azobenzenes is shown to decrease the time to break off, whereas this effect is significantly lower for the trans-ones. We further analyzed these results more quantitatively and obtained dependence of force-extention curves on the pulling force constants and speeds. Those results will be discussed in detail in our presentation.

The authors thank Y Ueno (AIST), A Kakugo (Hokkaido Univ.), and A Kuzuya (Kansai Univ.) for useful discussions. Financial supports: Strategic Advancement of Multi-Purpose Ultra-Human Robot and Artificial Intelligence Technologies (NEDO, P15009), MEXT/JSPS KAKENHI (JP17H00769), and Platform Project for Supporting Drug Discovery and Life Science Research (BINDS-AMED, JP18am0101112).

[1] Keya, J.J., *et al.*, DNA-assisted swarm control in a biomolecular motor system, *Nature Communications*, 453(9), 1-8, 2019.
[2] Asanuma H., Ito T., Komiyama M. Photoregulation of the formation and dissociation of a DNA duplex by using the *cis-trans* isomerization of azobenzene. *Angew. Chem. Int. Ed. Engl.*, 38, 2393–2395, 1999.
[3] Cluzel, P., *et al.*, DNA: An Extensible Molecule, *Science*, 271(9), 792-794, 1996.

Formulating R&D Guidelines for Molecular Robotics

Naoto Kawahara[1]
n-kawaha@med.kyushu-u.ac.jp

Ryuma Shineha[2]
r_shineha@seijyo.ac.jp

Akihiko Konagaya[3]
kona@dis.titech.ac.jp

[1] Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan
[2] Seijyo University, Seijo 6-1-20, Setagaya-ku, Tokyo 157-8511, Japan
[3] Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8503 Japan

Keywords: R&D, Molecular Robotics, Genetically Modified Organism, Cartagena Protocol on Biosafety, Containment Measures, Biological Control, Physical Control, Chemical Management

Thus far, we have established the ethical principles of Molecular Robotics with our colleagues. Based on the principles, we will progress to formulate practical R&D guidelines for proper use of Molecular Robotics. For example, its composition and/or ingredient, physical and/or chemical properties, and environmental influence might be important factors in the R&D guidelines. Otherwise, its containment measure involving a certain chemical management might be discussed with reference to the Cartagena Protocol on Biosafety. What then are such the R&D criteria of Molecular Robotics? - We discuss practical frameworks or viewpoints for proper use of Molecular Robotics focusing on its technological characteristics concerning senses, motions and intelligence.

[1] U.S. Library of Congress, “Restrictions on Genetically Modified Organisms: Japan”, https: //www.loc.gov/law/help/restrictions-on-gmos/japan.php, 2015
[2] Ministry of Economy, Trade and Industry, Japanese Government, “Chemical Management”, https://www.meti.go.jp/policy/chemical_management/english/index.html, 2014
[3] PMDA, “Quality Evaluation of Antibody Drug”, https://www.pmda.go.jp/files/000206143.pdf, 2012
[4] UN, “The Globally Harmonized System of Classification and Labelling of Chemicals (F ourth Revised Edition)”, https://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev0 4/English/ST-SG-AC10-30-Rev4e.pdf, 2011.
[5] Comité Europeo de Normalización, “Laboratory Biorisk Management Standard (CWA 15 793)”, ftp://ftp.cenorm.be/PUBLIC/CWAs/wokrshop31/CWA15793, 2008
[6] WHO, “Laboratory Biosafety Manual (Third Edition)”, https://www.who.int/csr/resources/ publications/biosafety/Biosafety3_j.pdf, 2004
[7] UN Secretariat of the Convention on Biological Diversity, “The Cartagena Protocol on Biosafety”, http://bch.cbd.int/protocol, 2003.

Programmable DNA reaction-diffusion system for a Voronoi pattern formation in hydrogel

Keita Abe[1] abe@molbot.mech.tohoku.ac.jp
Ibuki Kawamata[1] kawamata@molbot.mech.tohoku.ac.jp
Shin-ichiro M. Nomura[1] nomura@molbot.mech.tohoku.ac.jp
Satoshi Murata[1] murata@molbot.mech.tohoku.ac.jp

[1] Department of Robotics, Graduate School of Engineering, Tohoku University, Aramaki-aza-Aboba 6-6-01, Abobaku, Sendai, Miyagi 980-8579, Japan.

Keywords: Pattern formation, Reaction diffusion system, DNA logic gate

In the early stages of biological development, ordered patterns form as a process of morphogenesis. Some of ordered pattern such as skin pattern of fish and far pattern of mammals are reproduced by numerical simulation based on reaction-diffusion system [1]. Such biological phenomena and theory inspire engineers to build artificial self-organizing system. To realize desired reactions in solution, synthetic DNA which interacts each other based on its base sequence is studied as a programmable material. Some artificial pattern formation by synthetic DNA shows the capability to program reaction [2, 3], however, to program diffusion, which is as important factor for pattern formation as reaction, is not so focused on.

Here, we demonstrate a method to program both reaction and diffusion in hydrogel for pattern formation using synthetic DNAs. To program the reaction, a DNA logic gate is used and to immobilize some of the DNA into the hydrogel medium, acrydite modified anchor DNA [4] is employed. Using the DNA logic gate and the anchoring method, we have successfully demonstrated Voronoi pattern formation in hydrogel medium based on sites which is specified as the diffusion source of DNA. Additionally, the diffusion of DNA is modulated by tuning the reversible interaction between diffusible DNA and immobilized DNA [5]. By the diffusion modulation, the sites were given each “weight” for weighted Voronoi pattern formation [6].

As a next step, we simplify the reaction system and reduce the component DNAs. We have also demonstrated similar Voronoi pattern formation with using a smaller number of DNA. We think that the simplification is effective to reduce undesired interactions and extend the programmability. We will report the method of the programmable pattern formation and its latest result.

[1] Turing, A. M., *Philos. Trans. R. Soc., B*, 237(641), 37 – 72, 1952.
[2] Zadorin, A. S., Rondelez, Y., Gines, G., Dilhas, V., Urtel, G., Zambrano, A., Jean-Christophe G., Estevez-Torres, A., *Nat. Chem.*, 9(10), 990, 2017.
[3] Zenk, J., Scalise, D., Wang, K., Dorsey, P., Fern, J., Cruz A., Schulman, R., *RSC Adv.*, 7(29), 18032 – 18040, 2017.
[4] Kawamata, I., Hosoya, T., Takabatake, F., Sugawara, K., Nomura, S., Isokawa, T., Peper, F., Hagiya, M., Murata, S., 2016 *Fourth International Symposium on Computing and Networking (CANDAR)*, IEEE pp. 215-221, 2016.
[5] Rodjanapanyakul, T.,Takabatake, F., Abe, K., Kawamata, I., Nomura, S., Murata, S., *Phys. Rev. E*, 97, 2018
[6] Abe, K., Kawamata, I., Nomura, S., Murata, S., *Molecular Systems Design & Engineering*, 2019.

P8-08

Reconstitution of kinesin-based transport complex using DNA origami

Kodai Fukumoto[1] **Yuya Miyazono**[2]
kfukumoto@protein.osaka-u.ac.jp

Hisashi Tadakuma[1] **Yoshie Harada**[1]
tadakuma@protein.osaka-u.ac.jp yharada@protein.osaka-u.ac.jp

[1] Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan
[2] Graduate School of Frontier Science, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan

Keywords: DNA nanostructure, DNA origami, Molecular motor, Coordination

Intracellular transport is central to the cellular phenomena, where kinesins and other multiple motors orchestrate their function. Although it has been known that the intra- and inter-molecular coordination of kinesins render the long distance transport of the cargo[1-3], it remains elusive how the molecular layout (e.g. distance between individual motors) affect the collective transport. To elucidate the coordination mechanism, we have made a transport complex, in which kinesin motors were integrated onto DNA origami through covalent bond between tag-protein and its ligand (e.g. SNAP system). However, the low binding rate of the integration process hinders the yield of the artificial transport complex, which might be attribute to the negative charge of the DNA scaffold.

Here, we introduced positively charged Lys-tag to C-terminal of kinesin-SNAPf proteins. The Lys-tag allowed us to increase the binding rate of dimeric and monomeric kinesins onto DNA origami by 700 times (k_{on}=4x10^6 /M/s) and 300 times (k_{on}=6x10^5 /M/s), respectively. Furthermore, we succeeded in constructing artificial transport complex with 8 kinesin molecules. We found that the number of kinesin did not affect the speed of artificial transport complex, but affected the run length. We now investigate the effect of the molecular layout on the transport activity. These results suggest the potential of Lys-tag for DNA origami-based protein integration and transporter formation.

[1] Miyazono, Y., Hayashi, M., Karagiannis, P., Harada, Y. and Tadakuma, H. *Strain through the neck linker ensures processive runs: a DNA-kinesin hybrid nanomachine study*, EMBO J 29:93-106, 2010.
[2] Hendricks, A.G., Perlson, E., Ross, J.L., Schroeder, H.W., Tokito, M. and Holzbaur, E.L.F. *Motor Coordination via a Tug-of-War Mechanism Drives Bidirectional Vesicle Transport*, Current Biology, 20:697–702, 2010.
[3] Furuta, K., Furuta, A., Toyoshima, Y.Y., Amino, M., Oiwa, K. and Kojima, H. *Measuring collective transport by defined numbers of processive and nonprocessive kinesin motors*, PNAS, 110:501–506, 2013.

Gold nanoparticle-loaded DNA hydrogel microparticles for catalysis in aqueous phase

Daisuke Ishikawa[1]
daisuke-ishikawa@tmu.ac.jp

Masahiro Takinoue[2]
takinoue@dis.titech.ac.jp

Toru Murayama[1]
murayama@tmu.ac.jp

Masatake Haruta[1]
haruta-masatake@tmu.ac.jp

[1] Research Center for Gold Chemistry, Graduate School of Urban Environmental Sciences, Tokyo Metropolitan University, 1-1 Minami-osawa, Hachioji, Tokyo, Japan
[2] Department of Computer Science, Tokyo Institute of Technology, 4259-J2-36 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa, Japan

Keywords: DNA hydrogel, gold nanoparticle, catalysis, artificial cell

DNA is a biopolymer with designability that can create structures ranging in size from nanoscale structures such as DNA origami[1] to macroscale materials such as gels depending on the design of the base sequence. In particular, DNA hydrogels have been expected to be applied as smart materials such as sensing of specific ions by being complexed with metal nanoparticles, which are based on the use of coordination of metal ions to the bases and phosphate groups of DNA[2].

Here, we present DNA hydrogel particles with molecular permeability and catalytic function by utilizing the designability of the DNA. The DNA hydrogel particles were composed of Y-motif subunits consisting of three single-stranded DNAs with a palindromic sequence as a sticky end. Subsequently, gold nanoparticles (Au NPs) which have been known to show catalytic activity in water were formed in the DNA hydrogel particles by immersion and reduction of the hydrogel particles in $HAuCl_4$ aqueous solution followed by reduction of Au^{3+} ions (Figure 1). The DNA hydrogel particles grew larger when the cooling rate was reduced during annealing. The Au NPs formed and dispersed uniformly inside the DNA hydrogel particles were confirmed by HAADF-STEM and TEM observation. The size of Au NPs tended to decrease as the size of DNA hydrogels increased. Assuming that large size DNA hydrogel particles have high crosslink density, it suggests that the density inhibits the formation of Au NPs. Catalytic activity of the Au NPs in the DNA hydrogel particles was examined via the model reaction of the reduction of 4-nitrphenol to 4-aminophenol, which suggested that Au NPs worked as a catalyst.

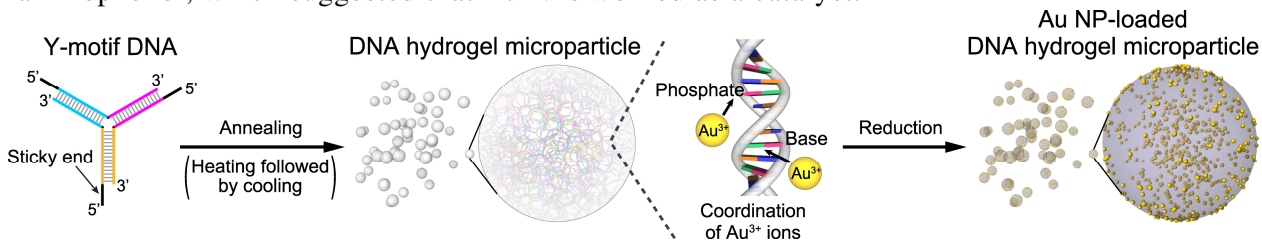


Figure 1. Scheme for preparation of the Au NP-loaded DNA hydrogel microparticles.

[1] D. Ishikawa, Y. Suzuki, C. Kurokawa, M. Ohara, M. Tsuchiya, M. Morita, M. Yanagisawa, M. Endo, R. Kawano, and M. Takinoue, DNA Origami Nanoplate-Based Emulsion with Designed Nanopore Function, *Angewandte Chemie International Edition*, 2019, accepted.
[2] I. Sissoëff, J. Grisvard, and E. Guillé, Studies on metal ions-DNA interactions: Specific behaviour of reiterative DNA sequences, *Progress in Biophysics and Molecular Biology*, 31:165–199, 1978.

Negating Latency for Fluid Interactions with Biomolecules in a Client/Server VR System

Gregory Gutmann [1]
gutmann@c.titech.ac.jp

Azuma Ryuzo [2]
azuma.r.ac@m.titech.ac.jp

Akihiko Konagaya [2]
kona@c.titech.ac.jp

[1] Tokyo Institute of Technology, J3-18, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan

Keywords: VR, Real-time, Simulation, GPGPU, Latency, Client/Server

We are working towards creating a virtual research and learning environment at the nanoscale, by utilizing VR and GPGPU computation. VR enables us to visualize and interact with atomic-scale phenomenon in a way that makes them seen life-size, perceptually centimeter to meter scale in VR [1]. GPGPU computing offers large amounts of computational performance; however, when using a single machine, performance limitations can still arise. Using a client and server approach can increase the available compute power, but it increases the latency, the round trip time of interactions to results. In this work, we are discussing our solution termed “Predictive Simulation”. By using this extrapolation based approach with a strong predictor we are able to significantly decrease the perceived latency while not placing a large compute load on the client like many current methods for negating latency do [4,5]. Our initial results show a 3x reduction in the discrepancy between the user’s live position and the results returning from the compute server. This has enabled live interaction in VR with larger-scale simulations than previously possible.

[1] Gutmann, G., Azuma, R., & Konagaya, A. (2018). A Virtual Reality Computational Platform Dedicated for the Emergence of Global Dynamics in a Massive Swarm of Objects. Journal of the Imaging Society of Japan,57(6), 647-653. doi:https://doi.org/10.11370/isj.57.647

[2] Azuma, R., Kishi, S., Gutmann, G., & Konagaya, A. (2018). All-atom molecular dynamics of film supported flat-shaped DNA origami in water. Chem-Bio Informatics Journal,18(0), 96-118. doi:10.1273/cbij.18.96

[3] Gutmann, G., Inoue, D., Kakugo, A., & Konagaya, A. (2017). Parallel Interaction Detection Algorithms for a Particle-based Live Controlled Real-time Microtubule Gliding Simulation System Accelerated by GPGPU. New Generation Computing,35(2), 157-180. doi:10.1007/s00354-017-0011-5

[4] Bernier, Y. W. (2001, March). Latency compensating methods in client/server in-game protocol design and optimization. In Game Developers Conference (Vol. 98033, No. 425).

[5] Aggarwal, S., Banavar, H., Khandelwal, A.: Accuracy in dead-reckoning based distributed multi-player games. In: SIGCOMM’04 Workshops, pp. 161–165. ACM Press, New York (2004)

Implementing Real Time Molecular Dynamics with in Haptic Molecular Modeling Environment

Yutaka Ueno[1]
uenoyt@ni.aist.go.jp

Arif Pramudwiatmoko[2]
pramudwiatmoko.a.aa@m.titech.ac.jp

Gregory Gutmann[2]
gutmann@c.titech.ac.jp

Akihiko Konagaya[2]
kona@c.titech.ac.jp

[1] Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology, 2-3-26 Aomi, Koto-ku, Tokyo 135-0064, Japan
[2] Tokyo Institute of Technology, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8502, Japan

Keywords: Molecular Robotics, Virtual Reality, Molecular Dynamics

With recent advances of virtual reality hardware and graphics rendering performance, molecular modeling environment have been improved also from novel software to support real time molecular dynamics. We have studied hand tracking 3D user interface for the molecular modeling software and introduced haptic rendering facility that allows user to feel touching and grasping molecules. It is useful in particular to modeling macromolecular complex systems used in molecular robotics application. Our studies started with existing physics engine, Open Dynamics Engine [1], and Unity game development suite [2]. Recently, novel simulation code with high performance rendering was introduced [3] to handle macromolecules with large numbers of atoms. Our molecular dynamics code is based on an elastic network model, a simplified force field which supports modeling works by hand. While recent physics engine simulate position based dynamics, we employed conventional force based simulation with Lenard-Jones potential and spring force allowing various interactions between molecular objects. Our implementation topics using graphics library DirectX12 will be discussed.

[1] Ueno, Y., Ito, S., Konagaya, A., Implementing a modeling software for animated protein-complex interactions using a physics simulation library. *Journal of Bioinformatics and Computational Biology* 12:1442003, 2014.
[2] Pramudwiatmoko, A., Tsutoh, T., Gutmann, G., Ueno, Y., Konagaya, A., Haptic Rendering Applied to Hand Tracking 3D User Interface for a Molecular Modeling Environment, *Proceedings of the 24th International Symposium on Artificial Life and Robotics*, 109-114, 2019.
[3] Gutmann. G., Inoue, D., Kakugo, A., Konagaya, A., Real-time 3D microtubule gliding simulation accelerated by GPU computing. *International Journal Automation and Computing*, 13:108–116, 2016.

Design and optimization of a branching structure for dendric DNA structure

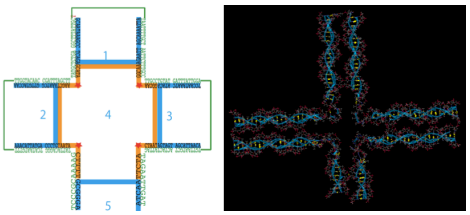
Takayuki Kobayashi[1] k714002@kansai-u.ac.jp
Yuki Mano[1] k200768@kansai-u.ac.jp
Ryuzo Azuma[2]
Yuichi Ohya[3] yohya@kansai-u.ac.jp
Yutaka Ueno[4] uenoyt@ni.aist.go.jp
Akihiko Konagaya[2] kona@c.titech.ac.jp
Akinori Kuzuya[3] kuzuya@kansai-u.ac.jp

[1] The Graduate School of Science and Engineering, Kansai University, 3-3-35, Yamate-cho, Suita, Osaka 564-8680, Japan
[2] School of Computing, Tokyo Institute of Technology, J3-25, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan
[3] Department of Chemistry and Materials Engineering, Kansai University, 3-3-35, Yamate-cho, Suita, Osaka 564-8680, Japan
[4] Artificial Intelligence Research Center, AIST, 2-4-7, Aomi, Koutou-ku, Tokyo 135-0064, Japan

Keywords: DNA Origami, Dendrimer, Molecular Dynamics

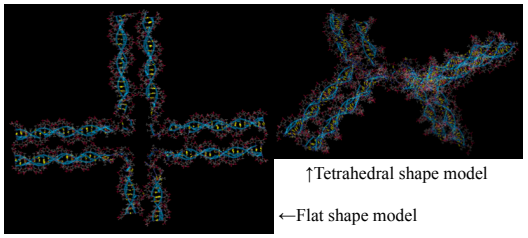
We considered about optimal design of DNA dendrimer: dendric DNA structure. By utilizing a GPU based simulation engine, we assessed flexibility of the branching structure of DNA origami object. Our results demonstrate that three-dimensional DNA origami objects need more flexibility than flat one to transform with electrostatic repulsion.

DNA origami object originate from the sequence design of staple strands, those are many short ssDNAs, so DNA origami method allows design objects without considering atomic details. But, it is important to assess structural properties of nano objects using atomic-level DNA origami structure modeling and simulations with atomic detail.



In this study, by molecular dynamics (MD) simulation, we evaluated flexibility of the DNA origami structure. It is already reported about flat one[1)], but it is not reported about three-dimensional one yet. According that, we consider the flexibility that is sufficiently to change geometry with electrostatic interaction.

First, the DNA origami object model was created by Maestro (Schrödinger's molecular modeling interface). Next, ions placement, solvation, and MD simulation by Desmond (Schrödinger's MD simulation program).



We created 2 models of flat-shape one and tetrahedral-shape one, and addressed conformational changes up to the order of nanoseconds simulation time. According to the results, the former keeps flat-shape, but latter similar to flat-shape. It means, those objects stabilized in flat-shape.

We expect that the structure stabilized in tetrahedral-shape with electrostatic interaction, since phosphate groups have negative charges. So, this structure is less flexible than we expected.

[1] H. Yan, *et al., Science* **2003,** *301,* 1882-1884

P9-01

Attempts to establish the in vitro BBB model suitable for drug development
— Comparative study of 2D rat model, 2D human cell line model, and 3D human cell line model—

Kimiko Kitamura [1,2] k.kitamura@nihs.go.jp
Yukari Shigemoto-Mogami [1,2] shigemot@nihs.go.jp
Kazue Hoshikawa [1,2] hoshikawa@nihs.go.jp
Agathe Figarol [3] a-figarol@chem.eng.osaka-u.ac.jp
Michiya Matsusaki [3] m-matsus@chem.eng.osaka-u.ac.jp
Tomomi Furihata[4] tomomif@toyaku.ac.jp
Seiichi Ishida [2] ishida@nihs.go.jp
Kaoru Sato [1,2] kasato@nihs.go.jp

[1] Laboratory of Neuropharmacology, Division of Pharmacology, National Institute of Health Sciences, 3-25-26, Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa, Japan
[2] Division of Pharmacology, National Institute of Health Science, 3-25-26, Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa, Japan
[3] Department of Applied Chemistry, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka, Japan
[4] School of Pharmacy, Tokyo University of Pharmacy and Life Sciences,1432-1 Horinouchi, Hachioji, Tokyo 192-0392 Japan

Keywords: Microphysiological system, blood-brain barrier, human cell line, permeability

Microphysiological systems（MPS）are new micro-engineered laboratory models that reproduce tissue phenotypes on a chip. It is expected that MPS will be models which reproduce systemized cell functions in *in vivo* and provide solutions to the industrialization of scientific findings. Limited number of *in vitro* blood brain barrier (BBB) models are available at present. We therefore are attempting the establishment of the BBB-MPS with high ability of extrapolation to human. A key challenge in this study is to determine the parameters reflecting BBB function. To this aim, we have selected some parameters by which we can determine whether the model is BBB-alike or not. The parameters include substrate permeability, trans-endothelial electrical resistance (TEER), the BBB functional protein expression, and intracellular localization of these proteins.

In this study, we compared these parameters collected from three different *in vitro* BBB models, i.e., trans-well 2D rat model (rat endothelial cells, pericytes, and astrocytes), transwell 2D human cell line model (human immortalized endothelial cells, pericytes, and astrocytes) and 3D human cell line model in which immortalized human endothelial cells reproduced the hollow vasculature in the extracellular matrix gel including immortalized human pericytes and astrocytes. We will report the characteristic of each model and will suggest possible application range of them.

P9-02

Small Compound-based Direct Reprogramming Using Large-scale Omics Data

Toru Nakamura[1] q238049t@mail.kyutech.jp
Michio Iwata[1] iwata121@bio.kyutech.ac.jp
Momoko Hamano[1] haman929@bio.kyutech.ac.jp
Ryohei Eguchi[1] q238008r@mail.kyutech.jp
Yoshihiro Yamanishi[1] yamani@bio.kyutech.ac.jp

[1]Department of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan,

Keywords: direct reprogramming, small compounds, pathway, transcriptome, machine learning, iPS cell, Big data

Direct reprogramming is a research field on direct conversion of fully differentiated mature cells into a variety of other cell types while bypassing an intermediate pluripotent state (e.g., iPS cell), and it has been receiving much attentions in regenerative medicine [1]. Direct reprogramming is known to have several advantages over the other reprogramming methods (e.g., iPS cell-mediated differentiation) in terms of high efficacy and safety.

Cell differentiation by direct reprogramming is determined by gene induction and gene knockdown of transcription factors (TFs). Virus-based gene induction of necessary TFs is a standard procedure for direct reprogramming, but there remain a risk of tumorigenesis because of virus. Therefore, the replacement of TFs by small compounds is highly desired in order to avoid the risk of tumorigenesis [2]. A computational approach is expected to promote the direct reprogramming research, but there is no information technology for small compound-based direct reprogramming.

In this study, we develop novel computational methods to predict small compounds that induce direct reprogramming for a variety of human cells. The prediction is performed by integrative analyses of small compound-induced transcriptome data, TF-perturbed transcriptome data, and differentiation-related pathway data.

[1] Sekiya S *et al*., Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors., *Nature* 2011;**475**(7356):390-3.
[2] H. Qin *et al*., Small molecules for reprogramming and transdifferentiation., *Cell Mol. Life Sci.* 2017;**74**:3553-3575.

P9-03

CNN can detect the sensitivity for radioresistance of cancer cells.

Masamitsu Konno[1], Masayasu Toratani[1], **Jun Koseki[1], Hideshi Ishii**[2]
mkonno@cfs.med.osaka-u.ac.jp

1 Graduate School of Medicine, Osaka University, 2-2 Yamadaoka Suita Osaka 565-0871, Japan

Keywords: CNN, Cancer, and Radioresistance

Artificial intelligence (AI) trained with a convolutional neural network (CNN) is a recent technological advancement. Previously, several attempts have been made to train AI using medical images for clinical applications. However, whether AI can distinguish microscopic images of mammalian cells has remained debatable. This study assesses the accuracy of image recognition techniques using the CNN to identify microscopic images. We also attempted to distinguish between mouse and human cells and their radioresistant clones. We used phase-contrast microscopic images of radioresistant clones from two cell lines, mouse squamous cell carcinoma NR-S1, and human cervical carcinoma ME-180. We obtained 10,000 images of each of the parental NR-S1 and ME-180 controls as well as radioresistant clones. We trained the CNN called VGG16 using these images and obtained an accuracy of 96%. Features extracted by the trained CNN were plotted using t-distributed stochastic neighbor embedding, and images of each cell line were well clustered. Overall, these findings suggest the utility of image recognition using AI for predicting minute differences among phase-contrast microscopic images of cancer cells and their radioresistant clones. SIGNIFICANCE: This study demonstrates rapid and accurate identification of radioresistant tumor cells in culture using artifical intelligence; this should have applications in future preclinical cancer research.

[1] Toratani, M., Konno, M. et. al. A Convolutional Neural Network Uses Microscopic Images to Differentiate between Mouse and Human Cell Lines and Their Radioresistant Clones. ***Cancer Research*** 78(23):6703-6707. (2018)

Media analysis of emerging sciences and technologies in Japan- implications for Molecular Robotics

Ryuma Shineha[1] r_shineha@seijo.ac.jp
Mikihito Tanaka[2] steman@waseda.jp

[1] Faculty of Arts and Literatures, Seijo University, 6-1-20 Seijo, Setagaya-ku, Tokyo 157-0066, Japan
[2] Graduate School of Political Science, Waseda University, 1-6-1, Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050, Japan

Keywords: Media Analysis, Stem Cell Science, GMO, Genome editing, Nanotechnology

Emerging Sciences and Technologies such as Stem Cell Science (SCR), GMO, Genome editing, and Nanotechnology has been recognized among the public through media discourses. For example media analysis on SCR in Japan showed the appearance of framing of “national promotion” and peripheralization of “ethical, legal, and social issues (ELSI)” particularly after the establishment of human iPS cell in 2007 [1]. Also on GMO controversies, many media discourses were analyzed [2-5].

To understand of dominant framings of these sciences and technologies will give us many insights to consider the relationship between Molcular Robotics (Molbot) and the society. This is important process to think about RRI of Molbot [6].

Thus, we conducted quantitative analysis and content analysis concerning newspaper articles on SCR, GMO, genome editing, and nanotechnology, over thirty years. We will show the time-lined change of main framings of these technologies.

[1] Ryuma, Shineha., Attention to Stem Cell Research in Japanese Mass Media: Twenty-Year Macrotrends and the Gap between Media Attention and Ethical, Legal, and Social Issues, *East Asian Science, Technology and Society: An International Journal*, 10(3): 229-246, 2016.
[2] Bauer M., Gaskell G. (eds) *Biotechnology-the Making of a Global Controversy.* Cambridge University Press, 2002.
[3] Bauer M., Gaskell G. (eds) *Genomics & Society Legal, Ethical & Social Dimension.* EARTHSCAN, 2006.
[4] Hibino, A., Nagata, M., Biotechnology in the Japanese Media? Comparative analysis of Newspaper Articles on Genetic Engineering between Japan and Europe. *Asian Journal of Social Psychology*, 9: 12-23, 2006
[5] Shineha, R., Hibino, A., Kato, K. Analysis of Japanese Newspaper Articles on Genetic Modification. *Journal of Science Communication*, 2: 1–8, 2008.
[6] Yoshizawa, G. et al., Responsible innovation in molecular robotics in Japan, *Chem-Bio Informatics Journal*, 18: 164-172, 2018.

P9-05

Potential concerns regarding designated ingredient containing food of Food Sanitation Law of Japan

Mitsuo Saito[1,2] mitsuosaito-kkr@umin.ac.jp
Yuzuru Hayashi[1] i_fumi_theory@tba.t-com.ne.jp

[1] Institute for Health Vigilance, 4-25-5, Tanashi, Nishi-Tokyo City, Tokyo 188-0011, Japan
[2] Japan Pharmaceutical Information Center, 2-12-15, Shibuya, Shibuya-ku, Tokyo 150-0002, Japan

Keywords: Designated ingredient containing food, Regulatory process, Health hazard potential, Food Sanitation Law of Japan

[Introduction] The Food Sanitation Law of Japan was amended on 13th June 2018, and new article 8 defined designated ingredient containing food that special caution should be required. It will be effective within two years, and the Ministry of Health and Welfare (MHLW) preparing for the ministerial ordinance to designation. In this presentation, we analyzed the regulatory process for designation including the evaluation process of the health hazard potential.

[Method] The committee materials were obtained from the MHLW and Food Safety Commission (FSC) website.

[Result] The working group meetings that select the candidate of designated materials were closed and no proceedings were opened to public. According to the discussion material of Subcommittee of Newly Developed Food (part of Pharmaceutical and Food Sanitation Council), questioner survey of physical deconditioning during the dietary supplement intake was main evidence of the risk evaluation for certain ingredient.

Some of the authors of the literature were also the member of Subcommittee of Newly Developed Food of the Food Sanitation Committee and Subcommittee of Newly Developed Food of FSC.

[Discussion] According to the Food Safety Basic Law, FSC is independent risk assessment body and overviews risk management bodies such as MHLW. Because of the lack of the specialist of the health food field, creating the evidence, analysis of the risk (FSC) and management of the risk (MHLW) are done by the same person/group.

To avoid the suspicion by the public and keep transparency, it is necessary to prepare the COI management rules in this field.

P9-06

Multi-functional analysis for applicability evaluation of human iPS cell-derived hepatocytes to DILI assays

Shinichiro Horiuchi hori542@nihs.go.jp
Yukie Kuroda y-kuroda@nihs.go.jp
Su-Ryang Kim kim@nihs.go.jp '
Ryuya Fujii r-fujii@nihs.go.jp
Yasunari Kanda kanda@nihs.go.jp
Seiichi Ishida ishida@nihs.go.jp

Division of Pharmacology, National Institute of Health Sciences,
3-25-26 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa, 210-9501, Japan

Keywords: human iPS cell-derived hepatocytes, drug induced liver injury, long culture, bile canaliculi, drug metabolism

[Purpose] Drug Induced Liver Injury (DILI) is the major factor leading to discontinuation of drug development, and it is important to identify DILI risk as early as possible during drug development process. The sustainable test system that can provide stable results is expected in DILI evaluation. However, DILI evaluation system using cryopreserved / primary human hepatocytes which is utilized generally at the early stage of drug development, is difficult to produce sustainably stable results because for the cell supply limitation and the differences among their donors. We have been studying the application of human iPS cell-derived hepatocytes (hiPSC-hep) to DILI assays in order to construct a system which can produce stable results repeatedly. In this study, we evaluated the functions of drug metabolism, fat storage, long-term culture and bile excretion required for DILI evaluation in hiPSC-hep. In addition, we curried out cytotoxicity test for typical compounds using hiPSC-hep.

[Methods] hiPSC-hep from vendor A was used for the experiment. The expressions of major CYPs (*CYP1A1*, *CYP1A2*, *CYP2C9*, *CYP2C19*, *CYP2D6*, *CYP3A4*) was compared with those in human cryopreserved hepatocytes (cryo-hep) by qPCR. Oleic acid was added to the medium and accumulation of lipid droplets in the cell was observed by Oil-Red O staining. Bile canaliculi were observed using fluorogenic substrates of biliary excretion transporters MRP2 or BSEP. In addition, long-term culture was performed using the long-term culture medium from vendor B, and cell morphology, CYP gene expression, and bile canaliculi formation were observed. Furthermore, the cytotoxicity to typical compounds was evaluated by ATP activity and released LDH activity, and was compared with cryo-hep.

[Results and Discussion] The expressions of *CYP1A1*, *CYP2C19*, and *CYP3A4* were at the same level to those of cryo-hep. The fat droplets were increased by addition of oleic acid and the fat storage ability was confirmed. By using the long-term culture medium from vendor B, the cells could be cultured for 28 days without the detachment. In addition, the expressions of CYPs were increased and the bile canaliculi were extended by long-term culture. Changes in ATP activity and released LDH activity by exposure to the compounds used in this study were similar to cryo-hep. These results indicated that hiPSC-hep from vendor A has various functions required for DILI evaluation and are tolerable for the long-term culture. Hereafter, we will collect data in cells of different lots and donors in order to construct a stable evaluation system.

P9-07

Evaluation of Cell Culture Profiling System with HepG2 Cell Culture Using Microphysiological Systems

Ryuya Fujii r-fujii@nihs.go.jp
Shinichiro Horiuchi hori542@nihs.go.jp
Yukie Kuroda y-kuroda@nihs.go.jp
Su-Ryang Kim kim@nihs.go.jp
Yasunari Kanda kanda@nihs.go.jp
Seiichi Ishida ishida@nihs.go.jp

Division of Pharmacology, National Institute of Health Sciences,
3-25-26 Tonomachi, kawasaki-ku, Kawasaki, 210-9501, Japan

Keywords: Microphysiological systems, HepG2, cell culture profiling system

[Background]

In recent years, in order to reduce the cost of evaluating the safety and the pharmacokinetics of compounds in drug development, there is a demand for the development of a technology capable of culturing cells while maintaining high functions *in vitro*. We focused on the cell culture profiling system that can measure 95 kinds of medium components including amino acids at once by LC-MS/MS as one of the methods to evaluate expression of cell functions and to their quality. HepG2 cells, hepatocyte-derived cancer cell, were cultured on a cell culture device, QuasiVivo, one of the commercially available MPS (Microphysiological systems) which mimicks the environment *in vivo*, and the effectiveness of quality control by the cell culture profiling system was evaluated.

[Method]

HepG2 cells were cultured in Dulbecco's Modified Eagle's Medium (4.5 g/L glucose DMEM supplemented with 10% fetal bovine serum and 1% antibiotic and antifungal solution) for 5 days and the culture supernatants were collected over time. After protein removal by addition of acetonitrile to the culture supernatant, LC-MS/MS (LCMS-8050, Shimadzu) was used to quantify the culture components. Cell culture profiling system (Shimadzu) was used for the culture medium profiling.

[Result]

One of the culture components strongly correlated with the cell number was found as a result of measuring the culture supernatant of HepG2 cells over time using the cell culture profiling system. When HepG2 cells were cultured with QuasiVivo (QV500, Kirkstall), and the medium component changes were detected compared with those of the control in static culture. We are continuing the HepG2 cell culture with other MPS and measuring their culture medium components, thus we would like to discuss the effectiveness of this method.

P10-01

Evidence for the utility of human induced pluripotent stem cell-derived neurons in safety pharmacology-fact data indicating the achievement of network activities

Kanako Takahashi[1] ktakahashi@nihs.go.jp
Kaori Chujo[1] chu-773@nihs.go.jp
Yasunari Kanda[2] kanda@nihs.go.jp
Aoi Odawara[3],
Ikuro Suzuki[3], i-suzuki@tohtech.ac.jp
Kaoru Sato[1] kasato@nihs.go.jp

1 Laboratory of Neuropharmacology, Division of Pharmacology, National Institute of Health Sciences, Kanagawa 210-9501, Japan
2 Division of Pharmacology, National Institute of Health Sciences, Kanagawa 210-9501, Japan
3 Department of Electronics, Graduate School of Engineering, Tohoku Institute of Technology, Miyagi, 982-8577, Japan

Keywords: Human induced pluripotent stem cell (hiPSC)-derived neurons, Neural networks, Micro-electrode array (MEA), Synchronized burst firing (SBF), Calcium signaling, Immunocytochemistry

Neural networks consisted of human induced pluripotent stem cell-derived neurons (hiPSCN) have great potentials in drug development. However, little data are available concerning reproducibility of stable hiPSCN networks and their specific characters. In this study, we reproduced hiPSCN neural networks by the specific culture condition (high density culture, long culture period) and characterized the resulting neural networks.

XCell neurons (XCell Science) are commercially available hiPSCN mimicking forebrain neurons, which are comprised of glutamatergic excitatory neurons and GABAergic inhibitory neurons. XCell neurons were seeded at 3.0×10^5 cells/cm^2 and cultured for 2-3 months. We first examined synapse formation immunoytochemically. We used *Scales*, optical clearing methods, owing to the poor transparency of the samples. The co-localization of presynaptic marker synapsin1 and postsynaptic marker PSD95 on MAP2-positive dendrites appeared at DIV 42. The number of synapsin1(+)PSD95(+) clusters was increased after DIV42-63. We also confirmed the neural networks spontaneous electrophysiological activities by micro-electrode array (MEA) systems. Synchronized burst firing (SBF) activity due to synapse transmission within networks was observed from DIV 42 and increased at DIV 56. The application of NMDAR antagonist AP5 and AMPA/kainate antagonist DNQX decreased SFB activities, respectively. Furthermore, we characterized these cells pharmacologically by fura-2 Ca^{2+} imaging. Most cells with neuron-like shape showed the responsiveness ($[Ca^{2+}]_{in}$ increase) to L-glutamate (L-Glu) from 14 DIV. L-Glu-induced $[Ca^{2+}]_{in}$ elevations were suppressed by AP5 and DNQX at 28 DIV, respectively. Spontaneous activity appeared at 14 DIV and thereafter. These activities were blocked by Na^+ channel blocker tetrodotoxin (TTX) and were enhanced by GABAA antagonist picrotoxin (PIC). At 90 DIV, 80.3 % hiPSCN were responsive to PIC. These data suggest that our culture protocol enables the stable reproduction of hiPSCN neural networks and the resulting human *in vitro* networks are promising tools for the drug development.

P10-02

Development of automatic analysis and quality control of mass spectrometry-based metabolome data

Yasuko Aita[1]
ait_18y@tokyo-med.ac.jp

Masahiro Sugimoto[12]
mshrsgmt@tokyo-med.ac.jp

[1] Research and Development Center for Minimally Invasive Therapies Health Promotion and Preemptive Medicine, Tokyo Medical University, Shinjuku, Tokyo 160-8402, Japan
[2] Institute for Advanced Biosciences, Keio University 246-2 Mizukami, Kakuganji, Tsuruoka, Yamagata, 997-0052, Japan

Keywords: Metabolomics, Quality control, AI

Various tools for processing measurement data of mass spectrometry with separation system, such as, capillary electrophoresis and liquid chromatography, has been developed as both academic and commercial software. The processes include data conversion, peak detection, alignment and peak matching among samples, peak interpretation, compound identification, and quantification. Ideally, a pipeline conducting all process providing enough processing quality is preferable while no software fulfill this requirement, so far [1-3].

Each process has various options, which make it difficult for automatic processing and require the optimization of these options based on the given data [4]. In general, users are required to repeatedly optimize the options based on the analyzed results using something evaluation function and visual inspections. However, this would reduce the reproducibility of the analyzed results.

Here, we have developed a novel tool named MasterHands that tries to reproduce the users' manual curation of metabolomic data. As a training data, manually curated and not curated data should be given. This software learns the difference between these data and tries to optimize the options, which yields the high quality analytical results. We evaluated the performance of the developed algorithm based on the peak area integrated with and without automatic option optimizing. The latter showed the peak area close to the ones curated by users. The calculation engine and GUI were implemented in C ++ and Java, respectively. Also, Python API was implemented for being controlled by scripting language.

[1] Sugimoto Masahiro, et al. "Bioinformatics tools for mass spectroscopy-based metabolomic data processing and analysis." *Current bioinformatics* 7.1 (2012): 96-108
[2] Sugimoto Masahiro. "Metabolomic pathway visualization tool outsourcing editing function." *Engineering in Medicine and Biology Society (EMBC)*, 2015 37th Annual International Conference of the IEEE. IEEE, 2015.
[3] Sugimoto Masahiro, et al. "MMMDB: mouse multiple tissue metabolome database." *Nucleic acids research* 40.D1 (2011): D809-D814.
[4] Sugimoto Masahiro, et al. "Differential metabolomics software for capillary electrophoresis -mass spectrometry data analysis." *Metabolomics* 6.1 (2010): 27-41.

P10-03

CASLcDB: Comprehensive Annotation of human SLc transporters DataBase

Hafumi Nishi[1]
hnishi@ecei.tohoku.ac.jp

Yuya Hanazono[1]
hanazono@sb.ecei.tohoku.ac.jp

Kazunori Yamada[1]
kyamada@ecei.tohoku.ac.jp

Hitoshi Yamagata[2]
hitoshi.yamagata@medical.canon

Kengo Kinoshita[1]
kengo@ecei.tohoku.ac.jp

[1] Graduate School of Information Sciences, Tohoku University, 6-3-09 Aramaki aza Aoba, Aoba-ku, Sendai, Miyagi 980-8579 Japan
[2] Advanced Research Laboratory, Canon Medical Systems Corporation, 1385 Shimoishigami, Otawara, Tochigi 324-0036 Japan

Keywords: solute carrier transporter, database, genetic variants

SoLute Carrier (SLC) transporter is a transmembrane protein family which mediates the transportation of various molecules and ions. The size of the family is extremely large (~ 500 proteins) and can be divided into more than 50 subfamilies, some of which are known to be responsible for biologically crucial pathways, such as glucose intake. Additionally, SLC transporters have been recognized as therapeutic targets, as some are implicated in Mendelian diseases and others are related to drug responses. Despite the importance in cell survival and human health, most of SLC proteins and subfamilies are not well-characterized.

Aside from experimental studies of individual subfamilies and proteins, some SLC databases, such as The Transporter Classification Database (TCDB) and iMusta4SLC, try to describe the whole picture of SLCs and provide their functional information. However, these databases mainly focus on biochemical features of transporters, namely substrates and transport modes, not their cellular functions. To gain deeper understanding of SLC transporters, both biological and biochemical information should be integrated into a single source.

Here we present CASLcDB, a comprehensive database of human SLC transporters, which aggregates various biological and biochemical information of SLC transporters. With the database, users can explore structural information, biological networks/pathways and evolutionary features of each SLC protein. Structural models are provided with genetic variations from major variant databases to help identify harmful mutations and frequencies of variants in population, while biological networks and pathways imported from multiple data sources illustrate the functional relations between SLC and non-SLC proteins in cells. Phylogenetic trees across different species demonstrate surprising evolutionary diversity of subfamilies. CASLcDB is available at https://caslc.sb.ecei.tohoku.ac.jp/

[1] César-Razquin, A., Snijder, B., Frappier-Brinton, T., Isserlin, R., Gyimesi, G., Bai, X., Reithmeier, R.A., Hepworth, D., Hediger, M.A., Edwards, A.M., Superti-Furga, G. A Call for Systematic Research on Solute Carriers. *Cell*. 162(3):478-487, 2015
[2] Lin, L., Yee, S.W., Kim, R.B., Giacomini, K.M. SLC transporters as therapeutic targets: emerging opportunities. *Nature Review Drug Discovery*. 14(8):543-546, 2015
[3] Saier, M.H., Reddy, V.S., Tsu, B.V., Ahmed, M.S., Li, C., Moreno-Hagelsieb, G. The Transporter Classification Database (TCDB). *Nucleic Acids Research.* 44(D1):D372-D379, 2016
[4] Higuchi, A., Nonaka, N., Yura, K. iMusta4SLC: Database for the structural property and variations of solute carrier transporters. *Biophysics Physicobiology*. 15:94-103, 2018

P10-04

Approach of the global QSAR modeling for fish acute toxicity and the future issue for improvement.

Koji Jojima[1] jojima-koji@ceri.jp
Yumi Akahori[2] akahori-yumi@ceri.jp
Hideki Miyaura[1] miyaura-hideki@ceri.jp

[1] Chemicals Assessment and Research Center, Chemicals Evaluation and Research Institute, Japan, 1600 Shimotakano, Sugito-machi, Kitakatsushika-gun, Saitama 345-0043, Japan
[2] Chemicals Assessment and Research Center, Chemicals Evaluation and Research Institute, Japan, 1-4-25 Koraku, Bunkyo-ku, Tokyo 112-0004, Japan

Keywords: Fish Toxicity, Quantitative Structure-Activity Relationship, Machine Learning

Quantitative Structure-Activity Relationship (QSAR) models for prediction of fish acute toxicity are one of methods that have been used as screening tools or alternatives to animal testing for ecological risk assessment in TSCA or REACH regulation. However, most of freely available QSAR models let their applicability domain limited for keeping predictability[1]. The goal of this study is to develop the QSAR model for fish acute toxicity with wider applicability domain keeping high predictability. For achieving this goal, we tried to construct the prediction model of fish acute toxicity from chemical descriptors (chemical property, fingerprints, substructure information etc.) using several machine learning methods. The fish 96-hr LC50 values for over 1,000 chemicals (including Class1~4 of Verhaar Classification) were collected from publically available databases and were used to model construction. The R^2 of the constructed model was 0.69 in cross-validation, and 0.65 in external validation. And the analysis of cross-validation result revealed the low predictability for highly reactive chemicals. Chemical reactivity was reported as one of the factors affect the aquatic toxicity; for example, RC50 (GSH) of Michael Acceptors were correlated with their IGC50 for Tetrahymena[2]. We confirmed that LC50 for fish was correlated with RC50 (GSH) than logP in reactive chemicals in the dataset. Therefore, the inclusion of quantitative parameter of chemical reactivity has potential to improve the global QSAR model.

References

[1] Bin-Le LIN, Chie YAMADA, Kikuo YOSHIDA, Situation and Challenge about in silico Technique Development for Predicting Ecotoxicity of Chemical Substances for Risk Tradeoff Analysis, *Japanese Journal of Risk Analysis*, 21(1), 23-31, 2011.
[2] David Asturiol, Andrew Worth, The Use of Chemical Reactivity Assays in Toxicity Prediction, *JRC Scientific and Technical Reports*, 2011. Retrieved July 29, 2019, from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC65567/lbna24870enn.pdf

P10-05

Attempt to Construct a Primitive Metabolic Network in Deep Hydrothermal Vent Environments

Ryo Hamano[1] 191x213x@stu.kobe-u.ac.jp
Yuto Kitajima[2] 1645008t@stu.kobe-u.ac.jp
Kohei Shimamura[1] shimamura@port.kobe-u.ac.jp
Shigenori Tanaka[1] tanaka2@kobe-u.ac.jp

1 Graduate School of System Informatics, Kobe University, 1-1, Rokkodai, Nada, Kobe 657-8501, Japan
2 Department of Computer Science and Systems Engineering, Kobe University

Keywords: Origin of Life, Abiogenesis, Deep Hydrothermal Vent, Metabolic Network, Chemical Reaction Network, Network Expansion Algorithm

In recent years, a variety of scenarios and models associated with origin of life (or abiogenesis) have been proposed. However, most discussions of the origin begin with the assumption that fundamental enzymes and genes for starting metabolism already had existed (e.g. [1]). How a non-biological chemical reaction network was evolved to the current metabolic network by acquiring robustness in reaction processes is thus still unclear.

In an attempt to fill the gap, we construct a primitive metabolic reaction network from non-biological chemical reactions. The constructive procedure is as follows:

1. The database is prepared from the literature by collecting the abiotic chemical reactions occurring in the deep hydrothermal vents as promising places for the birth of life. Most of the information sources are from the literature that reported simulation experiments for the hydrothermal vent environment. In addition, we consider the cases in which the hydrothermal vent environment is not assumed but the experimental conditions are similar to it as well.
2. Materials (such as carbon dioxide, ammonia, hydrogen sulfide, and olivine) which were abundant in the pre-biotic environment, are regarded as starting chemical seeds.
3. From the seeds, one performs *network expansion algorithm* [2] to obtain non-biological chemical reaction network, where the network expansions are iteratively carried out until products are not generated while referring to the chemical reaction database.

Finally, 165 chemical reactions occurred, and a network consisting of 64 chemicals (nodes) was created. By calculating the degree of each molecule contained in the constructed network, we have found that the glycine and formaldehyde are the hub molecules that play an important role in maintaining or growing the network of chemical reactions. As another characteristic, our network showed high modularity. This trend can also be identified in the metabolic networks of current organisms [3]. Conducting clustering for the constructed network resulted in the split of the generated metabolites into four domains. Interestingly, the chemical reactions constituting the four clusters were based on the iron-sulfur world hypothesis, Oro's experiments, a series of amino acid production experiments simulated for the hydrothermal environment, and serpentinization, respectively. Chemical reactions inspired by the different abiogenesis may be mutually compensated and integrated.

We thus believe that the tracking of the network expansion will lead to an understanding of how each characteristic of the network shows changes and it will also lead to an understanding of the entire evolutionary process of non-biological chemical reaction network.

[1] Sojo V., Herschy B., Whicher A., Camprubi E., Lane N., "The Origin of Life in Alkaline Hydrothermal Vents", *Astrobiology*, **16**, 2 (2016).
[2] Goldford J.E., Hartman H., Smith T.F., Segrè D., "Remnants of an Ancient Metabolism without Phosphate", *Cell*, **168**(6),1126 (2017).
[3] Ravasz E., Somera A.L., Mongru D.A., Oltvai Z.N., Barabási A.L., "Hierarchical Organization of Modularity in Metabolic Networks", *Science*, **297**, 1551 (2002).

P10-06

Computational Direct Reprogramming by Integrating Genome, Transcriptome and Epigenome Data

Ryohei Eguchi[1]
q238008r@mail.kyutech.jp

Momoko Hamano[1]
haman929@bio.kyutech.ac.jp

Michio Iwata[1]
iwata121@bio.kyutech.ac.jp

Toru Nakamura[1]
q238049t@mail.kyutech.jp

Shinya Oki[2]
soki@dev.med.kyushu-u.ac.jp

Yoshihiro Yamanishi[1]
yamani@bio.kyutech.ac.jp

[1] Department of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan,
[2] Department of Developmental Biology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

Keywords: direct reprogramming, transcription factors, epigenome, transcriptome, machine learning, iPS cell, Big data

Direct reprogramming is a research field on direct conversion of fully differentiated mature cells into a variety of other cell types while bypassing an intermediate pluripotent state (e.g., iPS cell). Direct reprogramming is known to have several advantages over the other reprogramming methods (e.g., iPS cell-mediated differentiation) in terms of high efficacy and safety, thus, it has been receiving much attentions in regenerative medicine [1].

Cell differentiation by direct reprogramming is determined by a set of transcription factors (TFs), but the experimental determination of the TF set is extremely difficult and costly. In fact, and there remain numerous cell conversions for which the associated TFs have not been identified. A computational approach for determining TFs is expected to promote the direct reprogramming research, but reliable computational methods with the objective have not been established.

In this study, we develop novel computational methods to predict TFs that induce direct reprogramming for a variety of human cells. The prediction is performed by an integrative analysis of genomic data (e.g., enhancers and promoters), transcriptome data (e.g., gene expression profiles on human cells) and epigenome data (e.g., ChIP-seq data) for TFs. We show the usefulness of the proposed methods on several direct cell conversions.

[1] S. Sekiya and A. Suzuki., Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* 2011;**475**(7356):390-3.