P1-01

Development status of ABINIT-MP program in 2018

Yuji Mochizuki[1,2] **Kota Sakakura**[3] **Yoshinobu Akinaga**[4] **Hiromasa Watanabe**[5]
fullmoon@rikkyo.ac.jp k-sakakura@ay.jp.nec.com akinaga@vinas.com hi-watanabe@hpc.co.jp

Kouichiro Kato[6] **Kaori Fukuzawa**[2,7] **Tatsuya Nakano**[8]
kouichiro.kato@mizuho-ir.co.jp k-fukuzawa@hoshi.ac.jp nakano@nihs.go.jp

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

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

The fragment molecular orbital (FMO) method [1-3] has been one of the most widely used schemes based on the fragmentations of molecular systems [4]. The popularity of FMO could be related with the usefulness of inter-fragment interaction energy (IFIE) or pair interaction energy (PIE), in particular for medicinal chemistry and biophysics. We have been developing ABINIT-MP as an original FMO program [3]. This program has several unique features, in comparison with other FMO programs, GAMESS-US [5] and PAICS [6]. In this poster presentation, we will summarize the current development status of ABINIT-MP (Open series [7]) in 2018. New functionalities such as energy decomposition analysis will be shown.

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

[1] Fedorov, D. G., Kitaura, K. ed., *The FRAGMENT MOLECULAR ORBIAL METHOD: Practical Applications to Large Molecular Systems*, CRC Press: Boca Raton, 2009.
[2] Fedorov, D. G.; Nagata, T.; Kitaura, K. Exploring Chemistry with the Fragment Molecular Orbital Method, *Phys. Chem. Chem. Phys.*, 14:7562-7577, 2012.
[3] Tanaka, S., Mochizuki, Y., Komeiji, Y., Okiyama, Y., Fukuzawa, K., Electron-Correlated Fragment-Molecular-Orbital Calculations for Biomolecular and Nano Systems, *Phys. Chem. Chem. Phys.*, 2014, 16:10310-10344, 2014.
[4] Gordon, M. S., Fedorov, D. G., Pruitt, S. R., Slipchenko, L. V., Fragmentation Methods: A Route to Accurate Calculations on Large Systems, *Chem. Rev.*, 112:632-672, 2012.
[5] Fedorov D. G., The fragment molecular orbital method: theoretical development, implementation in GAMESS, and applications, *WIREs, Comp. Mol. Sci.*, e1322, 2017.
[6] Ishikawa T., Ishikura T., Kuwata K., Theoretical study of the prion protein based on the fragment molecular orbital method, J. Comp. Chem., 30:2594-2601, 2009.
[7] Mochizuki Y., Sakakura K., Akinaga Y., Kato K., Watanabe H., Okiyama Y., Nakano T., Komeiji Y., Fukuzawa K., Tanaka S., Current Status of ABINIT-MP as a FMO Program and Related Works with Machine Learning, *J. Comp. Chem. Jpn.*, 16:119-122, 2018.

Ab Initio Molecular Simulations on Specific Interactions Between Amyloid-beta Peptide and Its Ligand

Tomoki Shinzato[1] shinzato@klab.cs.tut.ac.jp
Shogo Tomioka[1] tomioka@klab.cs.tut.ac.jp
Haruki Sogawa[1] sogawa@klab.cs.tut.ac.jp
Sergiy Shulga[2]
Yaroslav Blume[2]
Noriyuki Kurita[1] kurita@klab.cs.tut.ac.jp

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

Keywords: Fragment molecular orbital, Molecular dynamics, Alzheimer's disease, Aggregate, EGCG

The molecular pathogenesis of Alzheimer's disease (AD) is accompanied by senile plaques comprised of amyloid-beta peptides (Aβs) in a diseased brain [1]. Since Aβ has several hydrophobic amino acid residues, it can form strong aggregates in water due to the hydrophobic interactions between these residues, leading to a formation of Aβ fibrils [2]. Therefore, agents with a strong binding affinity to Aβ are expected to inhibit Aβ aggregation and be a potent inhibitor against AD pathogenesis [3].

The (–)-epigallocatechin-3-gallate(denoted hereafter as EGCG) is a potential inhibitor against Aβ aggregation [4], because EGCG consists of three aromatic rings, which can form π–π stacking contacts with the Aβ hydrophobic residues contributing to the formation of Aβ aggregates. However, the relevant mechanism of interactions between Aβ and EGCG is not elucidated. To clarify this mechanism, replica exchange molecular dynamics (MD) simulations as well as quantum chemical calculations were performed for Aβ hexamers with and without EGCG [5]. However, since a small fragment Aβ(16−22) hexamer was considered in this study, it was not clarified if the other residues in a full length Aβ(1−42) effect the interactions with EGCG or not.

In the present study, we investigated the specific interactions between Aβ(1−42) and EGCG, using *ab initio* molecular simulations based on protein-ligand docking, molecular dynamics and *ab initio* fragment molecular orbital (FMO) calculations and elucidated the specific interactions between a full length Aβ and EGCG. In addition, we performed similar molecular simulations for the other ligands, which will be effective for inhibiting AD pathogenesis. Based on the results simulated, we attempted to propose novel agents for inhibiting Aβ aggregations.

[1] F. M. Laird, *et al.*, *J. Neurosci.* 25 (2005) 11693–11709.
[2] S. W. Pimplikar, *Int. J. Biochem. Cell Biol.* 41 (2008) 1261–1268.
[3] A. J. Doig, *Curr. Opin. in Drug Disc. Devel.* 10 (2007) 533–539.
[4] A. V. Kristen, *et al.*, *Clin. Res. Cardiol.* 101 (2012) 805–813.
[5] S. T. Ngo, *et al.*, *J. Mol. Graph. & Model.* 76 (2017) 1–10.

P1-03

Efficient Implementation of FMO-RI-MP3 method in PAICS

Takeshi Ishikawa[1]
t-ishi@nagasaki-u.ac.jp

Kota Sakakura[2]
k-sakakura@ay.jp.nec.com

Yuji Mochizuki[3,4]
fullmoon@rikkyo.ac.jp

[1] Department of Molecular Microbiology and Immunology, Graduate School of Biomedical Sciences, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan.
[2] 1st Government and Public Solutions Division, NEC Corporation, 7-1, Shiba 5-chome, Minato-ku, Tokyo 108-8001, Japan.
[3] 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.
[4] Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan.

Keywords: Electron correlation, fragment molecular orbital method, resolution of the identity approximation, third-order Møller-Plesset perturbation theory, biomolecule

The third-order Møller−Plesset perturbation (MP3) theory using the resolution of the identity (RI) approximation[1] was combined with the fragment molecular orbital (FMO) method[2] to efficiently calculate a high-order electron correlation energy of biomolecular systems. We developed a new algorithm for the RI-MP3 calculation[3], which was implemented in PAICS[4,5]. After test calculations using a small molecule, the FMO-RI-MP3 calculations were performed for two biomolecular systems comprising a protein and a ligand. The computational cost of these calculations was only around 5 and 4 times higher than that of the FMO-RHF calculations. The error associated with the RI approximation was around 2.0% of the third-order correlation contribution to the total energy. On the other hand, the RI approximation error in the interaction energy between the protein and ligand molecule was insignificantly small, which reflected the negligible error in the inter fragment interaction energy. Considering the importance of the accurate evaluation of the molecular interaction in biomolecular systems, we can conclude that the FMO-RI-MP3 method is a highly potent tool for biological studies.

[1] M. Feyereisen, G. Fitzgerald, A. Komornicki, Use of approximate integrals in ab initio theory. An application in MP2 energy calculations, *Chem. Phys. Lett., 1993, 208, 359.*
[2] K. Kitaura, T. Sawai, T. Asada, T. Nakano, M. Uebayasi, Pair interaction molecular orbital method: an approximate computational method for molecular interactions, *Chem. Phys. Lett. 1999, 312, 319.*
[3] T. Ishikawa, K. Sakakura, Y. Mochizuki, RI-MP3 calculations of biomolecules based on the fragment molecular orbital method, *J. Comput. Chem., in press.*
[4] T. Ishikawa, T. Ishikura, K. Kuwata, Theoretical Study of the Prion Protein Based on the Fragment Molecular Orbital Method, *J. Comput. Chem. 2009, 30, 2594.*
[5] http://www.paics.net/

P1-04

Ligand chirality can affect histidine protonation of vitamin-D receptor: *ab initio* molecular orbital calculations in water

Yuta Terauchi[1] terauchi@klab.cs.tut.ac.jp
Rie Suzuki[1] r-suzuki@klab.cs.tut.ac.jp
Atsushi Kittaka[2], akittaka@pharm.teikyo-u.ac.jp
Midori Takimoto-Kamimura[3] m.kamimura@teijin.co.jp
Noriyuki Kurita[1] kurita@cs.tut.ac.jp

[1] Department of Computer Science and Engineering, Toyohashi University of Technology, 1−1 Hibarigaoka, Tempaku−cho, Toyohashi, Aichi, 441−8580, Japan
[2] Faculty of Pharmaceutical Sciences, Teikyo University, 2-11-1 Kaga, Itabashi, Tokyo, 173-8605, Japan
[3] Teijin Institute for Bio-Medical Research, Teijin Pharma Ltd., 4-3-2 Asahigaoka, Hino, Tokyo, 191-8512, Japan

Keywords: Chirality; Protonation of histidine; Molecular simulation; Fragment molecular orbital; Vitamin D receptor; Protein ligand interaction

Vitamin D is recognized to play important roles in the onset of immunological diseases as well as the regulation of the amount of Ca in the blood. Since these physiological actions caused by active vitamin D are triggered by the specific interaction between the vitamin D receptor (VDR) and active vitamin D, many types of compounds have been developed as potent ligands against VDR. In the recent experiment [1], the binding affinity between VDR and its ligand was found to depend significantly on the chirality of the ligand. However, the reason for the dependence has not been elucidated yet.

In the present study, we investigated the specific interactions between VDR and some ligands having different chirality, using *ab initio* fragment molecular orbital (FMO) calculations. The FMO results highlight that two histidine residues of VDR contribute significantly to the binding between VDR and ligand and that their protonation states can affect the specific interactions between VDR and ligand. We therefore considered other possible protonation states of these histidine residues and determined their most stable states, using the *ab initio* FMO calculations. The results illustrate the possibility that the difference in the chirality of a ligand can induce the change in protonation states of the histidine residues of VDR existing near the ligand. This finding provides an important warning that the protonation states of histidine residues existing near the ligand should be considered more precisely in the molecular simulations for investigating the specific interactions between protein and ligand.

[1] D. Sawada *et al.*, *Org. Biomol. Chem.*, 16 (2018) 2448−2455.

P1-05

Specific interactions between mycobacterial FtsZ and inhibitors derivatives: molecular docking and *ab initio* molecular simulations

Haruki Sogawa[1], Shogo Tomioka[1], Tomoki Shinzato[1], Pavel Karpov[2], Sergiy Shulga[2], Yaroslav Blume[2], Noriyuki Kurita[1]
kurita@cs.tut.ac.jp

1 Department of Computer Science and Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, Tenpaku-cho, Toyohashi, Aichi, 411-8580, Japan
2 Institute for Food Biotechnology and Genomics, National Academy of Sciences of Ukraine, 2a. Osypovskogo Str., Kyiv-123, 04123, Ukraine

Keywords: Cell division; Fragment molecular orbital; FtsZ; Inhibitor; Protein-ligand docking; Tuberculosis; zantrin Z3; ZZ3;

Tuberculosis (TB) is one of the most widespread infectious diseases and is caused by the bacillus Mycobacterium tuberculosis (Mtb). In the treatment of TB, various kinds of drugs such as isoniazid, rifampicin, pyrazinamide and ethambutol have been administered. However, there is a considerable potential for *Mtb* to have resistance against these drugs. In particular, *Mtb* having resistance against multiple drugs is called multidrug-resistant TB (MDR-TB), and the number of MDR-TBs is increasing rapidly. It is thus necessary to develop new anti-TB drugs targeting the most conservative proteins, which cannot be mutated easily [1]. As a candidate of anti-TB drugs, new compounds were proposed which suppress the growth of *Mtb* by inhibiting the division of *Mtb* cells [2]. These drugs target the cytoskeletal protein FtsZ (Filamentous temperature-sensitive Z), which plays an essential role in the cell division mechanism [3]. In addition, the treatment of TB should be completed in a short time, because prolongation of the TB treatment may cause new MDR-TBs. Therefore, the development of novel potent anti-TB drugs, which can shorten the treatment period, is required.

As a novel inhibitor against *Mtb* FtsZ, we here considered zantrin Z3 and ZZ3. Zantrin Z3 is an Mtb FtsZ inhibitor proposed by D.N. Margalit *et al* [2]. Based on chemical experiments. *In 2015, ZZ 3, an improved inhibitor based on zantrin Z3, was announced* by G. M. Nepomuceno *et al* [4]. However, since there are some ligand-binding pockets in Mtb FtsZ, the binding site of zantrin Z3 and ZZ3 on FtsZ and the specific interactions between these inhibitors and FtsZ are not elucidated yet.

In the present study, we investigated the specific interactions between *Mtb* FtsZ and some inhibitor derivatives, using *ab initio* molecular simulations based on protein-ligand docking, classical molecular mechanics optimization and *ab initio* fragment molecular orbital (FMO) calculation. Based on the FMO results, we attempted to reveal which curcumin derivative can bind more strongly to FtsZ. In addition, we elucidated which parts of FtsZ and inhibitors derivative are important for the specific interactions between them. The result will be useful for proposing novel anti-TB drugs based on zantrin Z3 and ZZ3 derivatives.

[1] P.A. Karpov, *et al.*, Sci. Innov., 12: 43−55, 2016.
[2] D. N. Margalit, *et al.*, PNAS, 101: 11821–11826, 2004.
[3] M. Osawa, *et al.*, Science, 320: 792–794, 2008.
[4] G. M. Nepomuceno, *et al.*, ACS Med. Chem. Lett. , 6: 308−312, 2015.

P1-06

Free energy analysis of β-sheet aggregation by molecular dynamics simulation

Keiichi Masutani[1] masutani-ke@cheng.es.osaka-u.ac.jp
Yu Yamamori[2] yu.yamamori@aist.go.jp
Kang Kim[1] kk@cheng.es.osaka-u.ac.jp
Nobuyuki Matubayasi[1] nobuyuki@cheng.es.osaka-u.ac.jp

[1] Division of Chemical Engineering, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyamacho, Toyonaka, Osaka 560-8531, Japan
[2] National Institute of Advanced Industrial Science and Technology, Artificial Intelligence Research Center, Tokyo, 135-0064, Japan

Keywords: Molecular dynamics simulations, Aggregation, Cosolvent, Free energy

【Introduction】

Amyloid has β-sheet conformation and aggregation is considered to be a cause of amyloidosis diseases such as Alzheimer's disease. In the field of protein engineering, the aggregation of the expressed protein often results in the inclusion bodies, which are not suitable for a large-scale protein expression. Since the aggregation of proteins is affected not only by the interaction between proteins but also by the interaction between a protein and the solvent, it is indispensable to analyze the energetic aspects of the intermolecular interactions explicitly. In this study, we focus on NACore, which corresponds to the 68th to the 78th residue of human α-synuclein made of 140 amino acids [1]. NACore is considered as a causal substance of Parkinson's disease and is the key region of aggregation. Molecular dynamics simulations of its monomer, octamer, and hexadecamer in a pure-water solvent and mixed solvents of water with urea and DMSO are carried out to study the relative stabilities of the aggregates.

【Results & Discussion】

Figure 1 shows the intra-aggregate structural energy (E^{intra}), the solvation free energy ($\Delta\mu$), and their sum ($E^{\mathrm{intra}}+\Delta\mu$) per monomer as a function of the degree of oligomerization. The E^{intra} becomes more negative when the degree of oligomerization is increased. That is, the intra-aggregate structure is more stable. In contrast, a more positive value of $\Delta\mu$ indicates that the interaction between the aggregate and the water molecules is less favorable. The stability is characterized by ($E^{\mathrm{intra}}+\Delta\mu$) per monomer. This becomes more negative as the degree of oligomerization increases. It means that the aggregate is more stable than the monomer. This can be explained in terms of the interaction of aggregate, which is larger than the effect of destabilization due to the interaction between the aggregate and the solvent water.

Figure 2 shows ($E^{\mathrm{intra}}+\Delta\mu$) per monomer in the pure-water solvent and mixed solvents of water with urea and DMSO at each degree of oligomerization. The stability of the protein in urea and DMSO mixtures can then be compared with that in water through the theory of chemical equilibrium, and it is shown that urea suppresses the tendency of aggregation by orders of magnitude in terms of the protein concentration.

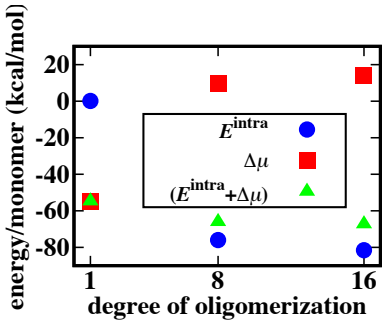


Figure 1:E^{intra}, $\Delta\mu$, and ($E^{\mathrm{intra}}+\Delta\mu$) as a function of the degree of oligomerization.

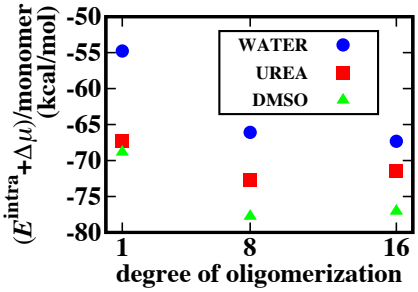


Figure 2:($E^{\mathrm{intra}}+\Delta\mu$) as a function of the degree of oligomerization in pure-water solvent and mixed solvents of water with urea and DMSO.

[1] J. A. Rodriguez, *et al.*, *Nature*, 525:486-490, 2015.

P1-07

Effect of Zn on Aβ Nonamer Aggregates: Molecular Dynamics and *Ab Initio* Molecular Orbital Calculations

Shogo Tomioka[1] tomioka@klab.cs.tut.ac.jp
Tomoki Shinzato[1] shinzato@klab.cs.tut.ac.jp
Haruki Sogawa[1] sogawa@klab.cs.tut.ac.jp
Sergiy Shulga[2]
Yaroslav Blume[2]
Noriyuki Kurita[1] kurita@klab.cs.tut.ac.jp

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

Keywords: Amyloid-β; Alzheimer's disease; Aggregate; Zn; Molecular simulation; Molecular dynamics; Fragment molecular orbital; Protein-protein interaction.

Alzheimer's disease (AD) is the most common disease in dementia [1]. It is recognized that aggregates of amyloid beta protein (Aβ) are deeply involved in the onset of AD [2]. Actually, Aβs produced from APP (amyloid precursor protein) form aggregates of various sizes, such as dimers and amyloid fibrils [3]. These aggregates were found to be self-organized in the experiment [4]. In addition, it was confirmed that the brains of AD patients have abnormally higher concentrations of metal ions such as Zn and Al compared with the brains of normal people [5], and the association between Aβ aggregation and these metal ions has been investigated widely [6].

In the present study, we investigated the influence of Zn ion on Aβ aggregation at atomic and electronic levels, using molecular dynamics (MD) and *ab initio* fragment molecular orbital (FMO) calculations in water. We employed Aβ nonamer with three-fold symmetry and analyzed the change in its structure induced by the addition of Zn ions by MD simulations. Moreover, the electronic states of the Aβ nonamer with and without Zn ions were investigated by *ab initio* FMO calculations to elucidate how the specific interactions between Aβ monomers in the nonamer are changed by the addition of Zn ions.

The MD simulations reveal that the interactions between the charged amino acid residues of Aβ monomers in the Aβ nonamer are significantly influenced by the addition of Zn ions to the Aβ nonamer. The electronic states of the Aβ nonamer are also affected by the Zn ions. In particular, the specific interactions between the charged residues of the Aβ monomers were found to be changed significantly by the Zn addition. Based on these results, we propose a model of an aggregation mechanism of the Aβ monomers, which provides useful information for proposing a novel compound of suppressing Aβ aggregations and contributing to the development of new AD therapeutic drugs.

[1] Scott, H., Anthony, F. J., *Dementia*, John Wiley & Sons Ltd, 2002.
[2] Selkoe, D. J., *Physiol. Rev*. 81 (2001) 741–766.
[3] Tycko, R., *Q. Rev. Biophys*. 39 (2006) 1–55.
[4] Antzutkin, O. N., Balbach, J. J., Leapman, R. D., Rizzo, N. W., Reed, J., Tycko, R., *Proc. Natl. Acad. Sci. U.S.A*. 97 (2000) 13045–13050.
[5] Miller, L. M., Wang, Q., Telivala, T., Smith, R., Lanzirotti, A., Miklossy, J., *J. of Structural Biology.* 155 (2006) 30–37.
[6] Lovell, M. A., Robertson, J. D., Teesdale, W. J., Campbell, J. L., Markesbery, W. R., *J. of the Neurolo. Sciences.* 158 (1998) 47–52.

P1-08

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

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

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

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

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

To identify binding sites of an HIV transactivator of transcription (Tat) protein and an amyloid-β (Aβ) peptide, the present study has investigated their docking, using molecular dynamics (MD) simulations.

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

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

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

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

P1-09

Ab initio Molecular Dynamics Simulation of Continuous Production of Organic Molecules in Alkaline Hydrothermal Vents

Yukimasa Muraki[1] 181p305p@stu.kobe-u.ac.jp
Kohei Shimamura[2] shimamura@port.kobe-u.ac.jp
Ichiro Mori[1] ichiro.mori@port.kobe-u.ac.jp
Shigenori Tanaka[2] tanaka2@kobe-u.ac.jp

[1] Graduate School of Science Technology and Innovation, Kobe University, 1-1, Rokkodai, Nada, Kobe 657-0013, Japan
[2] Graduate School of System Informatics, Kobe University, 1-1, Rokkodai, Nada, Kobe 657-8501, Japan

Keywords: Origin of Life, Alkaline Hydrothermal Vents, *Ab initio* Molecular Dynamics

Although there are several hypotheses attempting to interpret abiotic synthesis of organic molecules, such as formic acids and amino acids, in the non-reduced environment of the primitive Earth, the most of them lack the consideration in continuous production process. The continuous production of organic molecules would be essential for maintaining primitive life's metabolism. We therefore focus on the environment of alkaline hydrothermal vents, which is regarded as one of the most promising places as the origin of life [1]. Membranes consisting of FeS would be formed at the interface between seawater and hot water. Since the pH levels of the sea water and the hot water were estimated to be 6 and 10, respectively, i.e., a pH gradient was formed across the FeS membranes. In addition, the hot water contained a constant amount of H_2. In consideration of such an environment, a mechanism has been proposed in which electrons released from dissociation of H_2 under the pH gradient were supplied to CO_2 absorbed on FeS surface resulting in continuous production of organic molecules.

However, the catalytic mechanism of FeS membranes is not fully understood yet. In order to grasp its microscopic insights, we performed *ab initio* molecular dynamics simulation, where the environment of alkaline hydrothermal vents was modeled. The model contained two domains which are separated by a FeS slab: the primitive seawater domain consisting of H_2O, CO_2, and H_2S, and the hot water domain consisting of H_2O and H_2. A pyrite structure was used for the FeS slab because it was detected as the main component of the hydrothermal vents. To create a pH gradient, H_3O^+ and OH^- were added to the seawater and the hot water the domains, respectively.

In our simulation, a formic acid molecule was produced by the reduction of CO_2, where both of H_2O and H_2 were used as electron sources while H_3O^+ was used as H^+ source. OH^- was involved in H_2 and H_2O dissociation on the surface of the pyrite surface. We also found the S-S bond on the pyrite surface behaved like a disulfide bond (S-S) in current life systems. It has been reported that the S-S bond on pyrite surface would help redox reactions [2]. The state of S-S bond is cycled between oxidized (-S-S-) and reduced states (-SH HS-) in the redox reaction. In our simulation, the S-S bond was changed to reduced state form the oxidized one while the CO_2 molecule was reduced on the S-S bond. However, when the CO_2 was transformed into a formic acid, the bond state cycled back to the oxidized one. Therefore, it is considered that organic molecules such as formic acid would be continuously produced when membranes consisting of pyrite were formed in alkaline hydrothermal vents.

Since the computational cost of density functional theory (DFT) used in this simulation was expensive, the simulation model has to be kept very small (226 atoms). To further expand the model size, we are now trying to employ density functional tight binding (DFTB) [3]. Although DFTB would accelerate the simulation by introducing plausible parameters into DFT, the parameter for Fe-S is not available yet. In this poster session, the progress towards the simulation utilizing the DFTB method will be also presented.

[1] Sojo, V., *et al.*, *Astrobiology*, 16:181, 2016.
[2] Luther III, G., *Geochim. Cosmochim. Acta* 51:3193, 1987.
[3] https://www.dftb.org/

Is hydrophobic group in osmolyte hydrophilic? : A study by fragment based molecular theory

Yukina Uchino[1]
g1520307@edu.cc.ocha.ac.jp

Nahoko Kuroki[2]
g1670603@edu.cc.ocha.ac.jp

Hirotoshi Mori*[3]
mori.hirotoshi@ocha.ac.jp

[1] Department of Chemistry, Faculty of Science, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan
[2] Department of Chemistry and Biochemistry, Graduate School of Humanities and Sciences, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan
[3] Faculty of Core Research, Natural Science Division, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan

Keywords: Osmolyte, trimethylamine N-oxide (TMAO), *ab initio* effective fragment potential (EFP), hydrogen bond, hydrophobicity

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

Recently, in our group, we have focused on effective fragment potential (EFP) [1], which is an *ab initio* polarizable force field defined by a set of quantum chemical wave function of compact constituent molecules. Combining the EFP method with the MD technique, it is possible to perform AIMD level molecular simulation with much cheaper computational costs. Recently, we have reported that it is possible to predict excess molar quantities of mixtures of organic liquids or radial distribution of molecules in ionic liquids. [2,3] In this poster session, applying quantum chemical calculation and EFP theory to aqueous solution of trimethylamine N-oxide (TMAO), a well-known osmolyte [4], we will discuss hydrogen bonding network and hydrophobicity of TMAO in physiological condition.

[1] Gordon, M.S., Fedorov, D.G., Pruitt, S.R., and Slipchenko, L.V., Fragmentation Methods: A Route to Accurate Calculations on Large Systems, *Chemical Reviews*, 112(1):632-672, 2012.
[2] Kuroki, N., and Mori, H., Applicability of effective fragment potential version 2 – Molecular dynamics (EFP2-MD) simulations for predicting excess properties of mixed solvents, *Chemical Physics Letters*,694:82-85, 2018.
[3] Kuroki, N., and Mori, H., Effective Fragment Potential Version 2 - Molecular Dynamics (EFP2-MD) Simulation for Investigating Solution Structures of Ionic Liquids, *Chemistry Letters*, 45:1009-1011, 2016.
[4] Zou, Q., Bennion, B.J., Daggett, V., and Murphy, K.P., The Molecular Mechanism of Stabilization of Proteins by TMAO and Its Ability to Counteract the Effects of Urea, *Journal of American Chemical Society*, 124 (7):1192–1202, 2002.

P1-11

Study of regulation mechanism of agonistic / antagonistic activities of vitamin D receptor Ligand-Binding Domain

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

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

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

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

[1]Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehirotyou, Tsurumi-ku, Yokohama City 108-8639, Japan

Keywords: Molecular Dynamics simulation, Vitamin D receptor, pharmacophore

Vitamin D receptor (VDR) is one of the nuclear receptors (NR) and a significant target for drug discovery. For example, VDR agonists are used as therapeutic agents for osteoporosis. In contrast, although VDR antagonists are expected as therapeutic agents for marble bone diseases, its effective therapeutic agents have not been found.

The mechanism of the agonistic/antagonistic activities is not understood clearly on the basis of X-ray crystallography. According to X-ray crystal structures of NRs and other experimental results, the VDR ligand-binding domain (LBD) undergoes the conformational change upon ligand binding, and a local conformational change around helix 12 is key to regulating agonistic/antagonistic activities. However, all the crystal structures of VDR-LBD are almost identical, regardless of agonist/antagonist binding.

In order to elucidate the regulation mechanism of agonistic/antagonistic activities depending on VDR ligands, we performed pharmacophore analysis using crystal structures and solution structures generated by molecular dynamic (MD) simulations. MD simulations are capable of simulating structural fluctuation, and therefore, the dynamical pharmacophore model created by MD simulations involves structural fluctuation of both the ligand and VDR-LBD. In addition, we also examined solution structures of the antagonist bound form elucidated by a combination analysis of small-angle x-ray scattering experiment and MD simulations [1]. The structure exhibits a different conformation of helix 12. We created the dynamical pharmacophore models for both agonist and antagonist, and found each characteristic feature: In the agonist model, the hydrophobic interaction between the agonist and helix 12 is formed. In contrast, in the antagonist model, the hydrophobic interaction between the antagonist and helices 7 and 10/11 is formed, and the interaction between the antagonist and helix 12 is not observed. These results indicate that the regulation mechanism induced by agonist/antagonist may be originated from the different interaction patterns between the ligand and VDR-LBD.

[1] Anami, Y., Shimizu, N., Ekimoto, T., Egawa, D., Itoh, T., Ikeguchi, M., Yamanoto, K., *J. Med. Chem.*, 59, 7888-7900, 2016.

In silico protein-protein interaction analysis of axon guidance molecule semaphorin and receptor plexin

Erena Shimoji
i175414b@yokohama-cu.ac.jp

Tsutomu Yamane
tyamane@yokohama-cu.ac.jp

Toru Ekimoto
ekimoto@yokohama-cu.ac.jp

Terukazu Nogi
nogi@yokohama-cu.ac.jp

Meri Nagatomo
w165430g@yokohama-cu.ac.jp

Tsubasa Tanaka
w185502c@yokohama-cu.ac.jp

Mitsunori Ikeguchi
ike@yokohama-cu.ac.jp

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

Keywords: Semaphorin, Guidance molecule, Protein-protein interaction, Molecular dynamics simulation

Semaphorin has been known as an axon guidance molecule, secreted around cells. Plexin is a receptor of semaphorin, and their interactions are guidance cues for axons of neurons during development of the nervous system. Because the protein-protein interactions (PPI) of semaphorin and plexin have wide interaction surfaces, it is not trivial to understand the determinants of PPI. Semaphorin and plexin include "Sema domain" with high sequence similarity, and they form PPI through the sema domain.

We analyzed interactions of two complexes which are SEM6A-PLXA2 and SEM6D-PLXA1. Even though they have high sequence homology, they have different binding affinity[1]. Therefore, we analyzed the PPI using *in silico* approaches to recognize the feature of specific interactions. We previously calculated the binding free energy and to perform "*in silico* alanine scanning" [2,3], and identified the "hotspots" which are significant amino acids in binding.

Then, we analyzed structural dynamics for investigating interaction specificity. We have performed molecular dynamics simulations for four complexes, SEM6A-PLXA2, SEM6D-PLXA1, SEM6A-PLXA1 and SEM6D-PLXA2. To examine the interaction specificity, the exchange models, i.e., SEM6A-PLXA1 and SEM6D-PLXA2 were modeled by swapping semaphorins of SEM6A-PLXA2 and SEM6D-PLXA1. The exchange models are supposed to exhibit less affinity than wildtype models. We performed 1-μs molecular dynamics simulations for each model. Amino acids contacts of binding interface were analyzed to identify the important interactions of the complexes.

[1] B. J. C. Janssen, *et al*. Structural basis of semaphorin–plexin signaling. Nature 467 1118-1122, 2010.
[2] T. Kortemme, *et al*. A simple physical model for binding energy hot spots in protein–protein complexes. Proc. Natl. Acad. Sci. 99, 14116-14121, 2002.
[3] T. Kortemme, *et al*. Computational Alanine Scanning of Protein-Protein Interfaces. Sci. STKE 10, 2004.

Protein-ligand binding process studied by Markov state model

Toru Ekimoto[1] ekimoto@yokohama-cu.ac.jp
Mitsugu Araki[2] araki.mitsugu.6w@kyoto-u.ac.jp
Yuta Isaka[3] isaka@fbri.org
Tsutomu Yamane[1] tyamane@yokohama-cu.ac.jp
Yasushi Okuno[2] okuno.yasushi.4c@kyoto-u.ac.jp
Mitsunori Ikeguchi[1] 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] Graduate School of Medicine, Kyoto University, 53 Shogoin-Kawaharacho, Sakyo-ku, Kyoto 606-8507, Japan
[3] Center for Cluster Development and Coordination Research Development Group for in Silico Drug Discovery, Foundation for Biomedical Research and Innovation at Kobe, KCMI building 213, 6-3-5, Minatojima-Minamimachi, Chuo-ku,Kobe, 650-0047, Japan

Keywords: Molecular dynamics simulation, Markov state model, binding free energy, association rate constant, dissociation rate constant

Computational approaches are routinely used in drug discovery and lead-optimization processes. With advances in computational power and softwares, all-atom molecular dynamics (MD) simulations can be carried out easily, and (relative) binding free energy calculations with MD simulations become possible at present. Therefore, further developments of methodology using MD simulations are demanded for novel structure-based drug design, e.g., with considering dynamic structural change of proteins induced by a ligand binding and protein-ligand binding processes.

In this study, protein-ligand binding processes are analyzed by large amount of short MD simulations unified by Markov state model (MSM) [1]. In this framework, the protein-ligand binding process is described as a stochastic process among states under the assumption that the transition from the state depends only on the current state. From the transition matrix, the probabilities of staying each state in equilibrium and of transitions among states are determined, and therefore, the binding free energy (ΔG) and the association/dissociation rate constants (Kon/Koff) can be estimated by a ratio of the bound/unbound states and the mean first passage time between the two states. In addition, the MSM provides an explicit picture of the binding process, and structures on the process can be also obtained.

To assess the performance of MSM, we execute a lot of 100 ns MD simulations for three systems using K supercomputer (~200 μs in total), construct MSMs, and compare the estimated ΔG, Kon, and Koff to the experimental values. The three systems are Dihydrofolate reductase (DHFR)-inhibitor 1, DHFR-inhibitor 2, and a mutated DHFR-inhibitor 2 systems, and exhibit different Kon/Koff. The MSM is constructed by two steps: The first step is binding-pocket search by MD simulations which starts from dissociated positions of the protein and ligand. Using MSM constructed by the first step, bound states are determined, and additional MD simulations start from the bound states to sample ligand bindings into the binding pocket at suitable poses. The estimated Kon is in good agreement with the experimental value, and the estimated ΔG and Koff deviate ~2 kcal/mol and double-digit from the experimental ones, respectively.

This research was done in activities of the K supercomputer-based drug discovery project by Biogrid pharma consortium (KBDD). The results were obtained using the K-computer (project ID: hp150025 and hp160010, hp170036, hp180011).

[1] Husic, B.E. and Pande, V.S., Markov State Models: From and Art to a Science, *J. Am. Chem. Soc.*, 140:2386-2389, 2018.

Microsecond Molecular Dynamics Simulations of Medium Molecule Drugs and the docking with a target protein

Lisa Matsukura[1]
1833730007b@waka.kindai.ac.jp

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

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

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

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

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

Keywords: Molecular dynamics Simulation, Ligand Docking

Recently, Medium molecule drug, antibody drug and so on, have been developed as new drugs. Medium Molecule Drugs is usually the complex of a low weight molecular compound and a peptide, and have a ring shape and the medium size. The peptide portion has a specific sequence for the target protein, and the component of the low weight molecular compound also has specific for the target. Here, we focused on the Medium Molecule Drug for HSP90. The binding affinity of the peptide portion with HSP90 is low, while that of the complex with HSP90 is very high. The results of CD spectroscopy, however, suggested that both compounds have similar turn structure in the peptide regions[1]. To understand the differences between the affinities of them, we performed the microsecond molecular dynamics (MD) simulations of the peptide portion and the complex, and the docking them with HSP90. We have obtained the reasonable results that the both of the peptide portion and the complex compound have Type-I beta-turn around PRO. Moreover, the distribution of Type-I beta turn in the complex is much higher than that in the peptide portion, and the turn structure stabilizes the tertiary structure of the complex. The stability would be supported to the high docking affinity with HSP90.

[1] Keisuke Fukunaga, et al., Construction of a crown ether-like supramolecular library by conjugation of genetically-encoded peptide linkers displayed on bacteriophage T7, *Chemical Communications*, **50**:3921-3923, (2014)

Geometry Optimization Effect of Ligand on Protein-Ligand Binding Energy Calculation Using Fragment Molecular Orbital Method

Takao Otsuka[1] totsuka@riken.jp
Noriaki Okimoto[1] okimoto@riken.jp
Makoto Taiji[1] taiji@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research (BDR), Building B, 6-2-4 Furuedai, Suita, Osaka 565-0874, Japan

Keywords: electronic structure calculation, fragment molecular orbital method, protein-ligand binding energy, geometry optimization

Computational prediction of protein-ligand binding affinities is so important for rational drug design. One of the most popular strategy to obtain the protein-ligand binding energy by electronic structure calculations is to use the fragment molecular orbital (FMO) method [1]. The FMO method was developed as a way of calculating the QM electronic state of a whole molecular system through the fragment based type of strategy. Recently, we proposed an acceleration scheme for the binding energy calculation of protein-ligand system by using FMO method [2]. This scheme is based on a multilayer FMO method focusing on the protein-ligand interaction distance, reduces the computational costs, while maintaining accuracy in the evaluation of binding energy. We also reported an application study using this acceleration scheme and showed the high applicability of our scheme [3].

On the other hand, we still have some open questions for the protein-binding energy calculation using FMO method; as one of the examples, the effect of geometry optimization on protein-ligand system. (Although we have reported this kind of study by large-scale first-principles DFT calculation previously [4].) In this study, we report the effect of the geometry optimization on protein-ligand binding energy calculation using FMO method. In order to evaluate the difference of the binding energy clearly, only the ligand structure was optimized. The geometry optimization was performed by FMO with frozen domain and dimer (FDD) method [5]. The results in this study can compare with our previous study directly.

[1] Kitaura, K., Ikeo, E., Asada, T., Nakano, T., Uebayasi, M., *Chem. Phys. Lett.*, 313, 701, 1999.
[2] Otsuka, T., Okimoto, N., Taiji, M., J. Comput. Chem., 36, 2209, 2015.
[3] Okimoto, N., Otsuka, T., Hirano, Y., Taiji, M., ACS Omega, 3, 4475, 2018.
[4] Otsuka, T., Okimoto, N., Taiji, M., Bowler, D. R., Miyazaki, T., *J. Phys.: Conf. Ser.*, 454, 012057 (2013).
[5] Nakata, H., Fedorov, D. G., Nagata, T., Kitaura, K., Nakamura, S., J. Chem. Theory Comput., 11, 3053, 2015.

Development of the CHARMM force field for Cyclosporine A

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

Ryo Takahashi[1]
i140382e@yokohama-cu.ac.jp

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

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

[1] 1Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[2] RIKEN Medical Sciences Innovation Hub Program, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

Keywords: Cyclosporine A, CHARMM force field, Molecular dynamics simulation

Cyclosporine A (CsA), which is an 11-residue cyclic peptide, is an antibiotic found from a fungus of *Tolypocladium inflatum* and has been used as an immune suppressor. CsA shows biomembrane permeability, despite high molecular weight (M. w. ~ 1200). This biomembrane permeability is considered to be caused by the structural change between "open" and "closed" conformation, which are hydrophilic and hydrophobic structure, respectively. Also, CsA contains eight unnatural amino acids and a D-amino acid. Especially, four of eight unnatural residues are methylated backbone amide group (N-methyl amino acids), which play a role in the structural change of CsA. In the present study, we developed the CHARMM force field for CsA, especially we calculated the backbone conformational energy maps of N-methyl amino acids (alanine and glycine) through Quantum Mechanics (QM) calculation. These energy maps were entirely different from those of natural amino acids and consistent with conformations of N-methyl amino acids of crystal structures of CsA. Based on these results of energy maps, we generated the parameters for CMAPs energy term of CHARMM force field for N-methyl amino acids and performed 1μs molecular dynamics simulations of CsA in aqueous solution.

P1-17

Free energy based structural refinement

Yoshiaki Tanida
tanida.yoshiaki@jp.fujitsu.com

Azuma Matsuura
matsuura.azuma@jp.fujitsu.com

Fujitsu Laboratories Ltd., Atsugi, Kanagawa 243-0197, Japan

Keywords: X-ray crystal structure analysis, structure refinement, alchemical free energy calculation

In drug discovery, it is an important issue to accurately determine the binding structure of the ligand molecule and the target molecule, i.e., the complex structure. However, even using synchrotron radiation (SR) X-ray crystal structure analysis, since the thermal fluctuation is large in a solvent, a missing region often remains near the important binding pocket. On the other hand, current “state-of-the-art” free energy calculations based on alchemical transformation for ligand binding are widely accepted in computational ligand design and optimization. Consequently, modern computational free energy calculations are expected to be useful for building the precise model for ligand binding to the target molecule.

In this study, we demonstrate the ability of such free energy based modeling for complex structure, CK2/cdk2 (PDB: 4FKL).

In silico binding affinity analysis for phosphodiesterase-10A inhibitors.

Chisa Yuasa[1]
w175438e@yokohama-cu.ac.jp

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

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

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

[1] Graduate School of Medical Life Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama-shi, Kanagawa, Japan

Keywords: Ligand docking, free energy calculation, Molecular Dynamics simulation, residue mutation

Phosphodiesterase (PDE) hydrolyzes cyclic-AMP (c-AMP) and cyclic-GMP (c-GMP), which play important roles in signal transduction. PDE10A specifically expresses in brains, and PDE10A is considered as a drug target of schizophrenia. Although 24 drug candidates are developed, the relationship of their structures and drug activities has not been clear.

In order to understand the mechanism of the affinity of PDE10A and the ligands, we performed free-energy simulations for a series of binding affinities of the ligands and the selectivity of PDE10A among the PDE family.

First, we calculated binding free energies of 24 ligands and compared them with experimental value: IC_{50} using 4 approaches: the Glide docking, MM-GBSA, FEP+ and MP-CAFEE. The binding free energies calculated by FEP+ were correlated with IC_{50}. From these examinations, we found two important points for the high correlation: precise docking poses of ligands and account of hydration water molecules at the binding site.

Second, we calculated the binding-free-energy changes upon 68 single mutations of residues at the binding sites using FEP. Mutations were chosen from the sequence alignment of the PDE family. The free-energy simulations successfully detected key residues for selectivity of PDE10A among the PDE family.

[1] J. Kunitomo, M. Yoshikawa, M. Fushimi, A. Kawada, J F. Quinn, H Oki, H. Kokubo, M. Kondo, K. Nakashima, N. Kamiguchi, K. Suzuki, H. Kimura, and T. Taniguchi, Discovery of 1-[2-Fluoro-4-(1H-pyrazol-1-yl)phenyl]-5-methoxy-3-(1- phenyl-1H-pyrazol-5-yl)pyridazin-4(1H)-one (TAK-063), a Highly Potent, Selective, and Orally Active Phosphodiesterase 10A (PDE10A) Inhibitor. J. Med. Chem. 2014, 57, 9627−9643

[2] L. Wang, B. J. Berne and A. Friesner On achieving high accuracy and reliability in the calculation of relative protein–ligand binding affinities. Proc. Natl. Acad. Sci. USA 2012, 109, 1937-1942

[3] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel, and R. A. Friesner, OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. J. Chem Theory Comput. 2016, 12, 281-296

Structural basis for pH-dependent ferric ion coordination modes of TtFbpA, the periplasmic subunit of an ABC-type iron transporter from *Thermus thermophilus* HB8

Peng Lu[1] zjhzlupeng@hotmail.com
Yoshitaka Moriwaki[2] moriwaki@bi.a.u-tokyo.ac.jp
Tohru Terada[3] tterada@iu.a.u-tokyo.ac.jp
Kentaro Shimizu[2] shimizu@bi.a.u-tokyo.ac.jp
Koji Nagata[1] aknagata@mail.ecc.u-tokyo.ac.jp

Departments of [1]Applied Biological Chemistry and [2]Biotechnology, and [3]Agricultural Bioinformatics Research Unit, Graduate School of Agricultural and Life Sciences, UTokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo, 113-8657, Japan

Keywords: ABC transporter; Bicarbonate-carbonate equilibrium; Ferric ion binding protein; Ferric ion coordination mode; pH-dependent structural change; Quantum calculation

Iron is one of the vital elements in living organisms. While humans obtain iron primarily through the food they eat, bacteria have developed diverse mechanisms to obtain iron from different sources including Fe^{3+}-transferrin, Fe^{3+}-lactoferrin, Fe-hemoglobin, Fe^{2+} ion,and Fe^{3+}-siderophores depending on their living environments. Therefore, specific receptors and transport systems are required to transport iron from the environment to the inside of cell [1].

In *Thermus thermophilus* HB8, a Gram-negative bacterium, an ATP-binding cassette (ABC) transport systems, TTHA1628-1629-1630, is annotated to be involved in iron uptake and has been validated functionally. This typical iron-uptake ABC transporter consists of three subunits, namely the iron-binding subunit TtFbpA (TTHA1628) in the periplasm, the transmembrane subunit TtFbpB (TTHA1629) in the inner membrane bilayer and the nucleotide-binding subunit TtFbpC (TTHA1630) in the cytoplasm. As an importer, TtFbpB and TtFbpC subunits are highly conserved among other ABC transporters, however, TtFbpA subunit is exclusively conserved in Gram-negative bacteria and of great interests [2].

Previously, two iron coordination modes, six- and five-coordinated, have been observed in the crystal structures of TtFbpA (PDB: 3WAE and 4ELR) at pH 5.5 [3] and 7.5 [4], respectively. In this study, we compared these two crystal structures and found that the different ferric ion coordination modes result from the pH-dependent equilibrium between bicarbonate and carbonate ions and also the conformational change of TtFbpA between open and closed states. These observations are confirmed by the results of quantum calculations for various types of possible ferric iron coordination modes.

[1] Nikaido H (2003) Molecular Basis of Bacterial Outer Membrane Permeability Revisited. Microbiol. Mol. Biol. Rev. 67, 593–656.
[2] Rees DC, Johnson E & Lewinson O (2009) ABC transporters: the power to change. Nat. Rev. Mol. Cell Biol. 10, 218–227.
[3] Wang S., *et al*., (2014) A novel mode of ferric ion coordination by the periplasmic ferric ion-binding subunit FbpA of an ABC-type iron transporter from Thermus thermophilus HB8. Acta Crystallogr. Sect. D Biol. Crystallogr. 70, 196–202.
[4] Wang Q., *et al*., (2013) Structural insights into the coordination of iron by Thermus thermophilus HB8 ferric binding protein A Biochem Biophys Res Commun, 434(1):48-53.

How proteins recognize a phosphorylated amino acid: Comparative studies of an antibody and the other protein families

Raiji Kawade[1] kawade@protein.t.u-tokyo.ac.jp
Daisuke Kudroda[1] d-kuroda@protein.t.u-tokyo.ac.jp
Makoto Nakakido[1] nakakido@protein.t.u-tokyo.ac.jp
Jose Caaveiro[2] jose@phar.kyushu-u.ac.jp
Shigeru Okumura[3] info@abwizbio.com
Toshiaki Maruyama[3] info@abwizbio.com
Kouhei Tsumoto[1,4] tsumoto@bioeng.t.u-tokyo.ac.jp

[1], Department of bioengineering, school of engineering, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
[2] School of engineering, University of Kyusyu, 744 Motooka Nishi-ku, Fukuoka 819-0395, Japan
[3] Abwizbio Inc. 9823 Pacific Heights Blvd., Suite J, San Diego, CA 92121-4705
[4] Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

Keywords: Protein's phosphorylation, Molecular Dynamics simulations, Protein modeling

Protein phosphorylation is one of the major post-translational modifications that occur in the nucleus and cytoplasm of cells. Phosphorylation plays critical roles in a wide range of cellular processes such as signaling and on/off switching of interactions. Because of the important effects of phosphorylation on protein structure and function, the mechanisms by which phosphorylated amino acids mediate protein–protein interactions have been widely investigated [1]. Especially, molecular dynamics (MD) simulations have been adopted to analyze proteins' recognition mechanisms of phosphorylated amino acids [2]. However, most of the previous studies assumed that the protonation state of a phosphorylated amino acid was PO_3^{2-} despite the fact that at physiological pH, the phosphate group (pKa ~7) would exist as an equilibrium mixture of PO_3^{2-} and the singly protonated state (PO_3H^-).

To better understand the molecular mechanism of recognition of a phosphorylated amino acid, we performed MD simulations of 4 different natural protein families that recognize a phosphorylated peptide both in the PO_3^{2-} and PO_3H^- states. Our study suggested PO_3^{2-} was more preferable to PO_3H^- in the interactions due to the larger mobility and higher amount of charge of the phosphate group in the PO_3^{2-} state. Moreover, we evaluated the contribution of each residue to the recognition, and revealed that, among the basic residues, Arg is the more feasible residue than Lys to capture a phosphate group.

Finally, we also obtained a high-affinity (~20 nM) and high-selective monoclonal antibody against a phosphorylated antigen, analyzed its recognition mechanism based on *in vitro* and *in silico* techniques, and compared the results with those of natural protein families. Putting together, these comparative studies suggests strategies to design proteins that can recognize phosphorylated amino acids.

[1] M.B. Yaffe, et al., Phosphoserine/threonine-binding domains, *Current Opinion in Cell Biology*, 13:131–138, 2001.
[2] V.M. Anisimov, et al., Computational and experimental studies of the interaction between phospho-peptides and the C-terminal domain of BRCA1, *Journal of Computer-Aided Molecular Design*, 11: 1071-1084, 2011.

Non-empirical Coarse-grained Simulations for Lipid Bilayers and Proteins

Koji Okuwaki[1] okuwaki@rikkyo.ac.jp
Eiji Shinsho[2] m1813@hoshi.ac.jp
Hideo Doi[1] hideo-doi@rikkyo.ac.jp
Yuji Mochizuki[1, 3] fullmoon@rikkyo.ac.jp
Kaori Fukuzawa[2, 3] k-fukuzawa@hoshi.ac.jp
Etsuo Yonemochi[2] e-yonemochi@hoshi.ac.jp

[1] Department of Chemistry, Rikkyo University, 3-34-1 Nishi-Ikebukuro, Toshima-ku, Tokyo, 171-8501, Japan
[2] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa, Tokyo, 142-8501, Japan
[3] Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan

Keywords: FMO-DPD simulation, Lipid bilayers, Protein modeling

<Introduction> Recently, various researches and developments are actively conducted for designing lipids or proteins in the context of drug delivery systems (DDS) and sensors mimicking the cell membranes. For that purpose, detailed views of the structures of membrane (e.g. distribution, irregularity, and fluidity) should be useful, however such information is hardly obtainable through only experimental observations. Computational approaches, e.g. molecular dynamics (MD) and dissipative particle dynamics (DPD) [1], could thus be effective as complements. Recently, we have been developing a multiscale simulation method that the effective interaction parameters for DPD are to be evaluated by the fragment molecular orbital (FMO) [2] calculation, as the FMO-DPD scheme. [3,4] In this presentation, we report the results of FMO-DPD applied to the mixed lipid bilayers and the protein folding.

<Mixed lipid bilayers> DPD simulations were performed for a mixed lipid model of positively charged lipid (DOTAP) and neutral phospholipid (DPPC) used for DDS. As a result, we reproduced the experimental results that the two lipids were heterogeneously mixed and the shape of liposome became oblate spheroid. [5]

<Protein folding> As a touchstone along the structural reproduction of proteins, we performed non-empirical DPD simulation for chignolin, [6] the smallest protein consisting of 10 residues. As shown in the Figure 1, the backbone of hairpin curve shape was formed. We will also report the results of several proteins.

This research was supported from MEXT FS2020 project priority issue 6 and Grant-in-Aid for Scientific Research (16H04635).

[1] R. Groot et al., J. Chem. Phys. 107 (1997) 4423.
[2] "The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems", (2009, CRC).
[3] K. Okuwaki et al., J. Phys. Chem. B, 122 (2018) 338.
[4] H. Doi et al., Chem. Phys. Lett. 684 (2017) 427.
[5] E. Shinsho et al., J. Comp. Chem. Jpn., in press.
[6] S. Honda et al., Structure, 12 (2004)1507.

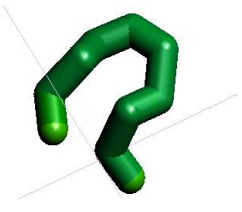


Figure 1. The DPD simulation of chignolin.

Energy analysis of mixtures of associated liquids and non-associated liquids using microcalorimetry and molecular dynamics simulation

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

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

Keywords: microcalorimetry, excess enthalpy, molecular dynamics

Molecular recognition and discrimination may be caused by the interactions among contacting surfaces of colliding molecules in solutions and mixtures. In particular, stereospecific interactions due to neighbouring surfaces may play the leading role in, for examples, enzyme-substance reactions, antigen-antibody reactions, some kinds of mechanisms of the senses of smell and taste, etc. The influences of stereospecific interactions in the liquid state have been determined by microcalorimetry and reported[1-5]. To understand interactions of associated liquids and non-associated liquids, the excess enthalpies of 1-butylamine + 1-butanol, 1-butylamine + 2-butanol, 2-butylamine + 1-butanol, 2-butylamine + 2-butanol, t-butylamine + 1-propanol, t-butylamine + 2-propanol, 2-butylamine + 1-propanol, 2-butylamine + 2-propanol, i-butylamine + 1-propanol, i-butylamine + 2-propanol, 1,2-propanediamine + 1,2-propanediol, 1,2-propanediamine + 1,3-propanediol, 1,3-propanediamine + 1,2-propanediol, 1,3-propanediamine + 1,3-propanediol, 1,2-propanediamine + 1-propanol, 1,2-propanediamine + 2-propanol, 1,3-propanediamine + 1-propanol, 1,3-propanediamine + 2-propanol, 1-propylamine + 1,2-propandiol, 1-propylamine + 1,3-propandiol, 2-propylamine + 1,2-propandiol, 2-propylamine + 1,3-propandiol have been measured at 298.15 K using a twin-microcalorimeter. All excess enthalpies were exothermic and large. In order to clarity that molecular structure effect on amine-alcohol interactions, molecular dynamics simulations were performed.

[1] Kimura, T., Suzuki, K., Takagi, S. Excess Enthalpies of o- and m- Isomers of Difluorobenzene, Dichorobenzene, Dibromobenzene,+ Methyl Methylthiomethy Sulfoxide or + Dimethyl Sulfoxide at 298.15 K, *Fluid Phase Equilib*., 136: 269-278, 1997.
[2] Kimura, T., Suzuki, K., Takagi, S. Excess Enthalpies of Alkylbenzene + Methyl Methyl-thiomethy Sulfoxide or + Dimethyl Sulfoxide at 298.15 K, *Thermochim. Acta*, 328: 55-64, 1999.
[3] Kimura, T., Matsushita, T., Kamiyama, T. Excess Enthalpies of Alkane-1-amines {$C_nH_{2n+1}NH_2$, n= 3 to n=8} + Methyl Methlthiomethyl Sufoxide, + Dimethyl Sufoxide at 298.15 K, *Thermochim. Acta*, 405: 129-139, 2003.
[4] Fujisawa, M., Kimura, T., Enthalpy and entropy changes on molecular inclusion of 1-heptanol into α- and β-cyclodextrin cavities in aqueous solutions
Thermochim. Acta, 416(1-2): 51-54, 2004.
[5] Kimura, T., Fujie. S., Yukiyama, T., Fujisawa, M., Kamiyama, T., Hatsumi, A., Enthalpy and entropy changes on molecular inclusion of pentane derivatives into α-cyclodextrin cavities in aqueous solutions, *J. Incl. Phenom. Macrocycl. Chem*. 70(3-4): 269–278, 2011.

In silico modeling of PAX8–PPARγ fusion protein with unknown three-dimensional structure in follicular thyroid neoplasm

Kaori Sakaguchi[1]
sakaguchi@eniac.scitec.kobe-u.ac.jp

Yoshio Okiyama[2]
okiyama@nihs.go.jp

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

[1] Department of Computational Science, Graduate School of System Informatics, Kobe University, 1-1, Rokkodai, Nada, Kobe, Hyogo 657-8501, Japan
[2] Division of Medicinal Safety Science, National Institute of Health Sciences, 3-25-26 Tonomachi, Kawasaki-ku, Kanagawa 210-9501, Japan

Keywords: Follicular neoplasm, Fusion protein, Protein modeling, Molecular dynamics simulation

Follicular thyroid neoplasms include benign follicular adenoma and follicular thyroid carcinoma. Notably, a gene fusion between paired box 8 (PAX8) and peroxisome proliferator-activated receptor γ (PPARγ) results in the production of a PAX8–PPARγ fusion protein (PPFP) in the tumor cells. PPFP is oncogenic and the PPARγ/PPFP ligand pioglitazone is highly therapeutic in a transgenic mouse model of PPFP thyroid carcinoma [1]. Furthermore, pioglitazone may be therapeutic in patients with PPFP [2]. Except for these issues, the functions of PPFP and its action mechanisms are not well understood and the three-dimensional (3D) structures of PPFP have not been reported so far. Thus, elucidating the 3D structures of PPFP is strongly required to understand the specific function and oncogenic mechanism of PPFP for drug discovery.

In this study, we propose plausible 3D structures of PPFP (see Figure below) by improving the modeling method reported at CBI 2017 meeting. We use the amino acid sequence of PPFP consisting of PAX8 (Met1–Ala396) and PPARγ (Met1–Tyr505) [1]. The PPFP that binds to a DNA sequence was constructed as follows: We first extended the structure from PAX8's paired domain and PPARγ at the both ends of PPFP whose crystal structures are known, and finally linked them. To validate the predicted PPFP models, we analyzed their dynamical stabilities with molecular dynamics simulation and investigated their various physicochemical characteristics.

[1] Giordano, T. J., *et al.*, Pioglitazone Therapy of PAX8-PPAR γ Fusion Protein Thyroid Carcinoma, *J. Clin. Endocrinol. Metab.*, 103: 1277–1281, 2018.
[2] Raman, P., and Koenig R.J., Pax-8-PPAR γ Fusion Protein in Thyroid Carcinoma, *Nat. Rev. Endocrinol*, 10:616–623, 2014

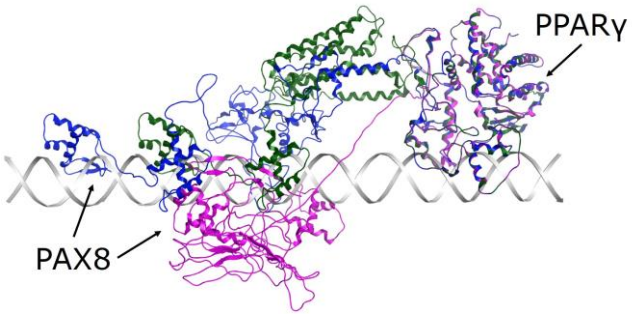


Figure: Superposition of three model structures of PPFP

FMO Study on the Effects of Phosphorylation in Janus Kinase (JAK) on Ligand Binding

Takahiro Hirata[1] thirata@eniac.scitec.kobe-u.ac.jp
Shusuke Tokutomi[1] tokutomi@eniac.scitec.kobe-u.ac.jp
Hirofumi Watanabe[2] hirofumi.watanabe@port.kobe-u.ac.jp
Kaori Fukuzawa[3] k-fukuzawa@hoshi.ac.jp
Shigenori Tanaka[1] tanaka2@kobe-u.ac.jp

[1] Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan
[2] Education Center on Computational Science and Engineering, Kobe University 7-1-48, Minatojimaminamimachi, Chuo-ku, Kobe 650-0047 Japan
[3] Department of Physical Chemistry, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa, Tokyo 142-8501, Japan

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

Immunosuppressive drugs such as steroids have been used for a long time, but their targets are broad and have many side effects. This is why molecular targeted drugs are expected, which act only on a specific molecule and can be ingested orally with less side effects than the classical inhibitor. Kinases are drawing attention as targets of these molecular targeted drugs, and among them, Janus kinase (JAK) inhibitors are noted in the treatment of rheumatoid arthritis.

There are two kinds of complex structures of JAK and ligands: a state in which its own tyrosine (TYR) residue is phosphorylated and that has a phosphorylated tyrosine (PTR) residue, and a state without it. However, the effect of PTR residues on interaction with ligands are not well known.

Here, the interaction energies between JAK family and ligand molecules were calculated by the Fragment Molecular Orbital (FMO) method, focusing on the Inter-Fragment Interaction Energies (IFIEs). In this study, JAK with PTR residues and that with TYR changed from PTR were comparatively analyzed. It was found that the IFIEs between the ligand and the residues which are important for ligand binding such as aspartic acid (ASP) and lysin (LYS) varied according to the presence or absence of PTR. In addition, sometimes the histidine (HIS) residue can be protonated and its conformation may be modified, so that the electrostatic interaction due to the negative charge of PTR influences many residues. The IFIE-sum of the phosphorylated structure and that of the non-phosphorylated structure were found to have a fair correlation, except for the discrepancy of IFIE between the PTR / TYR site and the ligand. Correlation between the inhibition constant K_i and the IFIE-sum was investigated, but it was difficult to obtain good correlation. Thus, there is room for further study as to whether and how to perform FMO calculations with or without PTR.

This study was performed in the framework of the FMO drug design consortium (FMODD). We thank Dr. Fumio Nakajima (Carna Biosciences, Inc.) for useful discussions. The results of FMO calculations were obtained using the K computer (project ID: hp170183 and hp180147).

P2-01

Design of Anti-Cancer Peptides with Counterpropagation Artificial Neural Networks

Miyabi Hishinuma[1] hishinuma.m.aa@m.titech.ac.jp
Francesca Grisoni[2,3] francesca.grisoni @pharma.ethz.ch
Claudia Neuhaus[2] claudia.neuhaus@pharma.ethz.ch
Gisela Gabernet Garriga[2] gisela.gabernet@pharma.ethz.ch
Jan A. Hiss[2] jan.hiss@pharma.ethz.ch
Masaaki Kotera[4] maskot@chemsys.t.u-tokyo.ac.jp
Gisbert Schneider[2] gisbert@ethz.ch

[1] School of Life Science and Technology, Tokyo Institute of Technology, 1-11-5, Midorigaoka, Meguro-ku, Tokyo, 152-0034, Japan
[2] ETH Zurich, Department of Chemistry and Applied Biosciences, Vladimir-Prelog-Weg 4, CH-8093 Zurich, Switzerland
[3] University of Milano-Bicocca, Dept. of Earth and Environmental Sciences, piazza della Scienza 1, 20126 Milano, Italy
[4] Department of Chemical System Engineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan

Keywords: Anti-Cancer Peptides, Machine Learning, Pharmacological Activity Prediction

Anti-Cancer Peptides (ACPs) are drawing attention as potential cancer therapeutics because of their pharmacological activity. The cancer cell selectivity of certain membranolytic ACPs originates from their ability to detect differences between the plasma membrane composition and structure of healthy cells and cancer cells, which means it is difficult for cancer cells to develop resistance against drugs.[1] However, the structure-activity relationships of ACPs are not fully understood yet. Effective computational approaches may help to identify new ACPs.[1] We demonstrated a rational strategy to design ACPs with counterpropagation artificial neural networks (CP-ANNs),[2] which is a predictive model based on self-organizing maps.[3] We constructed a virtual library of 1000 potentially alpha-helical peptides sequences using modLAMP, a Python library for generating antimicrobial peptides.[4] The generated peptides were subjected to virtual screening with CP-ANNs models developed to recognize ACPs. 21 peptides picked from the virtual library were chemically synthesized and tested for their cellular anticancer activity in bioassays.

[1] G. Gabernet, A. T. Müller, J. A. Hiss, G. Schneider, Membranolytic anticancer peptides. Med. Chem. Commun. 7:2232–2245, 2016.
[2] R. Hecht-Nielsen, Counterpropagation networks. Appl. Opt. 26:4979–4983, 1987.
[3] D. Ballabio, M. Vasighi, A MATLAB toolbox for Self Organizing Maps and supervised neural network learning strategies. Chemometrics and Intelligent Laboratory Systems. 118:24–32, 2012.
[4] A. T. Müller, G. Gabernet, J. A. Hiss, G. Schneider, modlAMP: Python for antimicrobial peptides. Bioinformatics (Oxford, England). 33:2753–2755, 2017.

P2-02

A newly developed method based on AI-oriented amino acid interaction mapping (AI-AAM) for efficient virtual scaffold hopping

Kyosuke Tsumura, Ph.D.[1]
kyosuke.tsumura@fujifilm.com

Shino Ohira, Ph.D.[1]
shino.ohira@fujifilm.com

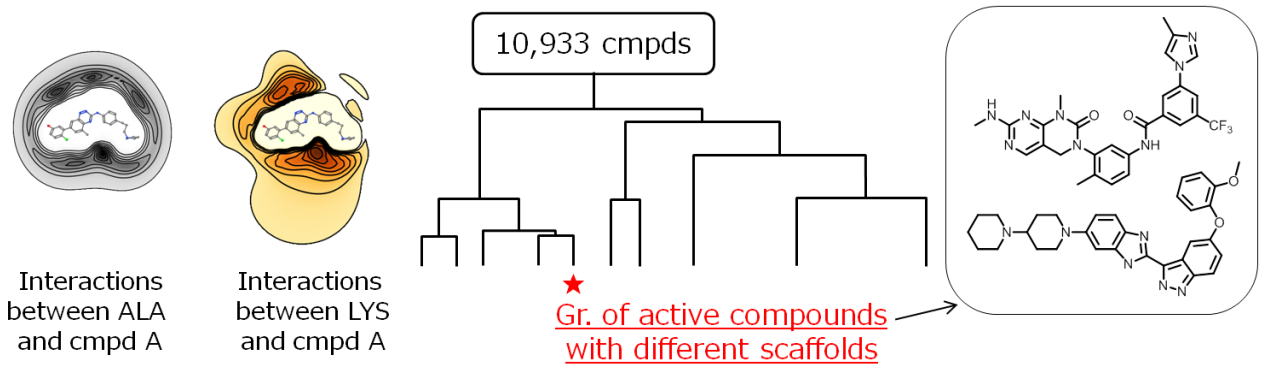
[1] Analysis Technology Center, R&D Management Headquarters, Fujifilm Corporation, 210 Nakanuma, Minamiashigara-shi, Kanagawa 250-0193, Japan

Keywords: Electronic structure calculation, Ligand docking, Protein modeling

When we acquired a seed or lead compound in drug discovery, the candidate often dropped out during in vivo/vitro testing. We present a new methodology called AI-AAM to obtain multiple candidate compounds to increase our success rate of new drug development.

For scaffold hopping, we needed a descriptor that could represent the binding affinities of a ligand to its target protein. In order to find the descriptor, we assumed that protein-ligand binding could be described as the set of interactions between a ligand and amino acids. Using quantum chemical and molecular dynamic simulations[1], we computed the interactions between a ligand and amino acids to create a new descriptor, Amino Acid Mapping[2] (AAM).

To understand the ability of AAM, we tested the new descriptor with public data sets (DUD-E: ABL1, hit ratio = 1.7%)[3]. After computing AAM on each compound, we had classified them based on Euclidean distance using Affinity Propagation[4]. Out of 221 groups, 183 hit compounds were efficiently classified into small number of groups. It was shown that the potency of binding affinities of a compound was related to AAM. When we compared the chemical structures in the “hit” group, we could find the completely different scaffolds. In addition to the public data sets, we applied AI-AAM to anti-cancer and antibiotic drug discovery projects and could find new scaffolds with the same order of IC_{50} or MIC. The hit ratio was about 10%. We developed AI-AAM which enabled the efficient ligand-based virtual scaffold hopping.



[1] (a) Gaussian 09, M. J. Frisch, et al., 2009, (b) AMBER 11, D.A. Case, et al., 2011.
[2] Patent pending.
[3] DUD-E, http://dude.docking.org/.
[4] Brendan J. Frey; Delbert Dueck (2007). "Clustering by passing messages between data points". *Science*. **315** (5814): 972–976.

P2-03

Structure Information Management System in Asahi Kasei Pharma

Kazufumi Ohkawa[1]
okawa.kf@om.asahi-kasei.co.jp

Kenichiro Takaba[1]
takaba.kb@om.asahi-kasei.co.jp

Kensuke Akaogi[2]
akaogi.kb@om.asahi-kasei.co.jp

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

Keywords: X-ray structure, FMO, database system

Nowadays, many laboratories have already accumulated a large number of in house X-ray data (protein-ligand complex structures), which aren't properly managed (assigned ID or stored in a database).

Adding that situation, Fragment Molecular Orbital (FMO) and Molecular Dynamics (MD) have become popular methods for understanding the nature of interaction between protein and ligand.

These structure related data are desirable to be integratedly managed, so we developed SIMS (Structure Information Management System) which enable researchers to enter data, track analysis report, and search protein-ligand interaction.

[1] Kitaura, K., Ikeo, E., Asada, T., Nakano, T., and Uebayasi, M., Fragment Molecular Orbital Method: an Approximate Computational Method for Large Molecules, *Chem. Phys. Lett.*, 313:701– 706, 1999.
[2] Salentin,S. *et al.*, PLIP: fully automated protein-ligand interaction profiler, *Nucl. Acids Res*, 43 (W1): W443-W447, 2015.
[3] Gert-Jan Bekker. *et al.*, Molmil: a molecular viewer for the PDB and beyond, *Journal of Cheminformatics*, 8:42, 2016.

P2-04

Implementation of protein-ligand docking engine sievgene_M for many- and multi-core processors

Takanori Sugihara[1,2] t-sugihara@aist.go.jp **Takashi Kurosawa**[3,2] **Hironori K. Nakamura**[4]

Tadaaki Mashimo[5,2] **Yoshifumi Fukunishi**[6,2] **Haruki Nakamura**[7]

[1] Japan Biological Informatics Consortium
[2] Technology Research Association for Next Generation Natural Products Chemistry
[3] Hitachi Solutions East Japan, Ltd.
[4] Biomodeling Research Co., Ltd.
[5] IMSBIO Co., Ltd.
[6] National Institute of Advanced Industrial Science and Technology (AIST)
[7] Institute for Protein Research, Osaka University

Keywords: Ligand docking, in silico screening, high-performance computing

We propose a method to accelerate in-silico screening and decrease the cost required to discover pharmaceutical candidate chemical compounds. We develop a multithreaded high-performance version of the protein-ligand docking engine for better performance. The new docking engine is called sievgene_M. By the subscript M, we mean that it is a version customized for the state-of-the-art many- and multi-core processors such as Intel's Xeon Phi 7210 and E5, respectively. It is easy to use sievgene_M in many- and multi-core environment [1,2]. One can just compile and execute it with the Intel's or GNU's Fortran compilers like any other application software.

In drug development, it is quite important to decrease labor to find candidate compounds. Usually, experimental methods are used for ligand screening. In those methods, one can analyze target protein-ligand systems in detail but the cost increases in proportional to labor.

In-silico screening is a useful automated method based on high-performance computing. In this method, we can evaluate millions of protein-ligand interactions for short time in an economical way. We explain the details of the method to parallelize the docking engine for many- and multi-core processors.

The docking engine sievgene_M is a part of the molecular simulation and analysis software suite myPresto. The newest version of myPresto has just been released through the Internet. Source codes of sievgene_M can be downloaded at the myPresto portal website and used for free [3].

[1] T. Sugihara, T. Kurosawa, H. K. Nakamura, T. Mashimo, Y. Fukunishi, and H. Nakamura, in preparation.
[2] Y. Fukunishi, Y. Mikami, H. Nakamura, J. Mol. Graph. Model. 24, 34-45, 2005.
[3] https://www.mypresto5.jp/

P2-05

Bandit Ensemble FMO for Protein-Ligand Binding Affinity Predictions

Kenichiro Takaba[1] takaba.kb@om.asahi-kaesi.co.jp
Kazufumi Ohkawa[2] okawa.kf@om.asahi-kasei.co.jp

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

Keywords: Fragment orbital method, Molecular dynamics, Machine learning, Binding free energy

Fragment orbital method (FMO) has emerged as a powerful computational tool in structure-based drug discovery. Pair interaction energy decomposition analysis (PIEDA) enables detailed analysis of protein-ligand interactions and many studies has shown that its interaction energies can be used to predict protein-ligand binding affinities [1,2]. However, the accuracy is not sufficient enough to be applied for lead optimization.

To overcome this difficulty, we introduce ensemble FMO where MD simulations are used to generate multiple protein-ligand complexes and FMO calculations are performed against ensembles of conformers with explicit water molecules. To assess ensemble FMO, the correlations between experimental and calculated binding affinities of two systems, internal project A and Pim1 kinase, were examined. Single conformer-based FMO results of Pim1 reported by Watanabe and coworkers were used as reference in this study [2].

The correlations between experimental pIC_{50} values and PIEDA binding energies which were calculated from MM-optimized single conformers under vacuum conditions were R^2=0.09 and R^2=0.53 for internal project A and Pim1, respectively. In contrast, the application of ensemble FMO was able to improve the correlation for internal project A (R^2=0.67), whereas the correlation was the same for Pim1 (R^2=0.53). If combinations of different PIEDA energies were allowed, then the best correlation between the experimental pIC_{50} values and ensemble FMO binding energies were R^2=0.76 and R^2=0.62 for internal project A and Pim1, respectively.

Furthermore, inspired by the work from Terayama and coworkers [3], we discuss the application of a machine learning method, Best Arm Identification, where unpromising compounds are avoided at an early stage and the limited amount of computational resources are allocated to more promising compounds. This approach is expected to maximize the performance of ensemble FMO where promising compounds are chosen from a set of candidate compounds within a limited amount of computational resource.

[1] Heifetz, A., *et al*., The Fragment Molecular Orbital Method Reveals New Insight into the Chemical Nature of GPCR-Ligand Interactions, *J. Chem. Inf. Model.* **2016**, 56, 159-172.
[2] Watanabe, C., *et al*., Theoretical Analysis of Activity Cliffs among Benzofuranone-Class Pim1 Inhibitors Using the Fragment Molecular Orbital Method with Molecular Mechanics Poisson-Boltzmann Surface Area (FMO+MM-PBSA) Approach, *J. Chem. Inf. Model.* **2017**, 57, 2996-3010.
[3] Terayama, K., *et al*., Machine learning accelerates MD-based binding pose prediction between ligands and proteins, *Bioinformatics* **2018**, 34, 770-778.

P2-06

Rapid and Accessible *In-silico* Macrocycle Design

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

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

Matthias Bauer[1]
matthias@cresset-group.com

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

[1]Cresset, New Cambridge House, Bassingbourn Road, Litlington, Cambridgeshire, SG8 0SS, UK

Keywords: Macrocyclization, Bioisostere replacement

An increasingly common medicinal chemistry technique is conformational restriction through macrocyclization, in order to attain higher affinity and selectivity for the target coupled to improved oral bioavailability.[1] Although the concept is simple, the execution is difficult: the proposed linker must have enough flexibility that it can join the proposed cyclization sites without introducing too much steric strain, but not so much that the entropic benefits of cyclization are lost. It must be synthetically feasible, must fit into the available space in the active site, and ideally should make favorable (or at least not unfavorable) interactions with the protein.

Macrocyclization can be seen as a special case of fragment linking, which in turn is a constrained form of fragment growing. Fragment linking strategies have been recently reported as a highly successful route to lead optimization.[2] In both linking and macrocyclization the design problem is to find a moiety which enforces the required geometry between the two link sites.

We present a modification of Cresset's bioisostere searching tool, Spark, to this problem. Traditional bioisostere searches specify a fragment to replace in the starting ligand and look for similar fragments in a database. Instead, Spark assesses bioisosteric replacements in product space, which allows more complex experiments. A modification of the Spark scoring method to include similarity to other ligands known to bind in the region of the linker, as well as constraints from the protein active site and known pharmacophoric requirements, allows sensible ranking of potential linkers.

Here we apply this technique to several data sets, including a set of recently-reported pyridine-based BRD4 inhibitors.[3] Results were obtained that are identical or very similar to reported BRD4 macrocycle inhibitors. Also, the distribution of generated linker sizes was in good agreement with experimental SAR data.

[1] Marsault, E.; Peterson, M. L. Macrocycles Are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery, *Journal of Medicinal Chemistry*, 54:1961-2004, 2011.

[2] Mondal, M.; Radeva, N.; Fanlo-Virgós, H.; Otto, S.; Klebe, G.; Hirsch, A. K. H. Fragment Linking and Optimization of Inhibitors of the Aspartic Protease Endothiapepsin: Fragment-Based Drug Design Facilitated by Dynamic Combinatorial Chemistry, *Angewandte Chemie International Edition*, 55: 9422-9426, 2016.

[3] Wang, L.; McDaniel, K. F.; Kati, W. M. Fragment-Based, Structure-Enabled Discovery of Novel Pyridones and Pyridone Macrocycles as Potent Bromodomain and Extra-Terminal Domain (BET) Family Bromodomain Inhibitors. *Journal of Medicinal Chemistry*, 60: 3828-3850, 2017.

P2-07

A web-based application for the visualization and exploration of the SAR *matrix*

Atsushi Yoshimori[1]
yoshimori@itmol.com

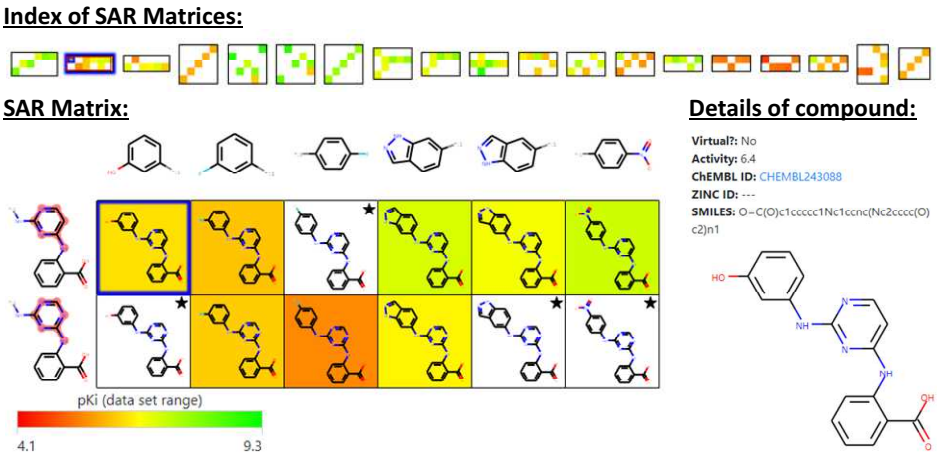
Tohru Tanoue[2]
tanoue.toru@infogram.co.jp

Jürgen Bajorath[3]
bajorath@bit.uni-bonn.de

[1] Institute for Theoretical Medicine, Inc., 4259-3 Nagatsuta-cho, Midori-ku, Yokohama, 226-8510, Japan
[2] Infogram, Inc., 2-17-19 Hakataekimae, Hakata-ku, Fukuoka 812-0011, Japan
[3] Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universität, Endenicher Allee 19c, D-53115 Bonn, Germany

Keywords: Structure-activity relationships, Matched molecular pair, Matching molecular series, SAR *matrix*

SAR *matrix* is designed to elucidate SAR information associated with groups of structurally related active compounds, which are extracted from large compound data sets [1]. The SAR *matrix* data structure identifies and organizes compound series produced by structurally analogous matching molecular series [1, 2]. SAR matrices contain not only active compounds but also many virtual compounds generated by the matched molecular pair formalism [3]. These virtual compounds are expected to be attractive candidates for hit-to-lead transformation and lead optimization. Accordingly, a chemically intuitive and interpretable user-interface for the SAR *matrix* method is expected to substantially aid in practical medicinal chemistry applications.



We are developing a web-based application for visualizing and exploring SAR *matrices* using modern javascript libraries. This new application supports medicinal chemists in SAR analysis and lead optimization.

[1] Wassermann, A.M., Haebel, P., Weskamp, N., and Bajorath, J., SAR matrices: automated extraction of information-rich SAR tables from large compound data sets, *Journal of Chemical Information and Modeling*, 52:1769-1776, 2012.
[2] Wawer, M., and Bajorath, J. Local structural changes, global data views: graphical substructure-activity relationship trailing. *Journal of Medicinal Chemistry*, 54:2944−2951, 2011.
[3] Kenny, P.W., and Sadowski, J., Structure modification in chemical databases. *In Chemoinformatics in Drug Discovery*, Oprea, T. I., Ed.; Wiley-VCH: Weinheim, Germany, 271−285, 2005

Peptide Docking Method using a Coarse-grained Potential

Sho Matsumoto[1] matsumoto@cch.cs.tut.ac.jp
Ryosuke Tsukamoto[1] tsukamoto@cch.cs.tut.ac.jp
Hitoshi Goto[1] gotoh@tut.jp

[1] *Computational Chemistry Laboratory, Computer Science and Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi, 441-8580, Aichi, Japan*

Keywords: Protein-peptide complex, Docking, Coarse-grained potential

A method for accurately evaluating the recognizing ability of a specific peptide to a target protein plays a crucial role in research and development for new drug discovery such as antibody or peptide medicines. In order to perform this *in silico* evaluation, it is necessary to find the peptide binding site on the surface of the target protein and evaluate binding energy between protein and peptide. However, the high affinity between them is possible to give a number of binding site candidates. In this presentation, we will introduce our peptide docking method that we have developed for accurately and quickly search of some plausible protein-peptide binding sites [1].

In this method, the binding site is scored by applying Krishnamoorthy potential [2] to tetrahedral four-body relationships between the protein surface and peptide described by is a coarse-grained protein model. We also adopted Defpol [3] and MSMS [4] methods that evaluate solvent contacting surfaces, to put search grid points for placing amino acid residues of peptides on the protein surface. As the exploring algorithm, we tried the elite tree search and Monte Carlo tree search. We used 28 known crystal structures of protein-peptide complexes as a standard set for verifying the performance of our methods. Furthermore, it is confirmed that the performance of our methods is comparable to that of AutoDock Vina, a famous docking system using all-atom model.

[1] R. Tsukamoto, H. Enomoto, H. Goto. 39th Symposium on Chemoinformatics, P11, 2016.
[2] C. Deutsch, B. Krishnamoorthy, Bioinformatics, 19, 1540-1548, 2003.
[3] C. S. Pomelli, J. Tomasi. J. Comput. Chem, 19, 1758-1776, 1998.
[4] M. Sanner, A. J. Olson, J. C. Spehner, Biopolymers, Vol 38, (3), 305-320, 1996.

P2-09

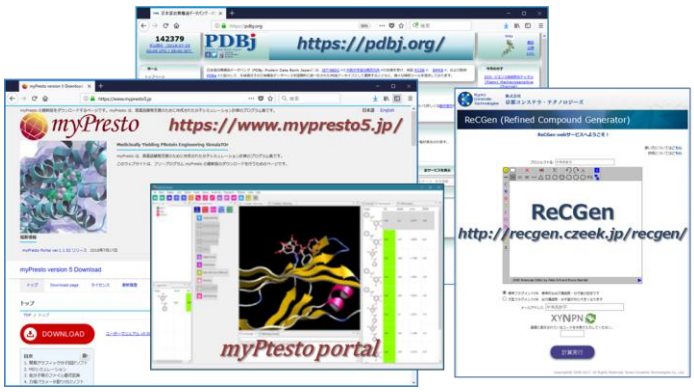
Current status of myPresto computer-aided drug development suite

Tadaaki Mashimo[1] mashimo-tadaaki@aist.go.jp
Yoshifumi Fukunishi[1,2] y-fukunishi@aist.go.jp
Takeshi Kawabata[3] kawabata@protein.osaka-u.ac.jp
Takashi Kurosawa[1,4] takashi.kurosawa.uh@hitachi-solutions.com
Chisato Kanai[1,5] kanai@k-ct.jp
Haruki Nakamura[3] harukin@protein.osaka-u.ac.jp

[1] Technology Research Association for Next Generation Natural Products Chemistry, 2-3-26, Aomi, Koto-ku, Tokyo 135-0064, Japan
[2] Molecular Profiling Research Center for Drug Discovery (molprof), National Institute of Advanced Industrial Science and Technology (AIST), 2-3-26, Aomi, Koto-ku, Tokyo 135-0064, Japan
[3] Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan
[4] Hitachi Solutions East Japan, Ltd., Kawasaki Ekimae Tower Riverk, 12-1 Ekimae-honcho, Kawasaki-ku, Kawasaki-shi, Kanagawa-ken 210-0007, japan
[5] Kyoto Constella Technologies Co., Ltd., 4th Flr., Kyozome Kaikan, 481 Tourouyama-cho, Nakagyo-ku, Kyoto, 604-8225, Japan

Keywords: myPresto5, GUI, PDBx/mmCIF, Molecular structure generation, ΔG estimation

"myPresto" [1] is a computer software suite for drug developments equipped with molecular dynamics (MD) simulation, molecular docking simulation, virtual compound screening, etc. myPresto is one of the few commercial-free MIT/free-BSD like computer programs and all the source codes are available with user's and programmer's guides. myPresto5 was released at Feb. 2018 and has been updated every month. In the version upgrade, a graphical user interface (GUI) program "myPresto portal" was newly added to myPresto. The Protein Data Bank has introduced new file format (PDBx/mmCIF) in addition to the conventional PDB format. To accept this new format, myPresto provides the file conversion tools. ReCGen (by Kyoto Constella Technologies Co., Ltd.) can generate new drug-like molecular structures based on compound databases. We renewed some programs including our compound database [2]. In addition, we have developed the docking-score QSAR method [3] which is binding-free energy (ΔG) estimation method using ChEMBL database. The docking-score-QSAR method approximate the ΔG values with a linear combination of docking scores against multiple protein pockets of the query compound in question. The accuracy was about Q^2 = 0.7-0.8 (average error = 1 kcal/mol) in 4-fold cross validation tests. The training session of myPresto will be opened [4].



[1] myPresto home page: https://www.mypresto5.jp/
[2] LigandBox database: http://www.mypresto5.com/ligandbox/cgi-bin/index.cgi?LANG=ja
[3] Fukunishi, Y., Yamashita, Y., Mashimo, T. and Nakamura, H. Prediction of Protein compound Binding Energies from Known Activity Data: Docking-score-based Method and its Applications, *Molecular Infomatics*, 37:1700120, 2018.
[4] See also <CBI2018 Sponsor's workshop> Biomodeling Research Co., Ltd. on Oct. 10 (Wed).

P2-10

Molecular Block Inserting: A Simple and Fast Algorithm for Efficient Generation of Ring Conformations

Hajime Sugiyama[1]
sugiyama.hajime@jp.fujitsu.com

Akihiko Ueda[1]
a.ueda@jp.fujitsu.com

Azuma Matsuura[1, 2]
matsuura.azuma@jp.fujitsu.com

Atsushi Tomonaga[1]
atomonaga@jp.fujitsu.com

[1] Fujitsu Limited, 1-2-70 Konan, Minato-ku, Tokyo 108-0075, Japan
[2] Fujitsu Laboratories Ltd., 10-1 Wakamiya, Morinosato, Atsugi, Kanagawa 243-0124, Japan

Keywords: Molecular Mechanics Calculation, Conformation Search

Macrocycles, which are more than 12-membered ring molecules, has been emerging modality as therapeutics agents, expanding drug target spaces which had not been covered with small molecules and large biotics. The novel modality hides the possibility of causing amazing success of drug development because of its flexibility and diversity being inherent in molecular size and ring topology.

In contrast with those advantages, the use of rational design for macrocycles consisting of structural based drug design techniques and molecular simulations has been somehow limited yet. Both greater degree of freedom and complicated ring closure condition are prone to make it difficult to predict their stable conformation, not to mention their binding structure to a target protein.

Here, the authors propose a simple and fast method to generate ring molecule conformations in a reasonable time. It is based on an asymptotic molecular progression; a molecular block is inserted between two bonded atoms of a ring molecule in a lower energy conformation, and then the entire structure of the molecule is energy-minimized. Finally that of the target molecule conformations is generated.

Besides an introduction to the method, a result of its application is also shown.

Comprehensive Protein-Ligand Interaction Analysis: FMO Calculation on the complexes of a Human Protease Renin and its Inhibitors

Yoichiro Yagi[1] yagi@ee.ous.ac.jp
Takatomo Kimura[2] kimura@konankako.co.jp
Chiduru Watanabe[3] chiduru.watanabe@riken.jp
Hirotomo Moriwaki[3] hirotomo.moriwaki@riken.jp
Yoshio Okiyama[4] okiyama@nihs.go.jp
Shigenori Tanaka[5] tanaka2@kobe-u.ac.jp
Teruki Honma[3] honma.teruki@riken.jp
Kaori Fukuzawa[6] k-fukuzawa@hoshi.ac.jp

[1] Department of Electrical and Electronic Engineering, Okayama University of Science, 1-1 Ridai-cho, Kita-ku, Okayama 700-0005, Japan
[2] Konan Chemical Industry Co., Ltd., 5-12 Nakagawa-cho, Takatsuki, Osaka 569-0066, Japan
[3] RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[4] National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan
[5] Department of Computational Science, Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe, Hyogo 657-8501, Japan
[6] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-Ku, Tokyo 142-8501, Japan

Keywords: Renin, Inhibitor, Fragment molecular orbital (FMO) calculation, Inter-fragment interaction energy (IFIE), Visualized cluster analysis (VISCANA)

It is well known that a human protease, renin, is the first component of the renin-angiotensin system in hypertension. In our continuous efforts, we have performed the fragment molecular orbital (FMO) calculation on the renin-inhibitor complexes at FMO2-MP2/6-31G* level by using ABINIT-MP/BioStation program [1]. We found that a strong correlation between the binding energies of renin and inhibitors and the activity value on 50% inhibitory concentration (IC_{50}) of the inhibitors [2,3]. The aims of this study are to classify inhibitors from a viewpoint of the interaction between renin and inhibitor, such as a value of interaction energy, interacting amino acid residues, etc., and to make good use of our computational methods to *in silico* drug design. In this research, we examine the feature of the inter-fragment interaction energies (IFIEs) between renin and inhibitors that obtained from FMO calculation on the 20 structures of renin-inhibitor complexes by using visualized cluster analysis (VISCANA) method implemented in BioStationViewer.

This research was done in activities of the FMO drug design consortium (FMODD). The results were obtained using the K computer (project ID: hp180147). Pair interaction energy decomposition analysis (PIEDA) calculation was carried out by using MIZUHO/BioStation software package.

[1] Tanaka, S., Mochizuki, Y., Komeji, Y., Okiyama, Y., and Fukuzawa, K., Electron-correlated fragment-molecular-orbital calculations for biomolecular and nano systems, *Physical Chemistry Chemical Physics*, 16:10310-10344, 2014.
[2] Yagi, Y., Kimura, T., Watanabe, C., Okiyama, Y., Tanaka, S., Honma, T., and Fukuzawa, K., Fragment Molecular Orbital Calculations on the Interaction and Binding Energies of Human Protease Renin and its Inhibitors, *CBI Annual Meeting 2016*, P1-08, 2016.
[3] Yagi, Y., Kimura, T., Watanabe, C., Okiyama, Y., Tanaka, S., Honma, T., and Fukuzawa, K., The Role of Water Molecules in Protein-Ligand Binding: Fragment Molecular Orbital Calculations on the Complexes of Renin with its Inhibitors, *CBI Annual Meeting 2017*, P2-08, 2017.

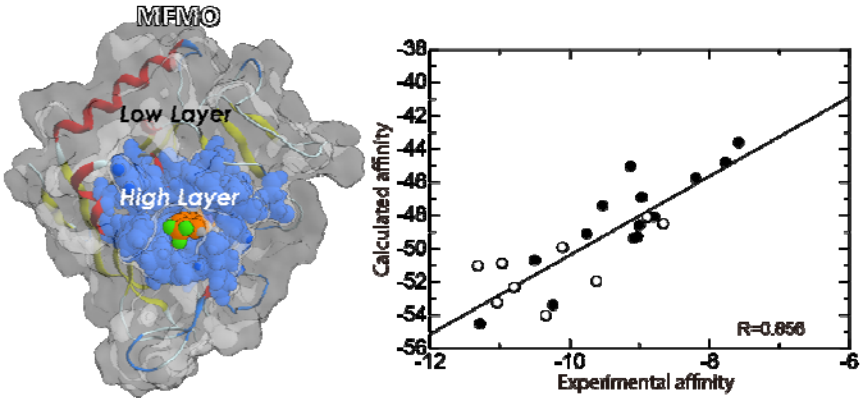
Use of the multilayer fragment molecular orbital method to predict the rank order of protein-ligand binding affinities: A case study using tankyrase2 inhibitors

Noriaki Okimoto[1] **Takao Otsuka**[1] **Yoshinori Hirano**[1] **Makoto Taiji**[1]
okimoto@riken.jp totsuka@riken.jp hirano@riken.jp taiji@riken.jp
[1] RIKEN, Center for Biosystems Dynamics Research, Building B, 6-2-4 Furuedai, Suita, Osaka 565-0874, Japan

Keywords: affinity prediction, molecular docking, molecular mechanics calculation, quantum chemical calculation

In computational drug discovery, ranking a series of compound analogues in the order that is consistent with the experimental binding affinities remains a challenge. Many of the computational methods available for evaluating binding affinities have adopted molecular mechanics (MM) based force fields, although they cannot completely describe protein–ligand interactions. In contrast, quantum mechanics (QM) calculations play an important role in understanding the protein-ligand interactions; however, their huge computational costs hinder their application in drug discovery.

In this study[1], we evaluated the ability to rank the binding affinities of tankyrase2 ligands by combining both MM and QM calculations. Our computational approach uses the protein-ligand binding energies obtained from a cost-effective multilayer fragment molecular orbital (MFMO) method[2,3] combined with the solvation energy obtained from the molecular mechanics-Poisson-Boltzmann/surface area method (MM-PB/SA) to predict binding affinity. This approach enabled us to rank tankyrase2 inhibitor analogues, outperforming several MM-based methods, including rescoring by molecular docking and the MM-PB/SA method alone. Our results show that this computational approach using the MFMO method is a promising tool for predicting the rank order of the binding affinities of inhibitor analogues.



References

1. Okimoto, N.; Ohtsuka, T.; Hirano, Y.; Taiji, M. ACS OMEGA 2018, 3(4), 4475-4485.
2. Fedorov, D. G.; Ishida, T.; Kitaura, K. J Phys Chem A 2005, 109(11), 2638-2646.
3. Otsuka, T.; Okimoto, N.; Taiji, M. J Comput Chem 2015, 36(30), 2209-2218.

Prediction of the synthetic accessibility of organic compounds using computational technology

Kentaro Takai[1]
takai.kentaro@jp.fujitsu.com

Seiji Fukui[1]
fukui.seiji@jp.fujitsu.com

Akihiko Ueda[1]
a.ueda@jp.fujitsu.com

Atsushi Tomonaga[1]
atomonaga@jp.fujitsu.com

[1] Fujitsu Limited, 1-2-70 Konan, Minato-ku, Tokyo 108-0075, Japan

Keywords: synthetic accessibility, retrosynthetic analysis, reaction scheme

To know the synthetic accessibility of the candidate compounds as drugs is important in drug design procedure. For computer-derived compounds which are often difficult to synthesize, it is especially important to know the accessibility before selecting more promising ones from the massive compounds. In our IT-based drug design (IT-Soyaku), compounds are designed using OPMF (**O**ptimum **P**acking of **M**olecular **F**ragments)[1] and the binding affinity between a target protein and a compound is evaluated using MAPLECAFEE (**MA**ssively **P**aral**LE**l **C**omputation of **A**bsolute Binding **F**re**E** **E**nergy)[2]. We can get a massive number of designed compounds, however it is difficult for synthetic chemists to select synthetically feasible compounds in a short time. Therefore, an automatic synthetic accessibility predicting system is highly expected (Figure 1).

Our developed system includes three functions, i.e., retrosynthetic analysis to decompose a target compound to starting compounds, reagent searching from commercially available or in-house compound DB, and scoring the synthetic accessibility. Criteria of compounds to be synthetically accessible are as follows, (i) all decomposed compounds exist in the DB. (ii) the score is below the allowable value determined by chemists empirically. In order to improve accuracy, we used not only reaction schemes defined by rearrangement data of atom bonding status before and after reactions but also limitation rules to applicability of the scheme to avoid inconvenience such as side reactions in the retrosynthetic analysis. In the system, we used our own method to decompose the compounds and to select reactions[3].

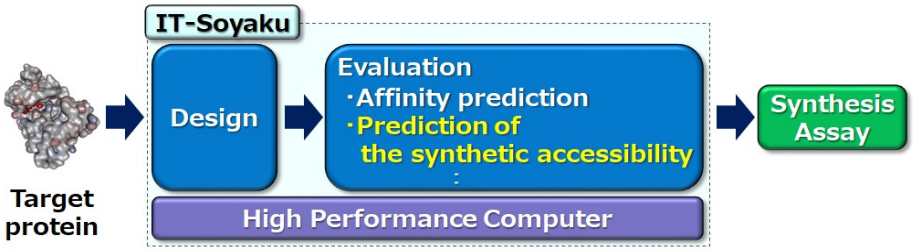


Figure 1. Flow of IT-Soyaku.

[1] T. Yamashita, A. Ueda, T. Mitsui, A. Tomonaga, S. Matsumoto, T. Kodama, H. Fujitani, Molecular dynamics simulation-based evaluation of the binding free energy of computationally designed drug candidates: importance of the dynamical effects, *Chem. Pharm. Bull.* 62:661-667, 2014.

[2] W. Mitchell and S. Matsumoto, Large-scale integrated super-computing platform for next generation virtual drug discovery, *Curr. Opin. Chem. Biol.* 15:553-559, 2011.

[3] A. Tomonaga and K. Kasai, Expression of reaction, reaction data base and recording medium of reaction information, Japanese Patent No. JP-10/218813-A (1998).

P2-14

A Concept of Automated Lead Optimization Method by Compound Property Enhancement and Learning to Rank

Nobuaki Yasuo[1,2] **Keisuke Watanabe**[1] **Naoki Arai**[1]
{yasuo, watanabe, arai}@cbi.c.titech.ac.jp

Hideto Hara[3] **Kentaro Rikimaru**[3] **Masakazu Sekijima**[1,4]
{hideto.hara, kentaro.rikimaru}@takeda.com sekijima@c.titech.ac.jp

[1] Department of Computer Science, Tokyo Institute of Technology, J3-23, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan
[2] JSPS Research fellow DC1, J3-23, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan
[3] Shonan Research Center, Takeda Pharmaceutical Company Limited, 2-26-1, Muraokahigashi, Fujisawa Kanagawa 251-8555, Japan
[4] Advanced Computational Drug Discovery Unit, Tokyo Institute of Technology, J3-23, 4259, Nagatsuta, Midori-ku Yokohama, Kanagawa 226-8503, Japan

Keywords: Lead optimization, Drug discovery, Machine Learning, Property enhancement

Lead optimization is an important step in drug discovery, where the chemical structures of drug candidates are modified to improve efficacy, selectivity, physicochemical properties, and toxicity. There is much demand for the automation of lead optimization due to the difficulty and the cost of the process.[1] Here we show a concept of automated lead optimization method, which outputs optimized compounds from input compounds. The algorithm is basically the iterative "exploration" and "evaluation" of compounds. The exploration is the process to find next candidate compounds based on the property enhancement system, which is previously reported by our group.[2] The evaluation is the process to score the candidates, which can be accomplished by the prediction of the drug-likeness. However, previous drug-likeness prediction is not optimized to distinguish the difference in the same series of compounds. In this study, we show the concept of automated lead optimization system and examined the evaluation module using learning to rank methods. The dataset is the actual compounds used in drug discovery projects, which have been previously synthesized at Takeda Pharmaceutical Company Limited. We predicted the order of synthesis of compounds. The advantage of this method is that all factors involved in the optimization were implicitly considered in the order of synthesis. We compared six models of learning to rank methods.

[1] S. M. Paul, D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg, and A. L. Schacht, How to improve R&D productivity: the pharmaceutical industry's grand challenge, *Nature reviews Drug discovery*, 9.3:203–214, 2010.
[2] N. Arai, Y. Shunsuke, N. Yasuo, R. Yoshino, M. Sekijima, Compound property enhancement by virtual compound synthesis, *Journal of Bioinformatics and Computational Biology* (In Proc. GIW2017), doi.org/10.1142/S0219720018400164, 2018.

Establishment of a Method for Analyzing Transcriptional Regulatory Mechanisms of Genes Grouped by Orthogonal Linear Separation Analysis (OLSA)

Shotaro Maedera[1] mdera0130@gmail.com
Tadahaya Mizuno[1] tadahaya@mol.f.u-tokyo.ac.jp
Setsuo Kinoshita[1] setsuo.kinoshita@gmail.com
Hiroyuki Kusuhara[1] kusuhara@mol.f.u-tokyo.ac.jp

[1] Laboratory of Molecular Pharmacokinetics, Graduate School of Pharmaceutical Sciences, The University of Tokyo, Japan

Keywords: bigdata analysis, *in silico* drug analysis, computational pharmacology

Recently, the technologies to obtain omics data have been advancing and studies utilizing omics data analysis have clarified various biological phenomena. Previously, we developed a new omics data analysis method, Orthogonal Linear Separation Analysis (OLSA), which contracts numerous variables of omics data to mathematical vectors including biological responses based on factor analysis. With OLSA, we analyzed the transcriptome data of breast cancer cell line (MCF-7) responses to 318 drugs obtained from Connectivity Map and contracted the cell responses to the 118 vectors. We attributed 47% of the vectors to biological annotations and there were vectors including autophagy induction effects, estrogen like effects, and others, which was consistent with literature reports and *in vitro* analysis. However, the underlying molecular mechanisms of vectors generated by OLSA are unclear because the method is unsupervised.

To understand the molecular mechanisms, we devised a method to discover transcription factors regulating the coordination of gene expression changes which characterize each vector. This method is divided into the following two steps: (i) extracting characteristic genes from the vector, and (ii) calculating the affinity between the DNA sequence around the transcription start site of the genes and transcription factors with 941 position weight matrices obtained from JASPAR automatically. We analyzed the 118 vectors with the method and detected 85 relationships between the vectors and transcription factors including HSF1, EGR, p53, and IRF1. The result was consistent with literature reports and especially the drugs which strongly changed the expression levels of coordinated genes corresponding to the p53-related vector were mostly reported to activate p53. We performed real time PCR and western blotting of MCF-7 stimulated with the drugs which were not reported to activate p53 but predicted to have a p53 activating effect, and several of them actually increased the expression levels of p53 activation markers.

We succeeded in constructing the automatic method to elucidate the underlying molecular mechanisms of vectors generated by OLSA. This pipeline will contribute to understanding pharmacological properties of drugs.

[1] Lamb J., Crawford E. D., Peck D., Modell J. W., Blat I. C., Wrobel M. J., Lerner J., Brunet J. P., Subramanian A., Ross K. N., Reich M., Hieronymus H., Wei G., Armstrong S. A., Haggarty S. J., Clemons P. A., Wei R., Carr S. A., Lander E. S., and Golub T. R.., *The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease.*, Science, 2006

Statistical analysis of inter- and intramolecular interactions for drug design based on FMO database

Chiduru Watanabe[1] chiduru.watanabe@riken.jp
Daisuke Takaya[1] daisuke.takaya@riken.jp
Shunpei Nagase[1] shunpei.nagase@riken.jp
Kikuko Kamisaka[1] kikuko.kamisaka@riken.jp
Kaori Fukuzawa[1,2] k-fukuzawa@hoshi.ac.jp
Teruki Honma[1] honma.teruki@riken.jp

[1] RIKEN Center for Biosystems Dynamics Research 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[2] School of Pharmacy and Pharmaceutical Sciences, Hoshi University 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan

Keywords: interaction energy database, FMO, ligand binding, molecular recognition

We developed “FMO database” [1] to collect >2000 entries of fragment molecular orbital (FMO) calculation [2] data using X-ray crystal structures, NMR structures, and MD snapshots obtained from the activities of FMO drug design consortium (FMODD) [3]. This database contains data of inter- and intramolecular interaction energies and atomic charge, etc., based on quantum calculations.

In this work, we statistically examined distribution of inter-fragment interaction energies (IFIEs) which were classified by fragment charge, type of amino acid residues, and energy components such as electrostatic, charge transfer, and dispersion interaction energies by pair interaction energy decomposition analysis (PIEDA). The purpose of this research is to reveal an essential interaction(s) for the recognition of ligand and protein by the statistical analysis of IFIEs. In addition, we analyzed the distributions of characteristic interaction energy values such as hydrogen bond and CH/π interactions. It is expected that criteria for interaction energy value to judge whether or not interaction that is essential for the molecular recognition and the construction of tertiary structure will be established. The FMO database is now opening to FMODD members; however it will be released to the public by the end of fiscal year 2018.

Acknowledgement

This research was done in activities of the FMODD [3]. The results of FMO calculations were obtained using the K computer (project ID: hp180147). PIEDA calculations were done by using MIZUHO/BioStation software package. This research was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP18am0101113. Finally CW and DT acknowledge JSPS KAKENHI Grant Number JP18K06619.

[1] URL of FMO database: http://drugdesign.riken.jp/FMODB
[2] Tanaka, S., Mochizuki, Y., Komeiji, Y., Okiyama, Y., Fukuzawa, K., Electron-Corrected Fragment Molecular Orbital Calculations for Biomolecular and Nano Systems, *Phys. Chem. Chem. Phys*. 16:10310-10344, 2014.
[3] http://eniac.scitec.kobe-u.ac.jp/fmodd/index.html

Development of FMO database and its recent updates

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

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

Keywords: FMO, interaction energy database, protein-ligand interaction.

In structure-based drug design, it is important to understand protein-ligand interactions such as hydrogen bonds, electrostatic interactions and the van der Waals interactions. For further understanding of more detailed interactions, our group has been focusing on the quantum mechanics (QM) which calculate the effects of donating and withdrawing electrons and can appropriately deal with the cation-π, CH-π and π-π interactions. Generally, QM calculation takes highly computational cost for applying biological macromolecules such as protein-ligand complex structures. Fragment molecular orbital (FMO) method [1] enables us to efficiently perform *ab initio* QM calculations by many-body expansion technique with fragmentation. The fragmentation technique of FMO method leads to calculate fragment-fragment interaction energies, which are useful for analyzing protein-ligand interactions. The pairwise fragment-fragment interaction energies are called as pair interaction energy (PIE) or inter-fragment interaction energy (IFIE). The FMO method is widely applied to analysis of the protein-ligand interaction. [2, 3]

Although IFIE data calculated by the FMO method provides rich information about the interactions and will be useful for researchers in other fields, FMO calculation settings such as preprocess of an input structure, parameter tunings are difficult for not experienced scientists. Many computer resources are needed to calculate many protein-ligand complexes. From the viewpoint of big data, it is expected that IFIE between each amino acid residue and a ligand molecule may be used as descriptors for statistical approaches such as linear regression and machine learning for protein-ligand interaction analysis. Accumulation of large scale of IFIE data is important for the purpose. In that situation, our group is developing FMO database [4], for providing IFIE data. The FMO database provides user friendly interfaces to access IFIE data and summarize IFIE and PIEDA data from ligand binding point of view. In this poster, we introduce recent research progress and development of the FMO database.

Acknowledgement

This research was done in activities of the FMO drug design consortium [5]. The results of FMO calculations were obtained using the K computer (project ID: hp180147). PIEDA calculations were done by using MIZUHO/BioStation software package. This research was partially supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP17am0101001. Finally DT and CW acknowledge JSPS KAKENHI Grant Number 18K06619.

1. Kitaura, K. *et al. Chem. Phys. Lett.* **1999**, *313*, 701-706.
2. Takaya, D. *et al. Bioorg. Med. Chem.* **2018** *submitted*
3. Watanabe, C *et al. J. Chem. Inf. Model.* **2017**, *57*, 2996-3010
4. FMODB URL: http://drugdesign.riken.jp/FMODB
5. http://eniac.scitec.kobe-u.ac.jp/fmodd/index.html

Subtype specificity analysis of estrogen receptor using fragment molecular orbital method

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

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

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

In order to clarify subtype specificity between the estrogen receptor (ER) α/β and ligands, we revealed the interaction between the ligand and surrounding amino acid residues using the Fragment Molecular Orbital (FMO) method [1]. Structural features of various ligands which affect subtype selectivity were revealed, based on similarity of binding mode with ER.

We calculated all the available PDB structures of the ER-ligand complex at the FMO2-MP2/6-31G* level. We analyzed the interaction energies (IFIE) and their energy component (PIEDA) between the ligand and surrounding amino acid residues. Solvent effects were considered by the FMO + MM-PBSA method [2]. We also performed clustering of ligands by similarity of interactions with receptor using the VISCANA analysis [3]. The ABINIT-MP program and K computer were used for FMO calculations, and Amber program was used for solvation free energy calculations.

By comparing the interaction between ligand and amino acid residues of ER, it was found that the interaction with mutated residues among subtypes (Met336Leu and Ile373Met) were different, and it was also suggested that such a difference affects the selectivity of subtypes. In addition, it was revealed that both hydrogen bond with Glu305 and CH/π interaction with Phe356 were essential for ligand binding of ERβ, and stable interaction with Met336 and His475 would be related to β selectivity. Such FMO based specification of residues with key interaction is considered to be useful for drug design.

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

[1] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayashi, Chem. Phys. Lett., 313:701–706, 1999.
[2] C. Watanabe, H. Watanabe, K. Fukuzawa, L.J.Parker, Y. Okiyama, H. Yuki, S. Yokoyama, H. Nakano, S. Tanaka, and T. Honma, J. Chem. Inf. Model., 57:2996-3010, 2017.
[3] S. Amari, M. Aizawa, J. Zhang, K. Fukuzawa, Y. Mochizuki, Y. Iwasawa, K. Nakata, H. Chuman, and T. Nakano, J. Chem. Inf. Model., 46:221–230, 2006.

Application of molecular dynamics simulation in drug design: case study of tankyrase2

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

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

Keywords: Molecular docking, Molecular dynamics, multiple receptor conformations

In early stage of drug discovery research, computational screening approach is widely used to find hit compounds. Molecular docking is one of the computational methods for the compound screening. This method is generally docks a flexible ligand to a specific (static) protein structure and predicts the binding pose and binding affinity for each ligand in a practical time. It is well known that proteins fluctuate dynamically under physiologic conditions, and this flexibility is essential for ligand binding. Therefore, the ability of molecular docking using a specific ligand binding is not always effective for prediction of binding pose and binding affinity. To resolve these problems, incorporation of protein flexibility into molecular docking is required. Molecular docking using multiple receptor conformations ("ensemble docking" [1-3]) is one of approaches to consider protein flexibility. These multiple receptor conformations can be obtained by X-ray crystallographic structures and/or structures sampled by molecular dynamics (MD) simulations. However, there still remains difficulty on proper generation and selection of multiple receptors used for ensemble docking.

In this study, we performed several MD simulations of tankyrase2 with and without ligand and investigated the conformational changes of the ligand binding pockets. As a result, we found that MD simulations of apo X-ray structures induced reduction of ligand binding pocket, on the other hand, those of holo structures showed different behavior of the binding pocket. Furthermore, we analyzed these characteristic binding pocket conformations and evaluated performance of molecular docking using these multiple conformations.

References:
[1] Totrov M. and Abagyan R., Current Opinion in Structure Biology, 2008, 18:178-184
[2] Barril X. and Morley S. D., J. Med Chem. 2005, 48: 4432-4443.
[3] Knegtel R. M. *et al.*, J. Mol. Biol. 1997, 266:424–40.

Bioisosteric conversion based on electrostatic potential

Shunpei Nagase[1] shunpei.nagase@riken.jp
Daisuke Takaya[1] daisuke.takaya@riken.jp
Chiduru Watanabe[1] chiduru.watanabe@riken.jp
Hitomi Yuki[1] hitomi.yuki@riken.jp
Yoshio Okiyama[1,2] okiyama@nihs.go.jp
Teruki Honma[1,3] honma.teruki@riken.jp

1 RIKEN Center for Biosystems Dynamics Research, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
2 Division of Medicinal Safety Science, National Institute of Health Sciences, 3-25-26 Tonomachi, Kawasaki-ku, Kanagawa 210-9501, Japan
3 RIKEN Cluster for Science and Technology Hub, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Keywords: quantum mechanics, electrostatic potential, electrostatic similarity, lead optimization

In lead optimization, a structure should be modified to improve its pharmacokinetic properties while maintaining the activity. Several studies have reported *in silico* screening using the electrostatic potential (ESP) obtained from a quantum mechanical (QM) calculation[1,2]. However, there are few researches on chemical structural conversion systems utilizing accurate ESP based on a QM method.

In this research, we newly developed Tscore that characterizes the similarity of ESP for this chemical structural conversion system. Each ESP was calculated from a substructure obtained by removing the maximum common structure, and expressed by grid data. A Tscore was calculated by comparing the ESPs at the same coordinates of two compounds. To demonstrate the advantage of this method, we evaluated R^2 between the ESP similarity and activity similarity of compound dataset assayed in the process of search for inhibitors of hematopoietic cell kinase (HCK)[3]. As a result, the R^2 value became higher in the case of the compounds over a certain threshold of Tanimoto Combo[4] that indicates both of the shape and color similarities of two substructures to be compared. It was shown that compounds with higher ESP similarities tend to have similar activities. In addition, we found that if a substructure with Tscore exceeding a certain threshold is selected in a chemical conversion, the decrease in activity can be suppressed. We expect this substructure transformation method based on the QM-ESP approach can reasonably optimize lead structures for drug design.

Acknowledgments: The results were obtained using the "HOKUSAI GreatWave" of RIKEN Head Office for Information Systems and Cybersecurity(ISC).

[1] Muchmore S. W., Souers A. J., and Akritopoulou-Zanze I., The use of three-dimensional shape and electrostatic similarity searching in the identification of a melanin-concentrating hormone receptor 1 antagonist, *Chem. Biol. Drug. Des.*, 67(2):174–176(2006).
[2] Bostrom J., Grant J. A., Fjellstrom O., Thelin A., and Gustafsson D., Potent fibrinolysis inhibitor discovered by shape and electrostatic complementarity to the drug tranexamic acid, *J. Med. Chem.*, 56(8):3273–3280(2013).
[3] Saito Y., Yuki H., Kuratani M., Hashizume Y., Takagi S., Honma T., Tanaka A., Shirouzu M., Mikuni J., Handa N., Ogahara I., Sone A., Najima Y., Tomabechi Y., Wakiyama M., Uchida N., et al., A Pyrrolo-Pyrimidine Derivative Targets Human Primary AML Stem Cells in Vivo, *Sci. Transl. Med.*, 5(181):181ra52(2013).
[4] Hawkins P. C. D., Skillman A. G., and Nicholls A., Comparison of Shape-Matching and Docking as Virtual Screening Tools, *J. Med. Chem.*, 50(1):74–82(2007).

Toward Rational State-Selective Stabilization of GPCR using Computational Protein Design Strategy

Masaya Mitsumoto[1,2] sanbon@ims.ac.jp
Ryosuke Nakano[3] nryosu1995@chiba-u.jp
Takeshi Murata[3,4] t.murata@faculty.chiba-u.jp
Nobuyasu Koga[1,2] nkoga@ims.ac.jp

[1] Exploratory Research Center on Life and Living Systems, National Institutes of Natural Sciences, 38, Nishigohnaka, Myodaiji-cho, Okazaki-shi, Aichi 444-8585, Japan
[2] Department of Structural Molecular Science, Graduate School of Physical Sciences, SOKENDAI, 38, Nishigohnaka, Myodaiji-cho, Okazaki-shi, Aichi 444-8585, Japan
[3] Faculty of Science, Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba-shi, Chiba, 263-8522, Japan
[4] PRESTO, Japan Science and Technology Agency, Kawaguchi Center Building, 4-1-8 Motomachi, Kawaguchi-shi, Saitama, 332-0012, Japan

Keywords: Protein design, GPCR, Intracellular loop 3, Stabilization, Structure infomatics

G-protein coupled receptors (GPCRs) are one of the major drug targets. The atomistic structures are important clue for structure-based drug design. However, obtaining the structures is difficult, due to the structural instability and fragility caused by flexible intracellular loop 3 (ICL3). To solve this problem, the GPCR structures were attempted to be stabilized by domain fusion [1], point mutation [2], antibody binding [3] etc. However, these approaches require many experimental trials and thus very time-and-cost-consuming. In addition, most of solved GPCR structures are in inactive state; active state structures are also important for drug design. Because of these reasons, rational approach to stabilize GPCR in the aimed state needs to be developed.

In contrast to GPCR, *de novo* designed proteins created by our group are extremely stable [4][5]; they formed stable structures even under over 100 degrees C. One of the reasons for the super high stability is possibly because of their loops; we designed the proteins with typical short loop types observed in naturally occurring proteins. By applying this protein design strategy, we computationally redesigned ICL3 from backbone level to stabilize GPCR in the aimed state, targeting adenosine A2A receptors for a proof of concept.

[1] Cherezov *et al.,* High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor, *Science*, 318:1258-1265, 2007.
[2] Vega *et al.,* Conformational thermostabilization of the beta1-adrenergic receptor in a detergent resistant form, *Proc. Natl. Acad. Sci.,* 105:877-882, 2008.
[3] Hino *et al*., G-protein-coupled receptor inactivation by allosteric inverse-agonist antibody, *Nature*, 482:237-240, 2012.
[4] Koga et al., Principles for designing ideal protein structures, *Nature*, 491:222-227, 2012.
[5] Lin *et al.*, Control over overall shape and size in de novo designed proteins, *Proc. Natl. Acad. Sci.,* 112:5478-5485, 2015.

Development of new ranking method based on MMGBSA for Virtual Screening

Naoya Asada[1]
naoya.asada@shionogi.co.jp

Yusuke Sakoh[1]
yusuke.sako@shionogi.co.jp

Kazuya Yasuo[1]
kazuya.yasuo@shionogi.co.jp

Kazunari Hattori[1]
kazunari.hattori@shionogi.co.jp

[1] Shionogi Pharmaceutical Research Center, SHIONOGI & CO., LTD., 3-1-1, Futaba-cho, Toyonaka-shi, Osaka 561-0825, Japan

Keywords: Virtual Screening, MMGBSA, pareto sort,

MMGBSA scoring method has been widely used for ligand virtual screening (VS) because of its moderate accuracy and throughput higher than other scoring methods using molecular dynamics [1]. In spite of its numerous applications for VS, it remains some limits such as the dependency on ligand's molecular weight and the lack of ligand deformation energy contribution.

To overcome above described limits, we developed a new ranking method based on MMGBSA for VS using pareto sort (PS) [2]. PS is one of the ranking algorithms, which has been applied for multi-parameter optimization problem. In this method, six descriptors which represent a variety of ligand properties related to affinity gain were calculated and total rank was computed using PS. We applied the developed method to six protein-ligand systems (MK14, PDE5A, FXa, PARP1 and two in-house targets). In 5 out of 6 systems, our method gave higher hit ratio of active compounds than that of the conventional MMGBSA method, and in remaining one system, hit ratio was almost equivalent in both methods. Furthermore, our method gave smaller molecules compared to the MMGBSA method, because our method overcame the limit of MMGBSA: the dependency on ligand's molecular weight.

In conclusion, we developed a new ranking method for VS, which could expect higher hit rate and molecules suitable for a starting point for drug discovery.

[1] Bernd Kuhn, Paul Gerber, Tanja Schulz-Gasch, and Martin Stahl., *J. Med. Chem.* **2005,** *48,* 4040-4048.

[2] Kalyanmoy Deb, Samir Agrawal, Amrit Pratap, T Meyarivan, *Parallel Problem Solving from Nature PPSN VI*, **2000**, 849-858.

Data Analysis Toolkits of Fragment Molecular Orbital Calculations to Visualize Interaction Energies Using the GUI Plugin for PyMOL

Takaki Tokiwa[1,2] t.tokiwa@dc.tohoku.ac.jp
Shogo Nakano[3] snakano@u-shizuoka-ken.ac.jp
Yuta Yamamoto[4] 17lb019f@rikkyo.ac.jp
Takeshi Ishikawa[5] t-ishi@nagasaki-u.ac.jp
Sohei Ito[3] itosohei@u-shizuoka-ken.ac.jp
Kaori Fukuzawa[6,7] k-fukuzawa@hoshi.ac.jp
Yuji Mochizuki[4,7] fullmoon@rikkyo.ac.jp
Hiroaki Tokiwa[4] tokiwa@rikkyo.ac.jp
Yasuteru Shigeta[8] shigeta@ccs.tsukuba.ac.jp

[1] Department of Chemistry, Graduate School of Science, Tohoku University, 6-3, Aoba, Aramaki, Aoba-ku, Miyagi 980-8578, Japan
[2] Department of Physics, Graduate School of Pure and Applied Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan
[3] School of Food and Nutritional Sciences, University of Shizuoka, 52-1 Yada, Suruga-ku, Shizuoka 422-8526, Japan
[4] Department of Chemistry, Rikkyo University, 3-34-1 Nishi-Ikebukuro, Toshima-ku, Tokyo 171-8501, Japan
[5] Graduate School of Biomedical Sciences, Department of Molecular Microbiology and Immunology, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan
[6] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-Ku, Tokyo 142-8501, Japan
[7] Institute of Industrial Science, The University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-8505 Japan
[8] Center for Computational Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan

Keywords: FMO calculation, Analysis Toolkit, PyMOL Plugin, Ruby, Python, Visualized energy

Data analysis and visualization toolkits for fragment molecular orbital (FMO) calculations were developed by Ruby and Python programs and named “AnalysisFMO” [1, 2]. AnalysisFMO toolkits can assist in analyzing the results obtained by FMO calculation program packages such as GAMESS-FMO, PAICS, and ABINIT-MP [3]. These toolkits include two tools: one is Ruby script to be able to extract inter-fragment interaction energies (IFIEs) or pair interaction energies (PIEs) from output file of the FMO calculation packages, and another is a Python [4] script for a PyMOL plugin [5] to visualize IFIEs or PIEs which displayed on protein structure in the PyMOL Viewer. These plugin scripts can analyze output data using two analysis modes: “All-pairs” and “Selected-pairs” modes. The former mode can analyze and visualize the residue–residue interactions such as domain interaction. The later mode, i.e., the script and plugin of a “Selected pair” mode, can visualize protein–ligand interaction such as protein-drug interactions.

[1] T. Tokiwa, S. Nakano, Y. Yamamoto, T. Ishikawa, S. Ito, K. Fukuzawa, Y. Mochizuki, H. Tokiwa, Y. Shigeta, *in preparation.*
[2] AnalysisFMO: http://dfns.u-shizuoka-ken.ac.jp/labs/proeng/custom20.html
[3] D. Fedorov and K. Kitaura eds., "The Fragment Molecular Orbital Method: Practical Applications to Large Molecular Systems", *CRC Press* (2009).
[4] M. F. Sanner, *J. Mol. Graphics Model.***17**, 57-61 (1999).
[5] PyMOL: http://www.pymol.org/ (accessed August 1, 2018).

Theoretical Interaction Analysis between Muscarinic acetylcholine receptors and Nobiletin by using fragment molecular orbital calculations

Kanade Shimizu[1]
18lb026b@rikkyo.ac.jp

Sundarama Arulmozhiraja[1,2]
raja@rikkyo.ac.jp

Yuta Yamamoto[1]
17lb019f@rikkyo.ac.jp

WOO Je-Tae[3]
jwoo@isc.chubu.ac.jp

Yasuo Watanabe[4]
yasuwat@yok.hamayaku.ac.jp

Shizuo Yamada[5]
yamada@u-shizuoka-ken.ac.jp

Hiroaki Tokiwa[1,2]
tokiwa@rikkyo.ac.jp

[1] Department of Chemistry, Rikkyo University, 3-34-1, Nishi-Ikebukuro, Toshima-ku, Tokyo, 171-8501, Japan
[2] AMED-CREST, Japan Agency for Medical Research and Development (AMED), 1-7-1, Ohte-machi, Chiyoda-ku, Tokyo, 100-0004, Japan
[3] Department of Biological Chemistry, Research Institute for Biological Functions, ChubuUniversity;1200 Matsumoto, Kasugai, Aichi 487–8501, Japan
[4] General Health Medical Center, Yokohama College of Pharmacy, 601 Matano, Totsuka, Yokohama, 245-0066, Japan
[5] Center for Pharma-Food Research, University of Shizuoka, 52-1 Yada, Suruga, Shizuoka, 422-8526 Japan

Keywords: GPCR, FMO calculation, Drug design

In recent years, G protein-coupled receptors (GPCRs) are intensively studied as drug targets. Muscarinic M_1-M_5 acetylcholine receptors (mAChRs) that belong to the GPCR family regulate many vital functions of the central and peripheral nervous systems. Although sequence similarity among the subtypes M_1 to M_5 is very high, only subtype M_1 receptor has been vigorously investigated as a target for the treatment of neurological disorders. Our recent study reveals that nobiletin, a flavonoid isolated from citrus peels, selectively binds and activates the M_1 receptor [1]; however, the structure of the complex (M_1 receptor with nobiletin) and the origin of nobiletin selectivity are unknown. In this study, we aim to elucidate the mechanism of action by which nobiletin, as a food ingredient, selectively binds to the M_1 receptor and to establish drug designing targeting M1 receptor. The structure of M_1 receptor ligated with nobiletin was constructed using homology modeling by following normal procedures. The crystal structure of the activated form of M_2 receptor (PDBID: 4MQT) was considered as the template for this purpose. The protein-ligand binding analysis were made by using *ab initio* Fragment Molecular Orbital (FMO) method at RI-MP2/cc-pVDZ level of theory.

[1] S.Onoue, S.Yamada, *et al., J. Pharm. Sci.,* vol. 11, no. 9, Sep. 2011.
[2] *Molecular Operating Environment (MOE), 2016.08*, Chemical Computing Group Inc.: 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2018.
[3] Ishikawa, T.; Kuwata, K., Fragment molecular orbital calculation using the RI-MP2 method. *Chemical Physics Letters,* **474,** 195-198, 2009.

Interaction analysis of selective JAK inhibitors by the FMO method

Yuma Handa[1] s141205@hoshi.ac.jp
Taiki Machida[1] s151233@hoshi.ac.jp
Shigenori Tanaka[2] Tanaka2@kobe-u.ac.jp
Takayuki Furuishi[1] t-furuishi@hoshi.ac.jp
Kaori Fukuzawa[1] k-fukuzawa@hoshi.ac.jp
Etsuo Yonemochi[1] e-yonemochi@hoshi.ac.jp

1 Hoshi University, 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan
2 Graduate School of System Informatics, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe 657-8501, Japan

Keywords: Janus kinase, Subtype selectivity, Ligand binding, Fragment Molecular Orbital (FMO) method, Interacting energy analysis

The Janus (JAK) family of cytoplasmic protein tyrosine kinases consist of four subtypes; JAK1, JAK2, JAK3 and TYK2. Since these subtypes are target receptors of specific disease, their selectivity becomes important in drug discovery. Clinically available JAK inhibitors have been Zeruzants® (tofacitinib) and Olmient® (baricitinib), where tofacitinib has no selectivity and baricitinib is selective for JAK1 and JAK2. In this study, FMO calculations [1,2] were performed to investigate molecular interactions between each subtype and two JAK inhibitors and to identify interaction for design of more potent inhibitors.

The Molecular operating environment (MOE) was used for structure preparation. The Amber10EHT force field was used for energy minimizations. FMO calculations were carried out at the MP2/6-31G* level by using ABINIT-MP software. Cocrystal structures between tofacitinib and JAK were obtained from Protein Date Bank (PDBID: 3EYG, 3FUP, 3LXK and 3LXN). For the complex structures between baricitinib and each subtype, docking study was performed. We calculated differences of energy component in the pair interaction energy decomposition analysis (PIEDA) [3,4] of each subtype. PIEDA calculation was done by using MIZUHO/BioStation software package.

For tofacitinib, there was no particular difference in binding energy and therefore selectivity was not observed. On the other hand, there were several amino acid residues with different interaction. Electrostatic interaction at glutamic acid (E883 in JAK1 numbering. The same has applied hereafter.), which was mutated to lysine in JAK2 and JAK3, was different due to the difference in charge of amino acid, and JAK 2 and JAK 3 were more stable. G884 of JAK1 is conserved in other subtypes, however, differences of dispersion interaction were observed at the position of the main chain due to the difference of next residues. L1010 of JAK1 was conserved in other subtypes, but there were also differences in dispersion interaction because the number of CH/π interactions was different. In addition, F958 of JAK1, which was mutated to tyrosine in other subtypes, were involved in important interactions, but they were not directly contributed in selectivity. These residues were considered to be important in the design of JAK inhibitors.

This research was done in activities of the FMO drug design consortium (FMODD). The results were obtained using the K computer (project ID: hp180147)

[1] D. G. Fedorov, T. Nagata, K. Kitaura, Phys. Chem. Chem. Phys. 14, 7562 (2012).
[2] S. Tanaka, Y. Mochizuki, Y. Komeiji, Y. Okiyama, and K. Fukuzawa, Phys. Chem. Chem. Phys., 16, 10310 (2014).
[3] D. G. Fedorov, and K. Kitaura, J. Comput. Chem., 28, 222, (2007).
[4] T. Tsukamoto, et al., J. Comput. Chem. Jpn., 14, 1 (2015).

Computational Analysis of Potential Compounds bound to GPR35 by using Fragment Molecular Orbital Calculations

Yuta Yamamoto[1] 17lb019f@rikkyo.ac.jp
Sota Nagashima[1] 15cc147j@rikkyo.ac.jp
Sundaram Arulmozhiraja[1] raja@rikkyo.ac.jp
Takuto Koide[2] bm18018@ns.kogakuin.ac.jp
Miyako Naganuma[2] s115059@ns.kogakuin.ac.jp
Tomohiro Watanabe[2] matz@cc.kogakuin.ac.jp
Kazuyuki Furuta[3] furutak@okayama-u.ac.jp
Tomoko Hashimoto[2] bm18018@ns.kogakuin.ac.jp
Osamu Ohno[2] ohno@cc.kogakuin.ac.jp
Satoshi Tanaka[4] tanaka-s@mb.kyoto-phu.ac.jp
Kenji Matsuno[2] matz@cc.kogakuin.ac.jp
Hiroaki Tokiwa[1] tokiwa@rikkyo.ac.jp

[1] Department of Chemistry, Rikkyo University, 3-34-1, Nishi-Ikebukuro, Toshima-ku, Tokyo, Japan
[2] Department of Chemistry and Life Science, Kogakuin University, 2665-1, Nakanochou, Hachiouji-shi, Tokyo, Japan
[3] Division of Pharmaceutical Sciences, Okayama University, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, 1-1-1, Tsushima-naka, Kita-ku Okayama, Japan
[4] Department of Biopharmaceutics, Kyoto Pharmaceutical University, 5, Misasaginakauchi-cho, Yamashina, Kyoto 607-8414, Japan

Keywords: Fragment molecular orbital calculations, Ligand docking, Protein modeling, and GPCR

In recent years, multi-scale molecular simulation has been wildly applied to study various biological systems like medicinal/clinical target proteins. We studied such targets by using multi-scale molecular simulation combined with *ab initio* Fragment Molecular Orbital (FMO) calculations[1,2]. In the FMO framework, the entire protein system is divided into single-residue fragments and the total properties of large molecular systems are derived in a many-body expansion by combining the properties of fragments.

G protein-coupled receptors (GPCRs) are intensively studied as drug targets. GPR35 has attracted as a target protein for various anti-allergic drug candidates; however, since activity of the agonists highly depends upon the target species, clinical experiments on animals are difficult. Recently, Lodoxamide and Bufrolin have been reported as GPR35 agonists that eliminate species differences[3], but their activity was low. Matsuno *et al.*, on the other hand, succeeded in the development of highly active compounds that eliminate species differences[4]. In this study, we have theoretically constructed the complex structure of the newly synthesized ligands bound to human/rat GPR35, and analyzed the interaction mechanisms using molecular simulations and FMO calculations.

[1] Yamamoto, Y. *et al.,* Molecular association model of PPARalpha and its new specific and efficient ligand, pemafibrate: Structural basis for SPPARMα. Biochem Biophys Res Commun 499, 239-245, 2018.
[2] Arulmozhiraja, S. *et al.,* Comparative Binding Analysis of Dipeptidyl Peptidase IV (DPP-4) with Antidiabetic Drugs - An *Ab Initio* Fragment Molecular Orbital Study. PLoS One 11, e0166275-1-15, 2016.
[3] MacKenzie, A. E. *et al.,* The anti-allergic mast cell stabilizers lodoxamide and bufrolin as the first high and equipotent agonists of human and rat GPR35. Mol Pharmacol 85, 91-104, 2014.
[4] http://nenkai.pharm.or.jp/138/pc/ipdfview.asp?i=2028

P2-27

Site Identification by Ligand Competitive Saturation (SILCS) reproduces experimental binding trends for 31 TrmD ligands

Clinton Threlfall[1] clinton@silcsbio.com **Sirish Kaushik Lakkaraju**[1] **Sunhwan Jo**[1]

Olgun Guvench[1] **Alexander D. MacKerell Jr**[1,2]

[1] SilcsBio, LLC. 8 Market Pl, Baltimore MD 21202 USA
[2] Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, 20 Penn St, Baltimore MD 21201 USA

Keywords: Site-identification, Ligand, GFE, FragMaps

Site-Identification by Ligand Competitive Saturation (SILCS) computational functional group mapping provides insights into the binding preferences of a target protein that can be used qualitatively and quantitatively to drive ligand design. SILCS is a robust structure-based approach that gives information-rich Grid Free Energy (GFE) FragMaps that encompass critical aspects such as protein flexibility and explicit solvation. Here we describe the use of the SILCS approach on tRNA methyltransferase (TrmD) and 31 ligands belonging to two series originally donated by GSK and made publicly-available through Community Structure-Activity Resource (CSAR) and the D3R Database. SILCS-MC sampling of ligands in the field of the FragMaps yields Ligand Grid Free Energy (LGFE) scores. SILCS scoring correctly predicts favorable vs. unfavorable modifications relative to a reference ligand (27/30 predictions correct). Additionally, SILCS FragMaps recapitulate functional group patterns of both series of ligands. This information can be used to drive design and optimization visually.

Consideration of excessive metabolites on dynamical analysis of bacterial secondary metabolic pathways

Daisuke Tominaga[1] Daisuke-tominaga@aist.go.jp
Hideo Kawaguchi[2] kawaguchi_h@port.kobe-u.ac.jp
Yoshimi Hori[2]
Tomohisa Hasunuma[2] hasunuma@port.kobe-u.ac.jp
Chiaki Ogino[3] ochiaki@port.kobe-u.ac.jp

[1] Computational Bio Big-Data Open Innovation Laboratory, National Institute of Advanced Industrial Science and Technology, 3-4-1 Okubo, Shinjuku, Tokyo 169-8555, Japan
[2] Graduate School of Science, Technology and Innovation, Kobe University, 1-1-3 Minatojima, Chuo ward, Kobe ciry 650-8586, Japan
[3] Graduate School of Engineering, Kobe University, 1-1-3 Minatojima, Chuo ward, Kobe ciry 650-8586, Japan

Keywords: Electronic structure calculation, Ligand docking, Protein modeling

Analyses on secondary metabolic pathways that include industrial useful metabolites are often done to optimize production of the useful materials by microorganisms. One of the mostly established analysis for that is Flux Balance Analysis (FBA), however, it can not deal with dynamic changes of concentrations of metabolites that are often essential for industrial production[1]. So that mathematical models for dynamics, such as ordinary differential equation (ODE) systems are introduced[2]. Some chemical molecules, such as ATP, NAD or NADPH and so on, are often omitted or ignored at constructions of ODE system models because of observation difficulties, or the models can not be stable for large systems that consist of much number of metabolites[3]. This omission can improve mathematical stability of the models, and reduce the experimental observation costs. The omission is reliable when dynamic changes of omitted metabolites are negligible because of its' excessiveness.
We tested the influence of the omission on reliabilities of ODE system models of dynamical changes of metabolite concentrations in log-phase of Escherichia coli cultivation.
Concentrations of 131 metabolites are measured at ten time points. Enzymic reactions are searched comprehensively for each metabolite by KEGG search, and we selected 22 out of 131 on which an ODE (the S-system form) can be applicable (number of relating reactions is sufficiently small comparing with the number of sampling time points). Dynamical changes of each metabolite are modeled in two approaches; with and without excessive molecules, ATP, ADP, AMP, NAD+, NADH, NADP+ and NADPH.
The results show that dynamic changes of concentrations of some metabolites can not be modeled with the omission of the excessive molecules, and including excessive molecules improves precision of models.

[1] Palsson, B. O., Systems Biology: Constraint-based Reconstruction and Analysis, Cambridge University Press, 2015.
[2] Palsson, B. O., Systems Biology: Simulation of Dynamic Network States, Cambridge University Press, 2011.
[3] Voit, E.O., Computational Analysis of Biochemical system: A Practical Guide for Biochemists and Molecular Biologists, Cambridge University Press, 2000.

Pathway prediction of natural products by reverse synthetic analysis

Tsubasa Matsumoto[1]
matsumoto.t.ap@m.titech.ac.jp

Kohei Amano[1]
amano.k.ac@bio.titech.ac.jp

Miyabi Hishinuma[1]
hishinuma.m.aa@m.titech.ac.jp

Kimito Funatsu[2]
funatsu@chemsys.t.u-tokyo.ac.jp

Masaaki Kotera[2]
maskot@chemsys.t.u-tokyo.ac.jp

[1] School of Life Science and Technology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan
[2] Department of Chemical System Engineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Keywords: Natural products, Pathway analysis, Reverse synthetic analysis

Considerable amounts of pharmaceuticals were originated from natural products. Drug discovery process typically requires 10-18 years and 200-300 billion dollars, and it is extremely important to decide desirable lead compounds from natural products. Organic synthesis had been almost the only means of drug discovery, especially for compounds with complex molecular skeletons. Recently, synthetic biology approach has been caught attention. The flow of the synthetic biology approach consists of the following three steps; metabolic pathway elucidation, genetic recombination, and clinical trial. In order to effectively produce valuable natural products in synthetic biology, it is essential to understand the relationship between natural products and their biosynthetic machinery[1]. However, the number of natural products whose biosynthetic pathways are unknown, and the number of unknown enzyme genes involved in the biosynthetic pathways are both increasing[2]. Therefore, a method for exhaustively connecting these two types of information has been desired.

In this study, we have developed a reverse synthesis-based method for predicting the biosynthetic pathway of various natural products for the purpose of extracting responsible biosynthetic genes. Compared with the predictive performance presented in last year[3], the current algorithm greatly improved the performance by utilizing biosynthetic building block information[4] to limit the available reaction types. The performance was verified with KEGG's secondary metabolic map, and we predicted the metabolic pathway of unknown compounds.

[1] Moriya Y, Yamada T, Okuda S, Nakagawa Z, Kotera M, Tokimatsu T, Kanehisa M, Goto S. Identification of Enzyme Genes Using Chemical Structure Alignments of Substrate-Product Pairs. J Chem Inf Model. 56(3):510-516, 2016.
[2] Tabei Y, Yamanishi Y, Kotera M. Simultaneous prediction of enzyme orthologs from chemical transformation patterns for de novo metabolic pathway reconstruction. Bioinformatics, 32, i278-i287, 2016.
[3] Matsumoto T, Amano K, Hishinuma M, Kotera M. Prediction of natural biosynthetic pathway by reverse synthetic analysis, *Chem-Bio Informatics Society (CBI) Annual Meeting 2017*, P3-04, 2017.
[4] Amano K, Matsumoto T, Hishinuma M, Funatsu K, Kotera M. Prediction of biosynthetic building blocks in complicated natural products, *Chem-Bio Informatics Society (CBI) Annual Meeting 2018*.

P3-03

Prediction of biosynthetic building blocks in complicated natural products

Kohei Amano[1]
amano.k.ac@m.titech.ac.jp

Tsubasa Matsumoto[1]
matsumoto.t.ap@m.titech.ac.jp

Miyabi Hishinuma[1]
hishinuma.m.aa@m.titech.ac.jp

Kimito Funatsu[2]
funatsu@chemsys.t.u-tokyo.ac.jp

Masaaki Kotera[2]
maskot@chemsys.t.u-tokyo.ac.jp

[1] School of Life Science and Technology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan
[2] Department of Chemical System Engineering, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Keywords: Natural products, Pathway analysis, Chemical graph, Building block

Many natural products are important sources of pharmaceuticals. The identification of their metabolic pathways is beneficial for natural drug discovery. We previously developed methods to deal with unknown metabolic pathways by predicting an enzyme gene or an enzyme protein that catalyzes a specified reaction[1,2]. However, it is still difficult to predict biosynthetic pathways of complicated natural products with consideration of several unknown intermediates. It is necessary for reverse synthetic analysis to identify the starting material of the target product in its biosynthetic pathway.

In this study, we have developed the program to automatically predict a combination of biosynthetic building blocks in complicated natural products. It enabled to reveal the basic metabolites constructing the target product, and showed better interpretive capacity than ever before[3].

Natural products were represented as chemical graphs converted from MDL Molfile. We have created the biosynthetic building block library by extracting frameworks from molecules that are important stating materials or intermediates in the metabolic pathways. RDKit's substructure matching judged whether the target product contained a biosynthetic building block in the library. Finally, possible combinations of biosynthetic building blocks were output. This program correctly predicted about 90% of compounds in biosynthetic pathways of secondary metabolites.

[1] Tabei Y, Yamanishi Y, Kotera M. Simultaneous prediction of enzyme orthologs from chemical transformation patterns for de novo metabolic pathway reconstruction. *Bioinformatics*, 32, i278-i287, 2016.
[2] Moriya Y, Yamada T, Okuda S, Nakagawa Z, Kotera M, Tokimatsu T, Kanehisa M, Goto S. Identification of Enzyme Genes Using Chemical Structure Alignments of Substrate-Product Pairs. *J. Chem. Inf. Model*. 56(3):510-516, 2016.
[3] Amano K, Matsumoto T, Hishinuma M, Kotera M. Prediction of Biosynthetic Basic Parts of Compounds with Complicated Structures, *Chem-Bio Informatics Society (CBI) Annual Meeting 2017*, P3-06, 2017.

P3-04

Morphology-based analysis of myoblasts for prediction of myotube formation

Kei Yoshida[1]
yoshida.kei@a.mbox.nagoya-u.ac.jp

Kiyoshi Ishikawa[2]
ishikawa.kiyoshi@ps.nagoya-u.ac.jp

Kei Kanie[1]
kanie-k@ps.nagoya-u.ac.jp

Kenji Omori[2]
omori-k@ps.nagoya-u.ac.jp

Ryuji Kato[1]
kato-r@ps.nagoya-u.ac.jp

[1] Graduate School of Pharmaceutical Sciences, Nagoya University, Furocho, Chikusa-ku, Nagoya, 464-8601, Japan
[2] Mitsubishi Tanabe Pharma Corporation, 1000, Shida-cho, Aoba-ku, Yokohama, 227-0033, Japan

Keywords: Cellular morphology, morphology-based analysis, myoblast differentiation, cGMP

The development of new drugs depends on the efficiency of drug screening. Phenotype-based screening has attracted interest due to its considerable potency for the discovery of first-in-class drugs[1]. In general, fluorescently-labeled imagery is the leading technique for phenotype-based screening[2], however there are growing requirements to understand total culture profiles, which are unclear from end-point assays.

In this study, we demonstrate the morphology-based cellular evaluation of unlabeled cells to advance myotube formation assays. One of our aims was to study the myogenic differentiation process in C2C12 cells to discern the differences between cellular responses to different medium conditions (serum concentrations and insulin dosages). Our results showed that predictive morphological profiles that strongly correlate with myogenic differentiation can be generated from myotube images, even in the confluent stage. Differentiation rate after 14 days can be quantitatively predicted with highest accuracy from images taken at 0–11.5 days. In addition, for the application of our morphology-based cellular evaluation of C2C12, the effect of cyclic guanosine monophosphate (cGMP) on myogenic differentiation was analyzed. Our results show that the quantitated morphological profile from non-labeled time-course images can be an effective descriptor for analysis of the myotube differentiation screenings.

[1] Moffat, J. G., Vincent, F., Lee, J. A., Eder, J. and Prunotto, M., Opportunities and challenges in phenotypic drug discovery: An industry perspective, *Nat. Rev. Drug Discov*, 16:531–543, 2017.
[2] Singh, S., Carpenter, A. E. and Genovesio, A., Increasing the content of high-content screening: An overview, *journal of biomolecular screening*, 19:640–650, 2014.

P3-05

Development of Image Cell Picker for Cancer Spheroid Analysis

Hirohito Kato[1] katou.hirohito@a.mbox.nagoya-u.ac.jp
Mayu Shibuta[1] shibuta.mayu@e.mbox.nagoya-u.ac.jp
Ryoko Kusakabe[2] ryokokskb@gmail.com
Kei Kanie[1] kanie-k@ps.nagoya-u.ac.jp
Yuji Matsui[2] hmatsui@md.tsukuba.ac.jp
Masumi Yanagisawa[3] yanagisawa@engineeringsystem.co.jp
Toshiyuki Kanamori[4] t.kanamori@aist.go.jp
Shinji Sugiura[4] shinji.sugiura@aist.go.jp
Ryuji Kato[1] kato-r@ps.nagoya-u.ac.jp

[1] Department of Basic Medicinal Sciences, Graduate School of Pharmaceutical Sciences, Nagoya University, Furocho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan
[2] Division of Gastroenterology, Faculty of Medicine, University of Tsukuba, 1-1-1, Ten-noudai, Tsukuba, Ibaraki 305-8575, Japan
[3] Engineering System Co., Ltd., 5652-83, Sasaga, Matsumoto, Nagano 399-0033, Japan
[4] Drug Assay Device Research Group, Biotechnology Research Institute for Drug Discovery, National Institute of Advanced Industrial Science and Technology (AIST), Central 5th, 1-1-1, Higashi, Tsukuba, Ibaraki 605-8565, Japan

Keywords: Image analysis, Cancer spheroid, 3D culture, Morphology

Recently, three dimensionally (3D) cultured cells, such as spheroid, has become one of the important disease models for drug screening. Compared to the two dimensionally culture cells, 3D cultured cells are known to show higher functions that can represent partial biological functions and structures[1]. Therefore, there are growing attentions to establish new assay platforms using spheroids for more functional drug screening. 3D cultured cancer spheroids derived from cancer/tumor cell lines had been expected to lead novel understanding of the complexity of tumor drug responses. However, since there are still rare techniques to understand the functions/responses of 3D cultured spheroids cultured in ECM-derived hydro-gels with high-throughput manner, new quantitative assay platforms have been expected.

Here, we introduce an automated Image Cell Picker to quantitatively screen/study 3D cultured cancer spheroids by combining label-free live morphology-based analysis with photodegradable hydrogel cell isolation technique[2,3]. From the aspect of computer-assisted quantitative analysis of cancer spheroids, we here discuss the effectiveness of using time-course morphological information of cancer spheroids in our system.

[1] Wenzel, C. et al. 3D high-content screening for the identification of compounds that target cells in dormant tumor spheroid regions. Cancer Research 75, doi:10.1158/1538-7445.am2015-317 (2015).
[2] Tamura, M., Yanagawa, F., Sugiura, S., Takagi, T., Sumaru, K., Matsui, H., and Kanamori, T.: Optical cell separation from three-dimensional environment in photodegradable hydrogels for pure culture techniques, Sci Rep, 4, 6 (2014).
[3] Shibuta, M. et al. Imaging cell picker: A morphology-based automated cell separation system on a photodegradable hydrogel culture platform. Journal of Bioscience and Bioengineering, doi:https://doi.org/10.1016/j.jbiosc.2018.05.004 (2018).

P3-06

Adding trans-omics analysis features to TargetMine

Tatsuya Kameyama[1,2] tatsuya@nibiohn.go.jp
Yi-An Chen[1] chenyian@nibiohn.go.jp
Yosui Nojima[1] yosui@nibiohn.go.jp
Takeshi Fujiwara[1] takeshi.fujiwara@nibiohn.go.jp
Tsuyoshi Esaki[1] tsuyoshi-esaki@nibiohn.go.jp
Lokesh P. Tripathi[1] lokesh@nibiohn.go.jp
Mari Ito[1] & mari@nibiohn.go.jp
Kenji Mizuguchi[1,2] kenji@nibiohn.go.jp

[1]Laboratory of Bioinformatics, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8, Saito Asagi, Ibaraki-shi, Osaka 567-0085, Japan
[2] Graduate School of Frontier Biosciences, Osaka University (Building A, B, C) 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan

Keywords: Omics analysis of disease, Clinical Bioinformatics, pathway analysis

A major problem in research on the mechanisms of complicated disease and drug target discovery is that despite the accumulation of various omics data, analysis methods associated with them have not yet been fully established.

To solve this problem, we have tried to establish a system for trans-omics analysis, with the capability of easily connecting and analyzing “three-layer” omics data, i.e., profiles of genes/transcripts, proteins and metabolites.

The integrated database TargetMine[1-3], developed in our laboratory, integrates more than 50 biological public databases including functional annotations for genes, transcripts, proteins and small molecule compounds and has the ability of data mining and statistical analysis.

In this study, we first enhanced TargetMine by incorporating relevant data to facilitate mapping compounds onto metabolic pathways. Then we automated the following analysis flow using the TargetMine API: ⑴Generating a KEGG[4] web site address, which maps known molecules onto the metabolic pathway. ⑵Ranking the disease specificity of the metabolic pathway with differentially expressed omics data by the enrichment analysis. ⑶ Referencing the molecules of other layers that will be involved in a metabolically regulated molecule in the public database. For example, the system can retrieve the associations between genes and enzymes and the relations between metabolites and enzymes. These automated features were applied to published mouse omics data and the results are consistent with the discoveries in the publication.

Our study shows that TargetMine is able to analyze metabolomic data efficiently, and it enables us to grasp relations between pathways. Moreover, it can analyze various omics data integrally.

[1] Chen, Y.A., Tripathi, L.P., Mizuguchi, K. (2011) TargetMine, an Integrated Data Warehouse for Candidate Gene Prioritisation and Target Discovery. *PLoS ONE* 6(3): e17844.
[2] Chen, Y.A., Tripathi, L.P., Mizuguchi, K. (2016) An integrative data analysis platform for gene set analysis and knowledge discovery in a data warehouse framework. *Database (Oxford)*. 2016 Mar 17;2016. pii: baw009.
[3]http://targetmine.mizuguchilab.org
[4]https://www.genome.jp/kegg/kegg_ja.html

P3-07

A Study of Transcription Factor Network to Distinguish the Difference of Normal and Cancer Cells

Bharata Kalbuaji[1]
kalbuaji.b.aa@m.titech.ac.jp

Konagaya Akihiko[1]
kona@c.titech.ac.jp

[1] School of Computing, Tokyo Institute of Technology, 4259 Nagatsuta, Mirodi, Yokohama, 226-8502, Japan

Keywords: cancer, gene regulatory network, transcription factor

Gene regulation hold an important role in understanding how cell behave, including in cancer development. Interaction that happen from gene regulation process can be modeled using a gene regulatory network. In this study, we use transcription factor binding information in the form of position weight matrices (PWM) to generate a gene regulatory network among transcription factors. We applied this technique to five cancer dataset: bile duct cancer (BDC), hepatocyte carcinoma (HCC), colorectal cancer (CRC), lung adenocarcinoma (LUAD), and osteocarcoma (OSSRC).

Although the transcription factor networks obtained from those five cancer dataset are so complicated, we found a common subnetwork among four cancer dataset: BDC, HCC, CRC, and LUAD. We cannot find any common subnetwork that covers OSSRC data. To further investigate the result, we use pathway analysis to understand what kind of pathway is affected from genes that appear in the common subnetwork.

We use two pathway database to check, KEGG pathway and WikiPathway database. From both pathway analysis, several important pathways in cancer development appear. Those are: adipogenesis, circadian rhytm, transcriptional misregulation in cancer, and pathways in cancer. Adipogenesis and circadian rhytm has been known to be involved in cancer development. We also suspect that there is similarity in cancer cell characteristic from four cancer cell because all cell line is rich in adipocyte. So, adipogenesis may hold an important role to discover common mechanism regarding cancer in adipocyte rich cell.

In conclusion, we believe that understand how transcription factor interact with each other hold a key to discover the mechanism in cancer. We also show that it is probable that several cancer cell from different cell line to share similar characteristic if they also share similar normal cell characteristic. This may lead to find a candidate for drug target that can be used for several cancers.

[1] Saurabh Sinha; On counting position weight matrix matches in a sequence, with application to discriminative motif finding, Bioinformatics, Volume 22, e454–e463, 2006.

[2] Xuhua Xia, “Position Weight Matrix, Gibbs Sampler, and the Associated Significance Tests in Motif Characterization and Prediction,” Scientifica, vol. 2012, Article ID 917540, 15 pages, 2012.

[3] Wang, S., Li, X., Xu, M., Wang, J., Zhao, R.C.: Reduced adipogenesis after lung tumor exosomes priming in human mesenchymal stem cells via TGFβ signaling pathway. Molecular and Cellular Biochemistry 435(1-2), 59-66 (2017).

[4] Amemori, S., Ootani, A., Aoki, S., Fujise, T., Shimoda, R., Kakimoto, T., Shiraishi, R., Sakata, Y., Tsunada, S., Iwakiri, R., Fujimoto, K.: Adipocytes and preadipocytes promote the proliferation of colon cancer cells in vitro. American Journal of Physiology-Gastrointestinal and Liver Physiology 292(3), 923-929 (2007).

[5] Savvidis, C., Koutsilieris, M.: Circadian rhythm disruption in cancer biology. Molecular Medicine 18(9), 1 (2012).

Modeling and simulations of the kinetics of antigens and antibodies towards personalized medicine for allergies

Mizuka Komatsu[1] m-komatsu@stu.kobe-u.ac.jp
Takaharu Yaguchi[1,2] yaguchi@pearl.kobe-u.ac.jp

[1] Graduate School of System Informatics, Kobe University, Rokkodai-cho 1-1, Nada-ku, Kobe, 657-8501, Japan
[2] JST PRESTO, K's Gobancho, 7, Gobancho, Chiyoda-ku, Tokyo 102-0076 Japan

Keywords: Allergy, Parameter estimation, Personalized medicine, Differential Algebraic Equation

Recently allergies are drawing attention as systemic diseases. Although whole-body simulations are needed for analysis of such diseases, most of mathematical models of allergies are built typically at the cellular or molecular level. In this talk, by application of physiologically based pharmacokinetic models, we propose a whole-body model of allergies, more precisely, of the kinetics of antigens and antibodies. Physiologically based pharmacokinetic models are models that have been developed in the pharmaceutical field for describing the kinetics of drugs at the whole-body level [1]. As a first step, we have developed a model [2] for an experimental data, thereby exploring how the model works in simulations of allergies. Through our work, we clarify the underlying difficulties and how to address them, which are shortly explained as follows.

Firstly, because the model describes antigen-antibody interactions, the systems of equations consist of differential algebraic equations, which are known to be hard to solve numerically. We show that the equations have a certain property, with which suitable numerical solvers can be designed. This work confirms the reliability of the simulations.

Secondly, we apply the experimental data from some subjects obtained by Husby et. al [3] to our model and estimate parameters for each subject. Shortly speaking, the parameter estimations here reduced to undetermined problems thus parameters are not determined uniquely but form a manifold, which makes the estimations hard. Against the challenge, we propose a method of parameter estimations which is of multi-stage multi-start type to obtain multiple possible parameters.

Thirdly, we propose a method of analysis for the obtained parameters. These parameters are intuitively “samples” from the manifolds, each of which is uniquely determined by each subject. Therefore, by the analysis, the features of the manifolds corresponding to each subject are explored, which must be helpful for investigation of individual physiological or immunological varieties between subjects. Our result indicates that development of personalized medicine for allergies would become possible by our model and simulations.

[1] Peters, S. A., *Physiologically-based pharmacokinetic(PBPK) modeling and simulations : principles, methods, and applications in the pharamaceutical industry*, Wiley, NJ, 2011.
[2] Komatsu, M. and Yaguchi, T., Mathematical modeling of kinetics of antigens and antibodies based on experiments toward analysis of the mechanism of allergies, submitted.
[3] Husby, S., Jensenius, J. C. and Svehag, S.-E., Passage of undegraded dietary antigens into the blood of healthy adults. Quantification, estimation of size distribution, and relation of uptake to levels of specific antibodies, *Scand. J. Immunol.*, 22 83-92, 1985.

P3-09

Three-dimensional Loop Fragment Analysis of Proteins Focused on Neighboring Ligands

Hiroaki Kato [1]
kato @ hiroshima-cmt.ac.jp

Akihiro Hirama [2]
a133364 @ edu.tut.ac.jp

[1] Department of Distribution and Information Engineering, National Institute of Technology, Hiroshima College, 4272-1 Higashino, OsakiKamijima-cho, Hiroshima 725-0231 Japan
[2] Department of Computer Science and Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi, Aichi 441-8580 Japan

Keywords: loop fragment, neighboring ligand, fragment library, structural similarity, protein motif, EF-hand

Many proteins express their functions by interacting with other compound ligands or proteins. It is well known that particular structural feature called motif is closely related to the function. In the present work, we have constructed the protein 3D fragment library focused on the loop region and the neighborhood ligands. It has also applied to the structural feature analysis using known motif patterns.

In the present work, the protein structural data taken from Protein Data Bank (PDB) were employed with corresponding secondary structure information by DSSP program. A loop region is defined as a segment connecting two consecutive secondary structure (helix or strand) that composed four or more residues. We have focused on a peptide fragment that contained the target residue in loop region and its n neighboring residues in N-, C-terminal sides. The loop fragment is characterized with coordinates of the residues, distance matrix between residues, secondary structure information and fragment size (FS=2n+1). In addition, in the bounding sphere, a center is gravity coordinates of each Cα of a fragment, and a radius is the longest Cα distant from the center. The neighborhood ligands are defined as a hetero atom in the bounding sphere. The ligand binding residues are amino acids that bind to the neighborhood ligands. The loop fragment library have been constructed with the representative dataset (29,840 proteins).

We have selected 474 loop fragments from FS = 17 loop fragments library (2,655,070 fragments) with the condition of ligand binding residues number = 5. They were clustered by Ward's method. As the result, maximum cluster was composed of loop fragments that contained the EF-hand motif sequence pattern. In another cluster, we confirmed the calcium-binding site appeared on the N-terminal side of S100 protein family. Moreover, the other binding site was detected that both ends of loop region were strand region. The structural pattern of these loop fragments were different from the structure of the typical EF-hand motif.

In addition, fragment search trial was carried out using the above novel structural motif candidates as query. As the result, the loop fragments without the ligand information in PDB were also detected. Because of the radius of these loop fragments are larger than the radius of query, it is expected that they are binding site in the unliganded state. These results shows the potential applicability of my approach for structural data mining of proteins.

References:

[1] Lewit-Bentley, A., Rety, S., EF-hand calcium-binding proteins, *Current Opinion in Structural Biology*, **10**, 637-643 (2000).
[2] Donato, R., et al., Functions of S100 Proteins, *Current Molecular Medicine*, **13**, 25-57(2013).

Literature-based functional network predictor for Down-Syndrome

Kazunori Hamada[1] **Achyut Sapkota**[2]
j13431@kisarazu.kosen-ac.jp sapkota@ieee.org

[1] Department of Information and Computer Engineering, National Institute of Technology, Kisarazu College, 2-11-1, Kiyomidai-Higashi, Kisarazu, Chiba 292-0016 Japan

Keywords: Literature mining, Data mining, Down Syndrome, HSA21, Intellectual Disability

Down syndrome (DS), caused by the presence of an additional copy of a chromosome (No 21, i.e. HSA21 in human), is one of the major genetically induced disorders that significantly limit intellectual ability. It is also understood to be the cause of increased risk of other major diseases like congenital heart diseases, leukemia, hypotonia and Alzheimer's diseases [1-2]. Various mouse models are developed for the understanding of functions of the genes experimentally. There are about 200 genes associated with HSA21, however only the functions of few of them is understood [3]. An in-depth understanding of all the individual genes and their functions (isolated function or function as a group) is important for effective therapeutic discovery. Our current work is part of the integrated study approach to reveal the various genetic aspects of Down Syndrome on the basis of existing experimental and in-silico studies of genes, omics data and related diseases.

In this work, we developed an algorithm for the extraction and visualization of network of interrelated biological words that comprise diseases, genes, proteins and various symptoms etc. The abstracts of articles available in digital archive of biomedical and life sciences journal literatures (U.S. National Institutes of Health (NIH), PubMed Central) were screened iteratively based on set of search keywords created after every search iteration [4]. A filter that comprises the list of non-biological words prepared from selectively chosen non-biological literatures was applied to determine the search keywords for iterative operations. The components (the words) of final network were then scored using rule-based fuzzy logic [5]. The validity of the algorithm was tested with the well-known subnetwork consisting Down Syndrome mouse modes, genes related to HSA21 and (Ts65Dn, Ts1Cje), the human chromosome 21 (HSA21), the major region of study for Down Syndrome and its identical chromosome 16 (MMU16) in mouse and other related terms.

[1] Patterson, D., Molecular genetic analysis of Down syndrome. *Human genetics*, 126:, 195-214, 2009
[2] Wilcock, D.M., Schmitt, F.A. and Head, E., Cerebrovascular contributions to aging and Alzheimer's disease in Down syndrome. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1862: 909-914, 2016
[3] Nordquist, S.K., Smith, S.R. and Pierce, J.T., Systematic functional characterization of human 21st chromosome orthologs in Caenorhabditis elegans. *G3: Genes, Genomes, Genetics*, g3-200019, 2018.
[4] PubMed Central, https://www.ncbi.nlm.nih.gov/pubmed/
[5] Sapkota, A. and Ohmi, K., 2009., Error detection and performance analysis scheme for particle tracking velocimetry results using fuzzy logic. International Journal of Innovative Computing, Information and Control, 5(12B): pp.4927-4934, 2009

Orthogonal Linear Separation Analysis (OLSA): an Approach to Decompose the Complex Effects of a Drug to Understand Its Pharmacological Properties

Tadahaya Mizuno[1] tadahaya@mol.f.u-tokyo.ac.jp
Setsuo Kinoshita[1] setsuo.kinoshita@gmail.com
Shotaro Maedera[1] mdera0130@gmail.com
Takuya Ito[1] ito32539@gmail.com
Hiroyuki Kusuhara[1] kusuhara@mol.f.u-tokyo.ac.jp

[1] Molecular Pharmacokinetics, Department of Pharmaceutical Sciences, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Keywords: Profiling, Profile data analysis, OLSA, Connectivity Map, transcriptome

The response to a drug can be a complex of the entire biological responses to the perturbagen and to separate the complex effects of a drug into basic components is a prerequisite for a deep understanding of the pharmacological properties of drugs. Therefore, we investigated whether it is possible to decompose the complex effect of a drug into basic components described by variable patterns using profile data analysis, particularly focusing on factor analysis (FA). A concern of utilizing FA in profiling is that the centroid in the novel coordinate space has no biological meaning and varies among data sets, which means that the obtained factors (vectors) in such a situation may not correspond to consistent biological meanings. To address that concern, we have extended FA by using “response profiles” and “mirror data” of the examined data set. We have named this slightly modified FA as orthogonal linear separation analysis (OLSA) and have tested its performance in this study.

Transcriptome profile data of MCF-7-cells treated with 318 low molecular compounds were obtained from Connectivity Map and subjected to OLSA. Consequently, 118 factors contracting 11,911 genes were obtained. To examine the relevancy of those factors with biological responses, we investigated the property of each factor focusing on gene ontology of main genes consisting of the factor. Significant enrichment of the ontology in 65 of 118 factors were observed and similar results were obtained in two other data sets. One factor discriminated two Hsp90 inhibitors, geldanamycin and radicicol, while clustering analysis could not. Doxorubicin and other topoisomerase inhibitors were estimated to inhibit Na^+/K^+ ATPase, one of the suggested mechanisms of doxorubicin-induced cardiotoxicity. Based on the factor including PI3K/AKT/mTORC1 inhibition activity, 5 compounds were predicted to be novel inducers of autophagy, and other analysis including western blotting revealed that 4 of the 5 actually induced autophagy.

These findings indicate the potential of OLSA to decompose the effects of a drug and identify its basic components.

[1] Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.P., Subramanian, A., Ross, K.N., Reich, M., Hieronymus, H., Wei, G., Armstrong, S.A., Haggarty, S.J., Clemons, P.A., Wei, R., Carr, S.A., Lander, E.S., Golub, T.R., *The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease.*, Science, 313:1929-35, 2006.

Investigation of the common sequence-structural patterns in different folds through cross-profile comparison and molecular dynamics simulation

Yu Yamamori[1]
yu.yamamori@aist.go.jp

Kentaro Tomii[1]
k.tmoii@aist.go.jp

1 Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, 135-0064, Japan

Keywords: Cross-profile analysis, Molecular dynamics simulation, replica-exchange molecular dynamics

Structurally similar segments can be occurred in different protein folds. Combining the database analysis and molecular dynamics simulation, we investigated the evolutionary meaning of the existence of those segments. Previously, using cross-profile analysis, we found that profiles of some structural-segment clusters show the strong correlation between particular sequence profiles [1].

In this study, in order to validate the sequence-structure relationship, folding simulations of the peptides derived from two sequence profiles were performed using the replica-exchange molecular dynamics [2]. One of the sequence profiles is related to β-hairpin-like structure, and another is related to α-helix structure. For all of these simulations, 20 replicas were used and 50-ns simulations were performed for each replica. Water molecules are treated explicitly. The AMBER99SB force field was adopted for peptide, and the TIP3P model was used for water molecules.

The potential of mean forces along the radius of gyration of the peptides were calculated from trajectories of replica-exchange simulations (Figure 1). One of the structures corresponding to the global minima of each potential of mean force was picked up (Figure 1). We confirmed that the peptides can have stable folded structures corresponding to the structures expected from cross-profile analysis. We expected that the sequence-structure relationship can be discovered by cross-profile analysis can give a new insight to de novo protein design.

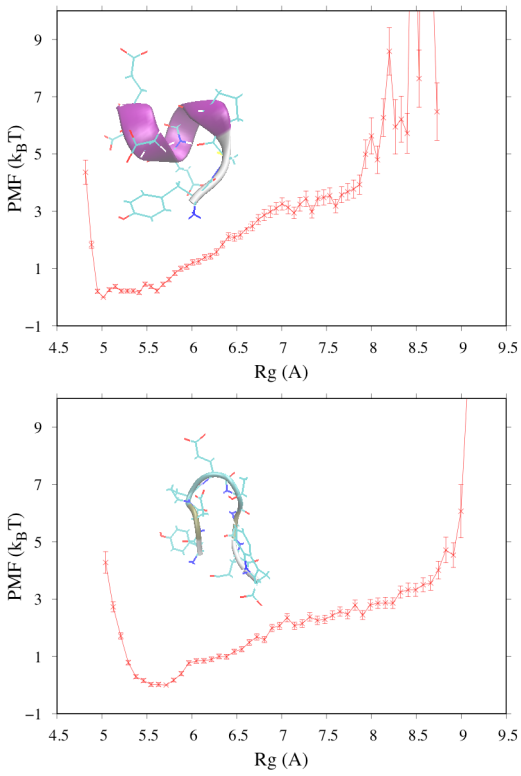


Figure 1 potential of mean force (PMF) along the radius of gyration of the peptides. Each snapshot structure is corresponding to the global minima of PMF.

[1] Kentaro Tomii, Yoshihiro Sawada and Shinya Honda *BMC Bioinfomatics* 2012 **13**:11
[2] Yuji Sugita and Yuko Okamoto *Chemical Physics Letters* 1999 **314** 141-151

Kampo drug repositioning: Analysis of the mode-of-action and prediction of new indications of Kampo medicines

Ryusuke Sawada[1]
rsawada@bioreg.kyushu-u.ac.jp

Michio Iwata[2]
iwata121@bio.kyutech.ac.jp

Masahito Umezaki[3]
masume@inm.u-toyama.ac.jp

Yoshihiko Usui[3]

Toshikazu Kobayashi[3]
koba.takaoka@zspec.net

Takaki Kubono[4]
m1661217@ems.u-toyama.ac.jp

Shusaku Hayashi[4]
hayashi@inm.u-toyama.ac.jp

Makoto Kadowaki[4]
makotok@inm.u-toyama.ac.jp

Yoshihiro Yamanishi[2,5]
yamani@bio.kyutech.ac.jp

[1] Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, Fukuoka 812-8582, Japan.
[2] Department of Bioscience and Bioinformatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan.
[3] Division of Chemo-Bioinformatics, Institute of Natural Medicine, University of Toyama, Toyama 930-0194, Japan.
[4] Division of Gastrointestinal Pathophysiology, Institute of Natural Medicine, University of Toyama, Toyama 930-0194, Japan.
[5] PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan.

Keywords: Kampo medicines, functional analysis, target prediction, Kampo repositioning

Kampo medicines are useful for treatment of multifactorial and chronic diseases. In this study we propose new computational methods for analyzing the mode-of-action and predicting new indications of Kampo medicines. Potential target proteins of the constituent compounds of Kampo medicines were predicted by docking simulations and machine learning methods based on large-scale omics data (e.g., genome, proteome, metabolome, interactome). For example, we predicted that daikenchuto would work for chronic inflammation-associated cancer, and we confirmed that the daikenchuto treatment indeed significantly suppressed the development of chronic colitis-associated colon cancer in a murine experimental model. Finally, we established KampoDB (http://wakanmoview.inm.u-toyama.ac.jp/kampo/), a novel database of Kampo medicines, which provides various useful scientific resources on Japanese traditional formulas Kampo medicines, constituent herbal drugs, constituent compounds, and target proteins of these constituent compounds. The current version of KampoDB contains 42 Kampo medicines, 54 crude drugs, 1230 constituent compounds, 460 known target proteins, and 1369 potential target proteins, and has functional annotations for biological pathways and molecular functions. KampoDB is useful for mode-of-action analysis of Kampo medicines and prediction of new indications for a wide range of diseases.

P4-01

Modelling Neonatal and MODY Diabetes in vitro Using iPS Cell-Derived Human Pancreatic Beta Cells

Filipa Soares[1], Marcus Yeo[1], Kalpesh Jhaveri[1], Masashi Matsunaga[1], Ludovic Vallier[1,2]

[1] DefiniGEN Ltd., Babraham Research Campus, Cambridge CB22 3AT, U.K.
[2] Cambridge Stem Cell Institute, University of Cambridge, U.K.
* Contact: enquiries@definigen.com

Keywords: iPS drug discovery platform, disease modelling, stem cell, regenerative medicine

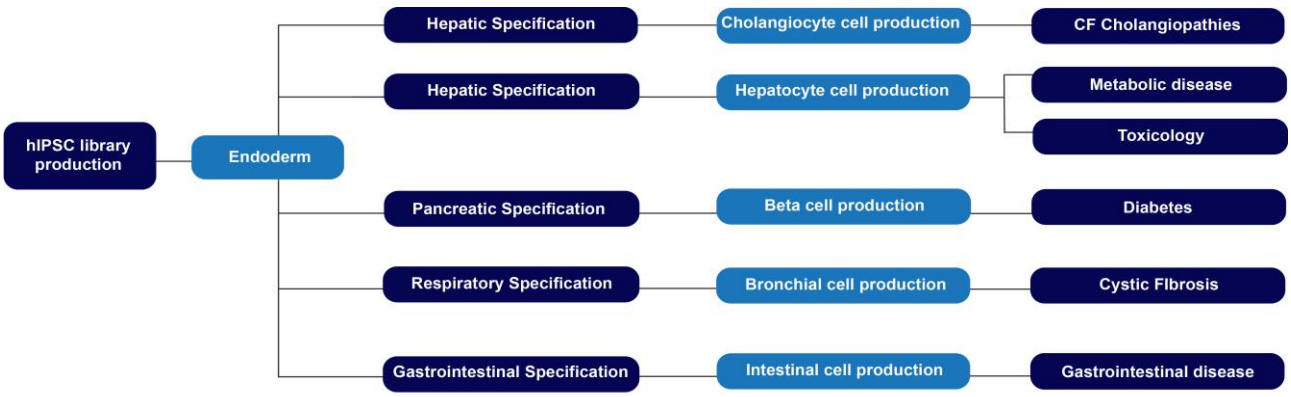
DefiniGEN is a spin-out company from the University of Cambridge which utilizes its directed differentiation platform OptiDIFF to provide high-functionality human cells to the drug discovery sector to enable the development of preclinical screens with improved predictivity of human health and efficacy. This platform has been combined with the CRISPR gene-editing platform by leading Cambridge life science company Horizon Discovery Ltd. The platform is generating monogenic IPS disease models for A1ATD, Familial Hypercholesterolemia, Neonatal and MODY diabetes which will be validated at both the genotypic and phenotypic level using an array of biochemical methodologies. Future work will focus on combined directed differentiation and genome-editing approaches to generate disease models for complex diseases such as diabetes type 2.

Advantages of combining IPSC direct differentiation with CRISPR technology

Human IPSC circumvent many of the ethical and commercial barriers associated with embryonic stem cells as they are derived from adult cells that have been reprogrammed to the stem cell state. Stem cells have two major properties:

1) they can form any cell type of the body.
2) they can self-renew indefinitely, producing a limitless supply of cells.

Genome editing and CRISPR/CAS9 technology are an efficient system to introduce mutations to healthy donor lines whereby isogenic control can be used to distinguish disease-relevant changes. Healthy human IPSC lines can be edited using CRISPR/CAS9 technology to create KI/KO mutants that will help study gene function and develop therapeutic approaches.



OptiDiff Platform. DefiniGEN is focussed on the production and disease modelling of related cell types including liver, pancreas, lung, intestinal and stomach cells.

P4-02

Development of an *in vitro* evaluation system for drug-induced hepatotoxicity using PXB-cells®

Mutsumi Inamatsu[1], **Mikaru Yamao**[1], **Yuko Ogawa**[1], **Chihiro Yamasaki**[1], **Yuji Ishida**[1,2], minamatsu@phoenixbio.co.jp

Chise Tateno[1,2]

[1] PhoenixBio Co., LTD. Research and Development Unit, 3-4-1, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-0046, Japan
[2] Research Center for Hepatology and Gastroenterology, Hiroshima University, 1-2-3, Kasumi, Minami-ku, Hiroshima 734-8551, Japan

Keywords: Toxicokinetics, ADMET, Regulating enzyme activity

Fresh human (h)-hepatocytes are considered to be the best *in vitro* model to study xenobiotic metabolism and hepatotoxicity. However, hepatocytes cannot be maintained for more than a week in two-dimensional cultures. Previously, we had shown that freshly isolated h-hepatocytes (PXB-cells®) from chimeric mice with humanized livers (PXB-mice®) can be maintained as a high-density culture (2.1 x 10^5 cells/cm^2), expressing major hepatic genes up to 3 weeks.

In physiological conditions, hepatocytes generate almost all adenosine triphosphate (ATP) in the mitochondria via aerobic respiration. However, hepatocytes in conventional culture systems generate the most ATP via cytosolic glycolysis because of high glucose content and limited oxygen supply. Such anaerobically poised cells are resistant to xenobiotics that impair mitochondrial function and are not suitable for mitochondrial toxicity evaluation. We examined the hepatotoxicity of mitochondrial complex-I inhibitor, rotenone, using PXB-cells® based on the reports that these hepatocytes were cultured in galactose-based medium (which substituted the conventional glucose based medium), and hyperoxia, to increase the susceptibility to mitochondrial impairment[1,2)]. To increase the supply of oxygen, we utilized Gas Permeable VECELL Plate (Vessel Inc.) [3)] and multi-gas incubators.

Lactate dehydrogenase assays showed that rotenone treatment decreased the viability of PXB-cells® in a dose-dependent manner. PXB-cells® cultured in galactose-based medium and hyperoxia showed the highest sensitivity to rotenone. Since this condition was maintaining expressions of several CYP mRNAs and CYP3A activity in PXB-cell® at least equivalent to conventional condition, compounds that show mitochondrial dysfunction after metabolism might be evaluated in this condition.

In conclusion, mitochondrial toxicity might be determined using PXB-cells® cultured in galactose-based medium and hyperoxia.

[1] Marroquin, L.D.,Hynes, J.,Dykens, J. A., Jamieson, J. D., and Will, Y., Circumventing the crabtree effect: replacing media glucose with garactose increases susceptibility of HepG2 cells to mitochondrial toxicants, *Toxicol Sci.*, 97: 539-547, 2007

[2] Liu C., Sekine, S., and Ito, K., Assessment of mitochondrial dysfunction-related, drug-induced hepatotoxicity in primary rat hepatocytes, *Toxicol Appl Pharmacol*., 302: 23-30, 2016

[3] Hamon M., Hanada, S., and Sakai, Y., Direct oxhgen supply with polydimethylsiloxane (PDMS) membranes induces a spontaneous organization of thick heterogeneous liver tissues from rat fetal liver cells *in vitro*, *Cell Transplant*, 21: 401-410, 2012

P4-03

Implementation of PRED subroutine of NONMEM 7 for versatile pharmacokinetic analysis using fast inversion of Laplace transform (FILT)

Ryota Jin[1] ryota-j@chiba-u.jp
Akihiro Hisaka[1] hisaka@chiba-u.jp

[1] Laboratory of Clinical Pharmacology and Pharmacometrics, Graduate School of Pharmaceutical Sciences, Chiba University, Chiba 260-8675, Japan

Keywords: FILT, NONMEM, Transfer function, Napp, Laplace transform

【Background and Objective】Fast Inversion of Laplace Transform (FILT) is a method which numerically solves various models defined by transfer function, i.e. Laplace transform of differential or partial differential equations. FILT was developed by Hosono et al[1] mainly for research in electrical engineering field and was applied to pharmacokinetic analysis by Yano et al[2]. Although FILT is precise and relatively efficient, and furthermore it is flexible then applicable to various pharmacokinetic analysis, available programs (MULTI(FILT)[2] and Napp[3]) are now outdated or inapplicable in Windows environment. In this study, we implemented a PRED subroutine of the latest version of NONMEM 7, a widely-used pharmacokinetic analysis program, by using FILT to solve the problem.

【Methods】We made a new PRED subroutine for implementation of FILT of the secondary order precision scheme using Fortran 95 which is now being adopted throughout in NONMEM 7. PRED is an internal model unit in NONMEM, To confirm the precision of calculation by the new PRED, we compared the simulation with results obtained by ADVAN (NONMEM predefined models) for the conventional calculation method and by Napp for the FILT method. The model used by ADVAN was 1-compartment model, and by Napp were an enterohepatic circulation model, a tank in series model, and a dispersion model.

【Results】The simulated values calculated by the new PRED agreed well with those calculated by ADVAN and Napp for the four models above, which confirmed that the new PRED was versatile and precise. By using FILT, it would be possible to expand potential of pharmacokinetic analysis of NONMEM 7.

[1] Hosono, T. Numerical inversion of Laplace transform and some applications to wave optics. *Radio Science* **16,** 1015–1019 (1981).
[2] YANO, Y., YAMAOKA, K. & TANAKA, H. A Nonlinear Least Squares Program, MULTI(FILT), Based on Fast Inverse Laplace Transform for Microcomputers. *Chemical and Pharmaceutical Bulletin* **37,** 1035–1038 (1989).
[3] Hisaka, A. & Sugiyama, Y. Analysis of nonlinear and nonsteady state hepatic extraction with the dispersion model using the finite difference method. *Journal of pharmacokinetics and biopharmaceutics* **26,** 495–519 (1998).

P4-04

Development of an informatics system for diversity of compound library

Yugo Shimizu[1] shimizu-yg@pha.keio.ac.jp
Masanori Osawa[1] osawa-ms@pha.keio.ac.jp
Kazuyoshi Ikeda[1] ikeda-kz@pha.keio.ac.jp

[1] Keio University, Faculty of Pharmacy, Division of Physics for Life Functions, 1-5-30 Shibakoen, Minato-ku, Tokyo 105-8512, Japan

Keywords: Library design, Structure-blinded analysis, New chemotype, Structural alert

Open innovation and collaborative drug discovery are expanding due to increasing the cost of drug development. Informatics approach is a key to analyze diversity of HTS compound library without disclosure of chemical structure. To expand collaborative drug discovery between industry and academia, the Drug-Discovery Innovation and Screening Consortium (DISC) of over 20 pharmaceutical companies in Japan has been established [1]. Our group is developing an in silico platform to analyze the consortium-based compound library without exact structure data.

We have updated our analysis system applied to larger compound libraries and developed informatics protocols to improve quality of compound library as pharmaceutical agents. The latest pharmaceutically developing compound data was collected from open drug discovery databases such as ChEMBL, PubChem, and SureChEMBL. Multiple structural alert (SA) filters were newly developed by combining different structural alert lists. The different criteria can be used for filtering whether the input compound is suitable as a pharmaceutical product or a chemical library compound. To search for chemical space suitable for new targets such as protein-protein interaction (PPI), compound library containing different chemotypes from conventional small molecules is required. We introduced a new filter to classify different chemotypes of three-dimensional shape based on the results of analysis of the relation between the existence probability of bioactive compounds and the drug-likeness parameters such as Fsp3. The filters would be useful for acquisition of new chemotypes as well as expansion of current chemical space.

In order to estimate the distribution of diversity in bioactive space, we created ligand-based target models by using several machine learning (ML) methods such as naïve Bayes, random forest, support vector machine, and gradient boosting. We extracted compounds and their bioactivity information from public databases such as ChEMBL and PubChem to train and test the models. Over 1,500 target models were trained for each target by employing both molecular properties and fingerprints as descriptors. The prediction performances of the ML methods were evaluated by cross-validation and were compared each other.

[1] https://www.amed.go.jp/program/list/06/03/001_02-01.html

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

P4-05

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 8. Web application and database consisting of curated public data and newly acquired experimental data

Hitoshi Kawashima[1] hkawashi@nibiohn.go.jp
Reiko Watanabe[1] reiko-watanabe@nibiohn.go.jp
Tsuyoshi Esaki[1] tsuyoshi-esaki@nibiohn.go.jp
Rikiya Ohashi[1,2] rikiya-ohashi@nibiohn.go.jp
Daisuke Satoh[3] daisuke.satoh@lifematics.co.jp
Chioko Nagao[1,4] chio@nibiohn.go.jp
Yayoi Natsume-Kitatani[1,4] natsume@nibiohn.go.jp
Kenji Mizuguchi[1,4] kenji@nibiohn.go.jp

[1] Laboratory of Bioinformatics, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Saito Asagi, Ibaraki, Osaka 567-0085, Japan
[2] Discovery Technology Laboratories, Mitsubishi Tanabe Pharma Corp., 2-2-50 Kawagishi, Toda, Saitama 335-8505, Japan
[3] Software Division, Lifematics Inc., 2-25 Kanda Suda-cho, Chiyoda, Tokyo 101-0041, Japan
[4] Laboratory of In-silico Drug Design, Center of Drug Design Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Saito Asagi, Ibaraki, Osaka 567-0085, Japan

Keywords: Pharmacokinetics, Database, Prediction models, Web application

We started an initiative “Development of a Drug Discovery Informatics System”, with support from the Japan Agency for Medical Research and Development (AMED) in collaboration with several other research groups to develop more accurate prediction systems for ADME and tox properties primarily targeting academic scientists.

We collected pharmacokinetic and physicochemical parameters from the public bioactivity database, ChEMBL. The data were curated according to a curation workflow and stored in the database with a curation flag. In addition to the public data, we have been acquiring both *in vitro* and *in vivo* experimental data using consistent protocols. In addition to the parameters collected previously, we collected efflux ratio of P-glycoprotein (P-gp), which is the major transporter in gut and brain.

We stored all the data in a PostgreSQL database and are developing a web application to view and analyze the database content. The database can be searched by key words, several IDs including ChEMBL ID, Standard InChI, Standard InChI Key and canonical SMILES, by drawing compound structure or by activity range of several parameters. Detail view of a compound shows the chemical structure, the activity of each parameter, the link to other databases and the position in the chemical space.

We are currently developing prediction models using these data as presented in other posters, and we plan to make it possible to have functions of database search of existing data and parameter prediction of new compounds with similar interfaces.

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

P4-06

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 9. Development of Regression Model of Unbound Fraction to Brain Homogenate from Chemical Structure

Tsuyoshi Esaki[1]
tsuyoshi-esaki@nibiohn.go.jp

Reiko Watanabe[1]
reiko-watanabe@nibiohn.go.jp

Rikiya Ohashi[1,2]
rikiya-ohashi@nibiohn.go.jp

Yayoi Natsume-Kitatani[1,3]
natsume@nibiohn.go.jp

Hitoshi Kawashima[1]
hkawashima@nibiohn.go.jp

Chioko Nagao[1,3]
chio@nibiohn.go.jp

Kenji Mizuguchi[1,3]
kenji@nibiohn.go.jp

[1] Laboratory of Bioinformatics, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8, Saito Asagi, Ibaraki, Osaka 567-0085, Japan
[2] Discovery Technology Laboratories, Mitsubishi Tanabe Pharma Corporation, 2-2-50, Kawagishi, Toda, Saitama 335-8505, Japan
[3] Laboratory of In-silico Drug Design, Center of Drug Design Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Saito Asagi, Ibaraki, Osaka 567-0085, Japan

Keywords: Brain homogenate, Brain distribution, Fraction unbound, Machine learning

Estimating the penetration of drugs into the central nervous system (CNS) in the early stages of drug development can accelerate the development of CNS drugs. It is necessary to predict the concentration of unbound drug in the brain ($C_{u,brain}$) for predicting the efficacy and toxicity of a drug. Fraction unbound in brain homogenate ($f_{u,b}$) is one of the most critical parameters to estimate the distribution of drug in the brain, but no available models to predict $f_{u,b}$ from chemical structure alone have been proposed. In the academic drug discovery environment, it is unfeasible to collect a large volume of new experimental data and although a considerable amount of measurements are available from public databases, but these databases typically contain $f_{u,b}$ under cover of unbound data measured using other tissues. In this study, we developed regression models to predict $f_{u,b}$ using curated data collected from literatures and an open database.

The experimental data were collected from literatures and ChEMBL, an open database. Our data set included more than 250 compounds. Three freely available software packages, CDK, Mordred and PaDEL-Descriptor, were used to calculate descriptors for the compounds. Three machine learning algorithms (Random Forest (RF), Support Vector Machine with Radial Kernel function (SVM) and Partial Least Square (PLS)) were employed for model development and the performance of each model was evaluated.

The RF and SVM model scored higher than PLS in both cross-validation and the evaluation on a test set The R2 scores for RF, SVM and PLS were from 0.52 to 0.56 in the cross validation, but R2 score of RF was the highest (R2 = 0.614) followed by SVM and one of PLS was less than 0.5 on the test compounds. The non-linear model (RF and SVM) may have higher predictive ability than the linear algorithm (PLS), since the binding of a compound to the brain tissues depends on numerous complex factors.

We presented an *in silico* model to predict $f_{u,b}$ without experimental procedures. Our study could open the possibility to access the development for all researchers interested in the brain target drugs

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

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 10. Prediction of renal clearance in human utilizing structural information

Reiko Watanabe[1] reiko-watanabe@nibiohn.go.jp
Tsuyoshi Esaki[1] tsuyoshi-esaki@nibiohn.go.jp
Hitoshi Kawashima[1] hkawashi@nibiohn.go.jp
Yayoi Natsume-Kitatani[1] natsume@nibiohn.go.jp
Chioko Nagao[1,3] chio@nibiohn.go.jp
Rikiya Ohashi[12] rikiya-ohashi@nibiohn.go.jp
Kenji Mizuguchi[1,3] kenji@nibiohn.go.jp

[1] Laboratory of Bioinformatics, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Saito Asagi, Ibaraki, Osaka 567-0085, Japan
[2] Discovery Technology Laboratories, Mitsubishi Tanabe Pharma Corporation, 2-2-50, Kawagishi, Toda, Saitama, 335-8505, Japan
[3] Laboratory of In-silico Drug Design, Center of Drug Design Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Saito Asagi, Ibaraki, Osaka 567-0085, Japan

Keywords: plasma protein binding, protein unbound, machine learning

Renal clearance is the prime route for the elimination of drugs with low to negligible metabolism and also of drug metabolites. It also represents the net result of glomerular filtration, active secretion and reabsorption. *In silico* approaches that variously employ physicochemical and molecular properties of compounds to quantify pharmacokinetic processes of interest are highly effective in the early stages of drug discovery because they are noninvasive and have high-throughput capacity. We had previously developed *in silico* models to predict the values of fraction unbound in plasma, which is an important determinant of drug efficacy in pharmacokinetic and pharmacodynamic studies. In this study, we have extended our approach to build models to predict renal clearance using the values of fraction unbound in plasma as calculated by our previous models. Experimental values of the renal clearance and fraction unbound in plasma were extracted from the literature and the ChEMBL database, and classification models using different machine learning methods were constructed to classify the compounds into three groups, namely, the reabsorption, intermediate and secretion types. Next, we generated regression models to predict the renal clearance, the prediction processes were examined by combining other prediction models such as classification models of elimination types or predicted value of fraction unbound in plasma. Our predictive models can be integrated into other pharmacokinetic modeling systems, which would be highly useful in academic drug discovery.

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

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

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 11. Prediction of sites of drug metabolism by CYP3A4 by molecular simulation

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

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

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

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

In this study, we present a new computational method (score function) to predict SOMs of compounds metabolized by CYP3A4. The new score function is composed of accessibility and reactivity scores. The accessibility scores for the sites (atoms) of compounds are estimated by ensemble docking simulation, while the activation energies of atoms estimated from the SMARTCyp software[3] are used as the reactivity scores in the method. We carried out a molecular dynamics (MD) simulation of apo type of CYP3A4 for 10 micro second to sample the receptor (pocket) structures for the ensemble docking. For the molecular docking, DOCK program with Grid score, which are optimized for CYP system [4], was used. We prepared 47 test set compounds metabolized by the CYP3A4 and 100,000 receptor snapshots (which correspond to the 10 micro second MD simulation) for the ensemble docking simulation. The top three ranked atoms by the new score function are predicted as possible SOMs for each compound in this study. We found the success rate of the predicted SOMs was 86 %, showing better result than those by the ensemble docking and SMARTCyp program alone.

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

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

P4-09

Development of a pharmacokinetics prediction system using multiscale integrated modeling: 12. Evaluation of the functionalities and the performance of the revised custom SoC LSI

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

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

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

We have been developing a series of special purpose computers for molecular dynamics simulations. The current system is the MDGRAPE-4[1]. The whole system consists of 512 custom System on Chip (SoC) LSI chips in 3D torus network. The target performance is to make it possible to simulate typical protein-ligand complex surrounded by water for 10 microseconds per day, which is two orders of magnitude faster than commercially available systems. One of the applications of long-term MD simulations of proteins is to predict the site of metabolism.

The MDGRAPE-4 hardware development completed in 2014, and the stabilization of operation and software development process succeeded for two years. However, the final performance does not reach the expected one. In the process of the development, we realized that the classical MD simulation of the typical system size has already reached the strong scaling limit of parallel computation and not only accelerating the computationally demanding part but also more specialized hardware for the type of computation is required to improve the performance.

To achieve the target performance, we have started the development of the improved system that is named MDGRAPE-4A. The system architecture is basically common with the MDGRAPE-4. The enhancement of the functionalities are as follows:

1) support of management and migration of atoms by the memory hardware

2) more sophisticated treatment of exclusion and reduced non-bonded interaction pairs in the special purpose pipelines

3) implementation SIMD functions and special instructions required in the MD simulations in the general-purpose cores

4) reduction of the latency and more flexible routing patterns in the network interface

5) support of a long-range electrostatic interaction calculation method that is suitable for the parallel systems with torus network topology

Currently the development process is under the evaluation of the revised LSI on the revised system board. The new system hardware will be completed in the end of 2018. In this presentation, we would like to show the expected performance of the revised system.

[1] Ohmura, I., *et al.*, MDGRAPE-4: a special-purpose computer system for molecular dynamics simulations, *Phil. Trans. R. Soc. A*, 372:20130387, 2014.

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

P4-10

Development of an informatics system for predicting cardiotoxicity: 4. Update of integrated cardiotoxicity database

Hitomi Yuki
hitomi.yuki@riken.jp

Tomohiro Sato
tomohiro.sato@riken.jp

Teruki Honma
honma.teruki@riken.jp

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

Keywords: Cardiotoxicity, Database, hERG, Ion channels, QT prolongation

The inhibition of hERG potassium channel is closely related to the prolonged QT interval [1], and to assess the risk could greatly contribute to the development of safer therapeutic compounds. To utilizing the recent increase of information about hERG inhibitors in public databases, we built integrated database for hERG blocking compounds. The merging procedure were reported in CBI annual meeting 2017. The database was released as “AMED Cardiotoxicity Database” [2] and the paper was also published [3]. You can find cardiotoxicity assay results quickly on the web by substructure or similarity compound searches. We will introduce the database contents, the analysis results, and future plans such as the scaffold difference between hERG actives and inactives, the user interface and assay results of other ion channels. We would like to discuss the usabilities and requests for the database.

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

[1] Picard, S., Goineau, S., Guillaume, P., Henry, J., Hanouz, J.-L. and Rouet, R., Supplemental studies for cardiovascular risk assessment in safety pharmacology: a critical overview, *Cardiovasc. Toxicol.*11:285–307, 2011
[2] http://drugdesign.riken.jp/hERGdb/
[3] Sato, T., Yuki, H., Ogura, K. and Honma, T., Construction of an integrated database for hERG blocking small molecules, *PLoS ONE*, 13(7):e0199348, 2018

Development of an informatics system for predicting cardiotoxicity: 5. Quantitative model for hERG blocking small molecules based on the integrated database.

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

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

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

Keywords: hERG, QT prolongation, Support vector regression

The inhibition of hERG potassium channel is closely related to the prolonged QT interval [1], and to assess the risk could greatly contribute to the development of safer therapeutic compounds. In the optimization stage of drug development, quantitative prediction of hERG inhibitory activity is crucial to design drug candidates without cardiotoxicity risk. Here, we developed a hERG regression model combining Support Vector Regression (SVR) and descriptor selection by Non-dominated sorting genetic algorithm (NSGA-II) based on the hERG integrated database reported as the largest dataset for hERG inhibitors in CBI meeting 2016. From the 291,219 compounds in the integrated database, 6,563 compounds with IC_{50} values were selected, and randomly separated into training set (70%) and test set (30%) to build a regression model and evaluate its prediction performance. To avoid overfitting by employing too many non-relevant explanatory variables, NSGA-II, which is a variation of genetic algorithm for multiple objective optimization, is used to select descriptors in order to maximize Q^2 and minimize RMSE in 5-fold cross validation, and minimize the number of used descriptors spontaneously. The prediction performance was then compared to those of ADMET predictor [2], commercial software providing various ADMET property predictions. The SVR model recorded R^2 of 0.595 and RMSE of 0.607 for test set, clearly exceeding those of ADMET predictor (0.134 and 0.960, respectively).

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

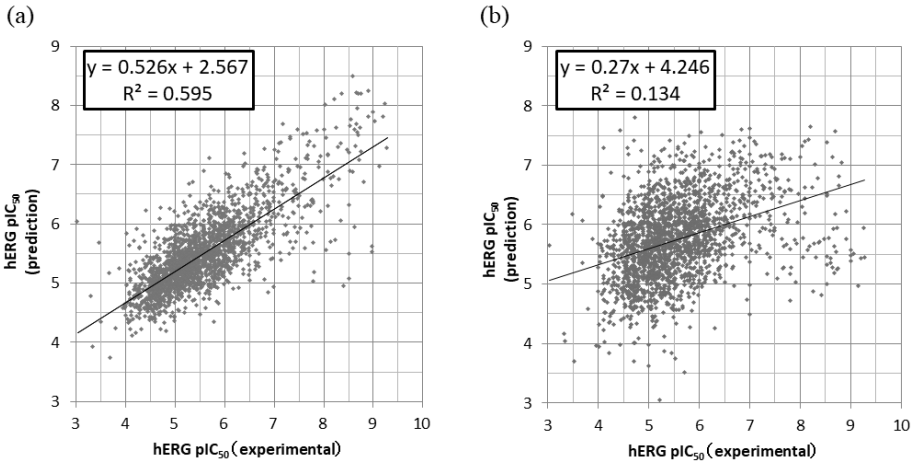


Figure 1. Prediction of hERG inhibition for test set using (a) AMED SVR model and (b) ADMET predictor (Horizontal axes: experimental pIC_{50} values, vertical axes: predicted pIC_{50} values).

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

[2] https://www.simulations-plus.com/software/admetpredictor/

P4-12

Use of Markov Chain Monte Carlo Method to Integrate In Vitro & In Vivo Data for Prediction of Drug Interactions Caused by Inhibition of Multiple CYP Species

Shizuka Hozuki[1]
hozuki0163@chiba-u.jp

Hideki Yoshioka[1]
afga7643@chiba-u.jp

Hiromi Sato[1]
hiromi-s@chiba-u.jp

Akihiro Hisaka[1]
hisaka@chiba-u.jp

[1] Clinical Pharmacology and Pharmacometrics, Graduate School of Pharmaceutical Sciences, Chiba University, 1-8-1, Inohana, Chuo-ku, Chiba, Chiba 260-8675, Japan

Keywords: Pharmacokinetics, Cytochrome P450, Drug-drug interaction, Bayesian inference, MCMC, WinBUGS

Prediction of AUC changes in drug-drug interaction (DDI) are not easy if involving CYP species are multiple or are uncertain with insufficient data. The objective of this study is to construct a systematic and exhaustive prediction method of DDIs involving multiple CYP species based on all available information (*in vitro* & *in vivo*) by using Bayesian inference which allows flexible error model.

Based on a static pharmacokinetic model taken account of appropriate error distributions, a framework was constructed to estimate substrates' contribution ratio (CR) and inhibitors' inhibition ratio (IR) of five CYP species (1A2, 2C9, 2C19, 2D6, 3A) by Markov Chain Monte Carlo method (software; WinBUGS1.4.3.). Model was constructed as the hierarchical Bayesian model. In addition to CR and IR, distribution volume of each inhibitor was estimated as a scaling factor between *in vitro* and *in vivo* data. The *in vitro* data of fm (fraction metabolized in vitro)/Ki and *in vivo* data of AUC changed ratio (AUCR) caused by DDI or polymorphism were collected from University of Washington's Metabolism and Transport Drug Interaction Database (DIDB®). We selected major 68 substrates and 32 inhibitors for the analysis in this study. Simulation analysis was extensively performed by using virtual data sets beforehand to confirm appropriateness and predictability of the method.

The simultaneous analysis of fm for 27 substrates, Ki for 32 inhibitors and AUCR for 201 drug pairs (9.2% of 2,176 all possible drug pairs) resulted in the prediction of AUCR of all possible drug pairs within reasonable estimation errors. In the real data analysis, the standard errors of fm, Ki and AUCR were estimated as 0.013 (normal distribution), 251%CV and 32.5%CV (lognormal distribution), respectively..

The present analysis allowed efficient use of observed data, which are limited in most cases, for proper management of DDIs. It may also contribute to appropriate classification of CYP substrates and inhibitors in the labeling. Furthermore, our simulation suggested that the present predictability of DDI would be improved noticeably if *in vitro* data (experimental fm/Ki) will be collected systematically in the future.

[1] Yoshiyuki Ohno, Akihiro Hisaka and Hiroshi Suzuki, General framework for the quantitative prediction of CYP3A4-mediated oral drug interactions based on the AUC increase by coadministration of standard drugs, *Clin Pharmacokinet.* 46(8):681-696, 2007

Individual data analysis of patients participated in clinical studies: relationship between longitudinal changes in cardiac functions and mortality risk in CHF

Sayuri Guro [1]
adwa1668@chiba-u.jp

Hideki Yoshioka [1]
afga7643@chiba-u.jp

Akihiro Hisaka [1]
hisaka@chiba-u.jp

[1] Clinical Pharmacology and Pharmacometrics, Graduate School of Pharmaceutical Sciences, Chiba University, Chiba, Japan.

Keywords: chronic heart failure, clinical trial data, cardiac function

Chronic heart failure (CHF) is a serious disease with a high mortality rate (five-year survival rate, 50-60%) [1], and its patients often experience longitudinal decline of cardiac physical functions, resulting in death. By using literature information of aggregated data (i.e. mean and sd), we previously found that the reduced mortality risk in CHF patients by treatment of various drugs was excellently associated with decrease in myocardial oxygen consumption (MVO_2), a cardiac load index. Further, the degree of their longitudinal changes was fluctuated by patient characteristics [2]. In this study, by using individual data of patients participated in clinical studies, we aimed to elucidate the relationship between the changes in cardiac functions and mortality risk in CHF patients in further detail focusing on heart rate (HR) and systolic blood pressure (SBP).

Based on approval of IRB of Chiba university, graduate school of pharmaceutical sciences, data of two randomized controlled trials in CHF patients (SOLVD Treatment and HF-ACTION) were obtained from *Biologic Specimen and Data Repositories Information Coordinating Center* (BioLINCC) [3]. SOLVD Treatment assessed the effect of enalapril (N=1285) vs placebo (N=1284), while HF-ACTION assessed the efficacy and safety of exercise training (N=1172) vs placebo (N=1159). We compared changes in HR and SBP from the randomization date between survivors and the deceased. For the deceased, the changes during 50 months before death was also compared.

We observed that SBP of the deceased was constantly decreased (~5 mmHg/year) after the randomization despite that SBP of the survivors was almost stable in both trials. Furthermore, our analysis showed that the decline of SBP in the deceased was more remarkable during two years before death. These suggests that the excessive decrease of SBP in the deceased should be distinguished from the therapeutic effect and could be a risk factor for death in CHF patients. For HR, no distinct difference between the deceased and survivors was observed in both trials. Assessing longitudinal change in cardiac functions may provide detailed understanding of the progression of CHF and its mortality risks.

[1] Shiba, N. and Shimokawa, H., Chronic heart failure in Japan: implications of the CHART studies, *Vasc Health Risk Manag*, 4:103-13, 2008.
[2] Guro, S., Takaoka, R., Hatakeyama, H., Hisaka, A., Model based meta-analysis to investigate therapeutic responses of β-blockers and ACE inhibitors in progression of chronic heart failure, *The 37th Annual Scientific Meeting of the Japanese Society of Clinical Pharmacology and Therapeutics*, 2016,1-O-12.
[3] https://biolincc.nhlbi.nih.gov/home.html

Model-based meta-analysis of the relationship between HbA1c change and urinary glucose excretion in subjects treated with six SGLT2 inhibitors aiming early prediction of efficacy in drug development

Ayana Ishikawa[1] aeta1660@chiba-u.jp
Hideki Yoshioka[1] afga7643@chiba-u.jp
Hiromi Sato[1] hiromi-s@chiba-u.jp
Akihiro Hisaka[1] hisaka@chiba-u.jp

[1]Laboratory of Clinical Pharmacology and Pharmacometrics, Graduate School of Pharmaceutical Sciences, Chiba University, Inohana, Chuo-ku, Chiba-City, Chiba, 1-8-1, Japan

Keywords: Prediction of drug efficacy, Biomarker, Model-based meta-analysis

In drug development, the drug efficacy is generally evaluated in Phase II study and later. If we can predict drug efficacy in Phase I, we may decide whether or not to continue development and may improve the efficiency of later clinical studies with proper dosage selection. In this study, we focused on six SGLT2 inhibitors used for pharmacotherapy of type 2 diabetes in Japan. In order to compare the efficacy of these SGLT2 inhibitors, the relationships between the daily urinary glucose excretion measured in healthy subjects in phase I studies and the HbA1c decrease evaluated in diabetes patients in later clinical studies were investigated by model-based meta-analysis (MBMA).

First, the dosages of six drugs were normalized based on the daily urinary glucose excretion (UGE) of healthy subjects in phase I single dose study described in the evaluation reports released from PMDA. Next, MBMA was performed by using the Emax model for the relation between the normalized dose and HbA1c changes from the baseline observed in 83 clinical studies collected from the literature. The model analysis was performed using a first-order conditional estimation with interaction (FOCE-I) implemented on NONMEM (v7.3.0).

As a result, urinary glucose excretion (UGE) was consistently elevated as the dose increased for all six SGLT2 inhibitors. It was confirmed that the daily UGE was approximately 50g at the efficacious clinical dose of these SGLT2 inhibitors. In addition, once their doses are normalized to the clinical dosage, the dose-response relationship for HbA1c decrease was almost comparable between these drugs. It was suggested that the clinical properties of six SGLT2 inhibitors are very similar, and the proper dose selection for later clinical studies can be made based on daily urinary glucose excretion which is evaluable even in phase I study.

Exploratory analysis of data in Phase I study relating the drug efficacy might be of help in improving the efficiency of later studies and selecting proper dosage as demonstrated in this study. Although this is a retrospective analysis, in the future, analogous prospective analysis might contribute to improvement in the efficiency of drug development.

Evaluation of correlation of newly developing noninvasive cell culture profiling system based on LC-MS/MS measurement of medium components to the functional changes during cell culture

Ryuya Fujii[1] **Kenichi Toyoda**[2] **Shinichiro Horiuchi**[1]
r-fujii@nihs.go.jp

Yukie Kuroda[1] **Su-Ryang Kim**[1] **Yasunari Kanda**[1] **Seiichi Ishida**[1]

[1]Division of Pharmacology, National Institute of Health Sciences,
3-25-26 Tonomachi, kawasaki-ku, Kawasaki, 210-9501, Japan
[2]Shimadzu Corporation,
1,Nishinokyo Kuwabara-cho, Nakagyo-ku, Kyoto 604-8511, Japan

Keywords: HepG2, LC-MS/MS, cell culture profiling system, medium components

[Background]
In recent years, *in vitro* cell-based assays with devices capable of culturing cells with high cell-function *in vitro* have been developed in order to reduce the cost for the drug-candidate safety evaluation in the early stage of drug discovery process. Liver has many functions for the catabolism of nutrients or xenobiotics, and development of a noninvasive measurement method is useful for the evaluation of the fundamental functions of hepatocyte (e.g., production of albumin and urea and expression of oxidase relating to drug metabolism) during cell culture on such devices. Therefore, application of a cell culture profiling system capable of quantifying 95 medium components and secretory metabolites simultaneously using an LC-MS/MS was examined. We have already reported an evaluation case in which differentiated and undifferentiated states of multifunctional stem cells were discriminated based on the measurement of changes in medium components (The Society for Biotechnology, Japan/Annual Meeting 2016). In this study, the application of this system to the analysis of hepatocyte functions was evaluated using a liver cancer cell line HepG2.
[Method]
HepG2 cells was cultured with high glucose DMEM supplemented with 10% fetal bovine serum and 1% antibiotic and antimycotic solution for 5 days, and the culture supernatant was collected over time. Acetonitrile was added to the culture supernatant to remove the protein, and then the medium components were quantitated according to the protocol previously established using LC-MS/MS (LCMS-8050, Shimadzu).
[Result]
LC-MS/MS analysis resulted in the detection of several medium components successfully. We would like to discuss the relation of changes in medium components to the HepG2 cell functions in the presentation.

Characterization of human hepatocytes isolated from chimeric mice with humanized liver (PXB-cells) by DNA microarray analysis for the evaluation of the applicability to cell-based drug safety tests

Shinichiro Horiuchi **Yukie Kuroda** **Su-Ryang Kim** **Ryuya Fujii**
Hori542@nihs.go.jp

Yasunari Kanda **Seiichi Ishida**

Division of Pharmacology, National Institute of Health Sciences,
3-25-26 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa, 210-9501, Japan

Keywords: PXB-cells, DNA microarray, drug safety test, drug metabolism

Evaluation of drug metabolism in humans is important for the safety test of drug candidate compounds. Human hepatocytes are mainly used in *in vitro* cell-based test at the early stage of drug development. However, there are some drawbacks, such as differences between lots caused by the individual donor variations or limited supply of each lot. Therefore, development of alternative cell source which circumvents these points is expected. When human hepatocytes are transplanted into the uPA/SCID mouse liver via the spleen, mouse hepatocytes are largely repopulated with the transplanted human hepatocytes. Stable supply of human hepatocyte derived from the same donor is possible by isolating hepatocytes from chimeric mice with humanized liver. We compared functions and characteristics among various hepatocytes by gene expression analysis to evaluate whether the human hepatocytes isolated from chimeric mice with humanized liver has characteristics similar to human hepatocytes.

PXB-cells (PhoenixBio, Co., Ltd., Hiroshima, Japan) were used as a human hepatocytes isolated from chimeric mice. Liver function-related 183 gene expressions was measured by Genopal focused DNA microarray (Mitubishi Chemical, Co., Ltd., Tokyo, Japan). The gene expression data obtained from PXB-cells, cryopreserved hepatocytes (cryo-hep), hepatocyte-like cells differentiated from human iPS cells (hiPSC-hep) and human liver cancer cells (HepG2) was compared by Principal Component Analysis (PCA), hierarchical clustering analysis and enrichment analysis.

When gene expression data of Genopal was analyzed by PCA, PXB-cell was located near cryo-hep and far from iPSC-HEP and HepG2 cell in principal component one (PC1) and PC2. In addition, PXB-cells was classified into the same cluster that includes cryo-hep by hierarchical clustering. These results shows that PXB-cells has properties closer to that of cryo-hep than to those of hiPSC-hep and HepaG2 cell. Genes of phase I drug metabolizing enzyme and hepatocyte-specific transporter were enriched in gene cluster highly expressed in PXB-cells and cryo-hep.

From the above, we consider that PXB-cells is the candidate for alternative cell source of human primary / cryo-preserved hepatocytes in *in vitro* cell-based drug metabolism test.

P5-01

AI for Chemistry Optimisation: Combining Machine Learning and Domain Knowledge

Matthew Segall[1] matt@optibrium.com
Sumie Tajima[2] tajima@hulinks.co.jp
Edmund Champness[1] ed@optibrium.com
Peter Hunt[1] peter@optibrium.com
Tamsin Mansley[1] tamsin@optibrium.com

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

Keywords: Multi-parameter optimisation, QSAR, Structural transformation, Evolutionary algorithm, de novo design

Artificial Intelligence involves the application of machine learning algorithms in the context of domain knowledge. In the case of compound design, this involves integration of information from multiple perspectives: understanding of structure-activity relationships (SAR), based on data from previously studied compounds; expertise from diverse fields to define the multi-parameter optimisation (MPO) objectives of a project; and knowledge of synthetic strategies that may be applicable to create the next round of compounds for investigation. All of these forms of knowledge can be captured and applied computationally: Machine learning methods can generate quantitative structure-activity relationship (QSAR) models to predict the properties of novel, virtual compounds; MPO methods capture the desired property criteria for a successful compound for a specific project and rigorously define an objective function to guide optimisation; and, evolutionary algorithms can be applied to explore optimisation strategies captured as structural transformations that reflect steps made in previous chemistry projects.

In this presentation, we will describe these methods and illustrate how they can be seamlessly combined to rigorously explore new, relevant compound ideas and prioritise those most likely to achieve a project objective. This approach can help to stimulate the search for new optimisation strategies and explore a much broader range of compounds than could be achieved based on a single chemist's or even a project team's experience. Example applications include the optimisation of compounds with a desired polypharmacology or selectivity profile and exploration of lead hopping strategies to overcome pharmacokinetic issues, while maintaining target potency.

[1] O. Obrezanova, M.D. Segall, *J. Chem. Inf. Model.*, 50(6):1053-1061, 2010.
[2] M.D. Segall, *Curr. Pharm. Des.*, 18(9):1292-1310, 2012.
[3] M.D. Segall et al., *J. Chem. Inf. Model.*, 51(11):2967-2976, 2011.
[4] I. Ujváry, J. Hayward, In N. Brown ed., *Bioisosteres in Medicinal Chemistry*, 2012.

P5-02

Development of Computer Aided Retrosynthesis system

Takayuki Serizawa[1]
serizawa.tf@om.asahi-kasei.co.jp

Kazufumi Ohkawa[1]
okawa.kf@om.asahi-kasei.co.jp

Takayuki Kochi[1]
kochi.tb@om.asahi-kasei.co.jp

[1] Laboratory for Medicinal Chemistry, Pharmaceuticals Research Center, AsahiKASEI Pharma corporation, 632-1 Mifuku, Izunokuni, Shizuoka 4102321, Japan

Keywords: Artificial Intelligence(AI), Machine Learning, Deep Learning, Retrosynthesis, Synthetic route prediction

Back ground:

Recently Artificial Intelligence is attracting attention not only Computer Vision but also Drug Discovery.

We have been using *de novo* molecular design system. The system can propose molecules but can't propose synthetic route. There is a problem that *de novo* molecular generator sometimes proposes synthetic infeasible compounds. To address this issue, we tried to develop Computer Aided Retrosynthetic Analysis system.

Results:

We developed Retrosynthetic AI with neural network and Monte Carlo Tree Search (MCTS) [1] with python and JavaScript as programming language. This approach is the same as Google's Alpha Go [2]. The AI can learn reaction pattern from more than a million of reaction data and proposes synthetic route of given molecules within several minutes. And we developed Web user interface for medicinal chemists.

[1] Marwin H. S. Segler, Mike Preuss, Mark P. Waller*: Nature. 2018 Mar 28;555(7698), pp. 604-610.*
[2] David Silver *et al*.: *Nature 2016 January 28 529, pp. 484–489.*
[3] CB Browne *et al*.: IEEE Transactions on Computational Intelligence and AI in Games (Volume: 4, Issue: 1, March 2012), pp. 1-43

P5-03

In silico Prediction of Severe Cutaneous Adverse Drug Reactions Based on the Adverse Event Reporting Database

Kazuyuki Ohya
c142017@ed.nagoya-cu.ac.jp

Kaori Ambe
ambek@phar.nagoya-cu.ac.jp

Masahiro Tohkin
tohkin@phar.nagoya-cu.ac.jp

Graduate School of Pharmaceutical Sciences, Nagoya City University,
3-1 Tanabe-dori, Mizuho-ku, Nagoya 467-8603 Japan

Keywords: JADER, Machine learning method, Stevens Johnson Syndrome, Toxic Epidermal Necrolysis.

Stevens Johnson Syndrome (SJS) and Toxic Epidermal Necrolysis (TEN) are idiosyncratic adverse drug reaction (iADR). Since iADR is hard to be predicted from animal experiments and clinical trials due to its high species specificity and low occurrence frequency, it is desired to establish a prediction method of iADR at the new drug development stage. Large-scale adverse drug effect data of post-marketing can grasp the relevance between medicinal drugs and various adverse reactions. Therefore, we focused on Severe Cutaneous Adverse Reactions (SCAR), such as SJS/TEN [1] based on adverse event reporting system, and built a predictive classification model by machine learning method using the drugs structure information.

The Japanese Adverse Drug Event Report database (JADER) [2] constructed by the Pharmaceutical and Medical Devices Agency (PMDA) has accumulated spontaneous adverse drug effects reports for more than a dozen years in Japan. SCAR-positive or negative drugs were defined from JADER database using the Proportional Reporting Ratios (PRR) [3] method and the number of reports. Chemical structure information on extracted drugs was obtained from PubChem [4] and molecular descriptor calculations were performed using Dragon 7 [5]. A predictive model was built by Deep Learning using the calculated molecular descriptors. In addition, in order to make more reliable prediction, applicability domain was defined [6]. As a result, we achieved classification of SCAR with approximately 70% prediction accuracy from only the structure information of drugs. This method is might be useful to improve efficiency of new drugs development.

[1] https://www.meddra.org/sites/default/files/guidance/file/smq_intguide_20_1_japanese.pdf
[2] https://www.pmda.go.jp/safety/info-services/drugs/adr-info/suspected-adr/0005.html
[3] Kubota K., Signal Detection from Spontaneous Reports: New Methods in MCA in the UK, FDA in the US and WHO, *Jpn J pharmacoepidemiol*, 6 (2) Dec 2001: 101-108.
[4] https://pubchem.ncbi.nlm.nih.gov/
[5] https://chm.kode-solutions.net/products_dragon.php
[6] OECD, *ENV/JM/MONO*, 2, 2007.

P5-04

Analyzing Deep Neural Networks on Molecular Activity Prediction using Cross Validation Approach

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

Shinji Hamada[1]
hamada@adsim.tut.ac.jp

Hitoshi Goto[1]
gotoh@tut.jp

[1] Computational Chemistry Laboratory, Graduate School of Computer Science and Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi, 441-8580, Aichi, Japan.

Keywords: QSAR(Quantitative Structure-Activity Relationship), Deep Neural Networks, Cross Validation, Drug Discovery

Deep Neural Networks (DNNs), which are neural networks composed of many hidden layers, are constructed by applying a deep learning algorithm using a large number of training data. DNNs technique is applied in various fields, and especially, in Merck Molecular Activity Challenge (MMAC), the prediction accuracy was equal to or higher than those of other conventional machine learning methods [1,2].

We have performed verification experiments to the equivalent DNNs using the training and evaluation data (descriptors and activity values) sets of various targets adopted by MMAC, and reproduced almost the same results [3]. From this verifications, we found that there are a few target data sets that show high prediction accuracy, whereas many data sets cannot enough achieve the expected learning effect. To elucidate the cause of these large differences and to derive some improvement strategies, we optimized hyperparameters of DNNs and analyzed feature quantities between training and evaluation data set of each target [4,5].

In this work, we have further tried various learning methods based on the k-fold Cross Validation approach and analyzed the differences with general learning algorithms used in deep learning and the effects on prediction accuracy. Comparison with the differences in the prediction accuracy among learning methods suggests that applications of the Cross Validation may suppress the over fittings for some targets.

This research is performed with the support of the Toyohashi University of Technology Leading Program.

[1] J. Ma; R. P. Sheridan; A. Liaw; G. E. Dahl; V. Svetnik. J. Chem. Info. Model. 2015, 55 (2), 263-274.
[2] Y. Xu; J. Ma; A. Liaw; R. P. Sheridan; V. Svetnik. J. Chem. Inf. Model. 57, 10, 2490-2504.
[3] Y. Kato; S. Hamada; H. Goto. Conference: The 2016 International Conference on Advanced Informatics: Concepts, Theory and Application (ICAICTA2016), At 2016-10-43-0081.
[4] Y. Kato; S. Hamada; H. Goto. SCCJ Spring Annual Meeting 2017, 2O06.
[5] Y. Kato; S. Hamada; H. Goto. SCCJ Spring Annual Meeting 2018, 2P08.

P5-05

Deep-learning of cancer stem cell morphology for anti-cancer stem cell molecule screening

Sakae Nishisako[1]
b0115176c0@edu.teu.ac.jp

Saori Aida[2]
aidasor@stf.teu.ac.jp

Hiroyuki Kameda[2]
kameda@stf.teu.ac.jp

Tomonari Kasai[1]
kasaitnr@stf.teu.ac.jp

Atsushi Sato[1]
atsato@stf.teu.ac.jp

Tomoyasu Sugiyama[1]
tsugiyama@stf.teu.ac.jp

[1] School of Bioscience and Biotechnology, Tokyo University of Technology, 1404-1 Katakura, Hachioji, Tokyo 192-0982, Japan
[2] School of Computer Science, Tokyo University of Technology, 1404-1 Katakura, Hachioji, Tokyo 192-0982, Japan

Keywords: Artificial intelligence, Cancer stem cell, miPS-LLCcm cells

The purpose of this study is to identify cancer stem cells (CSCs) by using artificial intelligence technology (AI). Recently, it is proposed that to suppression of CSCs could be important and effective for suppressing tumour growth.

One of the most effective cancer therapies is to use anti-cancer agent. However, these drugs are not always efficient to prevent tumour recurrences. This phenomenon is thought to be due to the presence of CSCs. The surviving CSCs after the therapy might form the tumours which are not suppressed. Thus, tumour could be reproduced unless eradicating CSCs. Conversely, tumour recurrence might not be occurred if CSCs were removed completely.

Inefficiency of killing CSCs by existing agents is not surprising because these agents were selected through cell assay using cancer cell lines which lost CSC characteristics. Unfortunately, there are no straightforward cell assays to measure the ability of affecting CSC characteristics of undifferentiated state. Here, we describe an automated method of identifying CSC based on the cell morphology of CSC characteristics.

Methods: We utilized mouse iPS-LLCcm cell line as a CSC model which was previously developed (PLoS ONE 2012, 7: e33544). Pair of cell images of phase contrast and Nanog-GFP reporter fluorescence was obtained using fluorescence microscopy equipped with digital camera system. The images were used as the input into deep-learning software Conditional Generative Adversarial Nets (CGAN).

Results and Discussion: miPS-LLCcm cells were grown on Petri dish coated with porcine gelatin. Almost 90% of cells showed Nanog-GFP fluorescence indicating the maintenance of CSC characteristics under the cell culture condition. Rest of cells showed no fluoresce indicating that these cells did not maintain the characteristics. Using these cell images as teacher, phase contrast bright-field images were examined for CGAN. It took many hours for the deep-learning. For example, over 24 hours were required to learn about seven thousands images. One of the bottle necks was the main memory size of hardware. After the deep-learning, AI distinguished undifferentiated CSCs from differentiated cells using phase contrast cell image in high correlation with the corresponding pair fluorescence cell image. Finally, we will develop an AI system that can evaluate whether a test cell is a CSC or not by using phase contrast bright-field cell image.

P5-06

Simulation study comparing non-linear mixed-effect modeling and machine learning: efficient integration of individual patient-level data from multiple clinical trials

Hideki Yoshioka[1]
afga7643@chiba-u.jp

Hiromi Sato[1]
hiromi-s@chiba-u.jp

Akihiro Hisaka[1]
hisaka@chiba-u.jp

[1] Laboratory of Clinical Pharmacology and Pharmacometrics, Graduate School of Pharmaceutical Sciences, Chiba University, Inohana 1-8-1, Chuo-ku, Chiba-shi, Chiba, 260-8675, Japan

Keywords: clinical trial data, machine learning, artificial intelligence

Today, individual patient-level anonymous data from thousands of clinical trials are available via several means such as ClinicalStudyDataRequest.com, which is currently the largest consortium organized by clinical study sponsors/funders (https://clinicalstudydatarequest.com). Furthermore, submission of electronic trial data for new drug approvals is recently obliged in various countries. Therefore, a comprehensive framework to integrate big data derived from numerous studies considering various covariates as well as complicated inter-study differences would be urgently required. Based on such outlook, we attempted to perform a simulation analysis to examine how the conventional approach using non-linear mixed-effect modeling works in such a situation in comparison with machine learning that are becoming the current major stream.

A total of ~10,000 patients, each of whom has ~100 feature records and belong to any of ~100 trials, were synthesized by Monte Carlo method. Assuming clinical trials that follow up drug effects on a target biomarker, a longitudinal change in the virtual biomarker was individually simulated according to the double exponential function (y=a*exp(b*exp(c*x))), where each of the three parameters (a, b, c) has both inter-individual variability (IIV) and inter-study variability (ISV) and ~5 covariates that are randomly selected from the features synthesized. To simplify the analysis, data were transformed into a time-independent format (i.e. change from baseline at the trial completion) and then analyzed by non-linear mixed effect modeling (NONMEM v7.3; FOCE-I) and machine learning (Python v3.7.0; scikit-learn). Results were comprehensively reviewed from the viewpoints mainly of covariate identification.

In NONMEM, the hierarchical structure of variability as well as a model function need be described as a control file by the user, and this is directly passed to the software. Consequently, in our analysis, the magnitudes of IIV and ISV were successfully separated, and all the covariates that were incorporated into the model during the data synthesis were identified with an extremely high significance level by the stepwise search. However, considering the execution time required drastically increased as the data got bigger, it would not be practical to explore significant covariates by this method when the number of features to be tested is huge. On the other hand, machine learning methods generally can handle data of numerous patients with more than hundreds of features easily compared to the conventional model-dependent approach. In the present analysis, we confirmed that even without any information to distinguish clinical trials which each patient belongs to, the random forest model explained data without distinct bias and correctly proposed all the pre-incorporated covariates as features with higher importance. However, in contrast to its usability, the inside structure was highly complex, so that we were able to rank features according to importance level, but not able to quantify their impact on outcomes as clearly as given by NONMEM.

Taken together, our analysis provided a better understanding of both aspects of advantages and disadvantages for the model-dependent approach and analyses using machine learning. To make full use of clinical trial data to be accumulated in the future, further efforts to incorporate theories or analytical hypotheses based on scientific evidences into machine learning models would be necessary.

P5-07

Auto Cell Image Classification System for Micronucleus Assay by Deep Learning

Hiroshi Ueda[1]
Hiroshi_Ueda@nts.toray.co.jp

Yoshinori Wakabayashi[2]
wakabayashi@by-hex.com

Ryuji Tanimura[1]
Ryuji_Tanimura@nts.toray.co.jp

Tomohide Masuda[1]
Tomohide_Masuda@nts.toray.co.jp

[1] Pharmaceutical Research Laboratories, Toray Industries, Inc. 6-10-1 Tebiro, Kamakura, Kanagawa, 248-8555, Japan
[2] BY-HEX LLP, 1-19-14 Shimizu, Suginami-ku, Tokyo 167-0033, Japan

Keywords: Micronucleus Assay, Deep Learning, Cell Image Classification

Micronucleus assay is a genotoxicity test method to observe "micronuclei" appearing due to chromosomal abnormality. Micronucleus assay is very labor-intensive, the experts of micronucleus assay visually check thousands of cultured cells by microscopic observation and classify the presence or absence of micronuclei. So, using deep learning, we have developed a system to calculate the micronucleus ratio by automatically classifying the presence or absence of micronuclei just by inputting the microscopic image.

In order to prepare the training set, we have picked out about 61,000 images that contains single cell from 48 microscopic images by using auto cell pick-out module “Kiridasi-kun” and then we have classified the picked-out single cell images to the micronuclei presence images or absence images by using user-assist GUI module “Bunrui-kun”. Using the prepared training set, we have created a classification model with the correct answer rate 92%. Genetic programming technique[1], which is an efficient method to obtain the optimum configuration of the network layers automatically, has been applied to decide the network configuration.

On the auto classification module “Hanbetu-kun” of cell images in micronucleus assay that implements the classification model above, just by inputting a microscopic image, one can obtain the micronucleus ratio by automatically classifying the presence or absence of micronuclei.

[1] Wakabayashi, Y., Mashimo, T., Wada, M., A Genetic Approach to Deep Leaning in Prediction of Molecular Properties, *Chem-Bio Informatics Society (CBI) Annual Meeting 2017*, P2-21, 2017.

Toward AI-based Molecular Force Fields

Koichiro Kato [1] **Tomohide Masuda** [2] **Hiroshi Ueda** [2]

Tomohide_Masuda@nts.toray.co.jp

Ryuji Tanimura [2] **Kanji Oshima** [3] **Satoshi Ono** [4]
Naoki Miyagawa [1] **Hideo Mizouchi** [1] **Chiduru Watanabe** [5]
Shunpei Nagase [5] **Kikuko Kamisaka** [5] **Teruki Honma** [5,6]
Kaori Fukuzawa [7] **Atsushi Tokuhisa** [8,9] **Ryo Kanada** [8]
Shuntaro Chiba [6] **Masateru Ota** [6] **Mitsunori Ikeguchi** [6,10]

[1] Mizuho Information & Research Institute, Inc., 2-3, Kanda-Nishikicho, Chiyoda-ku, Tokyo 101-8443, Japan
[2] Pharmaceutical Research Laboratories, Toray Industries, Inc., 6-10-1, Tebiro, Kamakura, Kanagawa 248-8555, Japan
[3] Biotechnology Research Laboratories, KANEKA CORPORATION, 1-8, Miyamae-cho, Takasago-cho, Takasago, Hyogo 676-8688, Japan
[4] Modality Laboratories, Innovative Reaearch Division, Mitsubishi Tanabe Pharma Corporation, 1000, Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan
[5] RIKEN Center for Biosystems Dynamics Research, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[6] RIKEN Cluster for Science and Technology Hub, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[7] School of Pharmacy and Pharmaceutical Sciences, Hoshi University, 2-4-41, Ebara, Shainagawa-ku, Tokyo 142-8501, Japan
[8] RIKEN Cluster for Science and Technology Hub, 6-3-5, Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
[9] RIKEN Center for Computational Science, 6-3-5, Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
[10] Graduate School of Medical Life Science, Yokohama City University, 1-7-29, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Keywords: artificial intelligence, machine learning, deep learning, molecular force field

The accuracy of molecular simulation such as molecular dynamics (MD) simulation is strongly dependent on the molecular force field. We have been developing artificial intelligence (AI) -based molecular force fields (AI models) which enable us to conduct various kinds of molecular simulations with quantum mechanics (QM) accuracy at molecular mechanics (MM) cost.

Concretely, two types of AI models have been worked on. One is constructed by learning the potential energy obtained by QM simulation without the conventional potential energy function based on physical chemistry description [1]. Since the model is designed to be differentiable, forces acting on each of atoms should be obtained by analytically differentiating the learned potential energy surface. This model is intended to be used as high-accuracy force field. The other is constructed by learning the environmental depending atomic charges obtained by QM simulation. Using fragment molecular orbital (FMO) method among QM simulations [2] allows us to prepare training data of atomic charges for protein systems with acceptable computational cost. “On-the-fly” atomic charges generated by the model lead to improve the accuracy of molecular simulations, since it is a part of parameters in the conventional molecular simulations.

We will discuss the prediction accuracy of the two models including the future applications.

Acknowledgments: This research has been conducted by cooperation of LINC (Life Intelligence Consortium) and FMODD (FMO Drug Design Consortium). The results of FMO have been achieved using the “K computer” (Project ID : hp180147). The authors are deeply grateful to them.

[1] J. Behler and M. Parrinello, *Phys. Rev. Lett.*, 98, 146401 (2007)
[2] S. Tanaka *et al.*, *Phys. Chem. Chem. Phys.*, 16, 10310 (2014)

P5-09

Effects of single mutations on STING activation

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

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

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

[1] Artificial Intelligence Research Center, Advanced Industrial Science and Technology, 2-4-7 Aomi, Koto-ku, Tokyo, 135-0064, Japan
[2] Medical Sciences Innovation Hub, RIKEN, 54 Shogoin Kawaharacho, Sakyo-ku, Kyoto, 606-8507, Japan
[3] High Pressure Protein Research Center, Institute of Advanced Technology, Kindai University, 930 Nishimitani, Kinokawa, Wakayama, 649-6493, Japan

Keywords: Molecular dynamics simulation, Feature extraction, Deep neural network, Conformational change by mutations

STING, a protein that plays an important role in innate immunity, induces type I interferon (especially IFN-ß) production when cell is infected with intracellular pathogens, such as viruses. The phosphorylation of C-terminal tail (CTT) of STING by TBK1, subsequent to the conformational change of CTT via ligand binding, leads to the IRF3-dependent IFN-ß production [1]. Recent studies, however, have revealed that several single mutations of STING, found in patients with autoimmune disorders, trigger constitutive IFN-ß productions, IRF3-indepedently, via CTT.

In this study, we tried to elucidate effects of conformational changes triggered by such mutations on constitutive activations, based on the molecular dynamics and feature extractions by deep neural network.

[1] Tsuchiya, Y., *et al.*, Ligand-induced ordering of the C-terminal tail primes STING for phosphorylation by TBK1, *EbioMedicine*, 9:87-96, 2016.

P5-10

Deep Learning-aided Label-free, Non-invasive Method for Live/dead Cell Discrimination and Counting

Yuta Kawamura[1], **Miho Hosoi**[1], **Mie Tsuchida**[1], **Makoto Hasegawa**[1], **Ryuzo Sasaki**[1, 2], **Shigeru Nemoto**[3], **Katsumi Kishimoto**[2], **Tamio Mizukami**[1, 2]

mizukami@nagahama-i-bio.ac.jp

[1] Nagahama Institute of Bio-Science and Technology, 1266 Tamura, Nagahama, Shiga, 526-0829, Japan
[2] Frontier Pharma, 1281-8 Tamura, Nagahama, Shiga, 526-0829, Japan
[3] Broad Band Tower, 2-1-6 Uchisaiwaicho, Chiyoda-ku, Tokyo 100-0011, Japan

Keywords: Deep learning, Imaging, Drug sensitivity test

Live and dead cell discrimination and counting are fundamental techniques used in various aspects of life science, medicine and drug discovery. At present, representative techniques include trypan blue dead cell staining, and measurement of the number of living and dead cells by a hemocytometer, but they have a disadvantage that it is difficult to measure multiple specimens. Furthermore, though it is possible to measure the number of viable or dead cells by image analysis of cells treated with a fluorescent reagent which stains viable or dead cells using high-content imaging equipment, it has a disadvantage that it is difficult to apply to cell types with complex morphology and besides expensive equipment is necessary. In addition, a common disadvantage of the above-mentioned techniques is necessity of invasive interruption of cell culture due to addition of staining reagents to cells.

Therefore, we devised and completed the cell viability identification and counting technology which solves the above problem. The present technology is aided by deep learning that enabled the image generation and position determination of cells as follows.

Firstly, we built a model that learned the relationship between the fluorescent microscopic image of the cells treated with the live and dead cell staining fluorescent reagents and the corresponding bright field image by deep learning. Without experimental work such as fluorescent labeling, it was possible to predict and generate pseudo-fluorescent cell images with high correctness from bright-field cell images input to the constructed model.

Secondly, we built another model (position determination model) to determine the position of the cell to be detected on the generated pseudo-fluorescent cell image. This model learned the relationship between the image created by manually striking the center position of the cell on the fluorescent cell image and the corresponding fluorescent cell image by deep learning. By inputting a pseudo-fluorescent cell image to this model, an image in which the center position of the cell on the fluorescent cell image was struck was generated.

By the separately prepared program, the number of detected representative points was automatically counted, and the cell number could be measured with high accuracy.

These findings will allow us to create the next-generation drug sensitivity test with such advantageous characteristics that are not only label-free and non-invasive, but also simple, fast, inexpensive, and versatile.

P5-11

Investigation of general physical observation method of monkeys using deep learning

Yosuke Ochiai, Tetsuo Kitamura,
ochiai.yousuke@mn.medience.co.jp kitamura.tetsuo@mh.medience.co.jp

Tomoharu Osada, Yoshinori Yamamoto,
osada.tomoharu@mg.medience.co.jp yamamoto.yoshinori@mg.medience.co.jp

Munehiro Nakagawa, Hiroyuki Izumi,
nakagawa.munehiro@md.medience.co.jp izumi.hiroyuki@mu.medience.co.jp

Kaname Takahashi,
takahashi.kaname@mm.medience.co.jp

Nonclinical Research Center, Drug Development Service Segment, LSI Medience Corporation.
14-1, Sunayama, Kamisu-shi, Ibaraki 314-0255, Japan

Keywords: Deep learning

[Introduction] Recent technological innovations have greatly influenced the development of new drugs. Deep learning is one of those innovations. With the use of deep learning, it has become realistically possible to utilize more experimental data for drug development than when it was conducted by humans alone. We have done this research with hopes that observation of animals, which had been carried out by humans, can be conducted for much longer periods of time with the utilization of deep learning. The final goal of this research is to create a model that can classify toxic behaviors of animals such as changes in activity, abnormal posture, ataxia and vomiting and/or retching from video recorded over long periods of time. As a preliminary step, we herein report classification of monkey postures using still images.
[Methods] Still images taken from monkeys' videos were used for classification. Still images of six monkeys were used as training data. For classification, postures that are typically taken by monkeys in addition to sitting and standing postures (including on four legs) such as walking and so on were labeled. The same still images were used for both training data and test data.
[Results] The accuracy of classification was 75%.
[Conclusion] Despite the need for improvement of the labels of model and classifications, it was possible to classify still images of sitting posture, standing posture, and walking style. In the future, training data will be increased, more models will be created and classification of still images of monkeys other than the ones acquired from the training will be attempted.

P5-12

Development of Template-Based Protein Structure Modeling Method Using Deep-Autoencoder with a Denoising Algorithm.

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

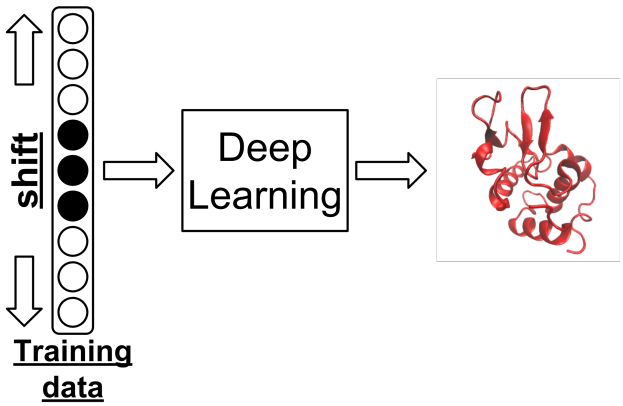
Mitsutaka Nemoto[1]
nemoto@waka.kindai.ac.jp

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

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

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

Information about the dynamics and tertiary structure of a target protein is essential to understand the function of the target protein and the drug discovery and so on. However, the tertiary structures of many proteins have not been solved, while the genome sequences have solved by DNA sequencing technique. Thus, a protein, which tertiary structure is unknown, have sometimes modeled by template structures of other similar proteins, which have tertiary structure data, e.g. Modeller [1] and SWISS - MODEL [2] etc. Recently, the technique of Deep Learning has been developed, and it works very well in an image recognition and so on. We, thus, have developed the template based structural modeling program with Deep-Autoencoder. The trajectories of the molecular dynamics simulations of a target protein were used as training data, which is considered the molecular fluctuations. Moreover, to ease the sequence mismatch between the target protein and template proteins, we have adopted a denoising algorithm, which shifts the coordinate of atoms, in the program.



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

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

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

Three-dimensional Quantitative Structure-Activity Relationship Analysis using Convolutional Neural Network

Hirotomo Moriwaki[1]
hirotomo.moriwaki@riken.jp

Yu-Shi Tian[2]
yushi-tian@phs.osaka-u.ac.jp

Norihito Kawashita[3]
nkawashita@life.kindai.ac.jp

Tatsuya Takagi[2]
ttakagi@phs.osaka-u.ac.jp

[1] RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
[2] Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita City, Osaka 565-0871, Japan
[3] Faculty of Sciences and Engineering, Kindai University, 3-4-1 Kowakae, Higashiosaka City, Osaka 577-8502, Japan

Keywords: 3D-QSAR, convolutional neural network

Quantitative structure–activity relationship (QSAR) techniques, especially those that possess three-dimensional (3D) attributes, such as the comparative molecular field analysis (CoMFA)[1] are frequently used in drug design and other related research domains since the late 1990s. However, the requirement for accurate multiple alignments of compounds increases the difficulties encountered in its use. This has led to the development of several techniques—such as VolSurf[2], GRIND[3], and Anchor-GRIND[4]—which do not require such alignments. We propose a new 3D-QSAR technique that makes use of molecular-interaction-field (MIF) grid potentials as inputs to a convolutional neural network. The model built by our new method has demonstrated higher discrimination accuracy compared to the conventional descriptor-based QSAR as well as Anchor-GRIND techniques. In addition, the method doesn't require a dataset consisting of a common chemical series, thus can apply to diverse datasets. The method is capable of providing useful information regarding the importance of individual atoms for enhancement of potency contained in the chemical dataset used in the analysis. In view of these advantages, the proposed method is expected to find wide applications in future drug-design operations.

[1] Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110* (18), 5959–5967.
[2] Cruciani, G.; Pastor, M.; Guba, W. VolSurf: A New Tool for the Pharmacokinetic Optimization of Lead Compounds. *Eur. J. Pharm. Sci.* **2000**, *11* (Supplement 2), S29–S39.
[3] Pastor, M.; Cruciani, G.; McLay, I.; Pickett, S.; Clementi, S. GRid-INdependent Descriptors (GRIND): A Novel Class of Alignment-Independent Three-Dimensional Molecular Descriptors. *J. Med. Chem.* **2000**, *43* (17), 3233–3243.
[4] Fontaine, F.; Pastor, M.; Zamora, I.; Sanz, F. Anchor-GRIND: Filling the Gap between Standard 3D QSAR and the GRid-INdependent Descriptors. *J. Med. Chem.* **2005**, *48* (7), 2687–2694.

P5-14

Text mining for expand coverage/combination therapy/adverse effect about immune checkpoint blockade

Kosuke Negishi[1]
k.negishi@elsevier.com

1 Life Science, Elsevier Japan K.K,1-9-15 Higashi-Azabu Minato-ku, Tokyo 106-0044 Japan

Keywords: Drug discovery, Natural language processing, Text Mining.

After marketed Nivolumab, some immune checkpoint blockade for anti-cancer agent have been developed. By this trend, pharmaceutical companies have considered these points.

1: expand coverage of efficacy
2: combination therapy with small molecule(s)
3: data collection about adverse effect

To solving these points, comprehensive journal information search and curation are effective. On the other hands, this is so hard to solve it. Because,

1: Data preparing from full text journal/PubMed/PubMed Central
2: Search/curation method from huge data contents (one of Big data)
3: Comprehensive journal information visualization

By this session, I would like to propose text mining solution to solve these points. It has natural language processing algorithm and comprehensive journal information from over 1400 title full text journal/PubMed/PubMed Central

ChemAtlas: chemical space extractor

Tetsuro KITAJIMA[1]
kitajima@isp.co.jp
Hitoshi KAMIJIMA[1]
kamijima@isp.co.jp
Hideo TSURU[1]
tsuru@isp.co.jp
Shinji YAMAMOTO[1]
ymt@isp.co.jp

[1] Research Institute of Systems Planning Inc., Kasuya Bldg. II 2nd floor, 23-23 Sakuragaoka, Shibuya, Tokyo 150-0031, Japan

Keywords: Deep Learning, chemical space

We are developing chemical space extractor "ChemAtlas". ChemAtlas can be used for visualizing chemical space, QSPR/QSAR task, and similar tasks.
Approaches that make use of intermediate layers of deep neural networks are reported to be effective for image recognition tasks and for tasks involving chemical compounds. Based on these reports, we decided to develop a new chemical space extractor based on the approach. ChemAtlas uses multitask deep learning and graph convolutional networks in obtaining intermediate states and aims to be useful for wide range of applications.
ChemAtlas has performance equivalent to ECFP 3072-bit hash in PubChem Assay prediction and its performance in SLogP prediction is greater than that of ECFP hash. (Shown in the figure 1)
We will report our results of chemical space visualization using ChemAtlas and the results of carrying out QSPR task.

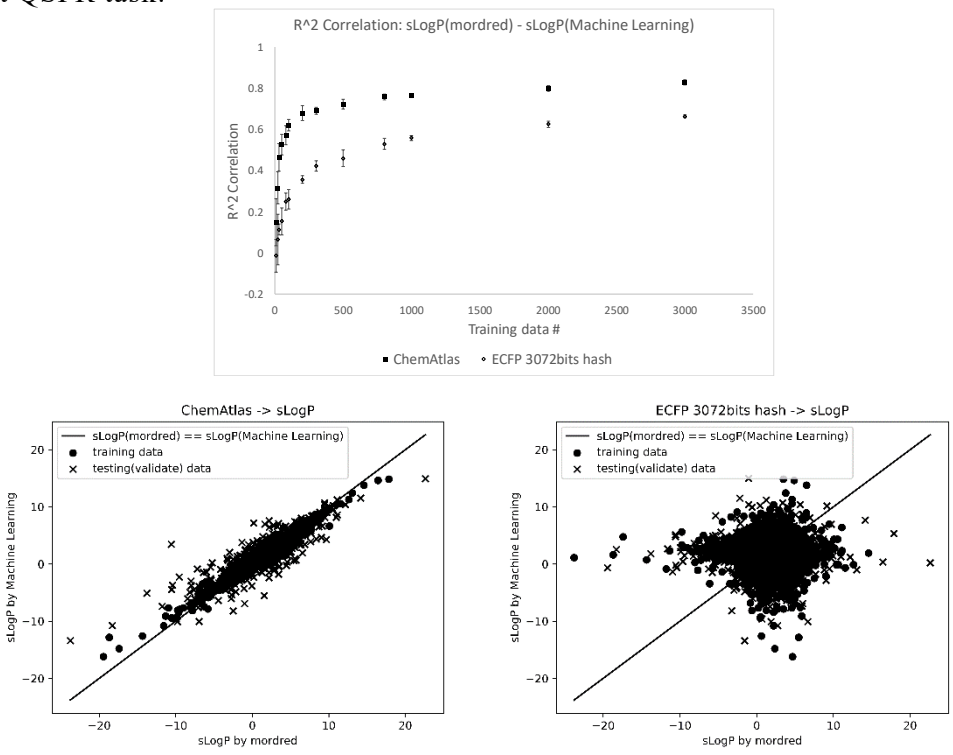


Figure 1 SLogP prediction: using ChemAtlas and ECFP

calculated by mordred - predicted by machine learning R^2 scores, plots

P5-16

Using domain-specific vocabulary to detect multiple-word phrases to improve word2vec embedding performance in Medical literature

Attayeb Mohsen[1]
attayeb@nibiohn.go.jp

Yi-An Chen[1]
chenyian@nibiohn.go.jp

Kenji Mizuguchi[1]
kenji@nibiohn.go.jp

[1] National institutes of Biomedical Innovation, Health, and Nutrition, 7-6-8, Saito-Asagi, Ibaraki City, Osaka 567-0085, Japan

Keywords: Word2vec, Deep learning, Medical literature.

Word2Vec[1] used neural network to represent text in a vector space. Each word is represented in a form of continuous vector. The values of these vectors are influenced by the context they are used in. Word2Vec representation can be used to understand the relations between words; It can also show the similarity between the words in usage and meaning.

Many medical terms consist of more than one word, for example, “Diabetes mellitus” and “Myocardial infarction.” These words combined can refer to a medical entity like a disease, sign, or procedure among others while using one word of them does not tell the same meaning.

Multiple-words phrases affect the training of word2vec model. There are different methods to detect these phrases to feed them to the machine-learning model as one word. In this study, we trained three models using PubMed abstracts (https://www.ncbi.nlm.nih.gov/pubmed/). The first model trained using one-word representation without considering multiple-words terms. In the second model, we used Phrase2Vec to detect them without human interference. The third model utilized a preprocessing step to convert multiple-word terms using a vocabulary list to one word by replacing the space with a hyphen. We assessed these three models using a set of random terms. We found out that using a vocabulary list to detect multiple-words phrases and concatenate them showed better performance among these three models.

[1] Mikolov, T., Sutskever I., Chen K., Corrado G, Dean J. *Distributed Representations of Words and Phrases and their Compositionality*. arXiv:13104546. 2013. http://arxiv.org/abs/1310.4546.

Development of a novel linear notation of chemical compounds for deep learning

Juanjuan Lu[1]
lu.j.ac@m.titech.ac.jp

Keisuke Yanagisawa[1]
yanagisawa@bi.c.titech.ac.jp

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

[1] School of Computing, Department of Computer Science, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan

Keywords: Variational auto-encoder, SMILES, Molecule generation, chemical compound

The variation of chemical compounds in this world is very large, and thus those we have already known are only a small part of them. Within those unknown part, there can be many drug-like molecules. In 2016, Gómez-Bombarelli *et al*.[1] proposed a variational auto-encoder model to generate drug-like molecules. They used SMILES[2] to represent molecules and let them be the input data of this model. However, in this model, it usually generated invalid molecules. To improve the model, Kusner *et al*.[3] proposed a grammar variational auto-encoder. In this model, they used an encoder to account the SMILES grammar, and the generation rate of valid molecules is twice higher than former one. However, even using such techniques, the generation rate is only 0.31. It is still low and not practical.

Through these researches we found that some rules of SMILES may cause harmful influences for generating valid molecules, and it is difficult to overcome if we continue to use SMILES. For example, changing one number in a fragment will generate invalid molecules. Thus, here we tried to create a new notation to represent molecules instead of SMILES. The notation is much easier to separate molecules into fragments, and then we use a sequence of fragments as the input data of variational auto-encoder to generate drug-like molecules.

[1] Gómez-Bombarelli, Rafael, et al. "Automatic chemical design using a data-driven continuous representation of molecules." *ACS central science* 4.2 (2018): 268-276.
[2] Weininger, David. "SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules." *Journal of chemical information and computer sciences* 28.1 (1988): 31-36.
[3] Kusner, Matt J., Brooks Paige, and José Miguel Hernández-Lobato. "Grammar variational autoencoder." *arXiv preprint arXiv:1703.01925* (2017).

Development of interatomic potential building package based on artificial neural network and an application for the alkaline deep sea hydrothermal vent environments

Kota Endo[1]
180x203x@stu.kobe-u.ac.jp

Kohei Shimamura[1]
shimamura@port.kobe-u.ac.jp

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

[1] Graduate School of System Informatics, Kobe University, 1-1, Rokkodai, Nada, Kobe 657-8501, Japan

We aim at the development of interatomic potential building package based on artificial neural network (ANN potential) for molecular dynamics (MD) simulation on extensive biomolecular systems. This is because the ANN potential is an empirical potential with the same accuracy as the *ab initio* MD (AIMD) method, [1] which can simulate chemical reactions including proton transfers that occur in living bodies. In the development procedure, a set of Atomic FingerPrint (AF) [1] which quantifies the characteristic local structure around each atom is constructed for each element. Using the AF sets as feature quantities, the learned ANN potential can completely reproduce the *ab initio* interatomic potential within the range of the training data. Combining AIMD simulation and learning ANN potential enables us to accelerate the MD simulation while maintaining high accuracy. [2] Specifically, we first learn ANN potential while conducting AIMD simulation, and then switch to ANN potential when it can predict reference energies with accuracy of more than 99%. On the other hand, if the simulation encounters an atomic structure with poor prediction accuracy, switch to AIMD simulation while learning ANN potential again. The unpredicted structure can be judged by the values of AF sets. [2]

However, since many of existing ANN potential building packages are targeted for CPU calculators, expensive large scale parallel computers are necessary. Therefore, we have developed a package using the Tensorflow library [3] executed on GPU calculators because GPU is cheaper and matrix calculation performed during learning is overwhelmingly faster than CPU.

We use our package to simulate the alkaline deep sea hydrothermal vent environment. [4] This environment is a strong candidate as the place of birth of life due to abundant electron sources and catalysts, where chemical reactions can be sustained. This environmental condition is essential to support metabolism and homeostasis of primitive life. We have already carried out AIMD simulations reproducing the hydrothermal vents environment, but problems arise in terms of calculation cost. Because of AIMD method with calculation cost of $O(N^3)$ (N is the number of atoms in the system), calculation time and the number of atoms are limited to several ps and ~1,000 atoms, respectively. We are therefore trying to execute a simulation in combination of the AIMD method with the ANN potential. We would like to report the simulation results at the poster presentation.

[1] Behler, J., *J. Phys.: Condens. Matter*, 26:183001, 2014.
[2] Botu, V., *et al.*, *Int. J. Quan. Chem.* 115:1074, 2015.
[3] https://www.tensorflow.org/guide/using_gpu
[4] Sojo, V., *et al.*, Astrobiology, 16:181, 2016.

Prediction of toxicity through the chemical space generated by deep learning

Yoshinori Wakabayashi[1]
wakabayashi@by-hex.com
Kenichi Naito[1]
naito@by-hex.com
Kazuyoshi Ikeda[2]
ikeda-kz@pha.keio.ac.jp

[1] BY-HEX LLP, 1-19-14 Shimizu, Suginami-ku, Tokyo 167-0033, Japan
[2] Keio University Faculty of Pharmacy Division of Physics for Life, 1-5-30 Shibakoen, Minato-ku, Japan

Keywords: deep learning, toxicity, chemical space, fingerprint, descriptor, ChEMBL

Exploring the chemical space spanned by various compound libraries is beneficial to understand the structure-activity relationship or to develop potential drugs. Chemical space maps are commonly constructed by a principal component analysis from high dimensional molecular descriptors such as fingerprints and physical properties. However, there seems to be no distinctive criteria about the choice of descriptors for deep learning by the conventional methods. In this work, we present a new method to automatically depict chemical space map from SMILES strings without complicated procedures of generation and selection of descriptors. A precursor layer of three nodes was appended just before the output layer on the DeepChem framework to retrieve the values for mapping chemical space. As a preliminary study, we applied this technique to the prediction of toxicity.
We will discuss validity of the prediction results and its applicability for bioactive chemical space using curated drug discovery data from ChEMBL.

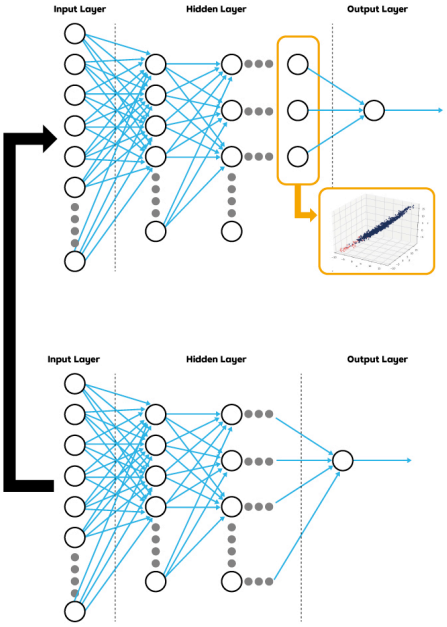


Diagram of our developed DNN method and chemical space

Performance Evaluation of Compound-Protein Interaction Prediction using Graph Convolutional Network

Hiroaki Iwata[1] iwata.hiroaki.3r@kyoto-u.ac.jp
Taro Kiritani[2] taro.kiritani@exwzd.com
Ryosuke Kojima[1] kojima.ryosuke.8e@kyoto-u.ac.jp
Shouichiro Ishida[1] ishida.shouichi.57a@st.kyoto-u.ac.jp
Takeshi Fujiwara[3] takeshi.fujiwara@nibiohn.go.jp
Mari Ito[3] mari@nibiohn.go.jp

[1] Graduate School of Medicine, Kyoto University, 53 Kawahara-cho, Syogoin, Sakyo-ku, Kyoto 606-8507, Japan
[2] ExaWizards Inc. 1-6-15, Hamamatsu-cho, Minato-ku, Tokyo, 105-0013, Japan
[3] National Institutes of Biomedical Innovation, Health and Nutrition Bioinformatics project, 7-6-8 Saito-Asagi, Ibaraki City, Osaka, 567-0085, Japan

Keywords: graph convolutional network, compound-protein interaction, ChEMBL database

The graph convolutional network (GCN) learns chemical and biophysiological properties of compounds based on their structural information [1,2]. We developed and evaluated performance of the GCN model for compound-protein interaction using bioassay dataset from the ChEMBL database[3]. The comparison between single-task model, and multitask model indicated the multitask approach was superior. We further conducted an accuracy assessment between the proposed GCN model, the conventional deep neural network model using compound descriptors [4], and the target prediction model that was developed at ChEMBL group[5,6]. Finally, we discuss an extension algorithm of the prediction model and its application to drug repositioning.

[1] Wu, Z., et al. MoleculeNet: a benchmark for molecular machine learning. *Chemical science*, 9.2: 513-530, 2018.
[2] DeepChem: Deep-learning models for Drug Discovery and Quantum Chemistry, http://github.com/deepchem/deepchem, accessed 2018-08-02.
[3] Gaulton, A., et al. The ChEMBL database in 2017. *Nucleic acids research*, 45.D1: D945-D954, 2016.
[4] Ma, J., et al. Deep neural nets as a method for quantitative structure–activity relationships. *Journal of chemical information and modeling*, 55.2: 263-274, 2015.
[5] Chen, B., et al. Evaluation of machine-learning methods for ligand-based virtual screening. *Journal of computer-aided molecular design*, 21.1-3: 53-62, 2007.
[6] chembl_target_predictions: https://github.com/chembl/chembl_target_predictions, accessed 2018-08-02.

P6-01

Universal Read-Across Approach To Predict Toxicities

Hirohisa Nagahori[1]
nagahori@sc.sumitomo-chem.co.jp

Masayuki Sato[1]
satom13@sc.sumitomo-chem.co.jp

[1] Environmental Health Science Laboratory, Sumitomo Chemical Co., Ltd., 1-98, Kasugadenaka 3-chome, Konohana-ku, Osaka 554-8558, Japan

Keywords: Structure-Activity Relationships, Animal Testing Alternatives, Read-Across Approach, Binomial Distribution, Validation Studies

To make read-across (RA) approach more reliable, robust, reproducible, easy, and high-throughput, the universal method was developed by using database and function in the QSAR toolbox. The approach consisted of three steps: negative and positive local category approaches, where the category was locally defined by functional group in the target substance, and analogue approach (Figure 1). Acute oral toxicity, bacterial reverse mutation, eye and skin irritation, and sensitization were practically predicted with clarification of mechanistic categories of the positive substances. Reliability in local category approaches was assessed by both the p-values calculated by cumulative distribution function of the binomial distribution, and similarity index (SI) of optimized k-nearest neighbors (k-NN). Applying these approaches to known 31 positives and 15 negative reference substances as the validation data set, the prediction was highly accurate with identification of known toxicity mechanisms. Then the toxicities of over 200 test and 50 in-house datasets were predicted and compared with the experimental results. High accuracy of generally >60% was kept in wide ranges of molecular weight (0 to 1200 Da) and lipophilicity (log K_{ow} -2 to ≥6). By considering the criterion for reliable prediction as the p-value of <0.05 and SI of ≥50%, the relatively high performance of prediction was obtained with accuracy of ≥95%, the coverage of ≥30%, and positive and negative likelihood ratio of ≥30 and <0.1, respectively, which suggested this approach was a practically useful even for screening and definitive diagnoses. The accuracy and coverage in all data sets and endpoints were high (80 to 100% and 20 to 60%, respectively), which revealed the wide applicability of the universal RA approach. The results suggested this universal RA approach was useful and reliable with wide-application to reduce animal toxicity studies.

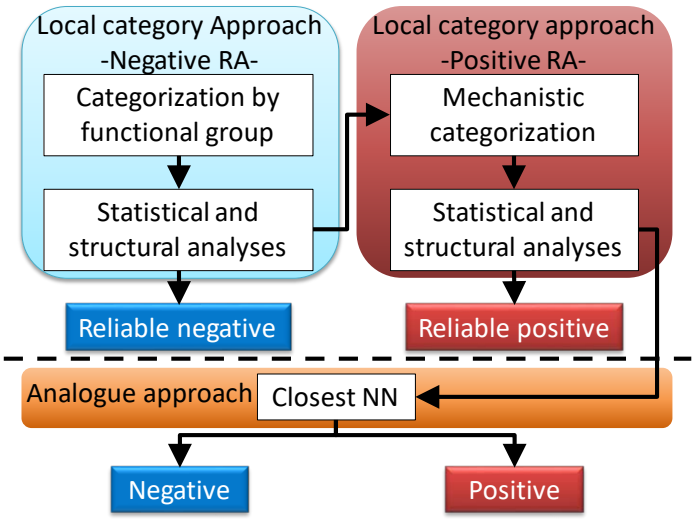


Figure 1. An illustration of universal RA approach.

P6-02

Application of association rule mining in adverse event reporting system

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

Haruna Hatahira[2]
125063hz@gmail.com

Misa Naganuma[1]
135049@gifu-pu.ac.jp

Junko Abe[1, 3]
junko.a822@gmail.com

Sayaka Sasaoka[4]
ssayaka3307@gmail.com

Ryogo Umetsu[1, 5]
umetsur.ayni@gmail.com

Satoshi Nakao[1]
145045@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] Department of Pharmacy, Kizawa Memorial Hospital, 590 Kobi-cho Shimokobi, Minokamo, Gifu, 505-8503, Japan
[3] Medical Database Co., LTD; 3-11-10 Higashi, Shibuya-ku, Tokyo 150-0011, Japan
[4] Department of Pharmacy, Hokushin General Hospital, Nishi 1-5-63, Nakano, Nagano, 383-8505, Japan
[5] Micron Inc., Nihonbashi Nishikawa Building, 1-5-3, Nihonbashi, Chuo-ku, Tokyo, 103-0027 Japan

Keywords: association rule mining, adverse event, spontaneous reporting system, JADER

Spontaneous reporting systems (SRSs) are useful for the detection of rare adverse events (AEs) and have been recognized as primary tools for pharmacovigilance that reflect the realities of clinical practice. Association rule mining has been proposed in as an analytical approach in order to study rare AE, and is a well-established method for discovering undetected relationships such as the possible risk factors between variables in huge databases.

Data were extracted from the Japanese Adverse Drug Event Report (JADER) database on the PMDA website (www.pmda.go.jp). The association rule mining was performed using the apriori function of the arules library in the arules package of the R software.

We demonstrate several association rules from the JADER database as follows: renal impairment related to the administration of platinum compounds, drug-induced gingival hyperplasia related to suspected drugs (immunosuppressants, calcium channel blockers, and anticonvulsants), and thromboembolic AEs related to combined estrogen-progestin preparations.

[1] Hasegawa S., Matsui T., Hane Y., Abe J., Hatahira H., Motooka Y., Sasaoka S., Fukuda A., Naganuma M., Hirade K., Takahashi Y., Kinosada Y., Nakamura M., Thromboembolic adverse event study of combined estrogen-progestin preparations using Japanese Adverse Drug Event Report database, *PLoS One*, 12, e0182045, 2017.

[2] Hatahira H., Abe J., Hane Y., Matsui T., Sasaoka S., Motooka Y., Hasegawa S., Fukuda A., Naganuma M., Ohmori T., Kinosada Y., Nakamura M., Drug-induced gingival hyperplasia: a retrospective study using spontaneous reporting system databases, *Journal of Pharmaceutical Health Care and Sciences*, 3, 19, 2017.

[3] Naganuma M., Motooka Y., Sasaoka S., Hatahira H., Hasegawa S., Fukuda A., Nakao S., Shimada K., Hirade K., Mori T., Yoshimura T., Kato T., Nakamura M., Analysis of adverse events of renal impairment related to platinum-based compounds using the Japanese Adverse Drug Event Report database, *SAGE Open Medicine*, 6, 1-11, 2018.

P6-03

Quantitative Structure-Activity Relationship (QSAR) analysis using deep learning based on Deep Snap, a novel molecular image input technique

Yasunari Matsuzaka matsuzys@my-pharm.ac.jp

Yoshihiro Uesawa uesawa@my-pharm.ac.jp

Department of Medical Molecular Informatics, Meiji Pharmaceutical University, Tokyo,

Keywords: DeepSnap, Tox21, Deep learning, QSAR, artificial intelligence

As for the safety assessment, a repetition dosage subacute toxicity test is used for 28 to 90 days in mammals. However, in general, the cytotoxicity in human could be caused by long-term exposure of various chemicals. Because it is necessary to perform the examination for a large number of chemical substances, high-throughput (HTP) assay and reduction in cost are required. However, prediction programs that precisely and highly evaluate the cytotoxicity for various chemical compounds are limited. Alternative methods based on in silico experiments focusing on computer prediction are essential for safety evaluation of chemical substances with high risk. Among them, Quantitative Structure-Activity Relationship (QSAR) analysis can predict physiological activity, toxicity, enzymatic reactions, receptor agonist/antagonist activities, and environmental fate, etc., from formulating establishment rule between chemical structure and its activity using the structural, quantum chemical, and physicochemical features represented as various numerical molecular descriptor [1-4]. More recently, deep learning, machine learning method designed to recognize pattern by multilayer neural network consisted of input, intermediate, and output layers, reported very high prediction accuracy, especially in the field of imaging and toxicology, and possess a possible function that can calculate feature values from molecular structure in QSAR analysis without human intervention [5-8]. However, there are some points to be improved due to increase prediction accuracy for application of the deep learning into QSAR analysis, that is, one is a systematic and suitable input for complicated data such as three-dimension structure of chemical compound, another is inadequate data volume for training data. To dissolve these tissues, we developed a novel QSAR model using deep learning based on three-dimension molecular imaging of chemical compounds [9]. Deep Snap, which is procedure capturing the omnidirectional snapshot, portray three-dimensional figures of chemical compounds using chemical structure drawing software, Jmol [10] based on the SDF file format developed by grant in Long-Range Research Initiative (LRI), Japan Chemical Industry Association (JCIA). To input the three-dimensional information into deep learning models without calculating structural descriptors, the three-dimensionally molecular structure is rotated in 45°increments on the X-, Y-, and Z-axes, and photographed for eight viewing directions. A total of 512 images were captured per one molecule and saved as PNG files due to be applied into deep learning model, where actual images of the three-dimensionally molecular structure were indicated as a ball-and-stick model for each atomic symbol with different colors representing different atoms. Using all imaging data obtained from approximately 7,000–9,000 different chemical structures on Tox21 data challenge 2014, a QSAR competition held by the National Institutes of Health in the United States [11], a deep learning model was constructed by AlexNet [12], a convolutional neural network implemented on the deep learning framework Caffe [13]. As results, the deep learning model using Deep Snap procedure successfully predicted chemical compounds that induce mitochondrial membrane potential disruption indicated by scoring an AUC value of 0.921 despite using unadjusted AlexNet, resulted in within top 10 at this competition [11]. We would use the novel prediction method using deep learning based on Deep Snap as one of QSAR analysis in AI-Substances Hazardous Integrated Prediction System (AI-SHIPS) project.

[1] Dougall, I. G. *J. Recept. Signal Transduct. Res.* **2001**, *21*, 117–137.
[2] Fang, H.; Tong, W.; Welsh, W. J.; Sheehan, D. M. *J. Mol. Struct.* **2003**, *622*, 113–125.
[3] Roy, K.; Das, R. N. *Curr. Drug Metab.* **2014**, *15*, 346–379.
[4] Silva, F. T.; Trossini, G. H. G. *Med. Chem.* **2014**, *10*, 441–448.
[5] LeCun, Y.; Bengio, Y.; Hinton, G. *Nature.* **2015**, 521, 436–444.
[6] Zhang, L.; Han, D.; Tan, J.; Zhu, H. *Drug Discov. Today* **2017**, *22*, 1680–1685.
[7] Pastur-Romay, L. A.; Cedron, F.; Pazos, A.; Porto-Pazos, A. B.; *Int. J. Mol. Sci.* **2016**, *17*.
[8] Mayr, A.; Klambauer, G.; Unterthiner, T.; Hochreiter, S. *Front. Environ. Sci*. **2016**.
[9] Uesawa, Y., "Prediction systems, prediction methods, prediction programs, learning model input data generation systems, and learning model input data generation programs", Pending Japanese patent application 2017, 129823. 2017.
[10] Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org/
[11] Tox21 data challenge 2014, National Center for Advancing Translational Sciences: https://tripod.nih.gov/tox21/challenge/
[12] ImageNet Classification with Deep Convolutional Neural Networks. Alex Krizhevsky, Ilya Sutskev, Geoffrey E. Hinton. Advances in Neural Information Processing Systems 25 (NIPS 2012), pp.1-9 (2012)
[13] Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; Darrell, T. *CVPR.* **2014**, 675–678.

Co-Creation and Communication for Real-Time Technology Assessment (CoRTTA) on Molecular Robotics

Ryuma Shineha[1] r_shineha@seijo.ac.jp
Ken Kawamura2[1] kawamura310@u.seijo.ac.jp
Go Yoshizawa[3] gy20@jcom.home.ne.jp
Daisuke Yoshinaga[4] sc.yoshinaga@gmail.com
Mikihito Tanaka[4] steman@waseda.jp

[1] Faculty of Arts and Literatures, 6-1-20 Seijo University, Seijo Setagaya-ku, Tokyo 157-8511, Japan.
[2] The center for social studies of science and technology, 6-1-20 Seijo University, Seijo Setagaya-ku, Tokyo 157-8511, Japan.
[3] Oslo Metropolitan University, Norway.
[4] Waseda University, 1-6-1 Nishiwaseda Shinjuku-ku, Tokyo 169-8050, Japan

Keywords: Real-Time Technology Assessment, Molecular Robotics, Biotechnology in Society

Our Co-Creation and Communication for Real-Time Technology Assessment (CoRTTA) project aims to extract agendas of discussion by media analysis and by the predictive estimation method (horizon scanning) regarding the Ethical, Legal and Social Issues (ELSI) of molecular robotics. Furthermore, through development of a "subject of discussion co-creation platform (NutShell)," in which high-tech information technology specialists as well as various stake holders participate, we formulate a "Real-Time Technology Assessment (RTTA)" system to swiftly focus social discussions in the relevant field.

In this paper, we would like to present the summary of Technology Assessment Note (TA-note) concerning molecular robotics. This TA-note is to encourage further discussion on ELSIs of molecular robotics by conducting focus groups workshop discussions and reviewing previous studies on ELSIs in fields of GMO, nanotechnology, regenerative medicine, genomics, and so on. Through these trials, we propose realization of a better agenda building process concerning ELSI and how feedback of the knowledge to the researchers at field sites should be channeled..

Boundary Work of Risk: A Case Study on a Molecular Robotics Laboratory in Japan

Ken Kawamura[1] kawamura310@u.seijo.ac.jp
Daisuke Yoshinaga[2] sc.yoshinaga@gmail.com
Go Yoshizawa[3] gy20@jcom.home.ne.jp
Mikihito Tanaka[4] steman@waseda.jp
Ryuma Shineha[5] r_shineha@seijo.ac.jp

[1] Center for Science, Technology and Society Studies, Seijo University, Seijo 6-1-20, Setagaya, Tokyo, Japan
[2] Waseda University, Japan
[3] Osaka University, Japan
[4] Waseda University, Japan
[5] Faculty of Arts and Literature, Seijo University, Seijo 6-1-20, Setagaya, Tokyo, Japan

Keywords: Molecular robotics, ELSI, Dual Use

The problem of dual use of technologies——the situation where it is unclear how to prevent misuse without foregoing beneficial applications——not only poses a real challenge for civil society in the effort to prevent harmful misuses but is also a cause of serious concerns for scientists working to develop such technologies, especially new inventions at the research and development stages. While scientists are seriously worried about potential misuses of their research outcomes, at the same time they are possibly hampered by the rigid regulations that severely limit their research activities. In this paper, I shed light on how scientists deal with dual-use concerns, based on an interview survey conducted on researchers working for a molecular robotics laboratory in Tokyo. Molecular robotics is an emerging discipline involving the design and creation of biological structures that serve various purposes by manipulating the order of nucleotides in DNA. In other words, it is the process of creating nanoscale robots utilizing the self-replicating capacity of DNA. Although the current accomplishments of the field remain experimental, the potential applicability of these structures is immense——as are the consequences of their misuse. Based on the analysis of the interview data, I show how the researchers in this molecular robotics laboratory organize their research projects while facing and coping with the dilemma posed by the potential risk of misuse and the concerns regarding excessive control. Especially I delve into their boundary work of risk——the practice demarcating the intrinsic risk of a given technology. This paper details a case study on how the researchers' own considerations regarding the ethical, legal, and social implications (ELSI) of technology shape their research practices.

[1] Gieryn, Thomas F., 1983, "Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists," *American Sociological Review* 48(6): 781-795, 1983.
[2] U.K. Parliamentary Office of Science and Technology, *Postnote* 340, 2009.
[3] Balmer, Andrew S, Jane Calvert et. al., "Taking Roles in Interdisciplinary Collaborations: Reflections on Working in Post-ELSI Spaces in the UK Synthetic Biology Community," *Science & Technology Studies* 28(3): 3-25, 2015.

Overview of Molecular Robot Ethics

Akihiko Konagaya[1]
kona@m.titech.ac.jp

Satoshi Murata[2]
murata@molbot.mech.tohoku.ac.jp

1 School of Computing, Tokyo Institute of Technology
2 Department of Robotics, Graduate School of Engineering, Tohoku University

Keywords: molecular robotics, molecular robot ethics

DNA synthetic technology opens the doors for not only genome science but also for molecular robotics which aims for bio-molecular artifacts with sense and intelligence [1]. In molecular robotics, DNA nano technologies such as DNA origami and DNA computing play an essential role in construction and integration of molecular channels, molecular actuators and molecular control circuits [2]. These technologies enable us to develop movable molecular robots with motor proteins and microtubules [3], light-induced peptide nanofibre growth [4] and nematic alignment of confined actin [5]. Emergence of such molecular robots raises new research issues in ethics with regards to molecular robot guideline and dual-use issues, to name a few [6]. In order to deal with these issues, we have started molecular robot ethics project [7] in collaboration with real time technology assessment project [8]. Related information are now available at molecular robot ethics site [9].

[1] Murata, S., *et al.*, M. Molecular Robotics: A New Paradigm for Artifacts. New Generation Computing 2013, 31 (1), 27–45.
[2] Hagiya, M., *et al.*, Molecular Robots with Sensors and Intelligence. Accounts of Chemical Research 2014, 47 (6), 1681–1690.
[3] Yusuke Sato, et al., Micrometer-sized molecular robot changes its shape in response to signal molecules, Science Robotics, 2, eaal3735 (2017).
[4] Hiroshi Inaba, et al., Light-induced propulsion of a giant liposome driven by peptide nanofiber growth, Science Report, 8, 6243 (2018).
[5] Shunske Tanaka, Kingo Takiguchi, Masahiro Hayashi, Repetitive stretching of giant liposomes utilizing the nematic alignment of confined actin, Communications Physics, 1, 18 (2018).
[6]http://en.molecular-robotics.org/2017/02/the-1st-international-symposium-on-molecular-robot-ethics/
[7] https://ristex.jst.go.jp/hite/en/community/project000291.html
[8] https://ristex.jst.go.jp/hite/en/community/project000290.html
[9] http://molecular-robot-ethics.org/en/

P6-07

Comparison of health hazard reporting of conventional food poisoning and designated ingredient containing food by Food Sanitation Law of Japan

Mitsuo Saito[1,2] mitsuosaito-kkr@umin.ac.jp
Yuzuru Hayashi[1] lin.rang@outlook.com
Takehiko Yajima[1] qzf11372@nifty.com

[1] Institute for Health Vigilance, 4-25-5, Tanashi, Nishi-Tokyo City, Tokyo 188-0011, Japan
[2] Japan Pharmaceutical Information Center, 2-12-15, Shibuya, Shibuya-ku, Tokyo 150-0002, Japan

Keywords: Designated ingredient containing food, Food poisoning, Health hazard reporting, Food Sanitation Law of Japan

[Introduction] The Food Sanitation Law of Japan was amended on 13th June 2018, and new article 8 defined designated ingredient containing food that special caution should be required. It will be effective within two years. In this presentation, we compared health hazard reporting requirements of conventional food poisoning and designated ingredient containing food.

[Method] Amended article and related documents were obtained from the website of MHLW and National Diet.

[Result] Current article 58 imposes the physician mandatory reporting of conventional food poisoning case to the public health center immediately. Adverse event reporting caused by the health food also follows this scheme. In contrast, new article 8 just requires physician, dentist, and pharmacist etc. (called as medical staffs) to grasp the health hazard caused by designated ingredient containing food, and collaborate the survey of local government if required. The new article 8 solely impose mandatory reporting to the local government on the whole seller of the by designated ingredient containing food, and request local government to inform MHLW.

[Discussion] In spite of the potential risk of designated ingredient containing food, the legal obligation of the physician is obviously mitigated. Under current regulation, destination of adverse event reporting of the Food for Specified Health Uses (Tokuho) and the Food with Functional Claims (Kinousei Hyoji Shokuhin) is Consumer Affairs Agency. However, the new article 8 does not mention the destination of the health hazard report if the designated ingredient containing food is also the Food for Specified Health Uses or the Food with Functional Claims. Hence if no coordination is taken, double administrative management may occur. Before the implementation of new article 8, careful designing of the system that is accurate and without omissions or duplications is expected.

Formulating Ethical Principles of Molecular Robotics (ver. 1.1)

Naoto Kawahara[1]
n-kawaha@med.kyushu-u.ac.jp

Ryuma Shineha[2]
r_shineha@seijyo.ac.jp

Akihiko Konagaya[3]
kona@dis.titech.ac.jp

[1] Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan
[2] Seijyo University, Seijo 6-1-20, Setagaya-ku, Tokyo 157-8511, Japan
[3] Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8503 Japan

Keywords: Ethical principles, Molecular robotics, Risk, Benefit, Safety, Security, Dual-use, Accountability, Transparency

Nowadays, it is an issue of extreme importance to establish an ethical framework with a new view of material, information and life according to a technological development. With ever-increasing progress in creativity and ingenuity of technology, new devices and systems appear continuously. However, there are concerns about the ethical scope of molecular robotics. In Japan, research and development of molecular robotics has been promoted, taking advantage of an important elemental technology concerning senses, motions and intelligence. More complicated configurations of systems in molecular robotics are conceivable, which will be applied to informatics, engineering, chemistry, biology and medicine in near future. In this context, we examined various ethical principles, codes, guidelines, etc. in the related fields. Then, we formulated the ethical principles of Molecular Robotics (ver. 1.1). We also request any person who engages in molecular robotics to comply with these principles.

[1] G7 Academies' Joint Statement, New Economic Growth: The Role of Science, Technology, Innovation and Infrastructure, https://royalsociety.org/~/media/about-us/international/g-science-statements/2017-may-3-new-economic-growth.pdf?la=en-GB, 2017.
[2] European Commission, Synthetic biology ethics: managing risks without limiting benefits, http://ec.europa.eu/environment/integration/research/newsalert/pdf/233na1_en.pdf, 2011.
[3] U.S. Presidential Commission for the Study of Bioethical Issues New Directions: The Ethics of Synthetic Biology and Emerging Technologies: Executive Summary and Recommendations, 2010.
[4] U.K. Engineering and Physical Sciences Research Council (EPSRC), Principles of Robotics, 2010.
[5] Peter Kemp, Jacob Dahl Rendtorff, The Barcelona Declaration: Towards an Integrated Approach to Basic Ethical Principles, *SYNTHESIS PHILOSOPHICA* 46: 239–251, 2008.
[6] U.S. National Society of Professional Engineers, Code of Ethics for Engineers, 2007.
[7] UNESCO World Commission on the Ethics of Scientific Knowledge and Technology (COMEST), Nanotechnologies and Ethics: Policies and Actions, 2007.
[8] WHO/CDS/EPR, Biorisk management Laboratory biosecurity guidance, 2006.
[9] Nicholas H. Steneck, Fostering Integrity in Research: Definitions, Current Knowledge, and Future Directions, *Science and Engineering Ethics* 12(1): 53-74, 2006.

P7-01

Refined method for designing functional siRNAs specific for any genes in mammalian cells

Yoshiaki Kobayashi [1]
yoshiaki-kobayashi@g.ecc.u-tokyo.ac.jp

Zhiyi Li [1,2]
lizhiyi16@mails.tsinghua.edu.cn

Kumiko Ui-Tei [1]
ktei@bs.s.u-tokyo.ac.jp

[1] Department of Biological Sciences, Graduate School of Science, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[2] Department of Biological Sciences, the University of Tsinghua, Haidian, Beijing, China

Keywords: RNA interference (RNAi), Small interfering RNA (siRNA), Off-target effect (OTE), Base-pairing stability, Chemical modification, Nucleic acid therapeutics

RNA interference (RNAi) is a mechanism of post-transcriptional gene silencing. In RNAi pathway, a small interfering RNA (siRNA) is loaded onto Argonaute (AGO) protein, and the guide strand of siRNA base pairs with a target mRNA with perfect complementarity of nucleotide sequence to repress its expression by cleavage by AGO protein.

We previously demonstrated that functional siRNA in mammalian cells satisfied four sequence rules [1]; (1) 5' end of the guide strand is A/U, (2) 5' end of the passenger strand is G/C, (3) 5' one-third region of the guide strand is A/U rich, and (4) no long GC stretch.
On the other hand, it is known that siRNA often exhibits off-target effect (OTE) which represses the expression of unintended mRNAs with partial sequence complementarities with the seed region (2-8 bases from the 5' end) of the guide strand. We revealed that the degree of OTE is correlated positively with the thermodynamic stability in base-pairing between the seed region of the guide strand and unintended mRNAs [2]. Therefore, it was expected that the siRNA with low base-pairing stability in the seed region is capable to avoid OTE. However, it was also revealed that functional siRNAs with reduced OTE are not selectable for about 6000 kinds of human genes due to such strict constraints for nucleotide sequences [3].

The purpose of this research is to design the refined method which is capable of selecting highly functional siRNA without off-target effect in mammalian cells. Therefore, we introduced chemical modifications or mismatches to regulate base-pairing stability, and to overcome the limitation for selecting siRNAs based on nucleotide sequences.

[1] Ui-Tei K, Naito Y, Takahashi F, Haraguchi T, Ohki-Hamazaki H, Juni A, Ueda R, Saigo K, *Nucleic Acids Research*, 32:936-948, 2004.
[2] Ui-Tei K, Naito Y, Nishi K, Juni A, Saigo K, *Nucleic Acids Research*, 36:7100-7109, 2008.
[3] Naito Y, Yoshimura J, Morishita S, Ui-Tei K, *BMC Bioinformatics*, 10:392, 2009.

Analyzing Co-author Networks to Search for Young Promising Researchers in Biological Science Fields

Hiroto Inoue[1] h-inoue@mail.dendai.ac.jp
Kenta Ishido[1] 18fmi02@ms.dendai.ac.jp
Yuri Suzuki[1] 15nc023@ms.dendai.ac.jp
Masanori Fujita[2] fujita.m.ai@m.titech.ac.jp
Takao Terano[2,3] tterano@computer.org

[1] Tokyo Denki University, 5 Senju Asahi-cho, Adachi-ku, Tokyo 120-8551, Japan
[2] Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8503, Japan
[3] National Institute of Advanced Industrial Science and Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan

Keywords: Big Data, Academic Literature Database, Co-author Network

Researchers are often evaluated their performance by various citation indices of their publications. However, popular ones such as h-index are not suitable for young researchers, because they monotonically increase and only reflect the past achievements [1]. Thus, to search for young promising researchers, novel quantitative measures are required for both academism and industries. For the purpose, using network analysis techniques, we examine a large database provided by JDreamIII [2] on biological science fields, which includes over 1.92 million literatures and 4.14 million authors during 2001 and 2015 years [3].

In this paper, we focus on co-author networks of the Japan Society for the Promotion of Science (JSPS) research fellows, who represent the excellent talents in Japan. We calculate the betweenness centralities of all the authors in the database. Then, the temporal changes of the proposed measure and h-index are compared with JSPS research fellows and the other ordinal researchers. As, depicted in Figure 1, the proposed measure is effective for the evaluation.

Our future work includes the development of a practical searching system based on the JSPS model under the support of the Life Intelligence Consortium (WG3-PJ08) [4].

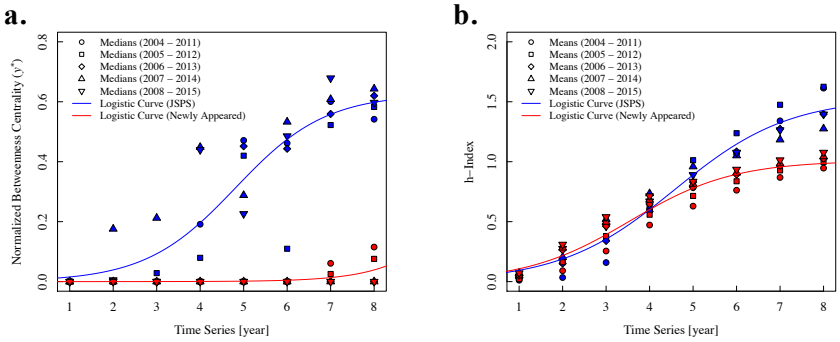


Figure 1. Growth models based on logistic regression analysis. The blue curves show JSPS models and the red curves show average young researchers who have careers as Ph.D. students, postdoctoral fellows, assistant professors, and others.

[1] R.N. Zare, Assessing Academic Researchers, *Angewandte Chemie International Edition*, 51:7338-7339, 2012.
[2] http://jream3.com (in Japanese)
[3] M. Fujita, H. Inoue and T. Terano, Searching Promising Researchers through Network Centrality Measures of Co-Author Networks of Technical Papers, *Proc. 41st Annual Computer Software and Applications Conference*, IEEE, 2017.
[4] https://rc.riken.jp/life-intelligence-consortium/wg/

Profiling of FFPE Tumor Samples by ChIP-Seq

Masato Yonezawa[1] **Sarantis Chlamydas**[1] **Adam Blattler**[1]
yonezawa@activemotif.com
Tiffany Yen[1] **Paul Labhart**[1] **Brian Egan** [1]

[1] Active Motif, Inc. Carlsbad CA, USA.

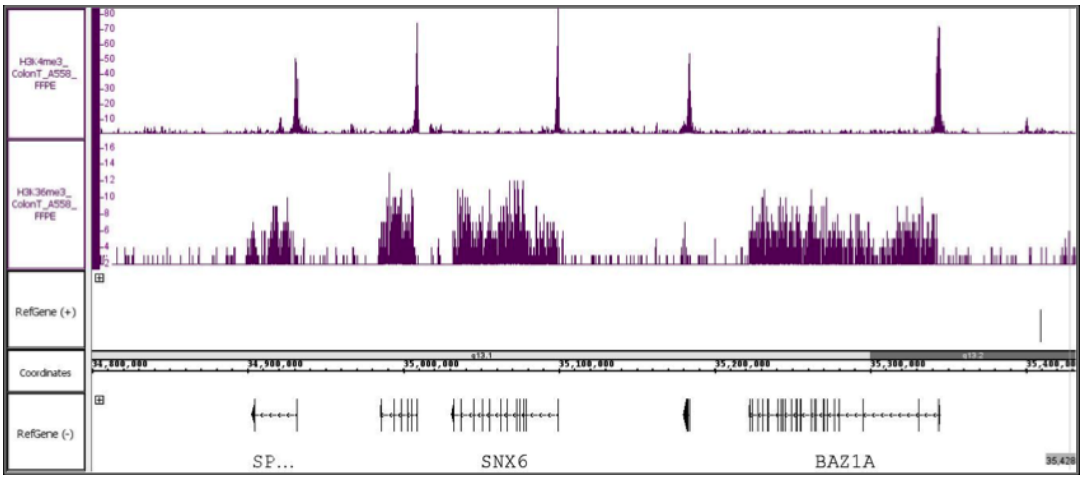
Keywords: Epigenome, ChIP-seq, FFPE, biomarker

Formaldehyde-fixed paraffin embedded (FFPE) specimens are a valuable resource for retrospective research on diseases and biomarker discovery. Clinical information, treatments and outcomes available from patient samples may reveal molecular mechanisms underlying diseases and diagnosis. Unlike fresh tumor samples, FFPE samples are more challenging for ChIP-seq (chromatin immunoprecipitation sequencing) analysis due to inefficiency in chromatin extraction arising from extensive crosslinking. However, recent advances at Active Motif have enabled the routine generation of high quality ChIP-seq data sets from limited amounts of FFPE tumor samples[1, 2]. In the present study, we show 1) chromatin extracted from FFPE samples work for profiling of histone modifications and transcription factors; 2) FFPE ChIP-seq profiles are highly concordant with those generated from matched frozen tumor samples; 3) FFPE ChIP-seq profiles reveal tumor specific patterns. Availability of our reliable ChIP-seq protocol for FFPE samples will bring advances in our understanding to diseases and in discovery of novel biomarkers.

[1] http://www.activemotif.com/catalog/1185/ffpe-chip-seq
[2] http://www.activemotif.com/catalog/937/chip-it-ffpe

H3K4me3

H3K36me3



ChIP-seq profiling using 20-year-old colon tumor FFPE block

P7-04

Novel mechanism of lung cancer stem cells growth by tobacco-specific nitrosamine NNK.

Naoya Hirata[1, 2] n-hirata@nihs.go.jp
Shigeru Yamada[1, 2] shyamada@nihs.go.jp
Yasunari Kanda[1] kanda@nihs.go.jp

[1] Division of Pharmacology, National Institute of Health Science, 3-25-26 Tonomachi, kawasaki-ku, Kanagawa, 210-9501, Japan
[2] Pharmacological Evaluation Institute of Japan, 3-25-22 Tonomachi, kawasaki-ku, Kanagawa, 210-9501, Japan

Keywords: lung cancer stem cell, NNK, ROS, Wnt

Epidemiological studies have suggested that tobacco smoking increases the risk of lung cancer. Tobacco specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) is one of the major carcinogenic components in tobacco smoking. However, the mechanisms by which NNK promotes lung cancer development are not fully understood. Growing evidence suggests that lung cancer is originated from cancer stem cells (CSCs), which have been identified by aldehyde dehydrogenase (ALDH) activity using various cell lines and clinical samples.

In the present study, we investigated the effect of NNK on CSC proliferation in human lung cancer cell line A549. We found that NNK increased ALDH-positive cells in a dose-dependent manner. To examine the downstream signaling, we focused on Wnt, Hedgehog and Notch signaling as self-renewal pathways. Stimulation with NNK induced the expression of Wnt target gene *Dkk1*. Notch target gene *Hes1* was also induced slightly. In contrast, NNK did not affect the expression of Hedgehog target gene *Gli1*. A selective Wnt signaling inhibitor PNU74654 blocked the NNK-induced *Dkk1* expression and ALDH-positive cells. In addition, NNK induced nuclear translocation of β-catenin. The Notch inhibitor DAPT had little effects. These data suggest that NNK increases CSCs via Wnt signaling in lung cancer.

We further investigated the mechanism by which NNK activates Wnt signaling. We found that NNK induced reactive oxygen species (ROS) production in A549 cells. The ROS scavenger N-acetylcysteine (NAC) inhibited the NNK-induced nuclear translocation of β-catenin and the expression of Wnt/β-catenin target gene *Dkk1*. NAC also inhibited the NNK-increased ALDH-positive cells. Since PI3K-Akt pathway has been reported to be involved in the nuclear translocation of β-catenin via ROS, we examined whether PI3K-Akt is involved in the NNK-induced CSC proliferation. NAC inhibited the NNK-induced phosphorylation of Akt. The selective PI3K-Akt inhibitor wortmannin inhibited the NNK-induced *Dkk1* expression and increase in ALDH-positive cells in A549 cells.

Taken together, these data suggest that NNK induces ROS production, activates PI3K/Akt and subsequently induces Wnt/β-catenin signaling, which results in expansion of lung CSCs. These findings might explain the development of lung cancer in cigarette smokers.

P7-05

Immune system simulation based on multi-agent model

[1)]Hiromichi Tsurui, [2)]Yuko Matsuda

tsurui@juntendo.ac.jp matsudayuko@gmail.com

[1)]Department. Pathology, Juntendo University School of Medicine.,

[2)]The University of Tokyo, Interfaculty Initiative in Information Studies, Graduate School of Interdisciplinary Information Studies

Keywords: immune-system, multi-agent system, object-oriented programing, Mathematica

We have been building an immune system simulation tool based on multi-agent model. The system is coded with Mathematica, especially based on its ability to describe in an object oriented manner. The system deals with T cells, B cells macrophages, dendritic cells, antibodies and bacteria as agents. Immune response in a lymph node, which is divided into 8 components like Harel's model, is chiefly treated. The system includes specific interaction between antigen – antibodies, and that between T cell receptor – epitopes presented on B cells based on a homology scoring matrix, blosum 80. The system could run more than 10000 agents simultaneously and could simulate bacterial infection and suppression.

P7-06

Simulation study of 3D-reconstruction of large biomolecule from the diffraction images obtained by X-ray free electron laser experiment

Miki Nakano[1] miki.nakano@riken.jp
Osamu Miyashita[1] osamu.miyashita@riken.jp
Slavica Jonic[2] slavica.jonic@upmc.fr
Florence Tama[1,3,4] florence.tama@riken.jp

[1] RIKEN Center for Computational Science, 6-7-1, Minatojima-minamimachi, Chuo-ku, Kobe, 650-0047, Japan
[2] IMPMC, Sorbonne Universités – CNRS UMR 7590, UPMC Univ Paris 6, MNHN, IRD UMR 206, 75005 Paris, France
[3] Department of Physics, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8602, Japan
[4] Institute of Transformative Bio-Molecules, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan

Keywords: X-ray free electron laser, Single particle analysis, Protein Structure

Single particle analysis using X-ray free-electron laser (XFEL) is a novel method to observe biomolecules under conditions close to nature. This bright and coherent beam allows us to obtain diffraction data without crystallization, and their short femtosecond pulses enables measurements without radiation damage.[1,2] However, resolving 3D structure from XFEL single particle diffraction data is still challenging because of the low signal to noise ratio and the lack of beam incidence angle information. Therefore, it is important to quantitatively assess how the experimental conditions such as the amount of data and their quality affect the expected resolution of the resulting 3D models. In this presentation, we show the results of our computational investigation using simulated diffraction patterns. We show that the resolution of the reconstructed 3D structure depends on the amount and quality of the diffraction images, and further discuss the requirement for experimental condition to restore molecular structures of a few nanometer resolution.[3]

[1] Hirata, K., Shinzawa-Itoh, K., Yano, N., Takemura, S., Kato, K., Hatanaka, M., Muramoto, K., Kawahara, T., Tsukihara, T., Yamashita, E., et al. Determination of Damage-Free Crystal Structure of an X-Ray-Sensitive Protein Using an XFEL. *Nat. methods*, 11:734–736, 2014
[2] Chapman, H. N., Barty, A., Bogan, M. J., Boutet, S., Frank, M., Hau-Riege, S. P., Marchesini, S., Woods, B. W., Bajt, S., Benner, W. H., et al. Femtosecond Diffractive Imaging with a Soft-X-Ray Free-Electron Laser. *Nat. Phys.*, 2:839–843, 2006
[3] Nakano, M., Miyashita, O., Jonic, S., Tokuhisa, A., and Tama, F. Single-particle XFEL 3D reconstruction of ribosome-size particles based on Fourier slice matching: requirements to reach subnanometer resolution. *J. Synchrotron Rad.* 25:1010–1021, 2018

P7-07

Image-based Morphological Analysis for Visualization and Optimization of Stem Cell Culture

Masaya Fujitani[1]
fujitani.masaya@h.mbox.nagoya-u.ac.jp

Kei Kanie[1]
kanie-k@ps.nagoya-u.ac.jp

Ryuji Kato[1]
kato-r@ps.nagoya-u.ac.jp

[1] Graduate School of Pharmaceutical Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan

Keywords: Cell morphology, Image analysis, Transcriptome analysis, Stem cell culture

By the advances of stem cell culture technology, various culture protocols have been reported for more cost effective and stable cell culture [1]. There are numerous protocols by the combination of culture media, plates, coating materials, and handling skills. However, it is difficult to find the most appropriate protocol since there are too many parameters to be evaluated. Therefore, commonly, once examined and succeeded protocol is used for long time and it is considered to be risky to examine new protocols. Such conservative protocol preservation is sometimes costly, and losing the chance to establish better protocol with new materials. Moreover, when a facility require new and optimized protocol for establishing cost effective original manufacturing process, or search for suitable protocol for new cells that have not yet been investigated in the past, we still need a laborious and costly examination process. In many cases, only the proliferation rate or endpoint assay is examined in such protocol exploration process.

Cellular morphology has long been known as an important indication to evaluate the cellular culture status. However, the rule of cellular morphology and its fitness to the culture condition has not been quantitatively defined. Commonly, the cellular morphology is evaluated ambiguously based on feeling and experiences.

We have been proposing morphology-based evaluation concept by quantifying cellular morphology as cellular fingerprint, and have verified the effectiveness of its application to both the cellular quality evaluation and its culture condition [2]. By converting the statistics of cellular morphological parameters as morphological fingerprint, we have found that delicate cellular response to the effect of culture condition can be recorded as quantitative data, and can be compared with bioinformatics analysis. In this poster, we show that such morphology-based analysis is effective to profile and compare the optimum culture conditions rapidly, feasibly, and cost effectively by using only the non-stained time-course microscopic images. In addition, we conducted transcriptome analysis to reveal the biological difference among such culture conditions. We here show the practical examples of phenotypic evaluation method of various types of culture conditions for stem cell.

[1] Chen, K., Mallon, B., McKay, R., Robey, P., Human Pluripotent Stem Cell Culture: Considerations for Maintenance, Expansion, and Therapeutics, *Cell Stem Cell*, 14(1):13-26, 2014

[2] Hiroto, S., Ichiro, T., Mai, O., Rumi, S., Kei, K., Yasujiro, K., Hiroyuki, H., Ryuji, K., Label-Free Morphology-Based Prediction of Multiple Differentiation Potentials of Human Mesenchymal Stem Cells for Early Evaluation of Intact Cells, *PLOS ONE*,9 (4): e93952, 2014

P7-08

Development of Mixed Reality-based Protein and Ligand Tertiary Structure Visualization System

Atsushi Koyama[1] Atsushi.k.ab@m.titech.ac.jp
Nobuaki Yasuo[1] yasuo@cbi.c.titech.ac.jp
Masakazu Sekijima[1,2] Sekijima.m.aa@m.titech.ac.jp

1 Department of Computer Science, Tokyo Institute of Technology, J3-23, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan
2 Advanced Computational Drug Discovery Unit, Tokyo Institute of Technology, J3-23, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan

Keywords: Mixed Reality, Visualization, Protein tertiary structure, Structure-based Drug Design

It is important to understand the tertiary structures of protein-ligand complexes for efficient drug discovery, since the cost of drug discovery is getting higher[1]. Many molecular visualization systems have been developed such as Pymol[2] and Jmol[3]. However, these systems do not fully represent the tertiary structures since they must project the 3D structures to 2D planes.

Recently, virtual reality technology has been rapidly progressed to obtain immersive experience, where enhanced 3D environments can be experienced as physically present in the virtual world. Mixed reality is also an emerging technique. In mixed reality experience, virtual world and real world are merged and co-exist so that we can interact both real and virtual object at the same time.

In this study, we developed HoloMol, which visualize 3D structures of proteins and ligands as it is. Microsoft HoloLens, a head-mounted mixed reality device is used. One advantage of HoloLens is that it is a see-through device that makes the communication easier. Also, it works without the connection to other computers. The 3D models can be manipulated and changed by hand gesture (Air-tap). Our application can show arbitrary complex by replacing models on containers at cloud. HoloMol can be downloaded at [URL] and will be released on Windows Store.

[1] Asher Mullard.: New drugs cost US$2.6 billion to develop. *Nature reviews Drug Discovery*, Vol. 13, No.12, p.877, 2015
[2] https://pymol.org/2/
[3] http://jmol.sourceforge.net

P7-09

Interactive Molecular Scale Soft Matter Simulation in VR

Gregory Gutmann[1] gutmann@c.titech.ac.jp

Akihiko Konagaya[1] kona@c.titech.ac.jp

[1] Tokyo Institute of Technology, J3-25, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan

Keywords: Soft Matter, VR, Simulation, Interactive

With recent efforts in gaining more understanding and control over emergent molecular scale dynamics and the challenges involved in doing so; there has been an interest in using simulation to aid in research. With this aim, we have been developing an Interactive VR simulation system to experiment with virtual molecules as an extension of our real-time particle simulation system for microtubule gliding assay [1]. The virtual molecules can be represented by dynamic ball and stick models which can be manipulated using a hands-on approach in VR. In other words, virtual hands created using a Leap Motion controller allow the user to touch and change the forms of virtual molecules. This has allowed us to simulate various deformable objects in our system. So far we have successfully simulated tube structures and simple objects; however, there are still challenges as well, both of which will be covered in this poster.

[1] Gutmann, G., Inoue, D., Kakugo, A., & Konagaya, A. (2017) *Parallel Interaction Detection Algorithms for a Particle-based Live Controlled Realtime Microtubule Gliding Simulation System Accelerated by GPGPU.* New Generation Computing, Vol. 35 No. 2 Springer Publishing.

P7-10

All-atom molecular dynamics simulation approach to film supported DNA origami in water

Ryuzo Azuma
azuma.r.ac@m.titech.ac.jp

Sae Kishi
kishi.s.aa@m.titech.ac.jp

Greg Gutmann
gutmann.g.aa@m.titech.ac.jp

Akihiko Konagaya
kona@c.titech.ac.jp

Tokyo Institute of Technology, J3-25, 4259, Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8503, Japan

Keywords: molecular robotics, DNA origami, NAMD2, Interface MD, TSUBAME3.0

DNA origami technology has been expanding its application fields such as molecular robotics [1-2]. The structure of a DNA origami object is essentially dependent on the sequence design of staples. Although this is one of the key benefits of the DNA origami method allowing nano-structure designs, atomic-level DNA origami structure modeling is also important to assess structural properties of crossovers [3]. Moreover, assessing the physicochemical properties of their structure has become increasingly important recently since there have been known facts about shape deformation dynamics under low ionic strength [4-5]. In this study, we focus on these dynamics of DNA origami structures using an all-atom MD simulation approach. The hallmark of our approach is to incorporate an inorganic supporting film, Muscovite mica film in this study, as a substrate. We developed a workflow for the combined system by utilizing caDNAno [6], Interface MD [7], and NAMD2 [8] with the latest refined force field AMBER PARMBSC1 [9].

We then observed continual stretching motion for a DNA origami in MD runs in the absence of cations. This result is reminiscent of the experimental observation for a DNA origami in a solution of EDTA, a chelating agent for cations [4]. We also show that under realistic water permittivity (dielectric constant >> 1), DNA strands may withstand denaturation and exhibits a regular pattern similar to argyle lattice. By taking histograms of Zp' (a measure of duplex forms) we analyze these characteristics more quantitatively.

Financial supports: Strategic Advancement of Multi-Purpose Ultra-Human Robot and Artificial Intelligence Technologies, NEDO; Platform Project for Supporting Drug Discovery and Life Science Research (BINDS), AMED.

[1] Murata, S., *et al.*, M. Molecular Robotics: A New Paradigm for Artifacts. New Generation Computing 2013, 31 (1), 27–45.
[2] Hagiya, M., *et al.*, Molecular Robots with Sensors and Intelligence. Accounts of Chemical Research 2014, 47 (6), 1681–1690.
[3] Bai, X. -c., *et al.*, Cryo-EM Structure of a 3D DNA-Origami Object. Proceedings of the National Academy of Sciences 2012, 109 (49), 20012–20017.
[4] Kim, H., *et al.*, Stability of DNA Origami Nanostructure under Diverse Chemical Environments. Chemistry of Materials 2014, 26 (18), 5265–5273.
[5] Kielar, C., *et al.*, On the Stability of DNA Origami Nanostructures in Low-Magnesium Buffers. Angewandte Chemie International Edition 2018.
[6] Douglas, S. M., *et al.*, Rapid Prototyping of 3D DNA-Origami Shapes with caDNAno. Nucleic Acids Research 2009, 37 (15), 5001–5006.
[7] Heinz, H., *et al.*, Thermodynamically Consistent Force Fields for the Assembly of Inorganic, Organic, and Biological Nanostructures: The INTERFACE Force Field. Langmuir 2013, 29 (6), 1754–1765.
[8] Phillips, J. C., *et al.*, Scalable Molecular Dynamics with NAMD. Journal of Computational Chemistry 2005, 26 (16), 1781–1802.
[9] Ivani, I., *et al.*, A. et al. Parmbsc1: A Refined Force Field for DNA Simulations. Nature Methods 2015, 13 (1), 55–58. *Nature methods*. 2016, 13 (1), pp 55-58.

Self-Assembly of a Flexible Multi-Joint Ring Motif

Shiyun Liu[1]
hikari@molbot.mech.tohoku.ac.jp
Ibuki Kawamata[1]
kawamata@molbot.mech.tohoku.ac.jp
Yuki Suzuki[2]
ysuzuki79r@molbot.mech.tohoku.ac.jp
Nomura M. Shin-ichiro[1]
nomura@molbot.mech.tohoku.ac.jp
Satoshi Murata[1]
murata@molbot.mech.tohoku.ac.jp

[1] Department of Robotics, Graduate School of Engineering, Tohoku University, 6-6-01, Aoba-yama, Sendai 980-8579, Japan
[2] Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, 6-3, Aoba-yama, Sendai 980-8578, Japan

Keywords: DNA origami, DNA nanotechnology, Self-assembly, Molecular robotics.

During the past few decades, the field of structural DNA nanotechnology has grown enormously [1]. It enables us to design a variety of self-assembled structures ranging from 2D to 3D crystalline structures to algorithmically self-assembled aperiodic structures [2-3]. The invention of DNA origami, which has greatly expanded the geometric complexity of DNA nanostructures and has accelerated research development in this area [4]. Making a large (mm-scale) planer lattice with no defect is one of remained problems of structural DNA nanotechnology. 2D crystal of DNA tiles or DNA origami motifs can grow up to several hundred micrometers scale but no more because it is difficult to avoid the interstices among the clusters.

Here, we propose a flexible DNA origami multi-joint ring motif, which has a potential to self-assemble a large lattice on a 2D substrate surface. We designed the motif as a DNA origami heptagon by a software called caDNAno [5]. The ring motif has an ability to connect with others through the hybridization of DNA strands, which are aligned symmetrically on all the edges. The edges of the motif are linked by flexible hinges, which allow the motif to deform into different shapes. If the motif stays rigid as a heptagon, it is impossible to cover 2D surface without gaps. Since the motif is flexible, it is expected to fill the interstices to cover the entire surface of the substrate. The geometrical design of the ring motif and some preliminary experimental results will be presented on the poster.

The proposed multi-joint ring motif can provide a novel possibility to create nanostructures of high complexity and high flexibility, which may have a capacity of self-recovery.

[1] Seeman, N. C, An overview of structural DNA nanotechnology. *Mol. Biotechnol*, 37: 246–257, 2007.
[2] Rothemund P. W. K, Papadakis N, Winfree E., Algorithmic Self-Assembly of DNA Sierpinski Triangles, *PLoS Biol* 2(12): e424,2004.
[3] Lin, C., Liu, Y., Rinker, S. and Yan, H., DNA Tile Based Self-Assembly: Building Complex Nanoarchitectures, *ChemPhysChem*, 7: 1641-1647, 2006.
[4] Rothemund, P. W. K, Folding DNA to create nanoscale shapes and patterns. *Nature*, 440: 297–302, 2006.
[5] Shawn M. Douglas, Adam H. Marblestone, Surat Teerapittayanon, Alejandro Vazquez, George M. Church, William M. Shih; Rapid prototyping of 3D DNA-origami shapes with caDNAno, *Nucleic Acids Research*, Volume 37, Issue 15: 5001–5006, 2009.

P7-12

Integrated gene logic-chip functioning in an artificial cell

Takeya Masubuchi [1,9], Masayuki Endo [2,3], Ryo Iizuka [4], Ayaka Iguchi [5], Dong Hyun Yoon [6], Tetsushi Sekiguchi [6], Hao Qi [1,7], Ryosuke Iinuma [1], Yuya Miyazono [1], Shuichi Shoji [5], Takashi Funatsu [4], Hiroshi Sugiyama [2,3], Yoshie Harada [3,8], Takuya Ueda [1] and **Hisashi Tadakuma** [1,3,8,]
tadakuma@protein.osaka-u.ac.jp

[1] Graduate School of Frontier Science, The University of Tokyo, [2] Department of Chemistry, Kyoto University, [3] Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University, [4] Graduate School of Pharmaceutical Sciences, The University of Tokyo, [5] Department of Nanoscience and Nanoengineering (ASE Graduate School), Waseda University, [6] Research Organization for Nano & Life Innovation, Waseda University, [7] Department of Chemical Engineering and Technology, Tianjin University, , [8] Institute for Protein Research, Osaka University, [9] Present address: Institute for Quantitative Biosciences, The University of Tokyo

Keywords: Gene expression, Synthetic Biology, Molecular Robots

In synthetic biology, the control of gene expression requires a multi-step computation of biological signals. However, the complexity of genetic circuits remains low because it is difficult to completely avoid crosstalk between genetic circuits. Here, we made an orthogonal self-contained device by integrating an actuator and sensors onto a DNA origami-based nano-chip that contains an enzyme, T7 RNA polymerase, and multiple target gene substrates. This gene nano-chip orthogonally transcribes its own genes, and the nano-layout ability of DNA origami allows us to rationally design the gene expression levels by controlling the intermolecular distances between the enzyme and the substrate genes. We further integrated logic gates so that the nano-chip responds to water-in-oil droplets and computes their miRNA profiles, which demonstrates that the nano-chip can function as a gene logic-chip. Our approach to component integration on a nano-chip may provide a basis for large-scale, integrated genetic circuits.

[1] Masubuchi T et al., Construction of integrated gene logic-chip. *Nature Nanotechnology* in press (doi: 10.1038/s41565-018-0202-3).

Design of a new reaction diffusion system for pattern formation in hydrogel medium

Keita Abe[1] abe@molbot.mech.tohoku.ac.jp
Ibuki Kawamata[1] kawamata@molbot.mech.tohoku.ac.jp
Shin-ichiro M. Nomura[1] nomura@molbot.mech.tohoku.ac.jp
Satoshi Murata[1] murata@molbot.mech.tohoku.ac.jp

[1] Department of Robotics, Graduate School of Engineering, Tohoku University, Aramaki-aza-Aboba 6-6-01, Abobaku, Sendai, Miyagi 980-8579, Japan.

Keywords: Pattern formation, Reaction diffusion system, DNA logic gate

Reaction diffusion mechanisms are widely found in early stages of biological development. It is thought to be one of the main mechanisms to form highly ordered patterns of various biological structures [1] [2]. Imitating these processes by programming an artificial reaction diffusion system has a large potential to engineer self-organization process for various applications. To achieve this, we need to manipulate the processes of “reaction” and “diffusion” independently. DNA is known for its capability to imitate arbitrary chemical reactions, thus is suitable as the material for artificial reaction diffusion systems [3]. Several DNA-based frameworks for artificial pattern formation have been reported so far [2] [4], however, little attention has been paid to control the “diffusion” term.

Here, we propose a technique to design a DNA-based reaction diffusion system with an ability to control the diffusion speed of DNA species. As a demonstration, a DNA logic gate system to form a bisector pattern is designed. Two types of DNA called *inputs A* and *B* diffuse from different sources placed in the gel medium. A bisector pattern is formed where the distributions of the species *A* and *B* overlap. The overlap is visualized by an AND gate pre-mixed in the gel, which outputs fluorescence if both of the *inputs A* “AND” *B* exist. The relative positioning of the bisector in between the two sources can be moved by tuning the diffusion speed of DNA [5]. In this diffusion modulation method, we can adjust the diffusion speed of the input DNA by changing concentration of additional DNA molecule called competitor.

Experimental results show that the bisector pattern was successfully formed and the position of the pattern was successfully controlled by the proposed reaction system with the diffusion modulation. This approach has a potential to scale up the molecular programming from nano- to micro- and larger scales, which provides us a useful tool to engineer self-organizing material for various application such as artificial organs and brains.

[1] Wolpert, Lewis, Cheryll Tickle, and Alfonso Martinez Arias. *Principles of development*. Oxford University Press, USA, 2015.
[2] Anton S. Zadorin, Yannick Rondelez, Guillaume Gines, Vadim Dilhas, Georg Urtel, Adrian Zambrano, Jean-Christophe Galas & André Estevez-Torres, “Synthesis and materialization of a reaction–diffusion French flag pattern.” *Nature chemistry*, 9(10), 990, 2017.
[3] Soloveichik, David, Georg Seelig, and Erik Winfree. “DNA as a universal substrate for chemical kinetics.” *Proceedings of the National Academy of Sciences* 107.12 (2010): 5393-5398.
[4] John Zenk, a Dominic Scalise,Kaiyuan Wang, Phillip Dorsey, Joshua Fern, Ariana Cruz, and Rebecca Schulman. Stable DNA-based reaction–diffusion patterns. *RSC Advances*, 7(29), 18032-18040, 2017
[5] Thanapop Rodjanapanyakul, Fumi Takabatake, Keita Abe, Ibuki Kawamata, Shinichiro M. Nomura, and Satoshi Murata. “Diffusion modulation of DNA by toehold exchange”. *Physical Review E*, 97(5), 052617, 2018.

Chain aggregation of liposomes through DNA Hairpin assembly

Yuto Otaki[1]
otaki@molbot.mech.tohoku.ac.jp

Ibuki Kawamata[1]
kawamata@molbot.mech.tohoku.ac.jp

Yuki Suzuki[2]
ysuzuki79r@molbot.mech.tohoku.ac.jp

Satoshi Murata[1]
murata@molbot.mech.tohoku.ac.jp

Nomura M. Shinichiro[1]
nomura@molbot.mech.tohoku.ac.jp

1 Department of robotics, Graduate School of Engineering, Tohoku University, 6-6-01 Aramaki - aza-Aoba, Aoba-ku, Sendai, Miyagi, 980-8579, Japan.
2 Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, 6-3 Aramaki - aza Aoba, Aoba-ku, Sendai, Miyagi, 980-8578, Japan

Keywords: DNA nanotechnology, Liposome, Drug delivery system (DDS)

Liposome, a capsule like structure formed by lipid bilayer, has been received attention as carriers for target-specific drug delivery [1]. Capability for recognition and binding to a target cell is based on specific binding between sensing molecules on liposomes and target molecules on the cell. However, there are some obstacles for recognition and binding, such as small number of target molecules and small size of target cell and steric hindrance caused by liposomes covering surface of target cell. Therefore, a new strategy, that signalize and amplify molecular recognition and binding events to overcome those obstacles, is needed.

DNA is known as a programmable molecule that offers methods for molecular detection [2]. Recently, molecular detection system with signal amplification, called Exponential hairpin assembly (EHA), was developed [3]. The system consists of four hairpin DNAs, forming exponentially growing dendritic structure in the present of target DNA.

Here, we propose a novel method, based on molecular robotics, to induce chain aggregation of liposomes around target via DNA. The method employs liposomes covered with two kinds of hairpin DNAs, two kind of other free hairpin DNAs and target DNA. Detection of the target is achieved by hybridization between hairpin on the liposomes and a target DNA, causing chain-branching reaction induced by EHA. The dendritic DNA structures of EHA are formed on the surface of liposomes and spread to other liposomes through the hairpin chain reaction, resulting chain aggregation around target. In experiments, we confirmed that target DNA triggered liposome aggregation. This method is expected to provide new functionality for various targeting strategies for DDS.

[1] Torchilin, Vladimir P. "Recent advances with liposomes as pharmaceutical carriers." *Nature reviews Drug discovery* 4.2 (2005): 145.
[2] Dirks, Robert M., and Niles A. Pierce. "Triggered amplification by hybridization chain reaction." *Proceedings of the National Academy of Sciences* 101.43 (2004): 15275-15278.
[3] Ma, Cuiping, et al. "Three-dimensional DNA nanostructures for colorimetric assay of nucleic acids." *Journal of Materials Chemistry B* 3.14 (2015): 2853-2857.

P7-15

Improved the conjugation yield between AuNP and thiol group in DNA origami by solution freezing

Shumpei Ishikawa k529897@kansai-u.ac.jp **Islam Md. Sirajul** **Naohide Akamatsu**

Yuichi Ohya ohya@kansai-u.ac.jp **Akinori Kuzuya** kuzuya@kansai-u.ac.jp

Department of Chemistry and Materials Engineering, Kansai University,
3-3-35, Yamate, Suita, Osaka564-8680, Japan

Keywords: DNA nanostructures, DNA origami, AuNP

We have previously developed a DNA origami motif with 9 wells in the structure (DNA Nanostick). In previous study, we succeeded in conjugating 4 AuNPs in the wells as designed (well number is 2, 4, 6, 8) in DNA Nanostick[1]. We used excess AuNPs such as 100 eq and long incubation such as 4 days for the reaction at 4°C. However, the yield of DNA Nanostick with 4 AuNPs was about 20%. This corresponds to the conjugation yield of 68.8% per well. Very recently, B. Liu and J. Liu have reported that the reaction yield was increased for conjugation of AuNP with thiol modified ssDNA by freezing the solution[2]. Upon decreasing the temperature, ice crystals are gradually formed, pushing AuNPs, DNA, and salt (Na^+, citrate ions) out to the "micropockets" to be concentrated, which significantly enhances the reaction rate. In this study, we applied this freezing method to DNA origami conjugation.

For conjugation of DNA Nanostick with AuNPs, we used 2 eq. of AuNPs and incubated them for 2 hours at -20℃ in a freezer or 2 min at -80℃ in dry ice. After that, excess AuNPs were removed by gel filtration, and the structure was observed using AFM in fluid. As a result, DNA Nanostick was found to be intact even the solution was frozen. Moreover, the yield of DNA Nanostick with 4 AuNPs (about 50%) was increased compare to that in the previous study. In the case of freezing at -20℃ for 2 hours (freezer method), the conjugation yield was 85.8% per well. In the case of freezing at -80℃ for 2 min (dry ice method), the conjugation yield was also increased (about 10%), however, it was lower than freezer method. In dry ice method, the velocity of ice crystal formed was too fast to concentrate solute not efficiently.

In summary, we succeeded in increasing the conjugation yield for reaction of DNA Nanostick with AuNP by freezing the solution. We also succeeded in reducing the cost that necessary amount of AuNP was reduced from 100 eq excess to 2 eq. The reaction time was also reduced from 4 days to 2 hours. Moreover, the conformation of DNA origami was not broken by freezing the solution.

[1] A. Kuzuya, M. Komiyama et al., *Small*, **2010**, *6*, 2664–2667.
[2] B. Liu, J. Liu, *J. Am. Chem. Soc.*, **2017**, *139*, 9471–9474.

Toward automatic-control of motility-protein crowds by DNA circuit

Shosei Ichiseki[1]
ichiseki@molbot.mech.tohoku.ac.jp

Ibuki Kawamata[1]
kawamata@molbot.mech.tohoku.ac.jp

Keya Jakia Jannat[2]
zakiazannat@yahoo.com

Satoshi murata[1]
murata@molbot.mech.tohoku.ac.jp

Akira Kakugo[2]
kakugo@sci.hokudai.ac.jp

Shinichiro M. Nomura[1]
nomura@molbot.mech.tohoku.ac.jp

[1] Department of Robotics, Graduate School of Engineering, Tohoku University, 6-6-01, Aramaki aza-aoba, Aobaku, Sendai, Miyagi 980-8579, Japan.
[2] Department of Chemical, Graduate School of Science, Hokkaido University, Kita10, Nishi8, Sapporo, Hokkaido, 060-0810, Japan.

Keywords: Motility proteins, molecular circuit, molecular programming, motion control

In the living organisms, various molecules work cooperatively, realizing complex reaction and various functions. For example, animal muscles realize macroscopic motion by cooperative work of myosins and actins. Recently, by learning from such living molecular systems, molecular robots have been developed by integrating molecular devices which have designed for realizing desired function [1][2].

Last year, J.J.Keya et al. demonstrated a study which switching assembly and disassembly of motility microtubules on kinesins by DNA hybridization[3]. In their work, photo-reactive DNA molecules switched the assembly/disassembly by alternately irradiating visible light and UV externally. Here, we aim to control such a collective behavior of kinesins-microtubules automatically by designed chemical reaction. In order to enable automatic output of desired DNA sequence, we adopted a DNA computer, universal strand generator (USG) [4]. This is the device which generates any sequence of ssDNA designed in advance at arbitrary time. We adjusted the experimental condition, such as temperature, buffer, ionic strength, etc. for coexistence of kinesin-microtubules griding assay with USG.The microtubules were modified with fluorescence dye and ssDNAs.

By using fluorescence microscopy, it was confiermed that ~70% of the microtubules were assembled by adding the ssDNA (for assemble). Then, we observed that the motility microtubules on kinesins assembled through USG output. Now we are seeking a suitable condition to conduct the automatic-disassemble process through USG.

This work would be useful not only for automatic control of motility-protein crowds but also the molecular robots control by the chemical signal, similar to the natural chemotaxis.

[1] S.Murata, et al. "Molecular robotics: A new paradigm for artifacts." *New Generation Computing* 31(1), (2013): 27-45.
[2] Y.Sato, et al. "Micrometer-sized molecular robot changes its shape in response to signal molecules." *Science Robotics* 2(4), (2017): eaal3735.
[3] J.J.Keya, et al. "DNA-assisted swarm control in a biomolecular motor system." *Nature communications* 9(1), (2018): 453.
[4]Team sendai. "Universal strand generator." *Biomod 2014*, http://teamsendai2014.github.io/

P7-17

Single molecule real-time observation of dynamically functioning DNA origami molecular machines

Yuta Yamasaki k122988@kansai-u.ac.jp **Naohide Akamatsu** **Ryosuke Watanabe**

Yuichi Ohya ohya@kansai-u.ac.jp **Akinori Kuzuya** kuzuya@kansai-u.ac.jp

Department of Chemistry and Materials Engineering, Kansai University, 3-3-35 Yamate, Suita, Osaka 564-8680, Japan

Keywords: Molecular Robotics, DNA nanostructures

We have reported nanomechanical DNA origami devices, DNA pliers[1] and DNA chopsticks that can transform in response to binding of target molecules. The former device is two-dimensional and the latter is three-dimensional. Although the latter showed more rigid nature and equivalent pinching rate to the target substrates, no remarkable advantage of DNA chopsticks was been observed. If the interaction between both structures and mica is different and DNA chopsticks is more mobile on mica than DNA pliers, it is advantageous in real time observation. In this study, we fix both structures on mica and add Na^+ as monovalent cations, to weaken the interaction between each structure and mica. Then, we investigated the difference in behavior of both structures.

Both structures were annealed and deposited to mica. Then, 1×TAE/Mg^{2+} buffer containing NaCl was added, and the system was observed with AFM. By counting numbers, the strength of the interaction between both structures and mica was compared. AFM observation revealed that, as the concentration of Na^+ increased, the number of DNA chopsticks became less than DNA pliers. It was found that the mobility of DNA chopsticks is higher than DNA pliers at a lower Na^+ concentration. We also attempted real-time observation by using high-speed AFM. After DNA chopsticks was to mica, NaCl was added during observation to give the structures mobility. Then, the target, streptavidin (SA) was added. As a result, selective closing was clearly observed upon capturing of SA. This suggested that DNA chopsticks could be applied to devices that can detect of target substrates in real time.

[1] A. Kuzuya, et al., *Nat. Commun*. **2011**, *2*, 499.

Making pores on biomembrane using DNA origami: Design and evaluation of φ12 nm artificial channel

Shoji Iwabuchi[1]
iwabuchi@molbot.mech.tohoku.ac.jp

Ibuki Kawamata[1]
kawamata@molbot.mech.tohoku.ac.jp

Yuki Suzuki[1]
ysuzuki79r@molbot.mech.tohoku.ac.jp

Satoshi Murata[1]
murata@molbot.mech.tohoku.ac.jp

Shin-ichiro M. Nomura[1]
nomura@molbot.mech.tohoku.ac.jp

[1] Molecular Robotics Laboratory, School of Engineering, Tohoku University, 6-6-01 Aoba Aramaki-Aza, Aoba, Sendai City, Miyagi 980-8579, Japan

Keywords: DNA nanotechnology, channel, nanopores, molecular transport, DNA origami

Development of a molecular robot, as a small and autonomous system that can work in solution, has attracted attention[1]. Closed-lipid bilayer membrane called liposome is mainly adopted as the robot body. The flexible membrane enables to deform in response to their internal molecular actuators with DNA logic gate[2].

However, since the lipid bilayer membrane prevents to pass hydrophilic molecules, there is an essential problem that the robot cannot respond to external signals or stimulations. The artificial channel on the lipid bilayer[3][4] is a solution to this problem. Though these channels are reported, it has been difficult to transport large molecule (such as dsDNA) or to control the selective transportation because the pore size of channels is small.

In this study, we designed an artificial channel with a large-size pore. It is made by DNA origami[5]. It is designed to bind to the bilayer membrane by cholesterol-modified DNA and to pass through the molecule by simple diffusion. To evaluate the function, giant unilamellar vesicles were used as the biomembrane model, and hydrophilic fluorescence dye was entrapped the outside (or inside) of the vesicles as the molecule to be transported. Then we mixed the artificial channel to the system and observed the transportation of the dye molecules directly by confocal microscopy.

As a result, the channel was attached to the GUV membrane and it is confirmed that the fluorescence dyes are transported through the channel. Now we are trying to increase the probability of transportation and to control the limitation of the size of the transportable molecule. The large pore can pass more kind of molecules, and the function of permeability control will lead to the realization of a molecular robot that can control its action by importing specific signal molecule.

[1] Hagiya, M., *et al*. "Molecular robots with sensors and intelligence." *Accounts of chemical research* , 47.6 : 1681-1690, 2014.
[2] Sato, Y., *et al.*, Micrometer-sized molecular robot changes its shape in response to signal molecules, *Science Robotics*, 2:eaal3735, 2017.
[3] Langecker, M., *et al.*, Synthetic lipid membrane channels formed by designed DNA nanostructures, *Science*, 338: 932-936, 2012.
[4] Göpfrich, K., *et al.*, Large-conductance transmembrane porin made from DNA origami, *ACS nano*, 10:8207-8214, 2016.
[5] Douglas, S. M., *et al.*, Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic acids research*, 37:5001-5006, 2009.