Kyoto University 21st Century COE Program

Knowledge Information Infrastructure for Genome Science

Kyoto University
21st Century COE Program
Genome Science
http://www.bic.kyoto-u.ac.jp/COE/

COE Administration Offices
Bioinformatics Center, Institute for Chemical Research, Kyoto University
Gokasho, Uji, Kyoto 611–0011, Japan
Tel. +81–774–38–3270 Fax. +81–774–38–3269
E–mail: coe@bic.kyoto–u.ac.jp

Graduate School of Pharmaceutical Sciences, Kyoto University
46–29 Yoshida–Shimo–Adachi–cho, Sakyo–ku, Kyoto 606–8501, Japan
Tel. +81–75–753–9274 Fax. 81–75–753–9275
E–mail: coe21@pharm.kyoto–u.ac.jp

Kyoto University
21st Century COE Program
Genome Science

Kyoto University 21st Century COE Program
Knowledge Information Infrastructure for Genome Science

Project Leader

Minoru Kanehisa, Director and Professor, Bioinformatics Center, ICR

Project Members

[Environmental Genomics]

Minoru Kanehisa, Director and Professor, Bioinformatics Center, ICR

Tatsuya Akutsu, Professor, Bioinformatics Center, ICR

Hiroshi Mamitsuka, Professor, Bioinformatics Center, ICR

Hiroaki Kato, Professor, Graduate School of Pharmaceutical Sciences

Masato Umeda, Professor, Institute for Chemical Research

[Chemogenomics]

Nobutaka Fujii, Professor, Graduate School of Pharmaceutical Sciences

Kiyoshi Tomioka, Professor, Graduate School of Pharmaceutical Sciences

Yoshiji Takemoto, Professor, Graduate School of Pharmaceutical Sciences

Motonari Uesugi, Professor, Institute for Chemical Research

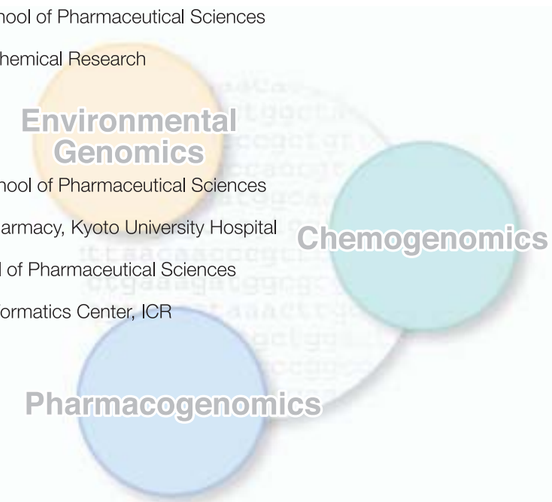
[Pharmacogenomics]

Gozoh Tsujimoto, Professor, Graduate School of Pharmaceutical Sciences

Ken-ichi Inui, Professor, Department of Pharmacy, Kyoto University Hospital

Shuji Kaneko, Professor, Graduate School of Pharmaceutical Sciences

Susumu Goto, Associate Professor, Bioinformatics Center, ICR



21st Century Center of Excellence Program
Knowledge Information
Infrastructure for Genome Science

Project Leader
Minoru Kanehisa
Bioinformatics Center, ICR, Kyoto University

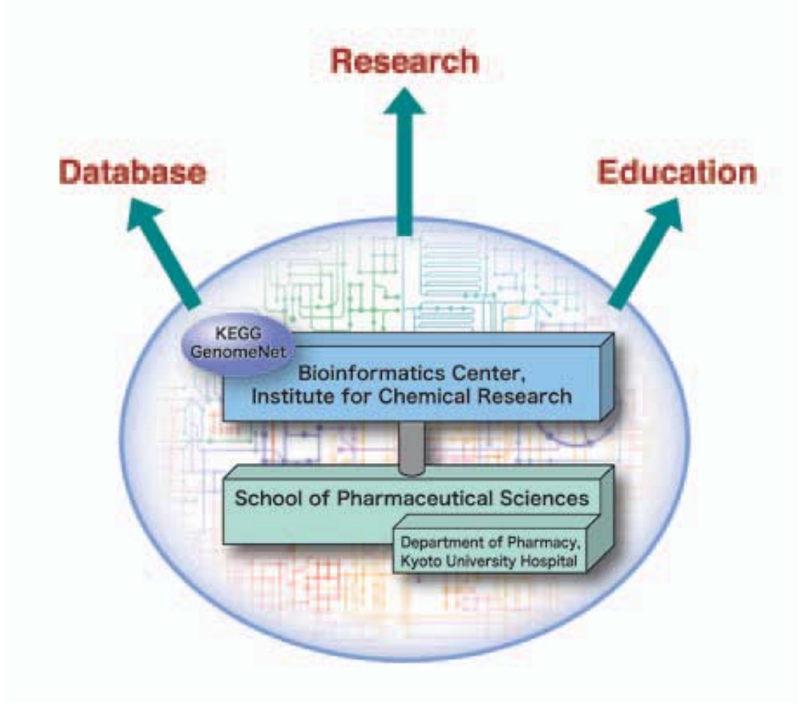
A grand challenge in the genome science of the 21st century is to computationally predict systemic functional behaviors of the cell, the organism, and the ecosystem from genomic and molecular information, especially for the purpose of medical, industrial, and other practical applications. This will require new bioinformatics, not simply for screening large-scale data, but rather for reconstructing the system and computing its interactions with the environment. While traditional genomics and other types of -omics approaches have contributed to our understanding of the genomic space of possible genes and proteins that make up the biological system, new chemical genomics initiatives will give us a glimpse of the chemical space of possible compounds and reactions that exist as an interface between the biological system and the natural environment.

The 21st Century Center of Excellence (COE) program entitled "Knowledge Information Infrastructure for Genome Science" is a joint venture between the Bioinformatics Center of the Institute for Chemical Research and the School of Pharmaceutical Sciences in Kyoto University, with support from the Ministry of Education, Sports, Science and Technology of Japan. The program aims at developing bioinformatics technologies to reconstruct the biological system and its environment by integrating our knowledge on the genomic space and the chemical space, which will enable practical applications including drug discovery and personalized medicine. In addition, a new education/training program is being introduced, integrating traditional bioinformatics with physics and chemistry and with medical and pharmaceutical sciences. The program also assists further developement of the KEGG resource (http://www.genome.jp/kegg/), by extending its PATHWAY collection to include wiring diagrams of both endogenous and exogenous molecules.

Promotion of interdisciplinary research in bioinformatics and chemical genomics

Database resources and computational services

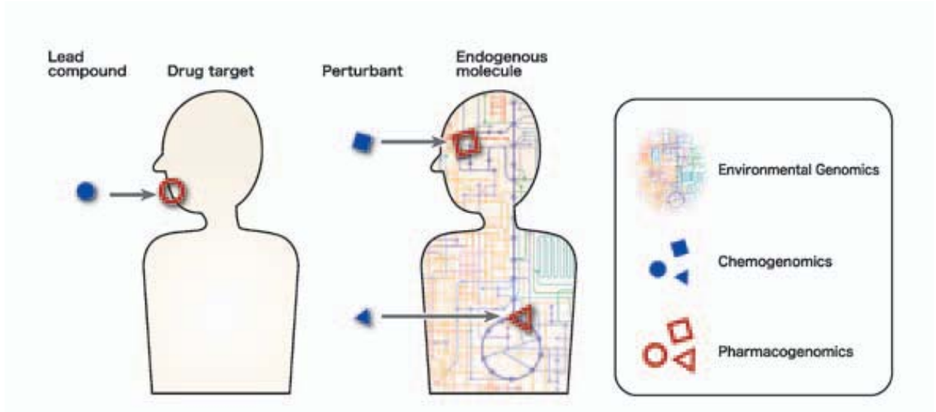
Education and training of young scientists




京都市酒蔵バイオVIL開所
「21世紀COEシンポ


Genome Science
ゲノム科学の知的情報基盤・研究拠点形成
～ ゲノムとケミストリーの融合を目指して ～
公開シンポジウム
「ゲノムからケミカルゲノムへ」

From Pharmaceutical Science to Biological Systems Science:



After completion of the Human Genome Project, various post-genomics approaches are being undertaken to utilize the genome information such as for drug target discovery and personalized medicine. In Fall 2003 the U.S. National Institutes of Health announced the Roadmap, which contained new chemical genomics initiatives for screening of useful chemical compounds including imaging probes and drug leads. Our COE program was initiated before the NIH Roadmap, but we had foreseen the importance of small chemical compounds. Our program consists of three research components: pharmacogenomics for drug target discovery, chemogenomics for drug lead discovery, and environmental genomics for understanding the relationship between genomic and environmental information, especially, drug targets and drug leads, in terms of the molecular interaction networks. Searching for a lead compound and its target is just one example of searching for a perturbant, its target, and any consequence of the perturbed biological system. Therefore, our approaches currently focused on pharmaceutical science can be generalized and applied to wider areas of biological sciences. We are developing bioinformatics technologies to integrate genomic and chemical information, and to decipher molecular wiring diagrams involving both endogenous and exogenous molecules. This will enable basic understanding of the biological systems and their interactions with the chemical environments, as well as new medical and industrial applications.

Bioinformatics Center

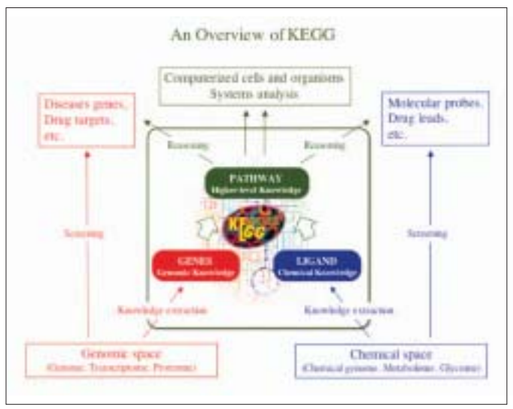
Institute for Chemical Research, Kyoto University

Professor
Minoru Kanehisa

Integration of Genomics and Chemistry by Reaction Networks

The macromolecular structures of DNAs, RNAs, and proteins are determined by template-based syntheses of replication, transcription, and translation, which represent the processes of genetic information transmission and expression as shown by the central dogma of molecular biology. In contrast, the structures of carbohydrates, lipids, and various secondary metabolites are determined, not by template structures, but by biosynthetic pathways. Such biosynthetic codes are far more complex than genetic codes, and our knowledge is still quite limited. Bacteria and plants are known to produce diverse substances, many of which have medical and pharmaceutical relevance, including antibiotics and crude drugs. With the complete genome sequences becoming available for an increasing number of organisms, it should in principle become possible to infer a complete set of biochemical substances produced by each organism, and furthermore to infer a set of xenobiotic substances that are degraded by an organism. We are thus developing new bioinformatics approaches for integrated analyses of genomic and chemical (environmental) information by organizing our knowledge on biosynthetic and biodegradation pathways, which will be used in practical applications of drug discovery and drug design.

Our current knowledge on the universe of enzymatic reactions is represented by the set of EC (Enzyme Commission) numbers. The EC numbers are given to distinguish enzymatic reactions (chemical information), but they are also utilized as identifiers of enzymes and enzyme genes (genomic information). This duality of the EC numbers makes it possible to link the genomic repertoire of enzyme genes to the chemical repertoire of endogenous and exogenous substances. We have thus far developed methods to compare chemical compound structures and to classify chemical transformation patterns in enzymatic reactions, which have been used to automatically assign EC numbers given chemical structures of substrates and products. These methods enable us to convert a set of chemical compound structures to a set of enzyme genes through a network of reactions, and vice versa. In other words, the chemcal information can be used to decipher the genome, and the genomic information can be used to predict chemical structures. The results of this research project is incorporated in KEGG, a database system developed by our group and used worldwide as a knowledge information infrastructure for genome science.



Bioinformatics Center

Institute for Chemical Research, Kyoto University

Professor
Tatsuya Akutsu

We develop efficient and flexible algorithms for bioinformatics. We also study mathematical aspects of systems biology. The research topics include inference and analysis of various types of biological networks, prediction and analysis of protein/RNA structures, statistical models for sequence analysis, and scale-free networks.

Mathematical Analysis of Structures of Biological Networks

Scale-free properties of biological networks have been extensively studied. We study how the topology of the scale-free networks is changed by the line graph transformation, where the line graph transformation relates two representations of a metabolic network: compound network and reaction network. The main result is that the degree distribution follows a power law $P(k) \propto k^{-\gamma+1}$ after the line graph transformation if the original network follows a power law $P(k) \propto k^{-\gamma}$. We also examined the degree distributions of compound networks and reaction networks using the KEGG database. The results suggest that a similar property holds for these networks.

Classification of Protein Sequences and Chemical Compounds Using Support Vector Machines

Support vector machines and kernel methods have been applied to various classification problems in bioinformatics. In order to apply kernel methods to bioinformatics problems, it is usually required to develop a kernel function which provides a kind of measure of similarity between two objects. Recently, we developed two kernels: the local alignment kernel for protein sequences, and an extended marginalized graph kernel for chemical compounds. The results of computational experiments on chemical compounds suggest that our method is much faster than the original marginalized graph kernel while keeping classification accuracy.

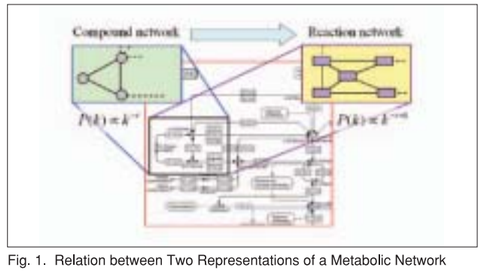


Fig. 1. Relation between Two Representations of a Metabolic Network

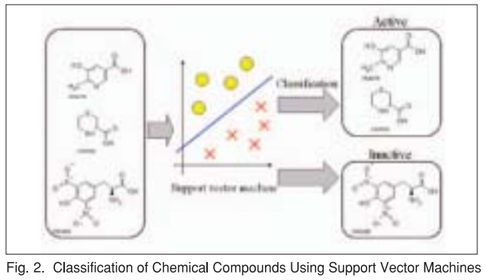


Fig. 2. Classification of Chemical Compounds Using Support Vector Machines

Environmental Genomics

Bioinformatics Center

Institute for Chemical Research, Kyoto University

Professor
Hiroshi Mamitsuka

With the recent advance of experimental techniques in molecular biology, the research of modern life science is shifting to the comprehensive understanding of biological mechanisms carried out by a variety of molecules. The focus of our laboratory is placed on biological phenomena that can be represented by biological networks such as metabolic pathways and signal transduction pathways. The research objective is to develop computational techniques based on computer science and/or statistics to systematically understand the principles of biological networks at the cellular and organism level. Considering the focus of the 21st Century COE program "Knowledge Information Infrastructure for Genome Science", we here show two typical examples of our research.

New Chemical Descriptor Based-on Graph Theory

The tree-width is, in graph theory, a metric for the complexity of a graph. More precisely, it is a measure indicating how a given graph is similar to a tree. For example, if a given graph of n nodes is a tree, its tree-width is one, and it will be two if the graph has a cycle. And it will be n-1 if the graph is a complete graph. We checked how we can use the tree-width as a chemical descriptor for a planar chemical structure of a given chemical compound. We focused on the chemical reactions by which the tree-width of a chemical compound can change. Our first finding was that these reactions tend to have particular EC numbers. We then found that the graph connectivity of an arbitrary pathway graph changes drastically when we remove these chemical reactions. These results imply that the tree-width can be a new chemical descriptor to elucidate the structural and biochemical properties of chemical compounds.

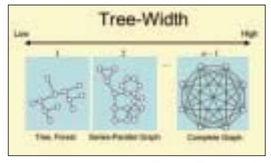


Fig. 1. Tree-Width

Estimation of Biologically Active Pathways with Microarray Data

A database of metabolic pathways is a collection of chemical reactions that appeared in the literature. This indicates that each reaction is not observed under a same cellular condition, implying that a long distance path might not be true under some condition. Motivated by this doubt, we build a method to find biologically active paths in a given metabolic pathways by using microarrays. We generate a hierarchical mixture of Markov chains using a given pathway and estimate its probabilistic parameters from given microarray experiments based on an EM (Expectation-Maximization) algorithm.
Experiments using real datasets showed that our method found paths actually used in terms of gene expressions as well as the correlation of chemical reactions located at distant parts in a pathway.

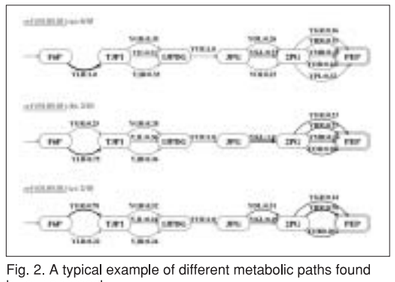


Fig. 2. A typical example of different metabolic paths found by our approach

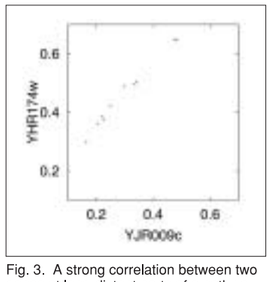


Fig. 3. A strong correlation between two genes at long-distant parts of a pathway

Graduate School of Pharmaceutical Sciences

Kyoto University

Professor

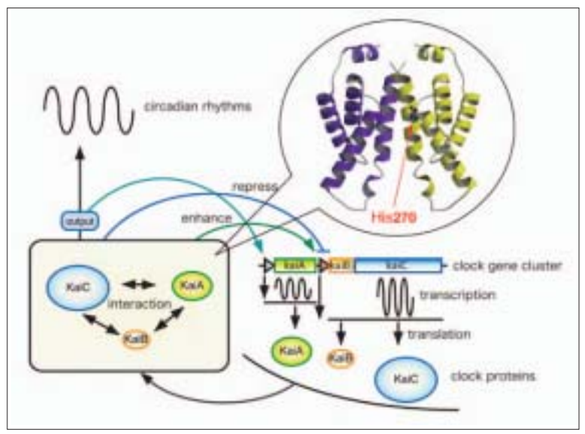
Hiroaki Kato

Structural Biology for Biological Clock

Structural biology is the study of the architecture and shape of biological macromolecules and what causes them to have the structures they have. This subject is of great interest to biologists, because macromolecules carry out most of the biological functions, and because typically they are folding into a specific three-dimensional shape to perform their functions. We are aiming to understand the function of important proteins that will be the key to the drug development by structural biology and to develop the research techniques of structural biology, especially the crystallization methods for membrane proteins and the ultra-high-resolution X-ray crystallographic techniques beyond 1 Å resolution using the synchrotron radiation of SPring-8. We believe that the underlying principle of the function is structure. We are not interested in the structure itself, but in the function. Thus, we are proceeding functional studies by various techniques from molecular biology to physical chemistry as well as protein crystallography.

The circadian clock exists in most organisms, and controls the daily activity cycles. In cyanobacteria, that appear on the earth as the first photosynthetic organisms, *kaiABC* is the clock gene cluster. It is known that the cluster consists of two operons, *kaiA*, whose product enhances *kaiBC* promoter activity, and *kaiBC*, which is repressed by the KaiC protein. We found that the C-terminal domain of KaiA is responsible for its dimer formation, binding to KaiC, enhancing KaiC phosphorylation and generating the circadian oscillations. To understand the circadian clock machinery at the atomic level, we solved the three-dimensional structure of the C-terminal clock oscillator domain of KaiA by X-ray crystallography. We also characterized in KaiA, a functionally essential residue, His270, and examined the roles of several aminoacid residues that have been reported previously to affect circadian rhythm.

Uzumaki et al., *Nat. Struc. Mol. Biol.*, **11**, 623-631 (2004)



Institute for Chemical Research

Kyoto University

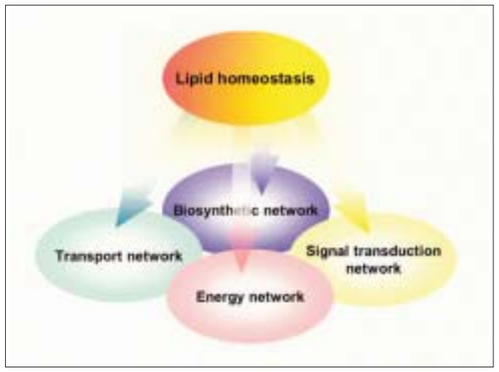
Professor

Masato Umeda

Lipid Homeostasis and Genome Network

Recent progress in lipid metabolome analyses have shown that more than 10,000 molecular species of lipids are present in biological membranes and a vast amounts of structural and functional information are stored in each lipid molecule. Although a primary function of membrane lipids is the formation of the permeability barrier of cells, they are enzymatically converted into a wide variety of bioactive lipid species such as prostagrandins, leukotriens, platelet-activating factor, cannabinoids, lysophosphatidic acid, inositol lipids, ceramide, playing a crucial role in "Signal transduction network" that control cellular metabolism and organism's activities. Lipid molecules are not tethered to the site of their synthesis, but are actively transported and assembled into different cellular membranes and body tissues. Specialized lipid sorting and transport machineries form a "Transport network of lipids", which distributes lipid molecules to various cellular membranes and tissues. During cell division, cells must double their components for daughter cells and extensive degradation and synthesis of lipids occur during membrane biogenesis. Although it is reasonable that this event should coincide with the synthesis of other cellular components such as DNA and proteins, the biochemical mechanisms of the cell-cycle regulated "Biosynthetic network of lipids" that coordinate the synthesis of every molecular species of lipids are largely unknown. Finally, the lipids are a major source of energy in the organisms. Some lipid metabolites play an important role as a fuel sensing and signaling molecule, forming an "Energy production and storage network", which controls energy homeostasis of the organisms.

It is becoming clear that the multi-layered regulatory networks of the lipid metabolism, transport and signal transduction, so called the "Lipid homeostasis", are formed at the cellular and whole organism levels. Dysfunction of the lipid homeostasis has been suggested to cause various diseases in the modern society, such as obesity, diabetes, atherosclerosis, cancer, infectious diseases, and neurological disorders. New advances in genomics and proteomics approaches have already begun to aid in the fight against the leading cause of mortality worldwide. We have undertaken molecular biology, cell biology, and behavioral genetics approaches to elucidate the role of molecular motion of membrane lipids in animal morphogenesis and thermoregulation. In this project, we would also employ bioinformatics approaches to understand the fundamental principles underlying the lipid homeostasis. Such an integrated approach also encompass the elucidation of novel targets for drug and gene therapy of the diseases.



Graduate School of Pharmaceutical Sciences

Kyoto University

Professor
Nobutaka Fujii

Genomic information-converging Medicinal Chemistry Based on Chemical Proteomics

Recent advance in genome science is providing us exponentially amplified drug targets. As such, there has been increasing upsurge in the development of innovative platform for genome/proteome-based drug discovery to facilitate the process thereof from a national standpoint. This project is intended to constitute a knowledge information infrastructure for a new frontier of chemogenomics-based medicinal chemistry beyond genomics and proteomics in fusion of genome science and chemistry. In order to embody the innovative axis for genome-lead drug discovery platform, we have been engaging the genomic information-converging approach in medicinal chemistry by integration of the following research programs.

1) Chemical characterization of drug targets based on comprehensive functional genomics

2) Downsizing of high molecular weight biologically active peptides/proteins using peptide-lead conformationally restricted templates for genome-lead drug discovery

3) Development of stereo-controlled synthetic process of peptide isosteres based on organo-metallic chemistry for "Peptide-Nonpeptide" link.

4) Chemical dissection of the aufbau principle of 7 transmembrane-G protein coupled receptor(7TM-GPCR) and daynamic outside-in signal transduction mechanism thereof

5) Reconstruction of 7TM-GPCR and pharmaceutical development thereof

Herein, we summarize the recent investigations focusing on the development of highly specific CXCR4 antagonist, the receptor of which is relevant to several problematic diseases (cancer metastasis, AIDS, rheumatoid arthritis, etc.), as a practical application of the above strategy (Fig. 1). Of note, the highly specific CXCR4 antagonist, as a chemical probe, served to disclose the many physiological and pathological functions of CXCR4 including its important role for the formation of the dark zone and light zone appeared in germinal center of lymph node (*Nature Immunology*, 2004, **5**: 943) A new method for the facile synthesis of membrane embedded peptides utilizing lipid bilayer-asisted chemical ligation has been also developed aiming at the chemical synthesis of CXCR4, which is classified to 7TM-GPCR family (*Chem. Commun*, **2004**: 607).

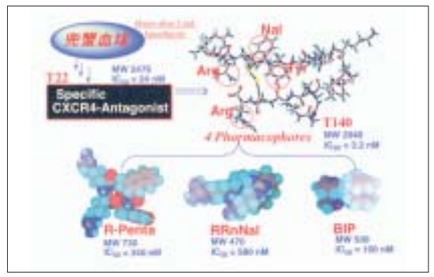


Fig. 1. Downsizing and Nonpeptidylation of High Molecular Weight Peptides/Proteins by Conformationally-restricted Drug Discovery Templates :Practice in Peptidic CXCR4 Antagonist

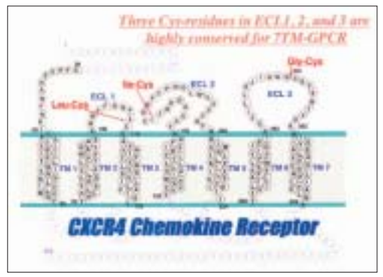


Fig. 2. Towards the Chemical Synthesis of 7TM-GPCR

Graduate School of Pharmaceutical Sciences

Kyoto University

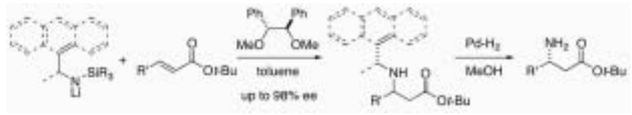
Professor
Kiyoshi Tomioka

Development of stereoselective asymmetric synthesis of biological information amines

The dynamic interaction of molecules results in an information function, and the rationalized accumulation of molecular network guarantees the generation, differentiation, and homeostasis of life. In our research, based on the concept that the initial process of the chemical reaction is assumed to be the recognition process of a molecule with a molecule, development of chiral molecules operative in activation and control of the reactive chemical species is the challenge. Asymmetric reaction and selective bond formation through dynamic molecular recognition is the practical target of our study. Highly efficient total synthesis focuses on nitrogen containing biological information molecules.

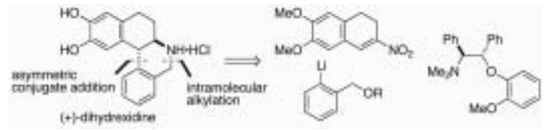
1. Asymmetric synthesis of β-amino acids by conjugate addition of a lithium amide

The ideal synthetic way to β-amino acids relies on the asymmetric addition of an amine. The lithium amide generated from trimethylsilylbenzylamine was convertible to a chiral amine nucleophile through complexation with a chiral ligand, and was enable the asymmetric conjugate addition to unsaturated carbonyl compounds with high selectivity.



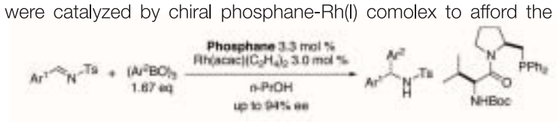
2. Asymmetric synthesis of nitrogen containing compounds by conjugate addition to nitroolefins

The nitrogen containing aromatics represented by dihydroxidine are the chemical family endowed with a full agonist activity to dopamine D1 receptor. Asymmetric conjugate addition of aryllithiums to nitroolefins was mediated by a chiral ligand to give a product with nearly complete enantiofacial selectivity 97% ee, which opened new efficient way to nitrogen containing heterocyclic compounds.



3. Asymmetric synthesis of amines by asymmetric alkylation of imines

The reaction of arylboronic acids with imines were catalyzed by chiral phosphane-Rh(I) comolex to afford the arylated amines with 94% ee.



Promising asymmetric synthetic ways to a variety of amines were developed.

Publication

(1) H. Doi, T. Sakai, et al. *J. Am. Chem. Soc.* **2003**, *125,* 2886.

(2) M. Yamashita, K. Yamada, et al. *J. Am. Chem. Soc.* **2004**, *126,* 1954.

(3) M. Kuriyama, T. Soeta, et al. *J. Am. Chem. Soc.* **2004**, *126,* 8128.

Graduate School of Pharmaceutical Sciences

Kyoto University

Professor

Yoshiji Takemoto

Creation of Biologically Functional Molecules that Specifically Interact with Proteases

With an aim of establishing the knowledge information infrastructure for genome science by means of organic synthetic chemistry, we are struggling with creation of artificial multi-functional molecules, which work like proteases, and development of catalytic asymmetric reactions and synthetic methods for synthetic compound libraries. Through these studies, novel drug temples as well as unknown pharmacophores that can interact multiply with biologically important macromolecules will be revealed.

1. Design and synthesis of artificial enzyme-like molecules

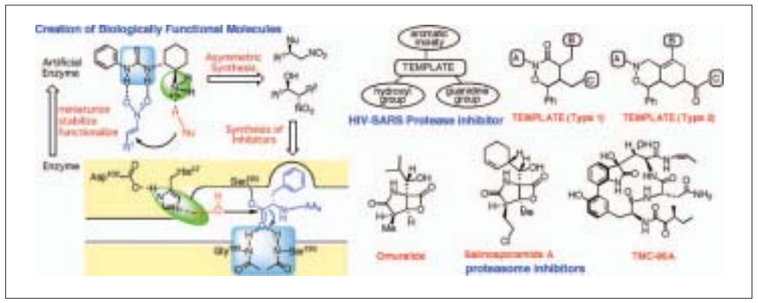
Taking into account the reaction mechanism of serine-proteases, we designed and synthesized various types of small molecules, which may possess enzymatic functions. As a result, the thiourea derivatives bearing a tertiary amino group were proved to have an excellent catalytic activity as a general acid as well as a general base like enzymes. In addition, we have succeeded in total asymmetric synthesis of (*R*)-baclofen and a NK-1 receptor antagonist (CP-99,994) using these thiourea catalysts.

2. Development of synthetic method for chiral heterocyclic drug templates

To develop effective inhibitors against HIV- and SARS-proteases, we designed two types of heterocyclic compounds (type 1 and 2) as a drug template, to which a hydroxy group, a guanidine, and an aromotic ring could be introduced at a different position, respectively. The preparation of these drug templates could be achieved stereoselectively by the Ir-catalyzed allylic substitution reaction, which we had already developed.

3. Total synthesis of natural products possessing inhibitory activity against proteasome

The proteasome has become an interesting target for the development of drugs against cancer and Alzheimer's disease. We are now struggling to develop an efficient synthetic method for the total synthesis of omuralide and TMC-95A. During these studies, we have already discovered two key reactions, In-mediated radical cyclization and Rh-catalyzed hydroamidation, for the key intermediates, (3*E*)-3-alkylidene-γ-lactames, of the targeted molecules.



Institute for Chemical Research

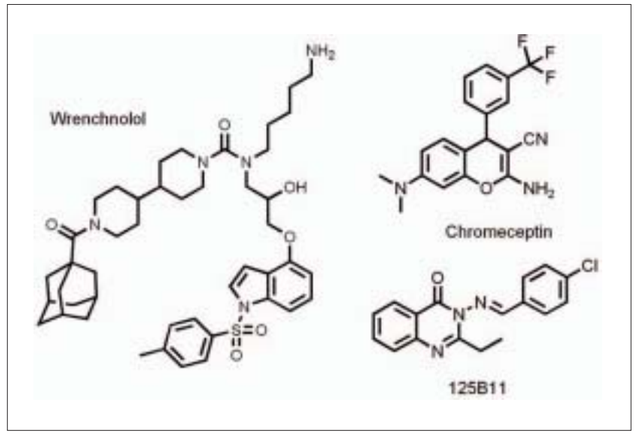
Kyoto University

Professor

Motonari Uesugi

Small-molecule initiated biology

In human history, small organic molecules have been utilized for improving human health and for revealing secrets of life. Discovery or design of small organic molecules with unique biological activity permits small-molecule-initiated exploration of biology and further understanding of human diseases. Our laboratory has been discovering small organic molecules that modulate transcription or differentiation to use them as tools to explore biology. Such chemistry-initiated biology is recently called chemical biology or chemical genetics, an emerging field of biology and medical sciences. Chemical biology is not drug discovery but basic biological research to understand human diseases or complex biological events. Nevertheless, chemical biology research may "catalyze" future drug discovery thorough the knowledge and technology gained in the small-molecule initiated basic research.



Graduate School of Pharmaceutical Sciences

Kyoto University

Professor

Gozoh Tsujimoto

Novel drug discovery based on the genomic information

About 3 billion base pairs of the human genome was completely sequenced. However, there still is a lot of works remained to do for understanding what kind of role for each gene to play. The genomic drug discovery science is the science field of discovery of the new drug, the medicine of the effect to be higher and the medicine with few side effects, using the genome information.

1. Functional genomics of G protein-coupled receptors (GPCR)

Information from genome sequencing estimated the existence of 700-800 GPCRs in the human genome: about 250 of GPCRs are identified as receptors for known ligands, and the rest are still orphan receptors (oGPCRs). Recognized for the potential of oGPCRs as targets of novel drug discovery, oGPCRs have attracted a tremendous level of attention in terms of continued identification of their endogenous ligands and elucidation of their physiological functions.

Fig.1. Cell-base screening system for ligand fishing

Attempts to study the targeted molecule distribution at the subcellular level have been limited by the lack of specific structural probes of GPCRs. Functional GPCR-GFP (green fluorescent protein) fusion molecules are important tools for optical measurement of biochemical and biophysical processes that are relevant to the signal transduction of GPCRs (Figure 1). Very recently, we successfully identified natural ligand (free fatty acid) for orphan GPCR GPR120. GPR120 localizes in the intestinal tract and regulates the secretion of the most potent insulin-secreting peptide hormone GLP-1; hence, this GPCR would be a good candidate for anti-diabetic drugs (Figure 2).

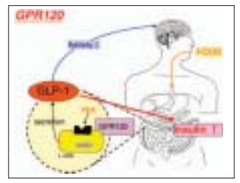


Fig. 2. Successful de-orphaning of GPR120

2. Functional genomics by microarray DNA Chip

A microarray is one of the most important basic technology for drug discovery from the aspect of genomics (Figure 3). The focus of genome research will be shifted to functional analysis of genes including the determination of precise transcript unit as transcriptome. The gene expression pattern (i.e. profile) of disease specific status could be obtained by DNA chip technology that makes it accelerate to find candidate molecules of drug target. We are interested in the construction of databases of gene expression data, gene expression analysis of disease model animals and human disease status (such as cancer, nephritis, diabetic nephropathy, etc), and finally discovering the candidate molecule of the effective novel drug target and revealing the mechanisms of human disorders.

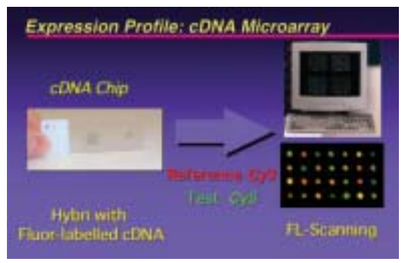


Fig. 3. Microarray

Department of Pharmacy

Kyoto University Hospital

Professor

Ken-ichi Inui

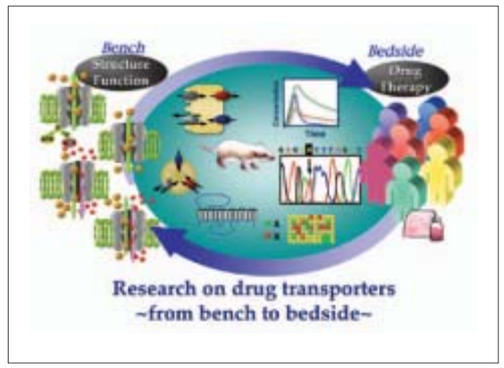
Laboratory and Clinical Studies of the Pharmacokinetics, Effectiveness and Toxicity of Drugs

We have systematically developed pharmacokinetic research from drug transport analyses based on the molecular levels to their clinical applications, i.e., functional and molecular characterization of drug transporters expressed in the intestine, kidney and liver, and the establishment of the personalized medicine based on the molecular information of drug transporters.

Tacrolimus has been widely used in many types of allograft recipients, but this agent features a wide intra- and interindividual pharmacokinetic variability, especially of bioavailability after oral administration. To establish the personalized immunosuppressant therapy, the relationship between the tacrolimus concentration/dose (C/D) ratio and the intestinal mRNA expression level of MDR1 or CYP3A4 in the recipients of living-donor liver transplantation (LDLT) were investigated. As a result, we have demonstrated that the intra- and interindividual variability of tacrolimus pharmacokinetics in allograft recipients is related to the intestinal expression level of MDR1, but not that of CYP3A4. This information is utilized for the determination of the initial dosage regimen of tacrolimus for LDLT recipients. For many drugs excreted into the urine, dosage regimens are usually adjusted according to creatinine clearance. However, creatinine clearance reflects only the glomerular filtration rate of normal kidney, and dosage schedule based on the creatinine clearance was reported to be inadequate for many drugs. To overcome this clinical problem, we have pursued the novel biomarker for renal drug excretion, and found that mRNA expression levels of organic anion transporter 3 (OAT3/SLC22A8) were well correlated with renal secretion of an anionic cephalosporin antibiotic cefazolin. Thus, the expression levels of renal drug transporters are suggested to be a useful marker to evaluate the renal function.

We have performed following research projects:

1) Functional and molecular characterization of drug transporters and its clinical application,
2) Pathophysiological roles of drug transporters,
3) Mechanisms of drug interaction,
4) Pharmacogenomics and personalized medicine.



Graduate School of Pharmaceutical Sciences

Kyoto University

Professor
Shuji Kaneko

Study of TRP channels expressed in the central nervous system

Among the 30 thousand human genes, number of 'druggable' genes are said to be 3 thousand. Also, the number of disease-modifying genes may be 3 thousand, and the overlap will be in a range of 6 to 15 hundred at the most. In the possible targets, GPCR and kinases are highly ranked, while the share of cation channel is only 2%. However, the percentages of druggable genes are kinase on top, GPCR the second, and cation channels are ranked as the third with a share of 5%. These estimates indicate our hope, however the genome-based drug discovery is far from the ion channel targets.

TRP (transient receptor potential) channels are new class of cation channel operated by various kinds of environmental stimuli, including temperature, pH, pressure (osmolarity), reactive radicals and receptor activation leading to phosphoinositide metabolism. Among the TRP superfamily, we are focusing on TRPC3/4/5 and TRPM2/7 channels specifically or abundantly expressed in the brain, which have great potentials as drug targets for CNS diseases.

Building Japanese-English ontology of life science terms

In parallel with the recent progress in life sciences, vast numbers of words for new substances and phenomena have been appearing. Since 1993, we have been analyzing English texts of medical journals selected mainly from the public MEDLINE database and collecting frequently-used terms together with their frequencies, concordances, typical usages, definitions and translations. The data were recorded in a versatile relational database and edited into several text-based, English-Japanese and Japanese-English dictionaries intended for public release. We have also developed convenient online and offline systems for our electronic dictionaries and are providing them on our homepage (http://lsd.pharm.kyoto-u.ac.jp/). The recent 2005 version of the online dictionary WebLSD contains 41,000 entries of English terms with their translations and definitions, 26,000 example sentences for 5,100 words, and on-demand concordance engine for all words. The access log statistics of the WebLSD server indicate more than 70,000 searches per day, and the number has been increasing exponentially year by year. Based on the dictionary, we are building a Japanese-English compatible thesaurus, which will be used in semantic web and meta-analyses of literature for knowledge finding.

Bioinformatics Center

Institute for Chemical Research, Kyoto University

Associate Professor
Susumu Goto

Database of Chemical Compounds and Glycans for Pharmacogenomics

We are developing a database of chemical structures with the information on reactions and pathways, which allows us genome-scale analysis such as expression data analysis. In this study, we mainly focus on the carbohydrate structure database GLYCAN and develop glycan related technologies for the pharmacogenomics applications.

1. Carbohydrate structure database: GLYCAN

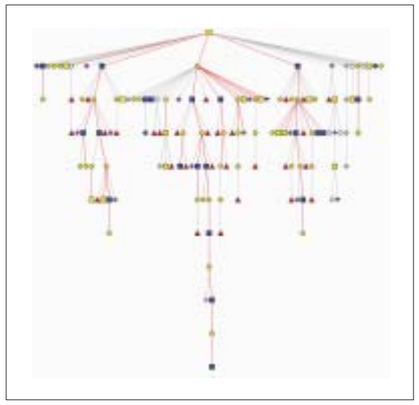
GLYCAN is a database of carbohydrate structures collected from research articles. The main portion of GLYCAN came from CarbBank, which is the most famous carbohydrate database but has been unfortunately discontinued due to the lack of funding. We reorganized the structures into newly developed KCF (KEGG Chemical Function) format, which is more suitable for structure-based applications, e.g. inputting new data and searching similar structures.

2. Composite structure map: CSM

Composite structure map is a map illustrating all possible variations of carbohydrate structures within organisms, where the edges and nodes represent glycosyltrasferases and carbohydrate structures (Figure). It is constructed as a bridge between carbohydrate functions and structures, and is able to display, for example, expression data of glycosyltransferases in a compact manner.

3. Carbohydrate structure input tool: KegDraw

KegDraw is an intuitive drawing tool for carbohydrate structures. KegDraw is a Java application, so it runs on locally in a platform-independent manner and it allows the drawing of not only glycan structures but also atomic coordination of chemical compounds. Glycan structures drawn in KegDraw can be used as queries to search against GLYCAN and other databases by KCaM structure search server and as inputs for the database updates.



Bioinformatics Education Programs

Bioinformatics is a broad, interdisciplinary field, and we provide two types of graduate education programs to meet the specific needs in different schools and departments. One is an advanced program for graduate students at the Bioinformatics Center, who would become highly trained specialists in both academia and industry. The other is an introductory program for graduate students at the School of Pharmaceutical Sciences, who can now select bioinformatics as a minor field.

Advanced program for specialists

In 2002 the Japanese Society for Bioinformatics (JSBi) announced its first version of the recommended education curriculum. Based on this curriculum, the Kyoto University Bioinformatics Center, in collaboration with the Human Genome Center of the University of Tokyo, initiated a series of courses, which are given simultaneously in the two centers using the video conferencing systems. These courses are also available on the Web through the e-learning system that includes WebCT and a video library. An additional component of our advanced education program is the collaboration with Boston University and Humboldt University Berlin. We exchange students for internship and jointly organize annual workshops. The advanced education program is mostly supported by the special coordination fund for promoting science and technology from the Ministry of Education, but the workshops have been organized under the 21st Century COE program.

Courses given in 2005

No.	Course	Content
0501	Biostatistics	Overview and introduction; Classification and discrimination; Information criteria; Bootstrap; Markov model; Kernel methods; Qualitative data analysis; QTL analysis
0502	Algorithms	Overview and introduction; Sequence alignment; Index structure for strings; Pattern discovery; Machine learning; Graph algorithms I; Graph algorithms II; 3D structure alignment
0506	Network Analysis	Gene network inference I; Gene network inference II; Reaction network prediction; Metabolic flux analysis; Technologies for modeling and simulation; Pathway modeling and simulation; Network motifs; Scale-free network
0503	Genome Data Analysis	Knowledge discovery from genomic data; Gene finding; Comparative genomics and evolution; Methods for molecular evolution analysis; Phylogeny inference; Evolution-based protein function analysis; Structure-based protein function analysis; Morphology and molecular evolution
0508	Systems Biology	Introduction to microarray analysis; Expression data clustering; Introduction to proteomics; Network analysis by comparative genomics; Simulation of biochemical reactions; Chemoinformatics; Glycome informatics; Reconstruction of biological systems
0525	Practical Bioinformatics	Database resources
0526	Practical Programming	Ruby programming; Perl programming; C programming
0527	Technical Writing	Presentation skills; Grant writing; Patent writing

Introductory program for non-specialists

For graduate students in the School of Pharmaceutical Sciences, bioinformatics is a practical discipline. The introductory program offers a practical guide to the databases and computational tools currently available. A selection of the courses based on the JSBi curriculum, which can also be taken in the main campus of Kyoto University through the video conferencing system, is used to approve credits for the bioinformatics minor. In addition to the graduate education, the 21st Century COE program has helped to introduce bioinformatics-related courses in the undergraduate education of pharmaceutical sciences.