P01-01(F-2)

Elucidation of binding interactions between neurotrophin receptor TrkA domain 5 and its binding peptide

Maho Takahashi[1] W205425d@yokohama-cu.ac.jp
Toru Ekimoto[1] ekimoto@yokohama-cu.ac.jp
Rika Suzuki[1] rks0723@yokohama-cu.ac.jp
Tsutomu Yamane[2] Tsutomu.yamane@riken.jp
Hideo Takahashi[1] hidtak@yokohama-cu.ac.jp
Mitsunori Ikeguchi[1,2] ike@yokohama-cu.ac.jp

[1] Yokohama City Univ., 1-7-29 Suehirocho, Tsurumi-ku, Yokohama-shi, Kanagawa, 230-0045, Japan
[2] R-CCS, RIKEN, 1-7-22, Suehirocho, Tsurumi-ku, Yokohama-shi, Kanagawa, 230-0045, Japan

Keywords: Molecular dynamics simulation, Ligand docking, Peptide modeling

The cell surface receptor tyrosine kinase A domain 5 (TrkAd5) is one of the neurotrophin receptors and specifically binds to nerve growth factor (NGF). Developments of NGF inhibitors targeting TrkAd5-NGF interaction are demanded for neurodegenerative diseases. As a candidate of the inhibitor, we discovered the TrkA-binding peptide 1 (TP1) using the phage display method. In this study, to proceed a rational design of the peptide ligands, such as TP1, the molecular recognition mechanism of TP1 was analyzed by a hybrid analysis of NMR method and all-atom molecular dynamics (MD) simulation. According to a chemical shift (CS) perturbation experiment, the TP1-binding site of TrkAd4 was identified. However, the structure of TrkAd5-TP1 complex was not solved, and the interaction mode between them was not understood structurally.

To clarify the binding pose of TP1, the binding mode was searched using conventional MD simulations and the enhanced sampling method, i.e., replica-exchange with solute tempering (REST) [1]. In the binding pose sampled by the conventional MD simulations, only one dominant pose was observed, and however, the binding pose could not explain the cause of the CS change observed in the edge region of the binding surface. In contrast, another new binding pose was sampled by REST, and the interaction pattern of the new pose could explain the CS data at the edge region. Consequently, two binding poses sampled by REST were stable, and their interaction modes were in good agreement with the experimental CS data. In both binding poses, the central part of TP1 was stuck in a hydrophobic hole at the binding surface, and we found that the rigidity of the central part, where the conformation was limited by proline, was important for the TP1 binding.

[1] M. Kamiya and Y. Sugita, Flexible selection of the solute region in replica exchange with solute tempering: Application to protein-folding simulations, *Journal of Chemical Physics,* **2018**, *149*, 072304-1-11.

P01-02(F-2)

QAEmap: Validation of low-resolution protein crystal structures using deep learning

Hiroaki Hata[1] hata-hiroaki@mki.co.jp
Miwa Sato[1] sato-miwa@mki.co.jp
Ikuko Miyaguchi[2] miyaguchi.ikuko@mv.mt-pharma.co.jp
Akiko Kashima[2] kashima.akiko@mv.mt-pharma.co.jp
Biao Ma[3] biao.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[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 650-0047, 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, Deep learning, X-ray crystallography

X-ray crystallography has been dominantly used to elucidate protein structures, which corresponds to ~90% of protein structures registered in the Protein Data Bank (PDB). However, resolutions of the protein structures are not always high enough for detailed analysis of molecular interactions. For drug discovery or simulation studies, the resolution of 2 Å or higher is necessary [1], while only ~40% of the PDB entities have resolutions higher than 2 Å [2].

In order to refine the low-resolution structures, we applied 3D-convolutional neural network for evaluation of protein structures determined by X-ray crystallography [3]. Our method, quality assessment based on an electron density map (QAEmap), estimates how well the structure fits the putative high-resolution structure by calculating a correlation between the coordinate structure to be evaluated and the electron density map of the correct structure. The low-resolution structure can be refined by modifying the structure to maximize the correlation.

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

[1] Ilatovskiy, A. V. & Abagyan, R. Computational Structural Biology for Drug Discovery. in *Structural Biology in Drug Discovery* 347–361 (John Wiley & Sons, Ltd, 2020). doi:10.1002/9781118681121.ch15.
[2] Burley, S. K. *et al.* Protein Data Bank: The single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* **47**, D520–D528 (2019).
[3] Miyaguchi, I. *et al*. QAEmap: A Novel Local Quality Assessment Method for Protein Crystal Structures Using Machine Learning. (2021) doi:10.21203/RS.3.RS-687363/V1.

P01-03(F-2)

Interaction analysis of cyclic artificial ion channel **AmFF**

Mayuko Nakagawa[1] w215421d@yokohama-cu.ac.jp
Tsutomu Yamane[2] tsutomu.yamane@riken.jp
Toru Ekimoto[1] ekimoto@yokohama-cu.ac.jp
Kohei Sato[3] koheisato@bio.titech.ac.jp
Kazushi Kinbara[3] kinbara.k.aa@m.titech.ac.jp
Mitsunori Ikeguchi[1,2] ike@yokohama-cu.ac.jp

[1] Yokohama City Univ., 1-7-29 Suehiro-chou, Tsurumi-ku, Yokohama-shi, Kanagawa,230-0045, Japan
[2] RIKEN, 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama-shi, Kanagawa,230-0045, Japan
[3] Tokyo Inst. of Tech., 4259 Nagatsuda-chou, Midori-ku, Tokyo,226-8503, Japan

Keywords: Quantum chemical calculation, Density Functional Theory, Artificial ion channel

The development of artificial molecular machines mimicking biomolecular functions has been advanced by Kinbara et al. For example, an artificial ion channel **2mer** [1] exhibits agonist/antagonist-dependent ion permeation, and an artificial ionophore **IMA** [2] transports anions across a membrane. To construct a guideline for designing artificial molecular machines rationally, we have analyzed the mechanism of the ion permeation of **2mer** and **IMA** from molecular dynamics simulations and intermolecular interaction analyses. Recently, the cyclic artificial ion channel **AmFF** has been developed by Kinbara et al. According to experimental results, **AmFF** transports cations across the membrane and has potassium selectivity. **AmFF** is a cyclic molecule consisting of a linear "core part" and a flexible PEG part, and both ends of the two core parts are connected by PEG part. The core part consists of three fluorinated aromatic rings connected by triple bonds, and then the core part is rigid. From experimental results on ion permeability, **AmFF** forms dimer conformation, and the pore formed by the dimer is stable because of the steady current flow.

To understand the pore structure of **AmFF** at atomic resolution and the interaction mode with cations and water in the pore, interaction analyses were performed using molecular modeling and quantum mechanics calculations. The dimer of **AmFF** corresponds to the tetramer of the core part. We developed several tetramer models: the four core parts were located symmetrically, and the orientations of the aromatic rings in the core part were set to facing or perpendicular. From structural optimization of each model, a tetramer model showed a stable conformation. Interestingly, the four core parts were not located symmetrically and had a structure that was offset by one aromatic ring. Comparing the stable tetramer model with the crystal structure of a compound similar to the core part, the tetramer model was in good agreement with the crystal structure. Using the tetramer model, the stability of the pore was examined by expanding pore size. The pore was stable even when the pore was expanded to the size of cations and water. In addition, to clarify stable spots of cations and water in the pore, single-point calculations were performed by placing a cation and a water molecule in the characteristic parts of the tetramer model. According to the energy profiles, the location of the cation was energetically stable at the central part of the pore and at the triple bond part near the entrance. In contrast, the energetically stable spots of the water molecule were at the places where the cation was not energetically stable, indicating that the stable spots of the cation and water molecules were complementary. In the future, we would like to analyze the dynamics by conducting QM/MM and developing a suitable force field that reflects the interaction mode.

[1] Muraoka, T.; Noguchi, D.; Kasai, S. R.; Sato, K.; Sasaki, R.; Tabata, V. K.; Ekimoto, T.; Ikeguchi, M.; Kamagata, K.; Hoshino, N.; Noji, H.; Akutagawa, T.; Ichimura, K.; Kinbara, K. A synthetic ion channel with anisotropic ligand response, *Nature Communications*, **2020**, 11(1), 2924.
[2] Mori, M.; Sato, K.; Ekimoto, T.; Okumura, S.; Ikeguchi, M.; Tabata, V. K.; Noji, H.; Kinbara, K. Imidazolinium-based Multiblock Amphiphile as Transmembrane Anion Transporter, *Chemistry – An Asian Journal*, **2021**, 16,147-157.

P01-04(F-2)

Application of visualization of the interfacial electrostatic complementarity (VIINEC) to antibodies targeting PD-1/PD-L1 interaction

Hiroki Ozono[1]
k8340102@kadai.jp

Takeshi Ishikawa[1]
ishi@cb.kagoshima-u.ac.jp

[1] Department of Chemistry, Biotechnology, and Chemical Engineering, Graduate School of Science and Engineering, Kagoshima University, 1-21-40 Korimoto, Kagoshima, Kagoshima 890-0065, Japan

Keywords: VIINEC, Protein–Protein interaction, Fragment molecular orbital method

Analysis of protein-protein interactions (PPI) is an important issue for the development of antibody drugs. Recently, we have developed a method named interfacial electrostatic complementarity (VIINEC) [1,2], which is based on the *ab initio* fragment molecular orbital (FMO) method. In this method, electrostatic complementarity at the protein–protein interface can be visually analyzed using the electron density and electrostatic potential (ESP) obtained from the FMO calculations for the complex. We expect that VIINEC becomes a useful tool for the rational design of antibody drugs.

In this study, we applied VIINEC to seven antibodies that inhibit the interaction between programmed cell death-1 (PD-1) and its ligand (PD-L1), which are known to be key molecules in the immunotherapy of cancer. Two antibodies (nivolumab and pembrolizumab) are targeting PD-1, and five antibodies (atezolizumab, avelumab, durvalumab, KN035, and BMS-936559) are targeting PD-L1. As an illustrative example, ESP maps of PD-1 and nivolumab at the interface of them in Figure 1-a, which clearly shows that a high degree of electrostatic complementarity is made. The electrostatic complementarity can be quantified using R^- value [2]. In this case, the R^- value was 0.9106, meaning that 91.06% of the total electrostatic interaction was attractive. Furthermore, the ΔR^- value [2], which is calculated from the R^- value, provides the degree of the contribution of each amino acid to the electrostatic complementarity. In Figure 1-b, ΔR^- values are represented for amino acid residues located at the interface together with the inter fragment interaction energies (IFIEs). The overall trend of the ΔR^- values was similar to that of the IFIE analysis. The results for the other antibodies will be given in our presentation, demonstrating the potential of VIINEC for rational antibody designs.

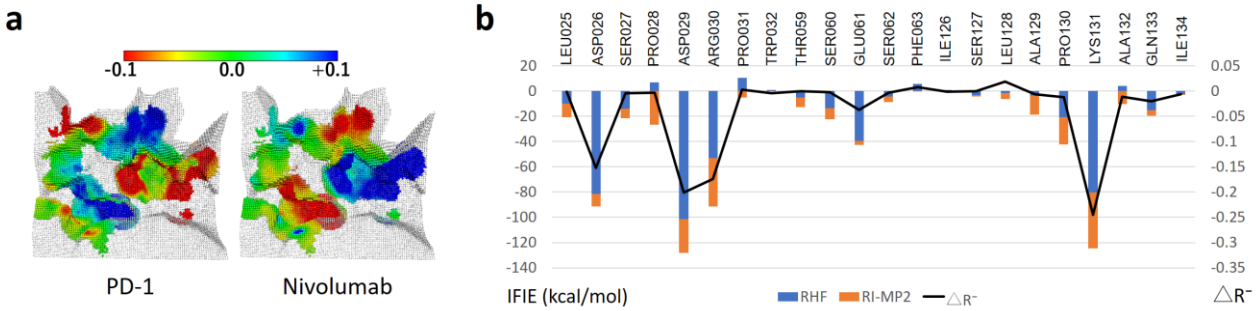


Figure 1: **a)** ESP maps for PD-1/nivolumab, **b)** ΔR^- values together with the IFIEs

[1] T. Ishikawa, A novel method for analysis of the electrostatic complementarity of protein-protein interaction based on fragment molecular orbital method, *Chem. Phys. Lett.*, **761** (2020) 138103.
[2] H. Ozono, T. Ishikawa, Visualization of the interfacial electrostatic complementarity: A method for analysis of protein–protein interaction based on ab initio quantum chemical calculations, *J. Chem. Theory Comput.* in press.

P01-05(F-2)

Advanced methods to predict conformers of cyclic peptides accurately by FRECCA

Sakiko Mori[1] mori.sakiko@fujitsu.com
Kentaro Takai[1] takai.kentaro@fujitsu.com
Seiji Fukui[1] fukui.seiji@fujitsu.com

Azuma Matsuura **Yusuke Miki**[1] **Atsushi Tomonaga**

[1] Fujitsu Ltd, 1-5 Omiya-cho, Saiwai-ku, Kawasaki-shi, Kanagawa, 212-0014, Japan

Keywords: Conformation analysis, Cyclic peptides

Cyclic peptides, one of the middle-sized molecules, represent attracting drug modality that combines the benefits of small molecules, *e.g.*, lower production cost and the potential to target intracellular proteins, and antibodies, *e.g.*, higher selectivity and lower toxicity. However, it is difficult to recapitulate the possible conformers of cyclic peptides because a systematic method to explore their large conformational space has not yet been established. In order to accelerate rational design of cyclic peptides, we are developing the conformation analysis technique for cyclic peptides named FRECCA (Fujitsu Ring Expansion method of Cyclic peptide Conformation Analysis) [1-3].

To confirm the performance of our technique, we analyzed 16 cyclic peptides of 4-11 residues (registered in Protein Data Bank [4] and Cambridge Structural Database [5]) and compared our results with those of Sindhikara *et al.* [6]. The minimum value of RMSD calculated for all heavy atoms to the measured structure was used as a measure of accuracy.

Figure 1 shows the extent to which obtained conformers by each method contain ones close to the measured structures. Our results of the minimum RMSD (labeled “FJ”) for all peptides were less than 2.0 angstroms with the measured structures. Details will be explained in the presentation.

We achieved the accurate conformation analysis for 16 peptides of 4-11 residues. For the actual use of the drug discovery process, we are accelerating our technique.

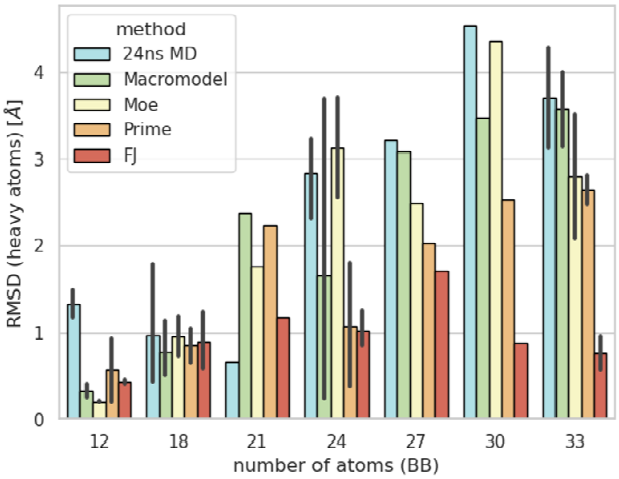


Figure 1: The minimum RMSD (heavy atoms) between obtained conformers by each method and 16 measured structures are summarized by the number of backbone atoms.

[1] Tomonaga, A.; Ueda, A.; Matsuura, A. *254th ACS National Meeting & Exposition* (2017/8)
[2] Tomonaga, A.; Sugiyama, H.; Ueda, A. WO2018/154789A1.
[3] Tomonaga, A.; Sugiyama, H.; Ueda, A. JP2019/179304A.
[4] https://www.rcsb.org/
[5] https://www.ccdc.cam.ac.uk/
[6] Sindhikara, D.; Spronk, S. A.; Day, T.; *et al.*, *J. Chem. Inf. Model.*, **2017**, *57(8)*, 1881-1894.

P01-06(F-2)

Comparison of molecular dynamics of cyclosporin A and cyclosporin E

Akari Ito[1] W205406f@yokohama-cu.ac.jp
Toru Ekimoto[1] ekimoto@yokohama-cu.ac.jp
Tsutomu Yamane[2] tsutomu.yamane@riken.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] Center for Computational Science, RIKEN, Yokohama 230-0045, Japan.

Keywords: Molecular dynamics, Cyclosporin, gREST

Cyclosporin A (CsA) is a naturally derived cyclic peptide used as an immunosuppressant drug. CsA has high cell-membrane permeability despite its large molecular size. In addition, compared to cyclosporin E (CsE), which is a metabolite of CsA, CsA has 10 times higher membrane permeability than CsE, although the only difference between them is in one amino acid: N-methylated VAL in CsA just changes to VAL in CsE [1]. To understand the mechanism of the membrane permeation of CsA and the difference in the membrane permeability between CsA and CsE, structural samplings during the membrane permeation event were executed using all-atom conventional molecular dynamics (cMD) simulations and the generalized replica-exchange with solute tempering (gREST) simulations [2]. In cMD simulations, several initial structures of CsA or CsE embedded in the membrane-water systems were used. Although no membrane-permeation events occurred during the timescale of several micro-seconds, we found that CsA and CsE tended to be located at the regions of lipid tail groups. In addition to the open and closed forms that observed in crystal structure, intermediate forms were also sampled during the cMD simulations.

To further sample the structural distribution of CsA, gREST simulations were performed with the center of mass of CsA fixed in water, at the membrane-water interface, and in the membrane near the lipid tail groups, which are characteristic locations found in the cMD results. In each location, the variety of conformations that CsA adopted was different, and the popular conformations appeared more frequently were also different. In water, the conformation of CsA showed a wide range of the structural distribution from open to closed-like forms. By contrast, in the interface and membrane, the variety of the conformation was restricted to the closed-like form.

[1] Ahlbach L C, Lexa W K, Bockus T A, Chen V, Crews P, Jacobson P M, Lokey S R, Future Med. Chem. 7, 2121-30 (2015).
[2] Motoshi Kamiya, Yuji Sugita, J Chem Phys, 149, 072304 (2018).

P01-07(F-2)

Dynamics of Orexin2 Receptor and G-protein Complex with Molecular Dynamics Simulations

Shun Yokoi[1] ce207021@meiji.ac.jp
Ayori Mitsutake[1] ayori@meiji.ac.jp

[1] Department of Physics, School of Science and Technology, Meiji University, 1-1-1, Higashi-Mita, Tama-ku, Kawasaki, Kanagawa, 214-8571, Japan

Keywords: GPCRs, OX2R, MD simulation, Complex system, Dynamics

Orexin2 receptor (OX2R) is classified as a class A G-protein-coupled receptor (GPCR) and belongs to the group of orexinergic systems that are composed of two peptide ligands (i.e., orexin-A and -B) and two GPCRs (i.e., OX1R and OX2R).[1] The OX2R is involved in the regulation of feeding behavior and sleep-wake rhythm to give some examples.[2-3] Such kinds of neurological processes are caused by GPCR activation. Through the activation of GPCRs by ligand binding, GPCRs change their conformation and then downstream signal transduction is carried out via a trimeric protein called G-protein that consists of three subunits (Gα, Gβ, and Gγ) and via G-protein-independent pathways through GPCR kinase-mediated phosphorylation and arrestin coupling.[4] Although the structure of the OX2R has been clarified by structural analysis,[5-6] the atomic-level mechanisms of GPCR activation, G-protein activation, and biased signaling remain unknown.

Here, we performed and analyzed several microsecond-scale molecular dynamics (MD) simulations of OX2R using AMBER MD software. We performed not only simulations of OX2R, but also modeled the complex of OX2R and G-protein and performed simulations of the complex. In the poster, we first show the results of the MD simulations and investigate the dynamics of OX2R and G-protein complex. Then, we discuss implications for the activation mechanism of OX2R and for the binding mechanism of G-protein.

[1] Sakurai, T.; *et al.*, Orexins and orexin receptors: A family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior, *Cell*, **1998**, *92*, 573–585.

[2] Leonard, C. S.; Kukkonen J. P. Orexin/hypocretin receptor signaling: a functional perspective, *Br. J. Pharmacol.*, **2014**, *171*, 294–313.

[3] Sakurai, T. The neural circuit of orexin (hypocretin): Maintaining sleep and wakefulness, *Nat. Rev. Neurosci.*, **2007**, *8*, 171–181.

[4] Hilger, D.; *et al.*, Structure and dynamics of GPCR signaling complexes, *Nat. Struc. Mol. Biol.*, **2018**, *25*, 4–12.

[5] Suno, R.; *et al.*, Crystal structures of human orexin 2 receptor bound to the subtype-selective antagonist EMPA, *Structure*, **2018**, *26*, 7–19.

[6] Chuan, H.; *et al.*, Structures of active-state orexin receptor2 rationalize peptide and small–molecule agonist recognition and receptor activation, *Nat. Commun.*, **2021**, *12*, 815.

P01-08(F-2)

Transport simulation of human LAT1-CD98hc complex

Natsumi Yoshida[1]
w215436e@yokohama-cu.ac.jp

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

Tsutomu Yamane[2]
tsutomu.yamane@riken.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] Center for Computational Science, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan.

Keywords: Molecular dynamics, Membrane protein, SLC transporter, Amino acid transporter, Conformational change, Substrate transport

The L-type amino acid transporter 1 (LAT1) has broad substrate recognitions and transports large neutral amino acids and their analogs across the membrane [1]. Because LAT1 is overexpressed in tumor cells and contributes to the supply of essential amino acids [2], the development of inhibitors is required for the treatment of cancers. Since LAT1 is also expressed in the blood-brain-barrier [3], understanding the transport mechanism will lead to the development of new drugs. Therefore, LAT1 is an important pharmaceutical target. LAT1 forms a heterodimer with CD98 heavy chain (CD98hc), and carries out a substrate transport. Upon the substrate transport, LAT1 is considered to undergo conformational changes between the outward-facing (OF) and the inward-facing (IF) states [4]. Recently, both the OF and IF conformations have been solved by cryo-EM [5-7], and however, the detailed mechanism underlying the transport coupled with the conformational changes has not been understood well.

To elucidate the transport mechanism coupled with the conformational changes, substrate transport simulations were performed using all-atom molecular dynamics (MD) simulations. We analyzed the transport mechanism coupled with conformational changes of LAT1. During the simulations, the substrate spontaneously moved to the intracellular side coupled with the conformational change.

[1] Uchino, H.; et al., Transport of Amino Acid-Related Compounds Mediated by L-Type Amino Acid Transporter 1 (LAT1): Insights Into the Mechanisms of Substrate Recognition. *Mol. Pharmacol.* **2002**, *61* (4), 729–737.
[2] Kanai, Y.; et al., Expression Cloning and Characterization of a Transporter for Large Neutral Amino Acids Activated by the Heavy Chain of 4F2 Antigen (CD98). *J. Biol. Chem.* **1998**, *273* (37), 23629–23632.
[3] Kageyama, T.; et al., The 4F2hc/LAT1 Complex Transports L-DOPA across the Blood-Brain Barrier. *Brain Res.* **2000**, *879* (1–2), 115–121.
[4] Drew, D.; Boudker, O. Shared Molecular Mechanisms of Membrane Transporters. *Annu. Rev. Biochem.* **2016**, *85*, 543–572.
[5] Lee, Y.; et al., Cryo-EM Structure of the Human L-Type Amino Acid Transporter 1 in Complex with Glycoprotein CD98hc. *Nat. Struct. Mol. Biol.* **2019**, *26* (6), 510–517.
[6] Yan, R.; Zhao, X.; Lei, J.; Zhou, Q. Structure of the Human LAT1–4F2hc Heteromeric Amino Acid Transporter Complex. *Nature* **2019**, *568* (7750), 127–130.
[7] Yan, R.; et al., Mechanism of Substrate Transport and Inhibition of the Human LAT1-4F2hc Amino Acid Transporter. *Cell Discov.* **2021,** *7* (1), 0–7.

P01-09(F-2)

Activation Pathway of IL-2-inducible T-cell kinase Explored by Tree-Search Molecular Dynamics

Yukina Nakai[1] i180482b@yokohama-cu.ac.jp
Toru Ekimoto[2] ekimoto@yokohama-cu.ac.jp
Tsutomu Yamane[3] tsutomu.yamane@riken.jp
Naoki Ogawa[2] w195501f@yokohama-cu.ac.jp
Kei Terayama[2] terayama@yokohama-cu.ac.jp
Mitsunori Ikeguchi[2,3] ike@yokohama-cu.ac.jp

[1] Yokohama City University, 22-2 Seto, Kanazawa-ku, Yokohama 236-0027, Japan.
[2] Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan.
[3] Center for Computational Science, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan.

Keywords: Molecular Dynamics, Conformational Change, Protein Kinase

Interleukin-2 (IL-2) -inducible T cell kinase (ITK) is a non-receptor tyrosine kinase that plays an essential role in T cell receptor signaling [1]. Because ITK also mediates the secretion of Th2 cytokines, including IL-4, IL-5, and IL-13, and regulates the development of effective Th2 responses during allergic asthma and parasitic infections, ITK is considered as a potential therapeutic target for the treatment of the Th2-mediated inflammatory diseases [1]. Several crystal structures of the kinase domain of human ITK in the inactive and active states have been solved so far [1], and show the structural features in the two endpoint states: the unfolded/folded conformations of A-loop and the inward/outward positions of αC-helix. However, the crystal structure alone does not provide the details of the activation mechanism coupled with the conformational changes between the endpoint states. In particular, the structural information of the intermediate state is necessary to understand the activation mechanism structurally.

To understand the activation mechanism of ITK, we studied conformational changes between the inactive and active states using all-atom molecular dynamics (MD) simulations. Because the time scale of the conformational change underlying the activation event exceeds the conventional MD simulations, an enhanced pathway-search method based on MD and a tree-search algorithm, termed the Tree-Search Molecular Dynamics simulation (TS-MD) [2] was employed for the pathway search during the activation event. We constructed evaluation functions that can describe the structural changes of αC-helix and A-loop, respectively, and conducted multiple rounds of TS-MD. First, we conducted TS-MD from the active conformation using an evaluation function that represents only the conformational change of αC-helix, and succeeded in obtaining the intermediate structure and the pathway to the intermediate structure, which has the structural features reported for other kinases. Second, TS-MD from the intermediate structure was conducted using an evaluation function that represents only the conformational change of A-loop. As a result, we succeeded in obtaining a pathway leading to a structure close to the inactive conformation. In conclusion, using TS-MD, the conformational changes during the activation event of ITK were successfully sampled.

[1] Kutach AK et al. Crystal structures of IL-2-inducible T cell kinase complexed with inhibitors: insights into rational drug design and activity regulation: crystal structures of ITK-inhibitor complexes. *Chemical Biology & Drug Design.* **2010**, 76(2), 154-163.
[2] Shin K et al. Enhancing Biomolecular Sampling with Reinforcement Learning: A Tree Search Molecular Dynamics Simulation Method. *ACS Omega*. **2019**, 4(9), 13853-13862.

P01-10(F-3)

Molecular dynamics simulations of $GABA_A$ receptor and general anesthetics

Shuya Nakata[1]
nakata@eniac.scitec.kobe-u.ac.jp

Yoshiharu Mori[1]
ymori@landscape.kobe-u.ac.jp

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

[1] Graduate School of System Informatics, Kobe University, Japan

Keywords: General anesthetics, Ion channels, Molecular simulation, Molecular dynamics

$GABA_A$ receptor is a pentameric ligand-gated ion channel that is a target molecule for several general anesthetics. Recently, the 3D structures of the $GABA_A$ receptor with the general anesthetics propofol and etomidate bound to the subunit interfaces at the transmembrane domain were determined by cryo-electron microscopy. (PDB ID: 6x3t, 6x3v) [1]. Comparison of the obtained structures provided valuable insights into the molecular mechanisms of general anesthetics [2], while the full picture of the mechanism of action is still unclear.

To investigate the effects of the anesthetic-bindings on the structural change of $GABA_A$ receptor during activation, we carried out 100 ns all-atom molecular simulations for a $GABA_A$ receptor in the POPC bilayer and explicit waters using the pre-open structure (PDB ID: 6x40) as initial structure. We investigated the following three conditions: (i) no anesthetic bound (ii) propofol bound (iii) etomidate bound. Anesthetics were inserted according to their binding structures, and multiple simulations (at least five for each condition) were performed by changing the initial compositions of the lipids. The CHARMM36m force field, the CGenFF force field, and the TIP3P model were used for protein, ligand, and water, respectively.

Computing the distribution of the pore radius using HOLE [3], we found that pore radius is increased around the 9' leucine residue which is called activation gate when anesthetics are bound. This is consistent with the observations in the anesthetic-bound structures obtained by the cryo-electron microscopy and may be related to the mechanism of anesthetic action. Subsequent analysis of the trajectories suggested that the binding of the anesthetics at the subunit interfaces affects the relative positioning of the subunits and contributes to the expansion of the pore. The details of the methods and the results will be shown in the presentation.

[1] Kim, J.J.; Gharpure, A.; Teng, J. et al. Shared structural mechanisms of general anaesthetics and benzodiazepines, *Nature*, **2020**, *585*, 303-308.

[2] Kim, J.J.; Ryan E. Hibbs. Direct Structural Insights into GABAA Receptor Pharmacology, *Trends Biochem Sci*, **2021**, *46*, 502–517.

[3] Oliver S. Smart et al. HOLE: A program for the analysis of the pore dimensions of ion channel structural models, *J Mol Graph*, **1996**, *14*, 354–360.

P01-11(F-3)

Free energy analysis of co-solvent effect on insulin dimer dissociation

Simon Hikiri
simon.hkr3i@cheng.es.osaka-u.ac.jp

Nobuyuki Matubayasi
nobuyuki@cheng.es.osaka-u.ac.jp

Division of Chemical Engineering Department of Materials Engineering Science Graduate School of Engineering Science Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560-8531 Japan

Keywords: Co-Solvent Effect, Theory of Liquid, Free Energy Calculation, Molecular Dynamics

We analyze the co-solvent effect on insulin dissociation from the viewpoint of free energy. Treating insulin as an all-atom model, we combine molecular dynamics simulation and energy representation theory [1] to analyze the free energy of the co-solvent effect. Correlation analysis of the interacting components of the free energy (van der Waals interaction, electrostatic interaction, etc.) are performed to identify the dominant components of the dissociation control.

The solvation free energy changes of insulin from dimer to monomer in pure water and in co-solvent solutions (a binary mixture of water molecules and co-solvent molecules) are calculated. Comparison of solvation free energy changes of insulin dissociation in pure water and in co-solvent solution are reveal the co-solvent effects and their physical origin. We focus on the acquisition of co-solvent effects from a fundamental point of view, and dealt with co-solvent molecules with simple structures. Based on the water molecule structure, we considered methanol with one of the hydrogens changed to a methyl group, ethanol with an ethyl group, phenol with a phenyl group and ethylene glycol with a hydroxyethyl group which is a unit of polyethylene glycol, as the co-solvent molecules to be calculated. And urea that actually exists in vivo is also considered as a co-solvent molecule. Detailed results will be reported on the day.

[1] Sakuraba, S.; Matubayasi N. Ermod: Fast and versatile computation software for solvation free energy with approximate theory of solutions *J. Comput. Chem.*, 2014, 35, 1592–1608.

P01-12(F-3)

MDContactCom: A tool to identify differences of protein molecular dynamics from two MD simulation trajectories in terms of residue-residue contacts

Chie Motono[1,2] c-motono@aist.go.jp
Shunsuke Yanagida[3] yanagita-shunsuke@mki.co.jp
Miwa Sato[3] sato-miwa@mki.co.jp
Takatsugu Hirokawa[1,4,5] t-hirokawa@md.tsukuba.ac.jp

[1] Cellular and Molecular Biotechnology Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tokyo 135-0064 Japan. [2] Computational Bio Big-Data Open Innovation Laboratory (CBBD-OIL), AIST, Waseda University, Tokyo 169-0072, Japan. [3] Bioscience & Healthcare Engineering Div.,Mitsui Knowledge Industry Co., Ltd, Tokyo 135-0064 Japan. [4] Division of Biomedical Science, Faculty of Medicine, University of Tsukuba, Ibaraki 305-0064, Japan. [5] Transborder Medical Research Center, Faculty of Medicine, University of Tsukuba, Ibaraki 305-8575, Japan.

Keywords: Protein, MD simulation, contact

Comparing results from multiple MD simulations performed under different conditions is important during the initial stages of analysis. We propose a tool called MD Contact Comparison (MDContactCom) that compares residue-residue contact fluctuations of two MD trajectories, quantifies the differences, identifies sites that exhibit large differences, and visualizes those sites on the protein structure [1]. Using this method, it is possible to identify sites affected by varying simulation conditions and reveal the path of propagation of the effect even when differences between the 3D structure of the molecule and the fluctuation RMSF of each residue is unclear. MDContactCom can monitor differences in complex protein dynamics between two MD trajectories and identify candidate sites to be analyzed in more detail. The software supports trajectories from five major MD engines (Amber, CHRMM, Desmond, GROMACS, and NAMD) as well as multiple pdb format. As such, MDContactCom is a versatile software package for analyzing most MD simulations.

We applied MDContactCom to analyze the MD trajectories of Cyclophilin A and its allosteric mutant [2,3]. MDContactCom detected residues in the pathways where the mutation effects propagate

MDContactCom is freely available for download on GitLab. The software is implemented in Python3. https://gitlab.com/chiemotono/mdcontactcom

[1] Motono, C. et al. MDContactCom: a tool to identify differences of protein molecular dynamics from two MD simulation trajectories in terms of interresidue contacts, *Bioinformatics*, **2021**, btab538. https://doi.org/10.1093/bioinformatics/btab538.
[2] Doshi,U. et al. Dynamical network of residue–residue contacts reveals coupled allosteric effects in recognition, catalysis, and mutation, *Proc. Natl. Acad. Sci.*, **2016**, *113*, 4735–4740.
[3] Holliday,M.J. et al. Networks of Dynamic Allostery Regulate Enzyme Function, *Structure.*, **2017**, *25*, 276–286.

P01-13(F-3)

Molecular Dynamics Simulations of DNA Exonuclease

Hiroki Otaki[1]
otaki@nagasaki-u.ac.jp

[1] Center for Bioinformatics and Molecular Medicine, Graduate School of Biomedical Sciences, Nagasaki University, 1-14 Bunkyo-machi, Nagasaki, 852-8521, Japan

Keywords: Molecular dynamics simulation, Network analysis, DNA exonuclease

The three prime repair exonuclease 1 (*TREX1*) gene encodes a nuclear protein (called TREX1) with 3′ to 5′ DNA-specific exonuclease activity. Mutations in *TREX1* gene are associated with a wide spectrum of disorders, including systemic lupus erythematosus (SLE), Aicardi–Goutières syndrome (AGS), familial chilblain lupus (FCL), and retinal vasculopathy with cerebral leukodystrophy (RVCL) [1]. Despite a lot of reports of mutations with the diseases, the molecular mechanism of how these mutations affect enzyme activity of TREX1 is still unclear.

In order to clarify the effect of mutations on the TREX1 system, we have modeled the complexes of TREX1 mutants with double-stranded DNA (dsDNA) in the presence of two Mg^{2+} ions and conducted molecular dynamics (MD) simulations. By using a network graph to visualize interaction network in the TREX1 system, we have found the differences in the interaction around the active site between the mutants and wild type (WT). For example, in the D130N mutant, the interaction between the magnesium ions and 3′-end of dsDNA becomes stronger than WT. This suggests that the mutant is unable to successfully release hydrolyzed nucleic acids, which may result in the accumulation of nucleic acids and the activation of the innate immune response [2]. The detailed results including other mutants will be discussed in this presentation.

This research is partially supported by Initiative on Promotion of Supercomputing for Young or Women Researchers, Information Technology Center, The University of Tokyo. This work was supported by JSPS KAKENHI Grant Number 19K16058.

[1] Rice, G. I.; Rodero M. P.; Crow Y. J. Human disease phenotypes associated with mutations in *TREX1*, *J. Clin. Immunol.*, **2015**, *35*, 235–243.
[2] Endo, Y.; Koga, T.; Otaki, H.; Furukawa K.; Kawakami A. Systemic lupus erythematosus overlapping dermatomyositis owing to a heterozygous *TREX1* Asp130Asn missense mutation, *Clin. Immunol.*, **2021**, *227*, 108732.

P02-01(F-3)

Insight into Allosteric ERK2 Inhibitors by using Fragment Molecular Orbital Method

Hajime Sugiyama[1] sugiyama.hajime.ma@m-chemical.co.jp
Mayu Yoshida[2] yoshida19@b.s.osakafu-u.ac.jp
Takayoshi Kinoshita[2] kinotk@b.s.osakafu-u.ac.jp

[1] Mitsubishi Chemical Corporation, 1000 Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-8502, Japan
[2] Osaka Pref. Univ., 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan

Keywords: ERK2, Allosteric site, Fragment molecular orbital method

Extracellular signal-regulated kinase 2 (ERK2) belongs to mitogen-activated protein kinase family, which regulates cell growth, differentiation, and a variety of other cellular responses by phosphorylating several substrates including signal transducer and activator of transcription 3 (STAT3), which has essential roles for normal glucose homeostasis [1].

The authors focused on ERK2/STAT3 pathway as a novel diabetes target as a result of text mining technology. Their previous studies confirmed that a peptide designed to inhibit ERK2 phosphorylation to STAT3 showed physiological effects *in vitro* and *vivo* [2]. A subsequent *in silico* screening rendered low-molecular-weight compounds with inhibitory activity [3].

We here investigated interactions between ERK2 and the *in silico* hit compounds. Fragment molecular orbital method and Pair interaction energy decomposition analysis extracted key interactions between the protein and the ligand from X-ray crystallography structures. Furthermore, this calculation conferred an energetically favorable ligand conformation.

These results would greatly enhance to develop highly potent and selective ERK2 inhibitors.

[1] H. Inoue, et al., *Nature Medicine* **10**: 168-174 (2004).
[2] T. Kinoshita, et al., *Chem. Biol. Drug Des.* **78**: 471- 476 (2011).
[3] T. Kinoshita, et al., *Bioorg. Med. Chem. Lett*. **26**: 955- 958 (2016).

P02-02(F-3)

Dynamic Residue Interaction Network Analysis of Secondary Mutations in Protease that Promote Drug Resistance in HIV-1

Ayaka Ojima s17a6025jk@s.chibakoudai.jp
Mohini Yadav s1879602ec@s.chibakoudai.jp
Norifumi Yamamoto norifumi.yamamoto@p.chibakoudai.jp

Department of Applied Chemistry, Faculty of Engineering,
Chiba Institute of Technology, Narashino, Japan

Keywords: Dynamic Residue Interaction Network, Drug Resistance

The human immunodeficiency virus (HIV) is the pathogen of the Acquired Immune Deficiency Syndrome (AIDS). AIDS has become a disease that can be controlled in the long term by anti-HIV drugs. However, there are serious concerns about the emergence of viral mutants that are resistant to anti-HIV drugs. Some amino acid mutations in HIV-1 protease promotes development of drug resistance caused by primary mutations, even without directly affecting drug efficacy against anti-HIV drugs. These mutations are referred to as "secondary mutations". In this study, we investigated the dynamic correlation between the drug binding site and its secondary mutation site in HIV-1 protease using dynamic residue interaction network (dRIN) analysis [1] based on molecular dynamics simulations.

[1] Yadav, M., Igarashi, M., Yamamoto, N.; Dynamic residue interaction network analysis of the oseltamivir binding site of N1 neuraminidase and its H274Y mutation site conferring drug resistance in influenza A virus, *PeerJ*, **2021**, *9*, e11552.

P02-03(F-3)

Dynamic Residue Interaction Network Analysis of Primary Mutations in Protease that Promote Drug Resistance in HIV-1

Ryoga Miyawaki s17a6107yj@s.chibakoudai.jp
Mohini Yadav s1879602ec@s.chibakoudai.jp
Norifumi Yamamoto norifumi.yamamoto@p.chibakoudai.jp

Department of Applied Chemistry, Faculty of Engineering,
Chiba Institute of Technology, Narashino, Japan

Keywords: Dynamic Residue Interaction Network, Drug Resistance

The human immunodeficiency virus (HIV) is the pathogen of the Acquired Immune Deficiency Syndrome (AIDS). AIDS has become a disease that can be controlled in the long term by anti-HIV drugs. However, due to prolonged treatment, there are serious concerns about the emergence of viral mutants that are resistant to anti-HIV drugs. Some amino acid mutations in the HIV-1 protease are known to directly and substantially reduce drug efficacy against anti-HIV drugs. These mutations are referred to as "primary mutations". In this study, we investigated the dynamic correlation between the drug binding site and its primary mutation site in HIV-1 protease using dynamic residue interaction network (dRIN) analysis [1] based on molecular dynamics simulations.

[1] Yadav, M., Igarashi, M., Yamamoto, N.; Dynamic residue interaction network analysis of the oseltamivir binding site of N1 neuraminidase and its H274Y mutation site conferring drug resistance in influenza A virus, *PeerJ*, **2021**, *9*, e11552.

P02-04(F-3)

Study of the Binding Mechanism of Inhibitor to InterLeukin-2 by Free Energy Calculations

Noritaka Inoue[1]
noritaka.inoue@schrodinger.com

[1] Schrödinger KK, Marunouchi Trust Tower North build. 13F, 1-8-1 Marunouchi, Chiyoda-ku, Tokyo,100-0005, Japan

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

Over the last 10 years, there has been tremendous progress in enhanced sampling algorithms for binding free energy calculations such as FEP/REST [1] and force field development [2] together with the advent of GPU architectures. In addition, these have been compiled with user-friendly GUI as a product FEP+ released by Schrödinger Inc. [3], which enables even non-experts to perform binding free energy calculations. As a result, highly accurate prediction of binding affinity between the ligand and the receptor has become a reality [4] and nowadays many pharmaceutical companies have adopted these methods in their drug designs.

The full physical mechanisms of the process from unbound state to bound state such as desolvation penalty or ligand strain are incorporated in the binding free energy calculation, so that it provides not only the prediction values but also the chance to interpret the binding mechanisms. The latter is also beneficial especially for difficult drug target proteins.

In this study we have applied FEP+ to InterLeukin-2(IL-2) complexed with its inhibitor to predict the binding affinity and detect the important residues that account for the ligand selectivity. IL-2 is an immunoregulatory cytokine and known as a difficult drug target because the inhibitor binds on the protein-protein interaction interface which has the inherent plasticity [5]. Our study has demonstrated that the FEP+ simulation can successively predict the binding affinity of IL-2 inhibitors (R^2 = 0.94, MUE = 0.82) and is also helpful to elucidate the binding mechanism.

[1] Wang, L.; *et al.*, Protein-Ligand Binding Free Energy Calculations with FEP, *Methods Mol Biol.*, **2019**, 201–231.

[2] Lu, C.; *et al.*, OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space, *J. Chem. Theory Comput.*, **2021**, *17*, 7, 4291–4300.

[3] https://www.schrodinger.com/products/fep

[4] Wang, L.; *et al.*, Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by way of a Modern Free Energy Calculation Protocol and Force Field, *J. Am. Chem. Soc.*, **2015**, *137*, 2695–2703.

[5] Thanos, C. D.; *et al.*, Potent Small-Molecule Binding to a Dynamic Hot Spot on IL-2, *J. Am. Chem. Soc.*, **2003**, *125*, 15280–15281.

P02-05(F-3)

Dynamic FMO analysis for RNA sequence specificity in inhibitor recognition of translation initiation factor complex

Yuma Handa[1] **Yusuke Kawashima**[1] **Takayuki Furuishi**[1] **Etsuo Yonemochi**[1] **Takuhiro Ito**[2] **Hironori Saito**[3,4] **Shintaro Iwasaki**[3,4] **Chiduru Watanabe**[5] **Teruki Honma**[5] **Kaori Fukuzawa**[1,6]
d2002@hoshi.ac.jp y-kawashima@hoshi.ac.jp t-furuishi@hoshi.ac.jp e-yonemochi@hoshi.ac.jp takuhiro.ito@riken.jp hironori.saito@riken.jp shintaro.iwasaki@riken.jp chiduru.watanabe@riken.jp honma.teruki@riken.jp 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] Laboratory for Translation Structural Biology, RIKEN Center for Biosystems Dynamics Research (BDR) W215 West Research Building,1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,Japan
[3] RIKEN Cluster for Pioneering Research (CPR) S205 Bioscience Bldg. 2-1, Hirosawa Wako,Saitama 351-0198, Japan
[4] Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba 277-8561, Japan
[5] Laboratory for Structure-Based Molecular Design, RIKEN Center for Biosystems Dynamics Research, Tsurumi-ku, Yokohama 230-0045, Japan.
[6] Tohoku University Graduate School 6-6-11 Aoba, Aramaki, Aoba-ku, Sendai-shi, Miyagi-ken 980-8579

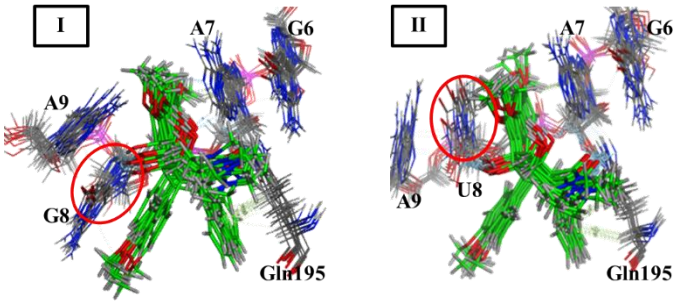
Keywords: eIF4A, RNA, Rocaglamide A, Fragment molecular orbital method, Molecular dynamics

Rocaglamide A (RocA), a natural compound found in *Aglaia* genus plants, has potent anticancer activity and acts as a translation inhibitor in the RNA-binding cavity of the initiation factor eIF4A. Although it is known to exhibit purine sequence specificity in RNA binding [1], the detailed molecular mechanism remains unclear. In this study, we analyzed the molecular recognition mechanism of elF4A by dynamic FMO calculations combined with classical MD and FMO methods. Using the quaternary complex structure of elF4A, RNA, AMP-PNP, and RocA (PDBID:5zc9) as the initial structure, we prepared wild-type and four mutant structures (G6U, A7U, G8U, and A9U) and performed MD calculations using AMBER14:EHT force fields of 50 ns for each complex. From each MD trajectory, 10 structures were extracted every 3 ns, and FMO calculations at the MP2/6-31G* level were performed.

The interaction energy analysis of RocA with the surrounding residues showed that the wild type was the most stable, with π/π interactions mainly with A7 and G8, and a hydrogen bond with Gln195, which were always maintained (Fig. I). In the G8U mutation (Fig. Ⅱ), U8 was inverted and formed an interaction with different sites of RocA while maintaining the stacking structure with A9; however, the interaction was weaker than the wild type. On the other hand, in the G6U, A7U, and A9U mutations, the fluctuation of A7 was increased, and the hydrogen bond and CH/π interaction between RocA and Gln195 and A7 could not be maintained, resulting in weaker interaction.

The FMO calculations were partially performed using the Fugaku supercomputer (HPCI project ID: hp210130). This study was partially supported by the AMED-BINDS project (grant number JP21am0101113).

[1] Shintaro. I. *et al.,* The Translation Inhibitor Rocaglamide Targets a Bimolecular Cavity between eIF4A and Polypurine RNA *Mol Cell. 2019;73(4):738-748.e9.*



Main Interaction around RocA (I:Wild Type, II:G8U mutant)

P03-01(F-4)

Development of *in silico* prediction method for idiosyncratic drug-induced liver injury

Saki Katayama[1]
c212710@ed.nagoya-cu.ac.jp

Sarara Doi[1]
c172035@ed.nagoya-cu.ac.jp

Takayuki Tachiki[1]
c182528@ed.nagoya-cu.ac.jp

Kaori Ambe[1]
ambek@phar.nagoya-cu.ac.jp

Masahiro Tohkin[1]
tohkin@phar.nagoya-cu.ac.jp

Toshimasa Jindo[2]
jindo.toshimasa@daiichisankyo.co.jp

1 Graduate School of Pharmaceutical Sciences, Nagoya City University,
3-1, Tanabe-dori, Mizuho-ku, Nagoya-shi, Aichi 467-8063, Japan
2 Daiichi Sankyo Co., Ltd.,
1-2-58, Hiromachi, Shinagawa-ku, Tokyo 140-8710, Japan

Keywords: idiosyncratic Drug-Induced Liver Injury, machine learning method,

Drug-induced liver injury (DILI) is an adverse drug reaction that is a serious problem in clinical practice, and many of them are idiosyncratic. Since it is difficult to detect the idiosyncratic DILI (iDILI) in animal experiments, it is strongly desired to develop a prediction method at the drug development stage. It is also expected to use machine learning modeling to predict the adverse drug reaction. We aimed to construct iDILI prediction methods based on the mechanism of toxicity by machine learning from the viewpoint of chemical structure and *in vitro* experimental data information of drugs in this study.

We made a dataset to classification DILI using Drug Induced Liver Injury Rank (DILIrank) which is mainly based on the information of the package insert [1], and a dataset to classification iDILI using LiverTox that provides up-to-date clinical patterns and management of liver injury [2]. First, we made a model to classification DILI, and then a model to classification iDILI. The performance of generated classification models was evaluated in terms of accuracy, sensitivity, specificity, and area under the curve (ROC-AUC) using 5 x 4 nested cross validation, maintaining the positive/negative ratio. We constructed classification models using a XGBoost algorism with molecular descriptors calculated by AlvaDesc [3] and *in vitro* experimental data information extracted from TOX21 [4].

The established DILI classification model showed 0.79, and iDILI classification model showed 0.69 as the average ROC-AUC.

In this study, we were able to construct an iDILI classification model that can provide more specific information such as the mechanism by using *in vitro* information.

This work was supported by research funds from Daiichi Sankyo TaNeDS grant program.

[1] Minjun Chen, Ayako Suzuki, Shraddha Thakkar, Ke Yu, Chuchu Hu, Weida Tong. DILIrank: the largest reference drug list ranked by the risk for developing drug-induced liver injury in humans. *Drug Discovery Today* Volume 21, Issue 4, April 2016, Pages 648-653.
[2] https://www.ncbi.nlm.nih.gov/books/NBK547852/
[3] https://www.affinity-science.com/alvadesc/
[4] https://ncats.nih.gov/tox21

P03-02(F-4)

Machine Learning-based Evaluation of Microscopic Images of Fukushima Patient-derived organoids

Kosuke Takagi[1]
kosuke.takagi@olympus.com

Motoki Takagi[2]
motokitakagi02@gmail.com

Kazuhito Goda[3]
kazuhito.goda@olympus.com

[1] Core Tech and Strategy, Olympus Corporation, 67-4 Takakura-cho, Hachioji, Tokyo,192-0033, Japan
[2] Medical-Industrial Translational Research Center, Fukushima Medical University, 1 Hikarigaoka, Fukushima,960-1295, Japan
[3] Biological Evaluation Technology, Olympus Corporation, 2-3 Kuboyama-cho, Hachioji, Tokyo,192-0023, Japan

Keywords: Machine-learning, Fukushima Patient-derived organoids, Microscopic images

Patient-derived organoids (PDOs) have a variety of potential applications in pre-clinical studies of the tumor response. PDOs have highly complicated multicellular structures, reflecting the genotypic and phenotypic diversity of their origins. Microscopic images enable us to capture various features and processes; however the complexity of PDOs increases difficulty of effective analysis with images. Here, we evaluate the microscopic images of Fukushima PDOs (F-PDOs) with a machine learning based approach, in order to extract effective information by reducing the complexity. At first, we extracted image features and assessed relations to other datasets with a convolutional neural network (CNN) trained for images and gene expressions. The result indicates that morphological features in images of PDOs correctly reflect the differences between origins. Also, the correlation to the gene expression indicated by the results suggests the possibility to construct an analyzing system that integrates information about PDOs' features including microscopic images of them. Finally, we discussed the predictability of the drug responses by integrating these data. It suggests the usefulness of imaging data in analyzing the drug response profiles of the tumor models.

P03-03(F-4)

A customizable benchmark dataset for evaluation of structure-based protein-ligand binding prediction

Xiaoyao Zhang[1]
zhang.x@cb.cs.titech.ac.jp

Mari Matsumura[1]
matsumura@cb.cs.titech.ac.jp

Takashi Ishida[1]
ishida@c.titech.ac.jp

[1] Department of Computer Science, School of Computing, Tokyo Institute of Technology, Room E505, West 8E building, W8-85 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan.

Keywords: Structure-based virtual screening, Machine learning, Deep learning.

With the development of deep learning as well as the growth of protein structure and compounds libraries, high performance and successful applications of protein-ligand binding prediction using machine learning methods have been reported. However, several studies [1] have shown that the performances of these machine-learning based binding prediction methods were overestimated due to the hidden bias and the insufficient data size of the evaluation datasets, such as DUD-E [2]. Moreover, the generalizability of these methods was not correctly evaluated.

Therefore, we here present a new dataset designed for a benchmark of machine learning based virtual screening methods. The dataset, named CUE, is based on ChEMBL activities data and can be customized in several aspects. With a loosened restriction on target selection, CUE has a relatively large amount of data that around 280 targets, 20K actives, and 15K inactives depend on the parameters selected. So that the dataset contains diverse target-ligand information, which provides more ability to assess the generalizability of structure-based methods. We also provide several customizable parameters such as affinities threshold for active ligand, to match different research needs.

To ensure that the data set is capable of machine learning based methods. Active and inactive ligands are filtered to remove bias. Targets with X-ray structures are chosen from single protein targets with low sequence homology. We also evaluated our dataset with weak classifiers. This shows that models can't learn to discriminate between actives and inactives on ligand information only.

[1] Sieg, J.; Flachsenberg, F.; Rarey, M. In Need of Bias Control: Evaluating Chemical Data for Machine Learning in Structure-Based Virtual Screening. *J. Chem. Inf. Model*. **2019**, *59*, *947–961*.

[2] Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, *6582– 6594*.

P03-04(F-4)

Design and Implementation of New Methods for Odor Molecule Generation

Sijie Zhang[1] zhang.s.ak@m.titech.ac.jp
Masaya Otake[2] masaya_otake@takasago.com
Hideki Miyazaki[2] hideki_miyazaki@takasago.com
Kunihide Hoshino[2] kunihide_hoshino@takasago.com
Takashi Ishida[1] ishida@c.titech.ac.jp

[1] School of Computing, Tokyo Institute of Technology. W8-85, 2-12-1 Ookayama, Meguro-ku, Tokyo,152-8550, Japan
[2] Corporate Research & Development Division, Takasago International Corporation. 1-4-11 Nishiyawata, Hiratsuka City, Kanagawa, 254-0073, Japan

Keywords: Molecular generation, Deep learning, Reinforcement learning, Genetic algorithm

Although computer-aided design (CADD) [1] has made great progress in the field of molecular generation [2, 3] and the advancements in deep learning have encouraged its application in CADD [4], it is still difficult for some specific features of molecular generation tasks, such as generation molecules with specified odors [5].

To solve this problem, we introduce two molecular generation methods: (1) A molecular generation method based on genetic algorithm (GA)[6], which decomposes the molecules into fragments and reconstruct the molecules by these fragments using GA; (2) A deep generative model with reinforcement learning and based on generative adversarial network (GAN)[7], for a specific odor label, the model could generate molecules with the specified odor. We employ the reinforcement learning to encourage the model to generate molecules that are more likely to have the specified odor.

We applied the model to the odor molecules dataset with certain odor labels and train the odor prediction model to evaluate our methods, to prove the effectiveness of the proposed methods.

This abstract is based on results obtained from a project, JPNP14004, commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

[1] Narayan, K. Lalit. Computer Aided Design and Manufacturing. New Delhi: Prentice Hall of India. 2008.
[2] Rao VS., Srinivas K., Modern drug discovery process: an in silico approach. *J. Bioinform. Seq Anal*, 2(5):89–94, 2011.
[3] Bajorath J. Integration of virtual and high-throughput screening. *Nat. Rev. Drug Discov*. 1(11):882–894, 2002.
[4] Chen H., Engkvist O., Wang Y., Olivecrona M., Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today. 2018.
[5] Sánchez-Lengeling, B., Wei, J.N., Lee, B.K., Gerkin, R.C., Aspuru-Guzik, A., Wiltschko, A.B. Machine Learning for Scent: Learning Generalizable Perceptual Representations of Small Molecules, ArXiv abs/1910.10685, 2019.
[6] Mitchell, Melanie. An Introduction to Genetic Algorithms. Cambridge, MA: MIT Press. 1996.
[7] Goodfellow, Ian; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua. Generative Adversarial Nets. Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014). pp. 2672–2680. 2014.

P03-05(F-4)

Improvement of Variational Auto-Encoder Design of Organic Molecules by Discriminator

Kyosuke Sato[1] Ptva0fz0@s.okayama-u.ac.jp
Masaaki Misawa[1] misawa@okayama-u.ac.jp
Kenji Tsuruta[1] tsuruta@okayama-u.ac.jp

[1] Graduate School of Natural Science and Technology, Okayama University., 3-1-1 Tsushimanaka, Kita-ku, Okayama-shi, Okayama,700-8530, Japan

Keywords: Molecular Design, SMILES, Deep Learning, Variational Auto-Encoders

Efficient search for optimum molecular structure contributes to an essential speedup of the development of organic devices and, in turn, to the improvement of their characteristics. In the present study, we focus on the deep learning variational auto-encoder (VAE) model, where molecules represented by SMILES strings can be efficiently converted to multivariable continuous space [1,2].

The VAE consists of two neural networks: an encoder and a decoder. The one-hot representation of SMILES is input to the encoder and mapped to the latent variable space. We use QM9 dataset [3,4] for training the VAE. After applying a molecular-mechanics method to obtaining 3D structures from SMILES, we can optimize physical properties of the molecules by other simulation methods such as density-functional-theory calculations even when there is not enough preexisting data set. The range of physical property space covered by the SMILES representation is thereby expanded and the data-driven optimization using Kernel Ridge Regression method can be performed within the search space [5].

However, there is a dead space where the SMILES string generated by the Decoder from the continuous variable space is chemically or grammatically incorrect, which degrades the efficiency of the optimization process. Therefore, we further improve the output rate of valid SMILES of decoder by introducing a “discriminator” attached to the VAE stream. The discriminator is the model to predict SMILES validity by calculating loss function for improvement of the VAE. Details of the method and the efficiency improvement will be discussed in the presentation.

[1] Gómez-Bombarelli, R.; *et al*., Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules, *ACS Central Science*, **2018**, *4*, 268-276
[2] Kusner, M. J.; Paige, B.; Hernández-Lobato, J. M. Grammar Variational Autoencoder, (2017). http://arxiv.org/abs/1703.01925
[3] 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, *Journal of Chemical Information and Modeling*, **2012**, *52*, 2864-2875.
[4] Ramakrishnan, R.; Dral, P. O.; Rupp M.; Von Lilienfeld, O. A. Quantum chemistry structures and properties of 134 kilo molecules, *Scientific Data*, **2014**, *1*, 1–7.
[5] Sato, K.; Tsuruta, K. Optimization of Molecular Characteristics via Machine Learning Based on Continuous Representation of Molecules, *Materials Science Forum*, **2021**, *1016*, 1492-1496.

P03-06(F-4)

Appropriate Evaluation Measurements for Regression Models

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

[1] The Center for Data Science Education and Research, Shiga University, 1-1-1 Banba, Hikone, Shiga 522-8522, Japan

Keywords: Regression model, in silico prediction, Data science, Model Evaluation

In recent years, accelerating the speed of finding seed compounds and reducing the cost of pharmaceutical research has become a necessity. The contribution of in silico drug discovery methods, which predict candidates as new drugs using physicochemical features and substructure fingerprints of compounds, is thus expected. Selecting the seed compounds without conducting experiments could reduce the time and cost required for drug development. However, estimating the characteristics of compounds in our body using a simple linear model alone is unsatisfactory because the effects and distribution of compounds are determined by the environment in our body and their interactions with other molecules. More complex models have been prepared to estimate compound characteristics with high predictive accuracy than simple models. Thus, it is increasingly important to correctly evaluate the predictive performance when selecting the models appropriate for research purposes.

The determinant coefficient, famous as R2, is one of the most prominent statistical measures for evaluating regression models. However, this measure cannot be used for nonlinear regression models. Furthermore, the determinant coefficient cannot be employed to assess the predictive accuracy of models. It is necessary to separate “goodness of fit” and “predictive accuracy” and objectively evaluate the predictive capability. The goodness of fit provides a statistical measure for the data points used in the fitting parameters and shows the explainability of the data used for training. The determinant coefficient is used to measure the goodness of fit of models.

On the contrary, predictive accuracy is a measure that shows how a model can predict new data correctly and shows the predictive capability of data that were not used for training. Predictive accuracy is only calculated against the observed and predicted values, and the number of descriptors is not considered. The mean squared error (MSE) thus is the most appropriate measure for calculating predictive accuracy.

This study employed published data to confirm the statistical measures regarding the differences between linear and nonlinear regression models. The difficulty of using the determinant coefficient is explained. In practical machine learning problems, the proper statistical measures were suggested under the following two conditions: MSE for cross-validation, and MSE along with correlation coefficients for the observed and predicted values of external test data.

Understanding statistical measures and using them appropriately is necessary. The suggested measures will support the practical selection of promising seed compounds and accelerate drug discovery.

This work was supported in part by a research grant for priority areas from Shiga University.

P03-07(F-4)

Effect of integrating HTS prediction outputs on ability of fingerprints to describe biological properties of chemicals

Shumpei Nemoto[1]
Nemo88@g.ecc.u-tokyo.ac.jp

Tadahaya Mizuno[1, †]
tadahaya@mol.f.u-tokyo.ac.jp

Hiroyuki Kusuhara[1, †]
kusuhara@mol.f.u-tokyo.ac.jp

[1] Laboratory of Molecular Pharmacokinetics, Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, Japan
[†] Author to whom correspondence should be addressed

Keywords: fingerprint, chemoinformatics, machine learning, high-throughput screening

Development of high-throughput screening (HTS) assay systems has led to the accumulation of a vast amount of data on biological properties of chemicals. Such biological properties have improved the accuracy of chemoinformatics tasks such as toxicity prediction. Many of machine learning approaches utilize the combination of chemical structures and HTS data for end-to-end purposes while collecting each model output and integrating them into the original fingerprints in a comparable manner would be multivariate that describes both structural and biological properties of chemicals and can be applied to a variety of chemoinformatics tasks. In addition, accumulation of such fingerprints and analyses of their differences are expected to lead to the discovery of new aspects of chemicals and the understanding of chemical spaces. However, the simple collection of machine learning model outputs of HTS data would not be desirable fingerprints because the activities of compounds in HTS data are generally imbalanced and the probability values obtained by machine learning are not scaled. In this study, we evaluated the effects of integrating model outputs of HTS data prediction with the original fingerprints, introducing SMOTE, a method of oversampling, in modeling, and calibration of each model on the ability of the generated multivariate to describe a wide range of biological properties of chemicals.

We obtained the ToxCast dataset as an HTS dataset and employed mol2vec fingerprint and XGBoost as model input and the machine learning method for each feature generation, respectively. Initially, 244 out of total 1569 assays were selected for feature candidates based on prediction accuracy. Each HTS data was trained with SMOTE[1], calibration, and the combination of them, and a multivariate transformer was generated by summarizing the predictors. The ability to describe biological properties of chemicals was evaluated by comparing the intra-dataset similarity of chemicals calculated in transcriptome profiles, transcriptome data of cultured cells treated with a chemical, and the generated multivariate sets of chemicals, based on a hypothesis that transcriptome profiles well reflect biological properties of chemicals. The transcriptome profiles were obtained from the connectivity map) database, and 292 chemicals derived from MCF7 cells were analyzed. From a macroscopic perspective, the integration was not effective in bringing the intra-dataset similarity in the generative multivariate sets closer to that in the transcriptome profiles. On the other hand, several pairs of chemicals (e.g., Monorden-Geldanamycin and Ionomycin-Thapsigargin), are similar in transcriptome profiles but not in structure, exhibited the closer patterns of the intra-dataset similarity by the integration of model outputs, and introducing SMOTE and calibration brought additional improvements. It should be noted that the picked up chemicals are not present in the HTS dataset used for training.

These results indicate that the ability of fingerprints to describe biological properties can be improved by integrating model outputs of HTS data prediction into the original fingerprints, and utilization of oversampling and calibration methods enhanced the ability, in several chemicals. Further analyses are necessary to clarify the properties of chemicals that benefit from this approach.

(1) Chawla, N. V.; Bowyer, K. W.; Hall, L. O.; Kegelmeyer, W. P. SMOTE: Synthetic Minority Over-Sampling Technique. *J. Artif. Intell. Res.* **2002**, *16* (Sept. 28), 321–357.

P03-08(F-4)

Development of a Prediction Method for Protein Cryptic Sites Using Machine Learning

Masahito Kumada[1]
5445197440@edu.k.u-tokyo.ac.jp

1 Master's Program in Medical Informatics and Biotechnology, Graduate School of Frontier Sciences, The University of Tokyo

Keywords: Structural Biology, Drug Discovery Support, Machine Learning

1. Background
Recently, it has been known that proteins contain hidden ligand (drug) binding sites called "cryptic sites", which are normally closed (apo-structure) but are formed when a drug binds to them (holo-structure), and are expected to be applied as new drug targets [1]. However, most of the cryptic sites discovered so far have been identified by chance by comparing the holo- and apo-structures of ligands and target proteins determined by structural biology analysis. If we can predict cryptic site containing proteins from their apo-structures, we will be able to discover new target proteins and develop new drug discovery research. Currently, efforts are being made to develop a method for predicting cryptic sites by using fragment molecules that induce cryptic sites as co-solvents in experiments and molecular dynamics simulations [1-3].

2. Objective
The purpose of this study is to develop a machine learning model to classify the presence or absence of cryptic sites using the protein structure of apo structure as input. We also attempted to evaluate the factors of cryptic sites from the generated machine learning model, and to find a guideline for developing new software to detect cryptic sites.

3. Results
The constructed dataset was divided into 175 training data and 36 test data. Random Forest, XGBoost, LightGBM, and SVM were used as machine learning models to train and compare the performance. As a result, SVM had the best performance, achieving F1 Score: 71.0% for the test data. We checked the surface structure around the protein pocket when the machine learning model answered correctly and incorrectly. As it stands now, when the model predicts incorrectly, the apo-structure has shallow indentations that could be cryptic sites, and the model misjudges other indentations. We also visualized the features that the SVM judged to be important in learning. The results show that the model considers the hydrophobicity score, alpha sphere density, and polarity score as important features. In the future, based on the above findings, we would like to customize Fpocket [4] to improve the accuracy of cryptic site detection.

[1] PeterCimermancicetal.,JMolBiol.428,709-719(2016)
[2] KimuraSRetal.,JChemInfModel.57,1388-1401(2017).
[3] *AntonijaKuzmanicetal.,Acc.Chem.Res.53,654-661(2020).*
[4] LeGuillouxetal.,BMCBioinformatics10,168(2009).

P04-01(F-5)

Rational Elucidation of Weak Interaction in Protein, A Model Study Utilized Copper Protein

Takamitsu Kohzuma[1] **Takahide Yamaguchi**[1] **Attila Taborosi**[1] **Seiji Mori**[1]
takamitsu.kohzuma.qbs@vc.ibaraki.ac.jp

Inst. Quantum Beam Sci., Ibaraki Univ., 2-1-1 Bunkyo, Mito, Ibaraki 310-8512, Japan

Keywords: Weak Interaction, X-ray crystal structure, Spectroscopy, Protein Stability, Electronic structure calculation

Noncovalent interactions (NCI) with the smaller interaction energy, < 10 kJ/mol originated from the dispersion force have been discussed as the important factor in many biological molecules [1]. The NCI plays key roles in the biological molecular recognition [2, 3].

The spectroscopic [4-7], electrochemical [4], crystallographic, and computational studies of a blue copper protein, pseudoazurin (PAz) and its Met16X (X = Phe, Leu, Val, Ile) variants gave clear functional meanings of NCI through the second coordination sphere. The stability and spectroscopic properties of Met16X PAz demonstrated significantly affected by the substitution of Met16. The combination analysis of X-ray structures and computational calculation in the Met16X mutants showed the side chains of Met16Phe, Met16Leu, Met16Val, and Met16Ile PAz are interacted with the His81 imidazole. The clear NCI of S-π/CH-π, π-π, double CH-π, and single CH-π were identified in the second coordination sphere of Wild Type, Met16Phe, Met16Leu, and Met16Ile PAz, respectively. The protein stability series of those Met16X mutants are also clearly reflected the strength of NCI in protein structure. Those rational NCI information through the chemistry of Met16X PAz will be expected to apply the protein design for the developments of bio-drugs.

[1] Watson, J. D.; Baker, T. A.; Bell, S. P.; Grann, A.; Levine, M.; Losick, R. *Molecular Biology of the Gene,* Cold Spring harbor Laboratory Press, 2015.
[2] Harel, M.; Schalk, I.; Ehret-Sabatier, L.; Bouet, F.; Goeldner, M.; Hirth, C.; Axelsen, P. H.; Silman, I.; Sussman, J. L. Quaternary ligand binding to aromatic residues in the active-site gorge of acetylcholinesterase, *Proc. Natl. Acad. Sci. U. S. A.*, **1993**, *90*, 9031–9035.
[3] Sussman, J. L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. Atomic structure of acetylcholinesterase from *Torpedo californica*: A prototypic acetylcholine-binding protein, *Science*. **1991**, *253*, 872–879.
[4] Abdelhamid, R. F.; Obara, Y.; Uchida, Y.; Kohzuma, T.; Dooley, D. M.; Brown, D. E.; Hori, H.; Interaction between Aromatic Ring and Copper Coordinated His81 Imidazole Regulates the Blue Copper Active Site Structure, *J. Biol. Inorg. Chem*., **2007**, *12*, 165-173.
[5] Yamaguchi, T.; Yano, J.; Vittal, Y.; Nihei, Y.; Togashi, H.; Szilagyi, R. K.; Kohzuma, T. The Allosteric Regulation of Axial/Rhombic Population in a "Type 1" Copper Site: Multi-Edge X-ray Absorption Spectroscopic and Density Functional Studies of Pseudoazurin, *Bull. Chem. Soc. Jpn*., **2015**, *88*, 1642-1652.
[6] Abdelhamid, R. F.; Obara, Y.; Kohzuma, T.; Alkaline transition of pseudoazurin Met16X mutant proteins: Protein stability influenced by the substitution of Met16 in the second sphere coordination, *J. Inorg. Biochem.* **2008**, *102*, 1373–1379.
[7] Yamaguchi, T.; Nihei, Y.; Sutherland, D. E. K.; Stillman, M. J.; Kohzuma, T. Stabilization of protein structure through π-π interaction in the second coordination sphere of pseudoazurin, *Protein Sci.* **2017**, *26*, 1921–1931.

P04-02(F-5)

Development of Machine-learning based Prediction model for hERG inhibitory activities by chemical compounds

Kazuaki Ishihara[1] 184001_tf@yok.hamayaku.ac.jp
Hideko Sone[1] hideko.sone@yok.hamayaku.ac.jp
Akihide Sumino[1] akihide.sumino@hamayaku.ac.jp
Kohsuke Hayamizu[1] k.hayamizu@hamayaku.ac.jp

[1] Yokohama Univ. Pharm., 601 Matano-cho, Totsuka-ku, Yokohama, Kanagawa, 245-0066, Japan

Keywords: hERG, machine learning, molecular descriptor, QSAR

The human-ether-à-go-go-related gene (hERG) encodes potassium channel is one of the major critical components associated with QT interval prolongation and arrhythmia called Torsades de Pointes. Several drugs such as astemizole, terfenadine and vardenafil have been withdrawn from the market because of the inhibition of cardiac activity. Recently, it has been reported that only chemicals do not inhibit hERG activity, but also do natural chemicals. Therefore, hERG is a key target as part of preclinical toxicity screening in drug discovery work. Previously, we reported the results of a machine learning classification model using a dataset of hERG inhibition on 3721 compounds [1]. However, the model was insufficient in sensitivity as hERG prediction, and needs to be more sensitivity. To address this issue, we study with the improvement model by combining data preprocessing and discrimination model using a larger dataset. For the dataset, hERG inhibition assay data was extracted from an integrated database consisting of CheMBL, GOSTAR, NCGC, and hERG Central, and assay results for 13994 compounds were used [2]. After preprocessing descriptor selection and parameter tuning, discrimental models were built with using artificial neural network (ANNE), random forest (RF), xgboost, support vector machine (SVM) and naïve Bayes. The model constructed with xgboost was the most fitted. The sensitivity, specificity, MCC, and ROC-AUC in the ensemble model with the top 5 models were 0.90, 0.82, 0.72, and 0.93, respectively. In addition, the importance of molecular descriptors was calculated by SHAP [3, 4] to explain the chemical feature for the hERG inhibitory activity, and VSA_EState2, MolLogP, TPSA, fr_C_O, and BalabanJ were selected as the top five molecular descriptors. The result suggests that selection of chemical descriptors was the most important to extract the features of the hERG inhibitory compound.

[1] Ishihara et al., 65th Conference, The Pharm. Soc. of Japan Kanto Branch, 2021.
[2] Sato T, Yuki H, Ogura K, Honma T. Construction of an integrated database for hERG blocking small molecules. PLoS One. 2018 Jul 6;13(7):e0199348.
[3] Lundberg, S.M., Erion, G., Chen, H. et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2, 56–67 (2020).
[4] Lundberg, Scott M. and Lee, Su-In. A Unified Approach to Interpreting Model Predictions. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017 Dec. Pages 4768–4777.

P04-03(F-5)

Analyzing Deep Neural Networks on Molecular Activity Prediction and Characteristics of the Created Network

Yoshiki Kato[1]
katoh@cch.cs.tut.ac.jp

Shinji Hamada[1]
hamada.shinji.hs@tut.jp

Hitoshi Goto[1]
gotoh@tut.jp

[1] Toyohashi Univ. of Tech., 1-1 Hibarigaoka, Tempaku-cho, Toyohashi, Aichi, 441-8580, Japan

Keywords: Quantitative Structure-Activity Relationship, Drug Discovery, Deep Neural Networks, Machine Learning

We are developing our own molecular activity prediction system to further improve the prediction accuracy. We are optimizing the hyperparameters (fixed parameters that do not change with training) of a large number of Deep Neural Networks (DNNs), and analyzing the characteristics of DNNs when they are used for this task [1,2].

Our predictions use 15 datasets of the Merck Molecular Activity Challenge [3,4], and we have created prediction networks for each. In a previous paper, results were presented mainly for hyperparameter comparisons, and found that mini-batch size (how much information about individual input molecules is learned when training) was particularly effective in improving predictions [1].

Therefore, we increased the number of conditions for the hyperparameters that were found to be important, and also analyzed the prediction network that was created. As a result, we found that the accuracy was improved when the mini-batch size was further reduced. However, decreasing this parameter means increasing the frequency of error back propagation. This is a trade-off for the increase in computation time, so we will discuss this point as well.

This work was partly supported by JSPS KAKENHI Grant Number JP20J12786 from the Japan Society for the Promotion of Science. The numerical calculations were carried out on the TSUBAME3.0 supercomputer at Tokyo Institute of Technology.

[1] Y. Kato; S. Hamada; H. Goto, Validation Study of QSAR/DNN Models Using the Competition Datasets, Molecular Informatics, Volume 39, Issue 1-2, 1900154, 2020

[2] Y. Kato; S. Hamada; H. Goto, Analyzing Deep Neural Networks on Molecular Activity Prediction using Cross Validation Approach, *CBI Annual Meeting 2018*

[3] J. Ma; R. P. Sheridan; A. Liaw; G. E. Dahl; V. Svetnik, Deep neural nets as a method for quantitative structure-activity relationships, J. Chem. Info. Model. 2015, 55 (2), 263-274

[4] Y. Xu; J. Ma; A. Liaw; R. P. Sheridan; V. Svetnik, Demystifying Multitask Deep Neural Networks for Quantitative Structure-Activity Relationships, J. Chem. Inf. Model. 57, 10, 2490-2504

P05-01(F-5)

Molecular basis of allosteric ERK2 inhibitors discovered by *in silico* screening

Mayu Yoshida[1] **Hajime Sugiyama**[2] **Haruna Nagao**[3] **Masaaki Sawa**[3]
Yoshida19@b.s.osakafu-u.ac.jp

Takayoshi Kinoshita[1]

[1] Osaka Pref. Univ., 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan

[2] Mitsubishi Chemical Corporation, 1000 Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-8502, Japan

[3] Carna Biosciences, Inc., 1-5-5 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047 Japan

Keywords: ERK2, Allosteric site, MAPK, Crystal structure, Selectivity

Extracellular signal-regulated kinase 2 (ERK2) controls vital physiological processes such as cell proliferation and differentiation via phosphorylation of various proteins including downstream kinases and transcription factors. ERK2 is a drug target for many diseases such as cancers. The disruption of ERK2 phosphorylation of signal transducer and activator of transcription 3 (STAT3) is efficacious against type 2 diabetes mellitus [1]. To obtain highly selective inhibitors, we payed attention to the allosteric substrate binding site (KIM site) but not the ATP binding site of ERK2. An inhibitor (PEP) with the KIM-site-binding sequence of STAT3 exerts inhibitory activity and positive effect on *db/db* mouse, a type 2 diabetes model [2]. Further, crystal structure showed that PEP bind to the KIM site of ERK2. Subsequently, an *in silico* screening provided the low molecular weight inhibitors (Hit-1 and -2) which are competitive with PEP [3].

To deeply understand molecular basis of these compounds, we evaluated them by biochemical experiments and structural dissections. The ELISA assays revealed that Hit-1 selectively inhibited ERK2 against p38α MAPK, but Hit-2 inhibited ERK2 and p38α MAPK at a comparable level. Crystal structures showed that Hit-1 and 2 bound to the KIM site with distinct manners. These results coincides with the PEP-competitive behavior of Hit-1, 2. Comparison of the binding manners illustrates the difference between these inhibitors in ERK2 selectivity. Furthermore, molecular dynamics simulations of ERK2 and p38α MAPK provided a structural basis for the high selectivity of Hit-1. The finding of this study would greatly enhance to develop highly potent and selective ERK2 inhibitors.

[1] H. Inoue, et al., *Nature Medicine* **10**: 168-174 (2004).
[2] T. Kinoshita, et al., *Chem. Biol. Drug Des.* **78**: 471- 476 (2011).
[3] T. Kinoshita, et al., *Bioorg. Med. Chem. Lett*. **26**: 955- 958 (2016).

P05-02(F-5)

Development of FMODB and recent updates, 2021

Daisuke Takaya[1,*] **Chiduru Watanabe**[1,2] **Kikuko Kamisaka**[1]
Yuma Handa[3] **Kazuki Watanabe**[4] **Kaori Fukuzawa**[3] **Teruki Honma**[1]

[1] RIKEN Center for Biosystems Dynamics Research [2] JST, PRESTO
[3] Hoshi University [4] Chiba University
*Corresponding author. E-mail: daisuke.takaya@riken.jp

Keywords: FMODB, database, FMO, IFIE/PIEDA, COVID-19

Understanding of bio-molecular interactions such as protein-ligand, protein-protein, and nucleic acid interactions are important for structure-based drug design. Our group has been focusing on not only molecular mechanics but also quantum mechanics (QM) which incorporate 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. The benefit of this fragmentation scheme is the availability of inter-fragment interaction energy (IFIE) and pair interaction energy decomposition analysis (PIEDA). The IFIE/PIEDA data provides useful information for analyzing protein-ligand interactions. To analyze IFIE data by a statistical approach, the accumulation of a huge number of FMO calculation data is essential. Therefore, we are continually updating the contents of FMODB [2] (Fig 1.) that 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 to analyze various molecular interactions in our plan. In this poster, we introduce new features of FMODB and recent applications for analyzing bio-molecules interactions of COVID-19 related biomolecules. [3]

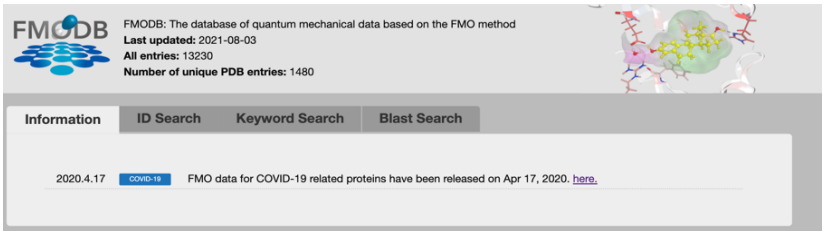


Fig 1. Top page of the FMODB (https://drugdesign.riken.jp/FMODB/)

Acknowledgement

This research was done in activities of FMO Drug Design Consortium, https://fmodd.jp/top-en/. The results of FMO calculations were partially obtained using the Fugaku supercomputer (project ID: 210130). 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 JP21am0101113. DT and CW acknowledge JSPS KAKENHI Grant Number 18K06619. Finally, CW acknowledges JST PRESTO grant (JPMJPR18GD).

1. K. Kitaura *et al.*, *Chem. Phys. Lett.* **1999**, *313*, 701-706
2. D. Takaya *et al., J Chem Inf Model* **2021**, *61(2),* 777-794
3. K. Fukuzawa *et al., accepted*

P05-03(F-5)

Visualization of Antigen Binding on DNA Origami Using Atomic Force Microscopy

Yuki Minamide[1]
k119802@kansai-u.ac.jp

Kentaro Nishiyama[3]
nishiyama.kentarou.83n@st.kyoto-u.ac.jp

Tomoya Niki[1]
k349342@kansai-u.ac.jp

Hiroki Akiba[2, 3]
hakiba@pharm.kyoto-u.ac.jp

Satoshi Nagata[2]
Satoshi-Nagata@nibiohn.go.jp

Haruhiko Kamada[2]
kamada@nibiohn.go.jp

Hiroaki Ohno[3]
hohno@pharm.kyoto-u.ac.jp

Akinori Kuzuya[1]
kuzuya@kansai-u.ac.jp

[1] Department of Chemistry and Material Engineering, Kansai University, 3-3-35 Yamate-cho, Suita-shi, Osaka, Japan
[2] Center for Drug Design Research, National Institutes of Biomedical Innovation, Health and Nutrition (NIBIOHN), 7-6-8 Saito-Asagi, Ibaraki-shi, Osaka, Japan
[3] Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshida-Shimo-Adachi-cho, Sakyo-ku, Kyoto, Japan

Keywords: DNA Origami, Antigen-antibody reaction, Molecular Imaging, AFM

Antibodies which are also known as immunoglobulins are Y-shaped protein used by the immune system. They specifically bind to antigen, and this reaction is widely used in medical diagnosis as "antigen test". There still are some problems in antigen test such as false positive error and false negative error. Requirement of relatively large amount of sample is another disadvantage compared to PCR test. In this study, we focused on DNA Origami.[1] DNA Origami is a technique to create any desired nanostructure by folding long single-stranded circular DNA strand called "scaffold" with the aid of more than 200 short single-stranded DNA called "staples". We have developed 2D DNA Origami device called "DNA Nanostick", which has nine nanometer-scales wells in 260 nm-long stick-like structure.[2] The wells can accommodate exactly one nano-sized molecules such as proteins to provide nanoarray of organic/inorganic nanoparticles. We attached an antibody fragment called Fab, which is a region on an antibody that binds to antigens,[3] to specific well of DNA Nanostick to develop "single-molecule antigen test" utilizing atomic force microscopy (AFM) imaging. Two kinds of Fab, those from Pembrolizumab and Nivolumab, clinically approved immunotherapy anti-cancer drugs,[4] were used to detect their specific target PD-1, immune checkpoint protein. Each azide-modified Fab was conjugated with dibenzocyclooctyne (DBCO)-modified staple strand via Cu-free click reaction, and immobilized at the central well. AFM imaging of this Fab-bearing DNA Nanosticks proved successful attachment of Fab to DNA Nanostick, and specific binding of PD-1 to the well as significant increase of the height around the well.

[1] P. W. K. Rothemund, *Nature* **2006**, *440*, 297-302.
[2] A. Kuzuya, *et al*., *ChemBioChem* **2009**, *10*, 1811-1815.
[3] D. A. Richards, *et al.*, *Chem. Sci*. **2017**, *8*, 63-77.
[4] A. B. M. L. A. Tavares, *et al.*, *New J. Chem*. **2019**, *43*, 7185-7189

P06-01(F-5)

Development of Biologically Interpretable Prediction Models for Drug-Induced Liver Malignant Tumors Based on the Activity of Molecular Initiating Events Using FAERS: a Self-Reported Adverse Drug Events Database

Kota Kurosaki d196955@std.my-pharm.ac.jp **Yoshihiro Uesawa** uesawa@my-pharm.ac.jp

Department of Medical Molecular Informatics, Meiji Pharmaceutical University, 2-522-1 Noshio, Kiyose, Tokyo 204-8588, Japan

Keywords: Quantitative structure–toxicity relationship, Computational toxicology, FAERS, Molecular initiating events, Adverse outcome pathways

LMTs (liver malignant tumors) have recently been identified as a type of drug-induced liver injury (DILI). DILIs are the most serious adverse drug events and can result in the withdrawal of pharmaceutical products or major regulatory changes. As a result, various quantitative structure–toxicity relationship (QSTR) models have been proposed to predict DILI based on the chemical structure to reduce drug discovery failure and cost. However, there is a lack of understanding of the method for evaluating LMT-inducing drugs to be included as an objective variable in the QSTR model and biological predictors for LMT-inducing drugs. As a result, no biologically interpretable LMT-inducing drugs prediction models have been developed. Previously, we proposed a method for drug annotation by assessing the risk of drug-induced LMTs using the Food and Drug Administration Adverse Event Reporting System (FAERS): a self-reported adverse events big database [1]. Moreover, we developed Toxicity Predictor [2]: a QSTR tool to predict the activity of molecular initiating events (MIEs) which is defined as the initial interaction between a molecule and biosystem [3], based on its chemical structure. Here, we hypothesized that activities of MIEs contribute to the prediction for LMT-inducing drugs, constructed the machine learning models to classify LMT-inducing drugs based on the activity of MIEs to build biologically interpretable *in silico* models. We constructed machine learning models based on the five algorithms (i.e., LightGBM, XGBoost, random forest, neural network, and support vector machine) and evaluated their predictive performances. Among the models, LightGBM outperformed the others. When we looked at the MIEs' contributions, we discovered that genomic instability, lipid accumulation, immune response, and estrogen function modulation were the most significant predictors of drug-induced LMT. These findings back up previous *in vitro* and *in vivo* research. Our findings could be useful in drug development, research, and regulatory decision-making. We will build more accurate and biologically interpretable DILI prediction models by increasing our understanding of biological predictors.

[1] Kurosaki, K. and Uesawa, Y. Molecular Initiating Events Associated with Drug-Induced Liver Malignant Tumors: An Integrated Study of the FDA Adverse Event Reporting System and Toxicity Predictions. *Biomolecules.* **2021**, *11*(7), 944.
[2] Kurosaki, K. and Uesawa, Y. A Toxicity Prediction Tool for Potential Agonist/Antagonist Activities in Molecular Initiating Events Based on Chemical Structures. *Int. J. Mol. Sci.* **2020**, *21*(21), 7853.
[3] Allen, T.E.; Goodman, J.M.; Gutsell, S. and Russell, P.J. Defining Molecular Initiating Events in the Adverse Outcome Pathway Framework for Risk Assessment. *Chem. Res. Toxicol.* **2014**, *27*, 2100–2112.

P07-01(F-6)

A gene set analysis method based on network summary values

Takashi Yoneya[1]
takashi.yoneya.sk@kyowakirin.com

[1] Kyowa Kirin Co., Ltd. Otemachi Financial City Grand Cube, 1-9-2 Otemachi, Chiyoda-ku, Tokyo,100-0004, Japan

Keywords: Gene set analysis, Network analysis, Transcriptome

High throughput transcriptome methods such as DNA microarray and RNA-seq are widely used for various studies and many data analysis methods were developed for such data. Gene set analysis (GSA) is a category of such methods. Malelki et. al. reviewed and classified the methods into three types, over-representation analysis (ORA), functional class scoring methods (FCS) and pathway topology-based methods (TP) [1]. Gene set enrichment analysis (GSEA) is one of the most popular GSAs [2] and categorized as an FCS.

We introduce a GSA method, gene set network analysis (GSNA). The difference between GSNA and popular GSA methods is whether it is univariate or multivariate, and the advantage of this analysis is that it can be scored by the similarity of fluctuation patterns within the group. Concretely, GSNA builds a gene network in each gene set by correlation coefficients and uses common network summary values for the gene set score.

We applied GSNA to time course data of vitamin D (VD) treated cell lines. GSNA detected the suppression of the “Interferon alpha beta signaling” gene set by the VD treatment. Negative correlations between VD and SLEDAI scores or interferon-alpha levels in SLE patients have been reported [3] and it indicated that GSNA is useful.

[1] Maleki, F., Ovens, K., Hogan, D. J., & Kusalik, A. J. (2020). Gene Set Analysis: Challenges, Opportunities, and Future Research. Frontiers in genetics, 11, 654.

[2] Subramanian A., Tamayo P., Mootha V. K., Mukherjee S., Ebert B. L., Gillette M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U.S.A. 102, 15545–15550.

[3] Mandal M, Tripathy R, Panda AK, Pattanaik SS, Dakua S, Pradhan AK, Chakraborty S, Ravindran B, Das BK. Vitamin D levels in Indian systemic lupus erythematosus patients: association with disease activity index and interferon alpha. Arthritis Res Ther. 2014 Feb 10;16(1):R49.

P07-02(F-6)

Comparison of gene co-expression inference methods based on biological validity and diversity

Iori Azuma[1]
azuma-iori@g.ecc.u-tokyo.ac.jp

Tadahaya Mizuno[1, †]
tadahaya@mol.f.u-tokyo.ac.jp

Hiroyuki Kusuhara[1, †]
kusuhara@mol.f.u-tokyo.ac.jp

[1] Laboratory of Molecular Pharmacokinetics, Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, Japan
[†] Author to whom correspondence should be addressed

Keywords: Gene Co-expression Network, Network analysis, Transcriptomics

Co-expression patterns derived from omics data such as transcriptomics form a complex network. The modules, sub-networks, in a gene co-expression network (GCN) contain not only information about grouping but also the relationships between individual genes, which indicates that a module in GCN retains more information than that of a gene ontology (GO) and contributes to the understanding of diseases. However, while many methods for estimating co-expression relationships indispensable for GCNs have been developed, their performances of extracting biologically meaningful modules remains to be evaluated and optimized in transcriptome data derived from clinical specimens. In this study, we proposed novel metrics to evaluate the biological validity and diversity of modules in human transcriptome data, and compared relationship inference methods to obtain biologically meaningful modules from human transcriptome data.

First, we evaluated prediction accuracy of the relationships between transcription factors (TFs) and their target genes to compare correlation coefficients and various machine learning methods in inference using the artificial benchmark data used in the DREAM5 competition. Those derived from human transcriptome data (TCGA glioblastoma and GSE4271) were compared as well, although ranking-based enrichment analysis was employed as an evaluation metric instead of prediction accuracy due to low reliability of negative examples in real world data. Next, the constructed inferences were converted to GCNs using planar filtered maximally graph (PMFG) and prediction accuracy was evaluated in the same way. Finally, we evaluated the biological validity and diversity of the modules extracted from the constructed GCNs with novel evaluation approaches based on ChIP-Atlas, a database of ChIP-seq data, and GO.

For the DREAM5 data, we achieved better prediction accuracy of TFs and their target genes relationships by using Random Forest and Elastic Net than using correlation coefficients. This shows the usefulness of machine learning methods in predicting the inference of GCN in these data sets, which agrees with previous works. Regarding human transcriptome data, ranking-based enrichment analysis using the relationships between TFs and regulatory genes in the ChIP-Atlas as ground truth data showed the consistent results with the artificial data sets. On the other hand, the correlation coefficient was superior in label-based enrichment analysis when the ground truth was divided into two labels, top and bottom.

Next, we transformed these inferences into PMFGs and gave the properties of complex network. For the constructed networks, we evaluated the enrichment of target genes in the vicinity of TFs, the shortest path distance (SPD) from TFs to their target genes. We found that the networks using the correlation coefficient corresponded better than those using machine learning methods such as Random Forest and Elastic Net in both cases. Furthermore, when the correspondence between the detected modules and GO was evaluated by the diversity of assigned GO, based on ontology structure, the network derived from the correlation coefficient inference reflected the most diverse ontology.

We proposed novel metrics to evaluate the biological validity and diversity of modules obtained from human GCN. The analyses using these metrics suggested that, unlike benchmark data such as artificial data, co-expression inference with correlation coefficients is more useful with regard to extracting biologically meaningful modules from human data than those with machine learning methods.

P07-03(F-6)

Analysis and visualization of single-cell RNA sequencing data by Partek Flow

Naofumi Seira[1], Ivan K. Lukic[2], Alison Hargreaves[2]
seira.naofumi@molsis.co.jp

[1] Life Science Department, MOLSIS Inc., 3-19-9 Hatchobori, Chuo-ku, Tokyo 104-0032, Japan
[2] Partek Incorporated., 624 Trade Center Blvd., Suite E Chesterfield, Missouri 63005, U.S.A.

Keywords: Next generation sequencing, Single-cell RNA sequencing, Partek Flow

In recent years, popularity of single cell based next generation sequencing has expanded at an exponential rate due to the ability of single-cell RNA sequencing (scRNA-seq) technologies to interrogate gene expression levels of individual cells. As such, scRNA-seq allows the user to elucidate cellular heterogeneity and cellular states of various cell types within a heterogenous populations. For these reasons, scRNA-seq is an indispensable tool in a numerous fields including cancer biology, immunology, neuroscience, and so forth. However, data analysis is challenging and poses a bottleneck for many research teams. Reasons include data size and complexity, computational requirements, volume of the data, insufficient availability of bioinformatics expertise, and existence of innumerable analysis tools with no “industry standard”.

Our solution for the challenge is Partek Flow [1], bioinformatics software which provides a complete, start-to-finish solution from data import to biological interpretation. Partek Flow supports data file types generated by all major vendors and does not require programming knowledge or a bioinformatics background to use. It combines powerful statistics with rich and interactive visualizations in an intuitive, visual interface (Fig.1) with context-sensitive tools to guide the user through the analysis process. Partek Flow can process scRNA-seq data from sequencing files down to gene-barcode matrix, normalize gene expression levels, remove batch effect and perform exploratory analysis by interactive 2D or 3D plots such as UMAP(Fig.2). Next, Partek Flow enables detection of novel cell types and their phenotyping by identification of biomarkers, offers various statistical tools (e.g. hurdle model or DESeq2), biological interpretation tools (pathways), and enables visualization of results by heat maps, violin plots, hierarchical pie charts and more. Moreover, cell differentiation can be assessed by trajectory analysis.

In this poster, we present the methods of scRNA-seq data analysis and visualization to identify different blood cell populations in Peripheral blood mononuclear cell samples in Partek Flow.

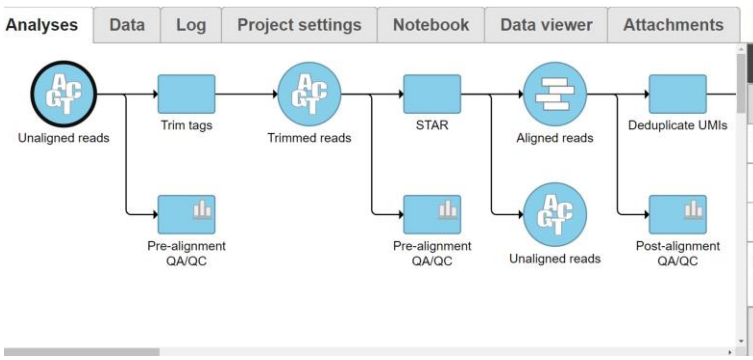


Fig.1. Visual interface in Partek Flow.

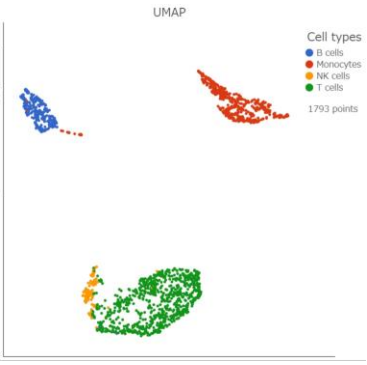


Fig.2. UMAP plot

[1] *Partek Flow;* Partek Incorporated; https://www.partek.com/partek-flow/

P07-04(F-6)

Predicting phase separation protein and its experimental condition using machine learning

Ka Yin Chin[1] w215418a@yokohama-cu.ac.jp
Shoichi Ishida[1] ishida.sho.nm@yokohama-cu.ac.jp
Tomohiro Sato[1] sato.tom.wa@yokohama-cu.ac.jp
Kei Terayama[1] terayama@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.

Keywords: Liquid-liquid phase separation, Phase separation protein, Machine learning, Biomolecular condensates

Membraneless organelle is generally formed by a physical phenomenon called liquid-liquid phase separation (LLPS), which has been intensively studied since the 2010s [1]. LLPS is mainly mediated through molecular interactions between phase-separation proteins (PSPs). In addition, environmental conditions such as pH and temperature are essential factors that can alter the behavior of LLPS [2]. Screening of PSPs and their conditions on a proteome-scale requires numerous experiments combining multiple conditions, which are very time-consuming and costly. Recently, various machine learning models have been developed to predict PSPs, and most of the models were trained based on protein sequences of PSPs and their biochemical features, regardless of the environmental conditions [3,4]. One of the models, Droppler [5], represented an attempt to predict the LLPS behavior using a protein sequence and given experimental conditions. However, its prediction accuracy was lower than other models, and Droppler requires input of experimental conditions, which are difficult to estimate. Here, we developed a machine learning model that can predict both PSPs and their experimental conditions. Our model was based on a Light Gradient Boosting Machine (LightGBM) strategy and was trained using protein sequences and experimental conditions obtained from LLPSDB [6]. As a result, our model showed higher prediction performance than Droppler. Furthermore, we calculated feature importances for our model and analyzed the relationship between PSPs and experimental conditions. We expect that our model would contribute to the efficient discovery of new PSPs and phase-separating conditions.

[1] Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular Condensates: Organizers of Cellular Biochemistry. *Nat. Rev. Mol. Cell Biol*, **2017**, 18 (5), 285–298.
[2] Gomes, E.; Shorter, J. The Molecular Language of Membraneless Organelles. *J. Biol. Chem.*, **2019**, 294 (18), 7115–7127.
[3] van Mierlo, G.; Jansen, J. R. G.; Wang, J.; Poser, I.; van Heeringen, S. J.; Vermeulen, M. Predicting Protein Condensate Formation Using Machine Learning. *Cell Reports*, **2021**, *34* (5), 108705.
[4] Sun, T.; Li, Q.; Xu, Y.; Zhang, Z.; Lai, L.; Pei, J. Prediction of Liquid-Liquid Phase Separation Proteins Using Machine Learning. *SSRN Journal,* **2020**.
[5] Raimondi, D.; et al., *In silico* prediction of *in vitro* protein liquid-liquid phase separation experiments outcomes with multi-head neural attention. *Bioinformatics*. **2021**, btab350.
[6] Li, Q.; et al., LLPSDB: a database of proteins undergoing liquid–liquid phase separation in vitro. *Nucleic Acids Res*. **2019**, *48*, D320-327.

P08-01(F-6)

Dynamical FMO Interaction Analysis of SARS-CoV-2 RNA dependent RNA polymerase and Remdesivir

Shuhei Miyakawa[1] **Koji Okuwaki**[1,2] **Yusuke Kawashima**[1]
s181232@hoshi.ac.jp okuwaki@rikkyo.ac.jp y-kawashima@hoshi.ac.jp
Yuma Handa[1] **Takayuki Furuishi**[1] **Etsuo Yonemochi**[1]
d2002@hoshi.ac.jp t-furuishi@hoshi.ac.jp e-yonemochi@hoshi.ac.jp
Koichiro Kato[2] **Kaori Fukuzawa**[1]
kato.koichiro.957@m.kyushu-u.ac.jp k-fukuzawa@hoshi.ac.jp

[1] Hoshi Univ., 2-4-41 Ebara, Shinagawa-ku, Tokyo, 142-8501, Japan
[2] Department of Chemistry and Research Center for Smart Molecules, Faculty of Science, Rikkyo Univ., 3-34-1 Nishi-ikebukuro, Toshima-ku, Tokyo 171-8501, Japan
[3] Kyushu Univ., 744 Motoka, Nishi-ku, Fukuoka, 819-0395, Japan

Keywords: SARS-CoV-2, RNA dependent RNA polymerase (RdRp), Molecular Dynamics, Fragment Molecular Orbital Method (FMO)

Remdesivir (Rem), one of the COVID-19 therapeutic agents, inhibits RNA elongation when it is located at the -3 position on the RNA recognized by RNA-dependent RNA polymerase (RdRp) of SARS-CoV-2 virus. In this study, we aim to elucidate the effect of different positions of Rem in the RNA sequence on the RNA strand elongation process.

Based on the cryo-electron microscopy structure of the RdRp-RNA-Rem complex (PDBID:7BV2), five structures were prepared by replacing the bases at positions +1 to -3 on the RNA with Rem/Adenine, and molecular dynamics (MD) calculations using the Amber program were performed for 50 ns each. Then Fragment Molecular Orbital (FMO) calculations were performed on the 50 extracted structures using the ABINIT-MP program to analyze the interaction around Rem.

MD calculations revealed that the structure at the end of the elongated strand (+1 position) is disordered when Rem is located at the position -3, and FMO calculations suggested reduction of its interaction energy by about 10%. On the other hand, the interaction of Rem at the -3 position with the surrounding residues were more unstable than that of Adenine. Although a local cation-π interaction was formed between the cyano-substituted sugar and LYS563, which is characteristic of the Rem structure, the interaction energy between Rem and whole complex was weakened. These results suggest that the dynamic behavior and stability of Rem at the -3 position may affect the arrest of RNA elongation.

This research was performed in the activities of the FMO drug design consortium (FMODD). The FMO calculations were partially performed using the Fugaku supercomputer (HPCI project ID: hp210130).

[1] Kitaura et. al.: Fragment molecular orbital method: an approximate computational method for large molecules. Chem. Phys. Lett. 313, 701-706 (1999).
[2] Mochizuki Y, Tanaka S, Fukuzawa K, eds.: Recent Advances of the Fragment Molecular Orbital Method: Springer, Singapore; 2021.

P09-01(F-6)

infiniSee – Similarity Searching in Ultra Large Chemical Spaces

Takashi Ikegami[1], **Shinji Amari**[1], **Ryoichi Kataoka**[1], **Marcus Gastreich**[2]
ikegami.takashi@molsis.co.jp

[1] MOLSIS Inc., 3-19-9 Hatchobori, Chuo-ku, Tokyo 104-0032, Japan
[2] BioSolveIT GmbH, An der Ziegelei 79, 53757 St. Augustin, Germany

Keywords: Chemical Space, Similarity Search, Virtual Screening, New Purchasable Compounds

infiniSee[1] is a Chemical Space navigation platform. Based on similarity, infiniSee finds molecules of interest by screening libraries or Chemical Spaces of *almost* infinite size. infiniSee opens the possibility to screen billions of compounds through its similarity search technology. Search in BioSolveIT partners' chemical spaces (WuXi's GalaXi, 2 billion molecules; Enamine's REAL Space, 20 billion molecules) and literature-based virtual chemical spaces (KnowledgeSpace, 10^{14} molecules) to obtain commercially available and synthesizable compounds. In version 3, OTAVA Chemical's CHEMriya has been integrated to further expand the chemical space with another 11 billion molecules generated by 30,000 building blocks and 44 reactions. Desired compounds of partners can be purchased and delivered to your laboratory within one month at the earliest, e.g. from Enamine.

The underlying concept of molecular similarity and on-the-fly solution generation use FTrees[2] and FTrees-FS[3] techniques, respectively (Fig. 1). Instead of searching already "assembled" molecules, infiniSee performs a combinatorial build-up of compounds from "fragments" in Chemical Spaces. The bonds between fragments are formed according to defined chemical reactions. It forms bonds between fragments according to defined chemical reactions and outputs unexpected similar compounds that match the chemical characteristics (similar to pharmacophore) of the query.

infiniSee has a graphical user interface that is easy for anyone to use (Fig. 2), and it can be easily combined with other calculations, such as docking simulation. We talk about an overview of infiniSee and examples of its applications.

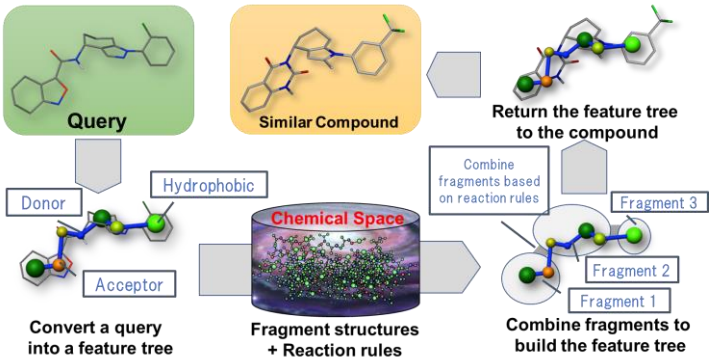


Fig.1. Similarity Search from Chemical Space by infiniSee

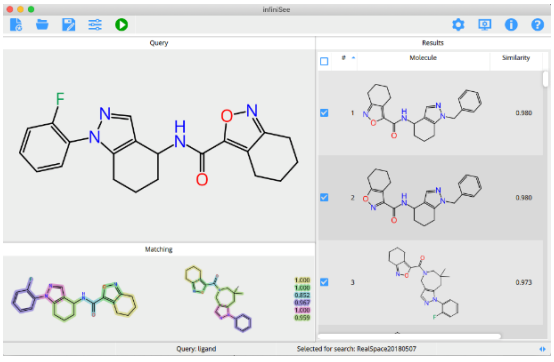


Fig. 2. infiniSee GUI

[1] *infiniSee ver. 3.0;* BioSolveIT GmbH; 2021, https://www.biosolveit.de/infiniSee/ (accessed 2021-7-27)
[2] Rarey, M.; Dixon, J. S. Feature trees: a new molecular similarity measure based on tree matching. *J. Comput. Aided Mol. Des.* **1998**, *12*, 471-490. DOI: 10.1023/A:1008068904628
[3] Rarey, M.; Stahl, M. Similarity searching in large combinatorial chemistry spaces. *J. Comput. Aided Mol. Des.* **2001**, *15*, 497-520. DOI: 10.1023/A:1011144622059

P09-02(F-6)

Development of SBDD Software Considering Empirical Scores with Hydrophobic Heteromeric Interactions to Reproduce the ChEMBL Database

Akihiro Masuda[1]
a15.783r@g.chuo-u.ac.jp
Hideaki Umeyama[2]
umeyama@bio.chuo-u.ac.jp
Takumi Hirao[2]
a18.ynm7@g.chuo-u.ac.jp
Mitsuo Iwadate[1,2]
iwadate@bio.chuo-u.ac.jp

[1] Department of Biological Sciences, Graduate School of Science and Engineering, Chuo University, 1-13-27, Kasuga, Bunkyo-ku, Tokyo,112-8551, Japan
[2] Department of Biological Sciences, Faculty of Science and Engineering, Chuo University, 1-13-27, Kasuga, Bunkyo-ku, Tokyo,112-8551, Japan

Keywords: Structure-based drug design(SBDD), Hydrophobic Interaction, Empirical Potential

The Structure-based drug design(SBDD) software was developed with both database searching and physical interactions, mainly hydrophobic interactions. It was developed using a combination of Pyhon and RDkit libraries in order to easily add new features to the existing ChooseLD program. It is a tremendously time-consuming task to recreate a large database, but time saving has been attempted. Many of the internal variables of the software have been optimized in the ChEMBL database. We examined the characteristics of these results in detail. We then performed a SCOP classification on the protein side and discussed when the application in this study would be more useful.

[1] Landrum G. RDKit http://www.rdkit.org
[2] Takaya, D., Takeda-Shitaka, M., Terashi, G., Kanou, K., Iwadate, M., and Umeyama, H., Bioinformatics based Ligand-Docking and in-silico screening, *Chem Pharm Bull(Tokyo)*, 56:742-744, 2008.
[3] Gaulton, A., Hersey, A., Nowotka, M., Bento, AP., Chambers, J., Mendez, D., Mutowo, P., Atkinson, F., Bellis, LJ., Cibrián-Uhalte, E., Davies, M., Dedman, N., Karlsson, A., Magariños, MP., Overington, JP., Papadatos, G., Smit, I., and Leach, AR. The ChEMBL database in 2017. *Nucleic Acids Res.*, 45:D945-D954, 2017.

P09-03(F-6)

Identification of bitter-related compounds against hT2R14 by structure-based virtual screening

Kohei Kuriki[1]
kuriki.kohei947@mail.kyutech.jp

Ryo Matsumoto[2]
ryo_matsumoto@ajinomoto.com

Chiori Ijichi[2]
chiori_ijichi@ajinomoto.com

Shunsuke Aoki[1]
aokis@bio.kyutech.ac.jp

[1] Depertment of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, japan
[2] Ajinomoto Co., Inc., 1-1, Suzuki-Cho, Kawasaki-Ku, Kawasaki-Shi 210-8681, Japan

Keywords: structure-based virtual screening, G-protein coupled receptor, bitter taste receptor, calcium imaging assay

Human recognizes tastes of sweet, umami and bitter through G-protein coupled receptor (GPCR) on the tongue [1-2]. The bitterness among these tastes is detected by about 25 human Taste 2 Receptors (hT2Rs) [3] and is related to the acceptance of foods and drugs. In addition, hT2Rs are expressed in extra-oral tissues and their physiological roles have received a lot of attention [4-5]. In this study, we focused on hT2R14 in hT2R family members, widely related to bitterness and expressed in extra-oral tissues such as bronchus and testis [6-8]. HT2R14 recognizes structurally a variety of chemical compounds and plays important physiological roles, thus hT2R14 is thought to be an attractive research target. We performed hierarchical structure-based virtual screening (SBVS) and subsequent experimental assay to identify the bitter and bitter blocker compounds of hT2R14. The binding affinities of 154,118 compounds against hT2R14 were estimated by rigid docking simulations in the first screening and further docking simulations with genetic algorithm in the second screening. The selected 10 compounds through the SBVS were experimentally evaluated the efficacy of hT2R14 activation by calcium imaging assay. As a result, we identified three novel bitter blockers and one bitter compound. The SBVS method is effective way for searching novel bitter-related compounds.

[1] Hoon, M. A.; *et al*., Putative Mammalian Taste Receptors: A Class of Taste-Specific GPCRs with Distinct Topographic Selectivity. Cell 1999, 96 (4), 541–551.
[2] Zhao, G. Q.; *et al.*, The Receptors for Mammalian Sweet and Umami Taste. Cell 2003, 115 (3), 255–266.
[3] Go, Y.; *et al*., Lineage-Specific Loss of Function of Bitter Taste Receptor Genes in Humans and Nonhuman Primates. Genetics 2005, 170 (1), 313–326.
[4] Trivedi BP. Neuroscience: hardwired for taste. Nature. 2012 Jun 20;486(7403):S7-9.
[5] Clark, A. A.; *et al*., TAS2R Bitter Taste Receptors Regulate Thyroid Function. FASEB J. 2015, 29 (1), 164–172.
[6] Behrens, M.; *et al*., The Human Taste Receptor HTAS2R14 Responds to a Variety of Different Bitter Compounds. Biochem. Biophys. Res. Commun. 2004, 319 (2), 479–485.
[7] Deshpande, D. A.; *et al*., Bitter Taste Receptors on Airway Smooth Muscle Bronchodilate by a Localized Calcium Flux and Reverse Obstruction. Nat. Med. 2010, 16 (11), 1299–1304.
[8] Governini, L.; *et al*., Expression of Taste Receptor 2 Subtypes in Human Testis and Sperm. J. Clin. Med. 2020, 9 (1), 264.

P09-04(F-6)

A computational study for the development of anti-prion compounds using the docking simulation

Daisuke Arimura[1] K2297171@kadai.jp
Takeshi Ishikawa[1] ishi@cb.kagoshima-u.ac.jp

[1] Department of Chemistry, Biotechnology, and Chemical Engineering, Graduate School of Science and Engineering, Kagoshima University, 1-21-40 Korimoto, Kagoshima, Kagoshima 890-0065, Japan

Keywords: Docking simulation, prion disease, anti-prion compounds

Prion protein (PrP) is a key molecule in prion diseases, including transmissible spongiform encephalopathies in humans and mammals. Two different conformations of PrP are currently recognized at least, cellular form (PrP^{C}) and scrapie form (PrP^{Sc}), and the conformational conversion from PrP^{C} into PrP^{Sc} is considered to cause prion diseases. We had attempted to develop new therapeutic agents for prion diseases by identifying new compounds that bind strongly to PrP^{C} and inhibit its conformational change [1-3]. Through these studies, however, it was found to be difficult to obtain a sufficiently high correlation between the results of docking simulation and experimental assays only by the simple comparison of them. Thus, it is necessary to make a new model for connecting them. In this study, to improve the correlation between the docking simulation and experimental results, we tried to make a model for ranking the compounds, which specialized for anti-prion compounds.

Firstly, all the three-dimensional structures of PrP^{C} deposited in Protein Data Bank were downloaded and carefully examined, resulting that we obtained 86 structures which are able to be used as the receptor for the docking simulations. Next, the docking simulations using these PrP^{C} structures were performed for 96 compounds whose binding affinity and anti-prion activity were reported in the previous studies [1]. In this study, AutoDock Vina [4] was used for the docking simulation. Finally, we made rankings of the compounds using different ways than the ranking by the typical docking simulation with a single receptor structure. As a result, it was found that a ranking method where not only the energies but also the docking structures of the compounds are considered could provide a more correlated ranking with the experimental assays.

[1] D. Ishibashi, et al., Structure-based drug discovery for prion disease by using a novel binding simulation, EBioMedicine, 9 (2016) 238.
[2] Y. Miyazaki, et al., Identification of alprenolol hydrochloride as an anti-prion compound using surface plasmon resonance imaging, Mol. Neurobiol. 56 (2019) 367.
[3] D. Ishibashi, et al., Novel compounds identified by structure-based prion disease drug discovery using in silico screening delay the progression of an illness in prion-infected mice, Neurotherapeutics, 17 (2020) 1836.
[4] O. Trott, et al., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry, J. Comput. Chem., 31, 455.

P09-05(F-7)

Evaluation of Estrogen Receptor-Ligand Binding Properties Using Dynamical FMO Analysis

Arisa Komiya[1] **Minami Suzuki**[1] **Yuma Handa**[1] **Koji Okuwaki**[1,2] **Takayuki Furuishi**[1] **Kaori Fukuzawa**[1] **Etsuo Yonemochi**[1]

S171094@hoshi.ac.jp s181125@hoshi.ac.jp d2002@hoshi.ac.jp okuwaki@rikkyo.ac.jp t-furuishi@hoshi.ac.jp k-fukuzawa@hoshi.ac.jp e-yonemochi@hoshi.ac.jp

[1] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-Ku, Tokyo 142-8501, Japan
[2] Department of Chemistry & Research Center for Smart Molecules, Faculty of Science, Rikkyo University, 3-34-1 Nishi-ikebukuro, Toshima-ku, Tokyo 171-8501, Japan

Keywords: Estrogen receptors, Fragment molecular orbital method, Molecular dynamics, protein-ligand interaction

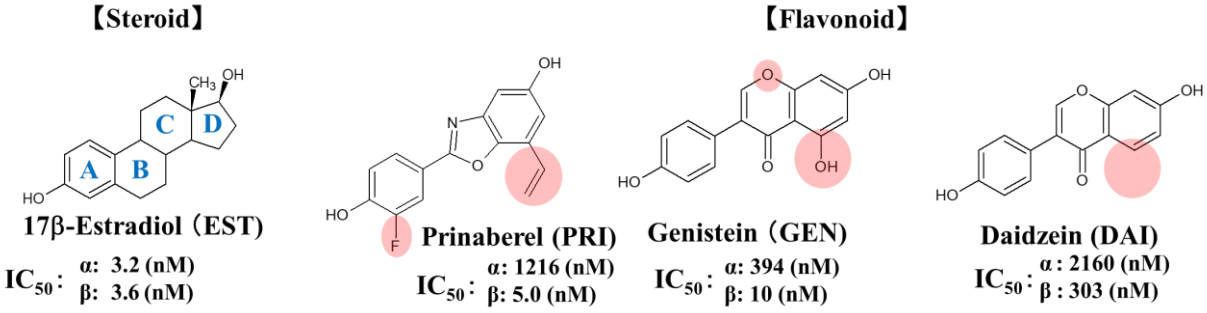
Estrogen receptors (ERs) have α/β subtypes. The endogenous hormone 17β-Estradiol (EST) binds to both ERα and ERβ, but flavonoid ligands such as Prinaberel (PRI), Genistein (GEN), and Daidzein (DAI) are known to be highly β-selective. PRI is about twice as likely to bind to ERβ as GEN, while DAI does not bind to ERβ as much. In this study, we use both molecular dynamics (MD) and the fragment molecular orbital (FMO) [1] methods to clarify the β-selectivity and binding properties of flavonoid ligands.

We performed 50 ns molecular dynamics (MD) simulations of eight ER-ligand complexes combining four ligands (EST, PRI, GEN, and DAI) and ERα/β using the Amber10:EHT force field. Subsequently, FMO calculations at the MP2/6-31G* level were performed for each 5ns structure of the MD calculation results to analyze ligand interaction energies and their components.

Comparing the ligand binding pocket sizes of ERα and ERβ, the volume of ERα was found to be 1.1 times larger than that of ERβ. The residues around the ligand were also more flexible in ERα. It was suggested that the flavonoid ligands are not readily bound to ERα because of a rotatable bond between A and C rings. The differences between the ligands were analyzed using inter-fragment interaction energies (IFIEs). The sum of IFIE between PRI and ERβ suggests that binding of PRI was the most stable, and the electrostatic interaction between the fluorine atom of PRI and Arg346 contribute to the stability. For GEN, the interaction of the hydroxyl group of the D ring with Ile376, Phe377, and the oxygen atom in the C ring with Thr299 was involved in the binding to ERβ. No specific interaction was observed for DAI with surrounding amino acid residues due to the absence of a vinyl or hydroxyl group on the D ring, suggesting that DAI is weakly bound to ERβ.

This research was performed in the activities of the FMO drug design consortium (FMODD).The FMO calculations were partially performed using the Fugaku supercomputer (HPCI project ID: hp210130).

[1] Mochizuki Y, Tanaka S, Fukuzawa K, eds.: Recent Advances of the Fragment Molecular Orbital Method: Springer, Singapore; 2021.



P09-06(F-7)

Interaction Analysis of GPCR in the Presence of Lipid Bilayer Using Fragment Molecular Orbital (FMO) Method

Taro Imahiro[1] S181025@hoshi.ac.jp
Yuma Handa[1] d2002@hoshi.ac.jp
Yusuke Kawashima[1] y-kawashima@hoshi.ac.jp
Koji Okuwaki[1, 2] okuwaki@rikkyo.ac.jp
Takayuki Furuishi[1] t-furuishi@hoshi.ac.jp
Etsuo Yonemochi[1] e-yonemochi@hoshi.ac.jp
Kaori Fukuzawa[1] k-fukuzawa@hoshi.ac.jp

[1] Hoshi University, 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan
[2] Department of Chemistry and Research Center for Smart Molecules, Faculty of Science, Rikkyo University, 3-34-1 Nishi-ikebukuro, Toshima-ku, Tokyo 171-8501, Japan

Keywords: G protein-coupled receptor (GPCR), Lipid bilayer, CHARMM-GUI, Fragment Molecular Orbital (FMO) Method, PIEDA, Protein-Ligand Interaction

G protein-coupled receptors (GPCRs), which are membrane proteins, regulate a variety of physiological functions *in vivo* and are an important family of protein targeting for drug discovery. Since GPCRs function in lipid bilayer *in vivo*, it is essential to elucidate the interaction between GPCRs and drug candidates in the presence of lipid bilayer in structure-based drug design.

In this study, we used the fragment molecular orbital (FMO) method [1, 2] to analyze the interaction of the GPCR and the lipids that make up the bilayer and the ligand. The human β2-adrenergic GPCR (PDB ID: 2RH1) was modeled in POPC lipid bilayer using CHARMM-GUI [3], relaxed, and then we carried out FMO calculations at the MP2/6-31G* level using Fugaku. We also analyzed the interaction around the receptor by partitioning the interaction energy components.

The interaction between the phospholipid heads and the GPCR showed strong electrostatic interactions with charged side chain amino acids such as Arg, Lys, Tyr, and Gln, while the phospholipid tails was sterically sandwiched by Leu of the GPCR, where strong dispersion interactions were observed. In addition, the interaction between the ligand and the GPCR was enhanced for Ser, while the interaction was decreased at the place of Asp and Asn due to the presence of the lipid bilayer.

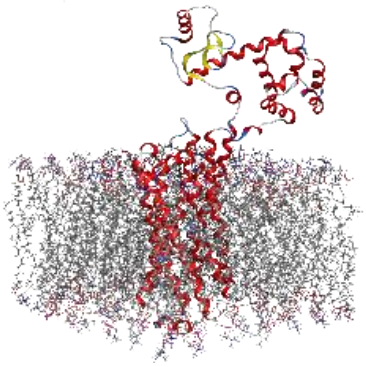


Fig. 2RH1 modeled in lipid bilayer

This research was performed in the activities of the FMO drug design consortium (FMODD). The FMO calculations were partially performed using the Fugaku supercomputer (HPCI project ID: hp210130)

[1] Kitaura *et al*. (1999): Fragment molecular orbital method: an approximate computational method for large molecules. Chem. Phys. Lett. 313, 701-706.
[2] Mochizuki *et al*. (2021): Recent Advances of the Fragment Molecular Orbital Method: Springer, Singapore.
[3] S. Jo *et al*. (2008): CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. J. Comput. Chem. 29:1859-1865.

P09-07(F-7)

Evaluation of binding property of nitrocatechol inhibitors with COMT by the fragment molecular orbital method

Satoko Shirai[1] S181114@hoshi.ac.jp
Ami Yamamoto[1] s161273@hoshi.ac.jp
Takayuki Furuishi[1] t-furuishi@hoshi.ac.jp
Etsuo Yonemochi[1] e-yonemochi@hoshi.ac.jp
Katsuki Takebe[2] takebek@dent.osaka-u.ac.jp
Narikazu Uzawa[2] uzawa@dent.osaka-u.ac.jp
Hiroshi Iijima[3] iijima.hiroshi@nihon-u.ac.jp
Kaori Fukuzawa[1] k-fukuzawa@hoshi.ac.jp

[1] Hoshi Univ.,2 4 41 Ebara, Shinagawa-ku, Tokyo, 142-8501, Japan
[2] Osaka Univ.,1 8 Yamadaoka, Suita-shi, Osaka,565-0871, Japan
[3] Nihon Univ.,7 7 1 Narashinodai, Funabashi-shi, Chiba, 274-8555, Japan

Keywords: Catechol-*O*-methyltransferase, Fragment Molecular Orbital (FMO) Method, structure-based drug design (SBDD), binding affinity, pharmacophore

COMT inhibitors for the treatment of Parkinson's disease are known to be hepatotoxic due to the nitrocatechol (NC) skeleton. Development of inhibitors with non-NC skeletons is expected, but has not yet realized. In this study, we have evaluated the effects of nitro substituents on the binding and reactivity of COMT inhibitors by structure-based drug design (SBDD) using quantum chemical calculation to provide information for the design of non-NC inhibitors.

Molecular modeling was performed for six COMT-inhibitor complexes based on their crystal structures. First, the AMBER 10: EHT force field was used to optimize hydrogen atoms and water molecules. Second, the positions of hydrogen atoms in the side chains of K144 and E199 were optimized at the B3LYP/6-31G* level using small model structure. Third, fragment molecular orbital (FMO) [1] calculations at the MP2/6-31G* level were performed for the whole COMT-inhibitor complexes.

In the COMT-inhibitor complexes, M40, D141, K144, D169, N170, and E199 commonly interact with the catechol skeleton of the inhibitors. Among them, D141, D169, and N170 play the role of Mg-coordinated ligands. W38, P174, L198 and M201 interact with the non-nitro functional groups of the inhibitors. Some inhibitors form CH/π interactions with W38 and L198, and dispersion interactions with M201. P174 exhibits various types of dispersion interactions depending on the inhibitor. In addition, charge distribution analysis showed that the sum of the charges of six carbon atoms of the catechol ring varied with the number of the nitro group substitutions. Comparison of the net charge of the inhibitors with that of substrate suggested that the electron-withdrawing nature of the nitro group was essential for the inhibition of the enzymatic reaction. Substitution of the nitro group with a fluorine atom resulted in a charge distribution similar to that of nitrocatechols, suggesting that a fluorine atom can be a substitute for nitro groups.

This research was performed in the activities of the FMO drug design consortium (FMODD). FMO calculations were partially performed using the Fugaku supercomputer (HPCI project ID: hp210130). This study was partially supported by the AMED-BINDS project (grant number JP21am0101113).

[1] Mochizuki Y, Tanaka S, Fukuzawa K, eds.: Recent Advances of the Fragment Molecular Orbital Method: Springer, Singapore; 2021.

P09-08(F-7)

Inhibitor discovery targeting UHRF1 by MD simulations and biochemical analyses

Yuki Shibahashi[1] w195418g@yokohama-cu.ac.jp
Toru Ekimoto[1] ekimoto@yokohama-cu.ac.jp
Satomi Kori[1] w195502g@yokohama-cu.ac.jp
Tsutomu Yamane[2] tsutomu.yamane@riken.jp
Kyohei Arita[1] aritak@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] Center for Computational Science, RIKEN, Yokohama 230-0045, Japan.

Keywords: in Silico high-throughput screening, Molecular dynamics, UHRF1

The accumulation of epigenetic alternation is one of the major causes of tumorigenesis. Because aberrant DNA methylation patterns cause genome instability and silencing of tumor suppressor genes in various types of tumor, drug discovery targeting the DNA methylation-regulating factors has a potential to develop a cancer treatment drug. In mammals, ubiquitin-like containing PHD and RING finger domain 1 (UHRF1) plays an essential factor for DNA methylation maintenance [1,2]. In addition, UHRF1 is overexpressed in various cancer cells, and down-regulation of UHRF1 reactivates the expression of tumor suppressor genes, suggesting inhibitor of UHRF1 is promising for drugs. UHRF1 consists of multi domains. In particular, the peptide binding groove in tandem Tudor domain (TTD-groove) functions as a binding platform for inter- or intra-molecular interactions, and the recruitment of UHRF1 to replication sites is triggered by the interaction of DNA ligase 1 di/trimethylated on Lys126 (LIG1K126me3) to the TTD-groove. Therefore, we focused on inhibiting the interaction between TTD-groove and LIG1K126me3.

To find candidate inhibitors, in-silico high-throughput screening cycle was developed and performed using over 200,000 compounds. This cycle consisted of docking simulations, molecular dynamics (MD) simulations, binding free-energy calculations, and dynamical pharmacophore screening. The binding affinity was evaluated by the docking score, MM-GBSA score, and the binding free energy, and the pose stability was evaluated by RMSD using three independent all-atom 10-ns MD simulations. Because it is generally not known which of the computational scores and/or indicators in the in-silico scheme correlate with the affinity evaluated experimentally, the top-ranked compounds were selected from multiple scores and indicators. As a results, we selected 130 candidates by this cycle, and succeeded in obtaining two hit compounds confirmed by experimental thermal stability assay and isothermal titration calorimetry. The hit compound was the top performer in three of the four indicators, i.e., all except for the docking score. In particular, the binding pose of the hit compound was quite stable in the 10-ns MD simulation. As a validation, we solved a crystal structure of the hit compound bound UHRF1 complex. The binding pose in the crystal structure was in good agreement with that observed in the MD simulations. In retrospective evaluation, the two hit compounds were evaluated as promising by MD-based metrics, and the hit rate among the compounds was 10.5%, which was higher than a commonly mentioned rate, 1%, indicating that MD simulation was useful to screen the compounds which bind to the target protein.

[1] Sharif, J.; *et al*., The SRA Protein Np95 Mediates Epigenetic Inheritance by Recruiting Dnmt1 to Methylated DNA, *Nature*, **2007**, *450*, 908-912.
[2] Bostick, M.; *et al*., UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells, *Science*, **2007**, *317*, 1760-1764.

P09-09(F-7)

Rational design of a helical peptide inhibitor targeting c-Myb–KIX interaction

Shunji Suetaka[1]
shunji-suetaka@g.ecc.u-tokyo.ac.jp

Yoshiki Oka[1]
oka-yoshiki333@g.ecc.u-tokyo.ac.jp

Tomoko Kunihara[1]
symbol.of.upbeat.life@gmail.com

Yuuki Hayashi[1]
hayashi@bio.c.u-tokyo.ac.jp

Munehito Arai[1,2]
arai@bio.c.u-tokyo.ac.jp

[1] Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo, 153-8902, Japan
[2] Department of Physics, Graduate School of Science, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo, 153-8902, Japan

Keywords: Peptides, Protein design, Intrinsically disordered protein, Drug discovery

The transcription factor c-Myb promotes the proliferation of hematopoietic cells by interacting with the KIX domain of CREB-binding protein. However, its aberrant expression causes serious diseases including leukemia [1]. Therefore, inhibiting the c-Myb–KIX interaction is a promising strategy for developing novel drugs. Here, we rationally designed a helical peptide inhibitor of the c-Myb–KIX interaction using the c-Myb transactivation domain (TAD) peptide fragment. Since the c-Myb TAD is intrinsically disordered and binds KIX via a conformational selection mechanism where helix formation precedes binding [2], stabilizing the helical structure is expected to increase the binding affinity to KIX. We introduced mutations into the c-Myb TAD peptide to stabilize the helical structure based on theoretical predictions using AGADIR. Three of the four initially designed peptides each had a different Lys-to-Arg substitution on the helix surface opposite the KIX-binding interface. Furthermore, the triple mutant with three Lys-to-Arg substitutions, named an RRR inhibitor, showed a high helical propensity and achieved a high affinity to KIX with a dissociation constant of 80 nM. Moreover, the RRR inhibitor efficiently competed out the c-Myb–KIX interaction. Interestingly, the helical contents of designed peptides and their binding affinities showed a good correlation. These results suggest that stabilizing the helical structure based on theoretical predictions, especially by conservative Lys-to-Arg substitutions, is a simple and useful strategy for designing helical peptide inhibitors of protein-protein interactions.

[1] Oh, I.H.; Reddy, E.P. The myb gene family in cell growth, differentiation and apoptosis, *Oncogene*, **1999**, *18*, 3017–3033.

[2] Arai, M.; *et al.*, Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding, *Proceedings of the National Academy of Sciences of the United States of America*, **2015**, *112*, 9614–9619.

P09-10(F-7)

PPAR subtype-dependent ligand binding investigated by docking simulation study

Yuka Yamashita[1] 20bs117j@shinshu-u.ac.jp
Miyuki Tanaka[2] m_tanaka@morinagamilk.co.jp
Eriko Misawa[2] e_misawa@morinagamilk.co.jp
Koji Umezawa[1,3] koume@shinshu-u.ac.jp

[1] Grad. Sch. of Sci. & Tech., Shinshu University, Minami-minowa 8304, Kami-ina, Nagano, 399-4598, Japan.
[2] Morinaga Milk Industry Co., Ltd., Higashihara 5-1-83, Zama, Kanagawa, 252-8583, Japan.
[3] Institute for Biomedical Sciences, Shinshu University, Minami-minowa 8304, Kami-ina, Nagano, 399-4598, Japan.

Keywords: Ligand docking, Nuclear receptor, Aloe sterol

Peroxisome proliferator-activated receptor (PPAR) is a nuclear receptor to regulate lipid and carbohydrate metabolism via gene expression. The three subtypes of PPAR are known and are named as PPARα, PPARδ and PPARγ. The PPARα, δ and γ contain a ligand-binding domain (LBD). The structures of LBDs are similar in PPARα, δ and γ, which have a pocket for ligand binding. A lot of ligands for PPAR LBDs were reported: endogenous compounds of fatty acids, synthetic compounds of drugs and food-derived compounds. The specific and non-specific compounds for PPARα, δ and γ have been developed. Although the complex structures of LBD-ligand have been determined experimentally, the recognition mechanism related to ligand specificity are still ambiguous.

Then, to elucidate the ligand binding specificity for PPAR LBDs, we performed docking simulation for all pairs of the PPAR-LBD structures and the ligands taken from the complex structures. The 55 PPARα, 44 PPARδ and 232 PPARγ structures in Protein Data Bank were used. We took total of 270 compounds from these complex structures, and we built these compound models for docking simulation. The results suggested that the predicted binding energy depend on the docking site. So, the exploring region for docking should be important to predict the correct specificity.

Furthermore, we investigated the compounds whose complex structures with PPAR LBD have not been determined. Then, we focused on sterols derived from *Aloe vera*. The *Aloe* sterols are assigned into two groups of lophenol and cycloartanol. The lophenol can activate PPARγ and PPARα more strongly than the cycloartanol do. The molecular models of lophenol and cycloartanol were built and were used for docking with the PPAR-LBD structures. The results show that the predicted binding energy of lophenol is lowest in complex with PPARγ. The docking pose of lophenol was different in PPARα, δ and γ. The details of docking poses and energy will be discussed by comparing natural and synthetic compounds.

P11-01(F-4)

Gentamicin-induced hearing loss: a toxicological study using drug−gene network analysis

Mizuki Tanaka[1]
165050@gifu-pu.ac.jp

Kiyoka Matsumoto[1]
165079@gifu-pu.ac.jp

Riko Satake[1]
175058@gifu-pu.ac.jp

Yu Yoshida[1]
175127@gifu-pu.ac.jp

Misaki Inoue[1]
175016@gifu-pu.ac.jp

Shiori Hasegawa[1]
135056@gifu-pu.ac.jp

Takaaki Suzuki[2]
t07tor5489@hi3.enjoy.ne.jp

Mari Iwata[3]
m-iwata@kifune.info

Kazuhiro Iguchi[4]
iguchi@gifu-pu.ac.jp

Mitsuhiro Nakamura[1]
mnakamura@gifu-pu.ac.jp

[1] Laboratory of Drug Informatics, Gifu Pharmaceutical University, 1-25-4, Daigaku-Nishi, Gifu, 501-1196, Japan
[2] Gifu Prefectural Government, 2-1-1 Yabutaminami, Gifu, 500-8570, Japan
[3] Kifune Pharmacy, 2-23-2 Hasuike, Yanaizu-cho, Gifu, 501-6103, Japan
[4] Laboratory of Community Pharmacy, Gifu Pharmaceutical University, 1-25-4, Daigaku-Nishi, Gifu, 501-1196, Japan

Keywords: Hearing loss, Gentamicin, Aminoglycoside, Drug−gene interactions

Many drugs, especially aminoglycoside (AG) antibiotics, can cause sensorineural hearing loss, leading to deafness. The objectives of the study were to elucidate the potential toxicological mechanism of AG-induced hearing loss through a drug–gene network analysis.

Our analysis demonstrated an association between several AGs and hearing loss using the Food and Drug Administration Adverse Event Reporting system (FAERS) database. We extracted Gentamicin (GEN)-associated genes (seed genes) and analyzed drug−gene interactions using the ClueGO plug-in in the Cytoscape software. We retrieved 17 seed genes related to GEN from the PharmGKB and Drug Gene Interaction databases. In total, 1018 human genes interacting with GEN were investigated using ClueGO. Through Molecular Complex Detection (MCODE) analysis, we identified 17 local gene clusters. The nodes and edges of the highest-ranked local gene cluster named “Cluster 1” were 30 and 433, respectively. According to the ClueGO analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG), Cluster 1 genes were highly enriched in “oxidative phosphorylation.” According to the ClueGO analysis using ClinVar, Cluster 1 genes were highly enriched in “mitochondrial diseases,” “mitochondrial complex I deficiency,” “hereditary hearing loss and deafness,” and “Leigh syndrome.” Our results demonstrated an association between several AGs and hearing loss. Drug−gene network analysis demonstrated that GEN may be associated with oxidative phosphorylation-associated genes, which may be genes associated with hearing loss.

We believe that our study makes a significant contribution to the literature because our findings may provide insight into the toxicological mechanism of AG-induced hearing loss and aid the development of drugs that prevent AG-induced hearing loss.

[1] Tanaka M, Hasegawa S, Nakao S, Shimada K, Mukai R, Matsumoto K, Nakamura M. Analysis of drug-induced hearing loss by using a spontaneous reporting system database. *PLOS ONE*, 2019, 14, e0217951

P12-01(F-5)

Degradation of DNA ribbons with complementary annihilation

Hiroki Sawada[1]
hiroki.sawada.q2@dc.tohoku.ac.jp

Ibuki Kawamata[1,2]
ibuki.kawamata@tohoku.ac.jp

Shin-ichiro M. Nomura[1]
shinichiro.nomura.b5@tohoku.ac.jp

Satoshi Murata[1]
satoshi.murata.a4@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.
2 Graduate School of Natural Science and Technology, Ochanomizu University, Otsuka 2-1-1, Bunkyouku, Tokyo 112-8610, Japan.

Keywords: Molecular self-assembly, Molecular Cybernetics, complementary annihilation, DNA tile

Various nanostructures in living organisms are constantly being replaced by their constituent molecules through metabolism. In other words, the molecules that assemble the nanostructures are always in danger of being damaged or degraded, but new molecules are synthesized to replace them, so that the nanostructures as a whole are always maintained in a healthy state. In order to achieve this function in artificial nanostructures, a method to decompose the nanostructure is required [1]. DNA tiles are a type of motif (building block) used for self-assembly of nanostructures [2]. The simplest DNA tiles are called “single stranded tiles” (SSTs), which can be used to assemble a variety of 2D or 3D nanostructures [3]. SSTs are composed of four domains, which combine with domains of other SSTs to form a desired structure. For instance, by repeatedly joining a limited number of different SSTs, it is possible to form one-dimensional structures such as tubes and ribbons [4].

Here, we propose a degradation method that utilizes the interaction between DNA nanostructures composed of SSTs (Fig. 1). We use DNA ribbons which is assembled by repeated binding of 7 types of SSTs. When wo types of ribbons designed to be complementary to each other meet, each SST in the structure forms a double strand with the complementary one in the other structure, and consequently the entire structure is degraded. We named this reaction “complementary annihilation”.

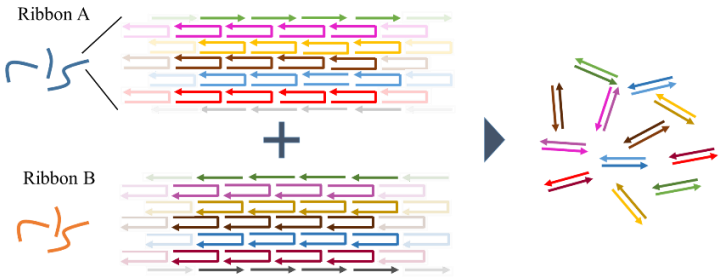


Fig. 1 Schematics of complementary annihilation. Arrows indicate DNA.

We evaluated the formation of two types of ribbon structures and the annihilation reaction between complementary ribbons using electrophoresis, real-time PCR, and fluorescence microscopy. We successfully confirmed that degradation based on the above principles does indeed occur. This method will be able to be used for decomposition of more complex structures in the future. If such a method of decomposition can be combined with a method of assembly, we will be one step closer to the artificial realization of functions such as biological metabolism.

[1] Green, L. N.; Subramanian, H. K.; Mardanlou, V.; Kim, J.; Hariadi, R. F.; Franco, E. Autonomous dynamic control of DNA nanostructure self-assembly, *Nature chemistry*, **2019**, 11, 510-520.
[2] Wei, B.; Dai, M.; Yin, P. Complex shapes self-assembled from single-stranded DNA tiles, *Nature*, **2012**, *485*, 623-626.
[3] Ke, Y.; Ong, L. L.; Shih, W. M.; Yin, P. Three-dimensional structures self-assembled from DNA bricks, *science*, **2012**, *338*, 1177-1183.
[4] Yin, P.; Hariadi, R. F.; Sahu, S.; Choi, H. M.; Park, S. H.; LaBean, T. H.; Reif, J. H. Programming DNA tube circumferences. *science*, **2008**, *321*, 824-826.

P13-01(F-7)

In silico-guided design of growth factor mutants with varied physicochemical parameters

Okada Yuga[1] okada-yuga112@g.ecc.u-tokyo.ac.jp
Akihiro Eguchi[1] 320akihiro0806@gmail.com
Daisuke Kuroda[1] dkuroda@g.ecc.u-tokyo.ac.jp
Ryosuke Ueki[1] r.ueki@chembio.t.u-tokyo.ac.jp
Shinsuke Sando[1] ssando@chembio.t.u-tokyo.ac.jp

[1] The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan

Keywords: In silico protein design, Growth factor, Signal transduction

Receptor tyrosine kinases (RTKs) are activated through the dimerization induced by growth factor and mediate the signal transduction with multifaceted effects such as cell proliferation, differentiation, and migration. The cell signaling mediated by RTKs has been attracting attention in regenerative medicines and pharmaceutical applications. Thus, it is important to understand the signaling mechanisms of RTKs at molecular levels. Heparin, a naturally occurring glycosaminoglycan, is associated with the regulation of RTK signaling. Previous studies proposed several possible roles of heparin in RTK signaling, such as promoting oligomerization of growth factors and RTKs [1] and improving the thermal stability and proteolysis resistance of growth factors [2,3]. To elucidate the role of heparin in the signal transduction, several growth factor mutants with enhanced or decreased heparin affinity, receptor affinity, and thermal stability have been designed based on the crystal structure analysis and consensus approaches [4-7]. However, these approaches still have limitations in designing growth factor mutants with the desired physicochemical parameters. In the present study, we aimed to design mutants of a growth factor with varied heparin affinity and thermal stability based on in silico design approach, which would allow the development of novel mutants with desired properties. In this presentation, the design outline of growth factor mutants, expression of the designed mutants, and evaluation of their physicochemical properties will be discussed.

[1] Harmer, N. J.; *et al.*, Multimers of the fibroblast growth factor (FGF)-FGF receptor-saccharide complex are formed on long oligomers of heparin, *Biochem. J.*, **2006**, *393*, 741–748.
[2] Nguyen, T. H.; *et al.*, A heparin-mimicking polymer conjugate stabilizes basic fibroblast growth factor, *Nat. Chem.*, **2013**, *5*, 221−227.
[3] Caldwell, M. A.; *et al*., Heparin stabilizes FGF-2 and modulates striatal precursor cell behavior in response to EGF, *Exp. Neurol.*, **2004**, *188*, 408–420.
[4] Koledova, Z.; *et al.*, Fibroblast growth factor 2 protein stability provides decreased dependence on heparin for induction of FGFR signaling and alters ERK signaling dynamics, *Front. Cell Dev. Biol.*, **2019**, *7*, 1–15.
[5] Zakrzewska, M.; *et al.*, Increased protein stability of FGF1 can compensate for its reduced affinity for heparin, *J. Biol. Chem.*, **2009**, *284*, 25388–25403.
[6] Zakrzewska, M.; *et al.*, Highly stable mutants of human fibroblast growth factor-1 exhibit prolonged biological action, *J. Mol. Biol.*, **2005**, *352*, 860–875.
[7] Huang, Z.; *et al.*, Uncoupling the Mitogenic and Metabolic Functions of FGF1 by Tuning FGF1-FGF Receptor Dimer Stability, *Cell Rep.*, **2017**, *20*, 1717–1728.

P13-02(F-7)

Application of Scilligence Software for Peptide Drug Discovery

Yasuhiro Shinozaki[1] shinozaki.yasuhiro@molsis.co.jp
Jonathan Buttrick[2] jonathan.buttrick@scilligence.com

[1] MOLSIS Inc., 3-19-9, Hatchobori, Chuo-ku, Tokyo 104-0032, Japan
[2] Scilligence Corporation, 625 Mount Auburn St., Suite 203, Cambridge, MA 02138, U.S.A.

Keywords: Data Management, HELM

Nowadays, modality in drug discovery has been diversified, and peptides, antibodies, oligonucleotides and their conjugates are targeted as pharmaceutical R&D in addition to conventional small molecules. For that reason, Scilligence Corporation [1] has developed products of informatics system that supports various drug modalities. In this poster presentation we would like to introduce an application case of Scilligence products for the research on peptide drug discovery.

In case of peptide drug discovery, properties of peptides (binding to the target, stability, etc.) are optimized by changing some amino acid residues to develop more efficient peptides. But it is an issue that this study takes a long time, because a huge number of experiments by combination of amino acid residues. This application case shows that introduction and customization of RegMol, one of Scilligence products, made generation and filtering of thousands of mutated peptides easier, and the efficiency of R&D was greatly improved as a result.

RegMol is the software for registration of sample information and assay data using the samples. It supports HELM (Hierarchical Editing Language for Macromolecules) [2] and can handle peptide easily. In addition, in this application case RegMol was customized as follows to suite the user's research:

- adding the enumerator that can quickly create a wide variety of peptides with varied mutations based on a template peptide
- configuring natural/unnatural amino acids and chemical monomers to be used with the enumerator
- configuring monomer scan, deletion scan and the use of multi-chain peptides to work with the enumerator
- integration with Pipeline Pilot to filter undesired peptides during enumeration process
- improving efficiency of CSV output of thousands of mutated peptides

[1] https://www.scilligence.com/web/
[2] Zhang, T.; Li, H.; Zi, H.; Stanton, R. V.; Rotstein, S. H. HELM: A Hierarchical Notation Language for Complex Biomolecule Structure Representation *J. Chem. Inf. Model.* **2012**, *52* (10), 2796–2806.