RDB was implemented on a UNIX workstation (e.g. Silicon Graphics OCTANE). We used an object oriented database management software ACEDB (A Caenorhabditis elegans DataBase, 2001; Dunham, Durbin, Thierry-Mieg & Bentley, 1994; Stein & Thierry-Mieg, 1998) as the base system. To modify data and insert new functions, PERL and/or C programs were integrated into ACEDB. As for the method of calculating the three-dimensional structure of ligands, we used Molecular Mechanics 2 (Allinger, 1977). The three-dimensional (3D) structural image is provided by computer commands that call visualization tools, such as Chime (Martz, 2002) or RasMol (Sayle & Milner-White, 1995).

Table 1. Tree class structure

Class	Contents	(Pointer)	(Contents in pop-up window)
RG	Membrane / Nuclear Receptor		LG List
LG	Large Group List		MG List / Group List
MG	Middle Group List		Group List
Group	Sequence similarity to the main species	[MulSq]	Multiple sequence alignment
	Species	species list	Receptor information
Receptor	Protein Data (Acc. No., No. of Seq.)	[PIR ref / SP ref]	PIR[#1] entry / Swiss Prot [#2]entry
		[PIR seq / SP seq]	□Transmembrane Region
			□DNA-Binding Region
			□Ligand-Binding Region
		[St 2D-pred]	2D Structure Prediction[#3]
	Sequence similarity to the main species	[MulSq]	Multiple sequence alignment
	3D data (Overlap region with PIR/SP)	[PDB ref]	PDB [#4]entry
		[St 3D-image]	3D image
	DNA Data (Accession No., No. of Seq.)	[GB ref]	GenBank[#5] entry
		[GB seq]	DNA Sequence
	Gene Data (Symbol, Aliases, Map pos.)	[GDB ref]	GDB[#6]entry
	SNPs information	[snp]	SNPs Collection System[#7] entry
	Drug information (Generic Name)	[drug]	Drug[#8] data
	Cell Signaling Networks information	[Signaling]	CSNDB[#9] entry
	Transcription Factor information	[Transfac]	Transfac[#10] entry
	Transcription Region information	[TRRD]	TRRD[#11] entry
	Binding Affinity information	[BindAff]	BADB[#12] entry

[#1] PIR: Protein sequence database (http://www-nbrf.georgetown.edu/pirwww/search/textpsd.html)
[#2] Swiss Prot: Protein sequence database (http://www.expasy.ch/sprot/)
[#3] 2D Struc. Prediction: Protein secondary-structure prediction, BCM Search Launcher, (http://searchlauncher.bcm.tmc.edu/seq-search/struc-predict.html)
[#4] PDB: 3D biological macromolecular structure database (http://www.rcsb.org/pdb/)
[#5] GenBank: DNA sequence database (http://www.genome.ad.jp/dbget-bin/www_bfind?genbank-today)
[#6] GDB: The Genome database (http://gdbwww.gdb.org/)
[#7] SNPs Collection System: An agent system for collecting SNPs data. (http://search.nihs.go.jp/snp/index.html)
[#8] Drug DB: Drug database (http://moldb.nihs.go.jp/moldb/)
[#9] CSNDB: Cell Signaling Networks Database (http://geo.nihs.go.jp/csndb/)
[#10] Transfac: Transcription factor database (http://transfac.gbf.de/TRANSFAC/index.html)
[#11] TRRD: Transcription Regulatory Region Database (http://wwwmgs.bionet.nsc.ru/mgs/)
[#12] BADB: Binding Affinity Database (http://moldb.nihs.go.jp/eddb/afdb/)

2.3 System configuration

In accordance with the architecture of ACEDB, all the information was stored as structured objects in tree forms. A tree can be arbitrarily extended in any direction as more information is gathered about a particular aspect of an object. Similar objects are grouped together within a "class". The class governs what can be stored in an object and how it is displayed and used. To make all the information about an object available, objects often contain labels (pointers) to other objects. They can also contain letters, numbers, objects and computer commands. The tree class structure in RDB was shown in Table 1. All information was integrated in the file for ACEDB.

The overall system configurations are shown in Figure 1. In the off-line system, we used a BLAST search (Altschul, Madden, Schaffer, Zhang, Zhang, Miller et al., 1997) and the MView program (Brown, Leroy & Sander, 1998) for studying sequence similarity. Sequences were passed to the BLAST search program and the result was modified by the MView program and stored in the RDB.

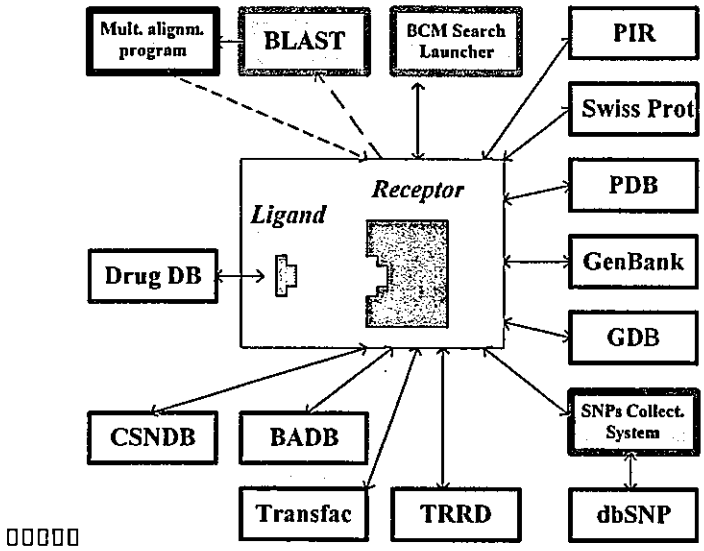

Figure 1. System configuration
↔ on-line linking; ⋯→ off-line data flow,
□ database, □ in-house database, □ program, □ system, □ in-house system
BLAST: Sequence similarity search (http://blast.genome.ad.jp/)
BCM Search Launcher (http://searchlauncher.bcm.tmc.edu/seq-search/struc-predict.html)
dbSNP: Single Nucleotide Polymorphism database (http://www.ncbi.nlm.nih.gov/SNP/)
Mult. align. program: Multiple alignment program (off-line)

2.4 Database sources

Receptor proteins were retrieved from the Swiss Prot and PIR databases on the Internet. The sources of the RDB are classified into three categories: (1) those that are collected from various references (basic data), (2) those that are retrieved from external on-line databases by the user's request (protein structure data, SNPs collection data, etc.), (3) those that were generated by some theoretical calculations (sequence similarity data, 3D structures of the ligand, etc.).

For the protein secondary-structure prediction, the corresponding amino acid sequence data was edited in a

relevant form and passed to another analytical system, for instance the BCM search Launcher (Smith, Wiese, Wojzynski, Davison & Worley, 1996). The relevant drug data, which includes Japanese and US drugs with CAS registry numbers, chemical structures and 3D structures are compiled in a Drug Database. Japanese drugs were retrieved from JAN (Japanese Accepted Names for Pharmaceuticals, n.d), and US drugs were from USP-NF (United State Pharmacopoeia – National Formulary, n.d). 3D structures of drugs were calculated using MM2.

2.5 Database contents

The numbers of LG (Large Group) in 'Membrane and Nuclear Receptor' are 36 and 5, respectively. At present, the total number of receptor proteins in the RDB is 1780. Each receptor protein has labels for the PIR / Swiss Prot entry, functional region and the secondary-structure prediction. The numbers of the DNA binding sites and ligand binding sites are 250 and 170, respectively. An aligned sequences chart for the different species was stored for the main receptors. There are 410 entries for 3D structure data in the RDB.

DNA sequences, which are translated into the receptor proteins, are available for each receptor protein. Gene data and SNPs information are included for most human receptor proteins. Data about drugs that bind to the receptors, is included as an example .

Cell signaling information is available only for human receptors (Takai-Igarashi, Nadaoka & Kaminuma, 1998; Takai-Igarashi & Kaminuma, 1999). Transcription information (Wingender, Chen, Fricke, Geffers, Hehl, Liebich et al., 2001; Kolchanov, Ignatieva, Ananko, Podkolodnaya, Stepanenko, Merkulova et al., 2002) covered all species. The binding affinity data was included only for endocrine disruptor related receptors (Kaminuma, Takai-Igarashi, Nakano & Nakata, 2000).

2.6 Automatic genetic variation data collection

An agent system of collecting Single Nucleotide Polymorphisms (SNPs) data on the Internet, was developed to search for and retrieve SNPs data related to those genes and proteins pre-registered in the system (Nakata, Takai-Igarashi, Nakano & Kaminuma, 2001b). The related gene names were previously input into the agent system and linked to IRDB. The position of any allelic frame-shift in the DNA sequence, the corresponding amino acid offset, and the converted amino acids are represented in the SNPs information.

3 DISCUSSION

IRDB was designed to be one part of the pharmaco-informatics infrastructure for genome-based personalized medicine (Kaminuma, Nakata, Nakano & Takai-Igarashi, 2001). A drug or its metabolite binding to target biomolecules, such as membrane receptors, cytoplasm enzymes, and nuclear receptors, triggers a series of reactions. Although these target molecules are not yet fully identified, it was estimated that nearly half of them are receptors (Drews, 1998).

For structure-based drug design, exact 3D structures of receptors and ligands are essential. Although only the 3D structures of a few receptors have been identified, theoretically predicted secondary-structures and aligned sequence- charts for different species are available in RDB. Although only endocrine disruptor related data is now included in BADB, much more experimental binding-affinity data are still required and theoretically calculated binding-affinity values could be included in the database in the future.

The signal pathways, which are the post-binding effects of the receptor and ligand, can be retrieved via CSNDB. The signal tranduction and transcription information may help in understanding the effects of various chemicals, such as drugs or environmental chemicals, on the living system via gene expression. SNPs data for receptors is essential for personalized medicine, in areas such as drug responses and common disease predisposition. We intend to include a link to OMIM in the near future, relating the receptor and ligand docking and the signal pathway flow. We expect this information to be useful in the basic research of drug design and for understanding living systems.

Our Receptor Database is open to the public. Access is not restricted by any firewall. By installing a free

visualization tools, such as Chime (Martz, 2002), the user can look at a three-dimensional image of the protein. No other tools are needed to look at the information in IRDB. The waiting time may be long for some sites, and sometimes there may be bad connections to the Analytical sites. To improve this, we intend to have the Analytical system on our site. Because of the huge number of amino acid sequences (in PIR and Swiss Prot), DNA sequences (in GenBank) and protein structural data (in PDB), we did not store them on our computer disk, but provided links to the original Web sites.

The whole IRDB system is constructed of many in-house sub-systems and independent systems. We do not intend to provide the software itself, because maintaining whole system would be very complicated and difficult.

4 ACKNOWLEDGMENTS

We appreciate for useful discussion with Dr. E. Wingender (GBF, Braunschweig), Dr. N. A. Kolchanov (ICG, Novosibirsk) and Dr. H. Toh (BERI, Osaka) on this subject. We are also grateful to those who provided sites on the Internet, which are relevant to our study. We also thank Mr. M. Hayakawa for PERL programming and Ms. S. Hasegawa for painstaking data input. This work was partly supported by Science Research Promotion Fund from Science and Technology Agency from 1996 to 1998.

5 REFERENCES

ACEDB (n.d) Homepage of A Caenorhabditis elegans DataBase. Available from: http://www.acedb.org/

Allinger, N. L. (1977) Conformational Analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *J. Am. Chem. Soc. 99*, 8127-8134.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped
BLAST & PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res, 25*, 3389-3402.

BLAST (n.d) Available from *NCBI, NIH* Web Site: http://www.ncbi.nlm.nih.gov/BLAST/

BCM Search Launcher (n.d) Available from *Baylor College of Medicine HGSC* Web site: http://searchlauncher.bcm.tmc.edu/

Brown, N. P., Leroy, C. & Sander, C. (1998) MView: a web-compatible database search or multiple alignment viewer.
Bioinformatics 14, 380-381

Drews, J. (1998) Technological Advances and a Paradigm Shift in Drug Research, In Drews, J. *In Quest of Tomorrow's Medicines* (pp 72-82), New York: Springer-Verlag

Dunham, I., Durbin, R., Thierry-Mieg, J. & Bentley, D. R. (1994) Physical mapping projects and ACEDB. In Bishop,
M. J. (Ed.), *Guide to Human Genome Computing* (pp.111-158), London: Academic Press

JAN (Japanese Accepted Names for Pharmaceuticals) Homepage of
Available from DCBI, NIHS (Japan) Website: http://moldb.nihs.go.jp/jan/index.html

Kaminuma, T., Takai-Igarashi, T., Nakano, T. & Nakata, K. (2000) Modeling of Signaling Pathways for Endocrine
Disruptors, *BioSystems. 55*, 23-31.

Kaminuma, T., Nakata, K., Nakano, T. & Takai-Igarashi, T. (2001) Pharmacoinformatics Infrastructure for

Genome-
based Personalized Medicine. *Chem-Bio Informatics Journal, 1,* 1-17.

Kolchanov, N. A., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Stepanenko, I. L., Merkulova, T. I., Pozdnyakov, M. A., Podkolodny, N. L., Naumochkin, A. N. & Romashchenko, A. G.. (2002) Transcription Regulatory Regions Database (TRRD): its status in 2002, *Nucleic Acids Res, 30,* 312-317.

Martz, E. (2002) Protein Explorer. Easy yet powerful macromolecular visualization. *Trends Biochem Sci. 27,* 107-109.

Martz (n.d) *Chime Resources.* Retrieved from the University of Massachussets Website: http://www.umass.edu/microbio/chime/

Nakata, K., Takai, T. & Kaminuma, T. (1999) Development of The Receptor Database (RDB): Application to The
Endocrine Disruptor Problem. *Bioinformatics. 15,* 544-552.

Nakata, K., Takai-Igarashi, T., Nakano, T. & Kaminuma, T. (2001a) Extension of The Receptor Database (RDB).
Chem-Bio Informatics Journal, 1, 115-119.

Nakata, K., Tokunaga, M., Toda,K., Takai-Igarashi, T. & Kaminuma, T. (2001b) An Agent System for Collecting SNPs
Data on the Internet. *Chem-Bio Informatics Journal, 1,* 120-123.

Smith, R. F., Wiese, B. A., Wojzynski, M. K., Davison, D. B. & Worley, K. C. (1996) BCM Search Launcher—an
integrated interface to molecular biology data base search and analysis services available on the World Wide Web.
Genome Res. 6, 454-462.

Sayle, R. A. & Milner-White, E. J. (1995) RASMOL: biomolecular graphics for all. *Trends Biochem Sci, 20,* 374-376.
http://www.umass.edu/microbio/rasmol/index2.htm

Stein L. D. & Thierry-Mieg, J. (1998) Scriptable access to the caenorhabditis elegans genome sequence and other
ACEDB databases. *Genome Res. 12,* 1308-1315.

Takai-Igarashi, T., Nadaoka, Y. & Kaminuma, T. (1998) A database for cell signaling networks. *J. Comp. Biol. 5,*
747-754

Takai-Igarashi, T. & Kaminuma, T. (1999) A pathway finding system for the cell signaling networks database. *In Silico*
Biology, 1, 129-146

USP-NF (United State Pharmacopoeia – National Formulary) Homepage of U.S. Pharmacopeia.
Available from: http://www.usp.org/

Wingender, E., Chen, X., Fricke, E., Geffers, R., Hehl, R., Liebich, I., Drull, M., Matys, V., Michael, H., Ohnhauser,
R., Prus, M., Schachere, F., Thiele, S. & Urbach, S. (2001) The TRANSFAC system on gene expression regulation,
Nucleic Acids Res, 29, 281-283.

[Frontiers in Bioscience 7, c68-73, July 1, 2002]

THEORETICAL APPROACH TO ENDOCRINE DISRUPTORS

Kotoko Nakata

Division of Chem-Bio Informatics, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo, 1508-8501, Japan

TABLE OF CONTENTS

1. ABSTRACT

Endocrine disruptors are now of scientific and public concern, because there is increasing evidence of their adverse effects on the health of an intact organism or its progeny and on changes in endocrine function. Although numerous substances have been identified as such chemicals, a huge number of chemicals remain to be tested for their endocrine disrupting capabilities. Because of the time and costs required for animal tests, some theoretical or computer-based method for screening this large number of chemicals is needed to reduce the numbers requiring animal testing. Improved quantitative structure activity relationship (QSAR) models were used for screening in combination with other approaches. New receptor-ligand docking simulations were being tested. There was good correlation between experimental and theoretical binding affinities. A database complex system being developed, which enables one to trace cellular signals triggered by the interaction of receptors with xenobiotic chemicals. Perspectives of computer-based screening methods are discussed.

2. INTRODUCTION

Rapidly increasing scientific evidence suggests that many of synthetic chemicals can interfere with normal hormone-like regulated biological processes to adversely affect development and/or reproductive function in wildlife and humans (1-14). These chemicals are called "endocrine disruptors (EDs)", because of their ability to interfere with endocrine systems. The World Health Organization (WHO) has defined EDs as exogenous chemical substances that alter the function(s) of the endocrine system and consequently causes adverse effects in an intact organism, or its progeny, or (sub)populations (http://www.who.int/pcs/emerg_site/edc_descr.html).

A large number of scientific projects related to endocrine disruptors in humans, laboratory testing, and wildlife species have been proposed (15). These studies are focused on (i) what is the mechanism by which the EDs modulate normal endocrine systems, and (ii) among the enormous number of existing chemicals, how we can screen for such EDs effectively.

As the number of suspicious EDs is estimated to be as many as 87000, it is essential to develop some theoretical or computer-based approach to pre-screen this large number of chemicals and reduce its number so that conventional wet lab testing or the so-called high through put screening (HTPS) can be applicable (http://www.epa.gov/oscpmont/oscpendo/). The QSAR approaches were tried for this purpose.

The development of endocrine-related databases system for advanced research is another approach. These databases include a potential endocrine database, a receptor database, a cell signaling database, etc. This paper does not intend to provide a complete research review on endocrine disruptors; only a short review of theoretical or computational approaches are described.

3. THEORETICAL APPROACH

3.1. EDSTAC report

In 1996, the United States Environmental Protection Agency (EPA) formed the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) to provide advice on how to design a screening and testing program for endocrine disrupting chemicals. The EPA developed a screening program by 1998 and implemented it by 1999. The EDSTAC report discussed the committee's recommendations for many aspects of the endocrine disruptor screening program, including details on priority setting and recommendations for potentially relevant exposure and effects data sources (http://www.epa.gov□/scipoly/oscpendo/history/finalrpt.htm). As QSAR methods have proven successful in molecular design and drug discovery (16), the EDSTAC considered QSAR as an important part of its priority setting process.

Shi *et al.* (17) developed an integrated system, which contains four sequential phases to predict the ability of chemicals to bind to the estrogen receptor (ER), for application to large data sets.

I. Lipinski "rule of 5"-type (18): simple rejection filters are provided to eliminate the chemicals that are most unlikely to bind the ER. Poor absorption or permeation is more likely when:

I-1. There are more than 5 H-bond donors (expressed as the sum of OHs and NHs)

I-2. The Molecular Weight is over 500;

I-3. The Log P is over 5 (MLogP is over 4.15);

I-4. There are more than 10 H-bond acceptors (expressed as the sum of Ns and Os)

Compound classes that are substrates for biological transporters are exceptions to the rule.

II. Three key 2D structural alerts, seven pharmacophore features, and the predictions of two classification models make use of K-nearest neighbor (KNN) and classification and regression tree (CART) methods.

III. QSAR models are used quantitatively to predict the activities of chemicals categorically predicted to be active in phase II.

IV. The phase II and III predictions are combined with other available information, such as human exposure level, environmental fate, and production volume, to determine a chemical's priority for testing.

3.2. Comparative molecular field analysis (CoMFA) Model

A number of QSAR models have been reported for ligand binding to the ER (19-25). Although these models yield good statistical results, they have limited applicability in predicting the ER-ligand binding affinity of chemicals that cover a wide range of structural diversity. Shi *et al.* (26) tried two types of QSAR models, CoMFA and hologram QSAR (HQSAR), for inclusion in phase III to quantitatively predict chemical binding to the ER. They used the relative binding affinities (RBAs) to the ER for 130 chemicals covering a wide range of structural diversity and concluded that CoMFA yielded the best QSAR models in terms of self-consistency and predictive ability of the test chemicals.

3.3. Modeling of signaling pathways

There are already a large number of chemicals that should be tested for their endocrine modulating capabilities. Some theoretical methods are needed for the first stages of this process, because of the time and costs required for wet lab testing. However, the conventional quantitative structure activity relationship (QSAR) approaches are of limited relevance to this problem, as these methods do not take into account the detailed mechanisms of biological molecular interactions. Kaminuma *et al.* (27) presented a prototype of an integrated database and a knowledge-based complex of chemical substances and biomolecules that can describe the internal signaling evoked by endocrine disruptors from gate-points to the endpoints. The main components of this database are a potential endocrine disruptor database, a receptor database, a cell signaling networks database, a transfactor database, and a binding affinity database based on modes of actions.

3.4. Endocrine Disruptor Structure Database (EDSD)

Kaminuma *et al.* (27) did a literature survey of potential endocrine disruptors among different categories of chemicals that included synthetic estrogens for medicine, phytoestrogens, pesticides, industrial chemicals, environmental pollutants, and metals and their compounds. Then, these chemicals were listed categorically (http://www.nihs.go.jp/hse/endocrine-e/paradigm/paradigm.html). From these preliminary lists of endocrine disruptors, the Endocrine Disruptor Structure Database (EDSD) was developed. EDSD includes such entries as chemical name, CAS registry numbers, synonyms, physiological properties, and two- and three-dimensional structural data that are important for predicting chemical properties and QSAR. The three-dimensional structures of EDs, were calculated by using Molecular Mechanics 2 (MM2)(28).

3.5. Binding affinity database (BADB)

For each of the potential endocrine disrupting chemicals the mode of action was surveyed and the targets were identified. Although EDs stimulate target organisms by various modes of action, the three basic modes of action are (a) interaction with extracellular binding proteins, (b) interaction with enzyme systems that metabolize hormones, and (c) interaction with hormone receptors (29). A

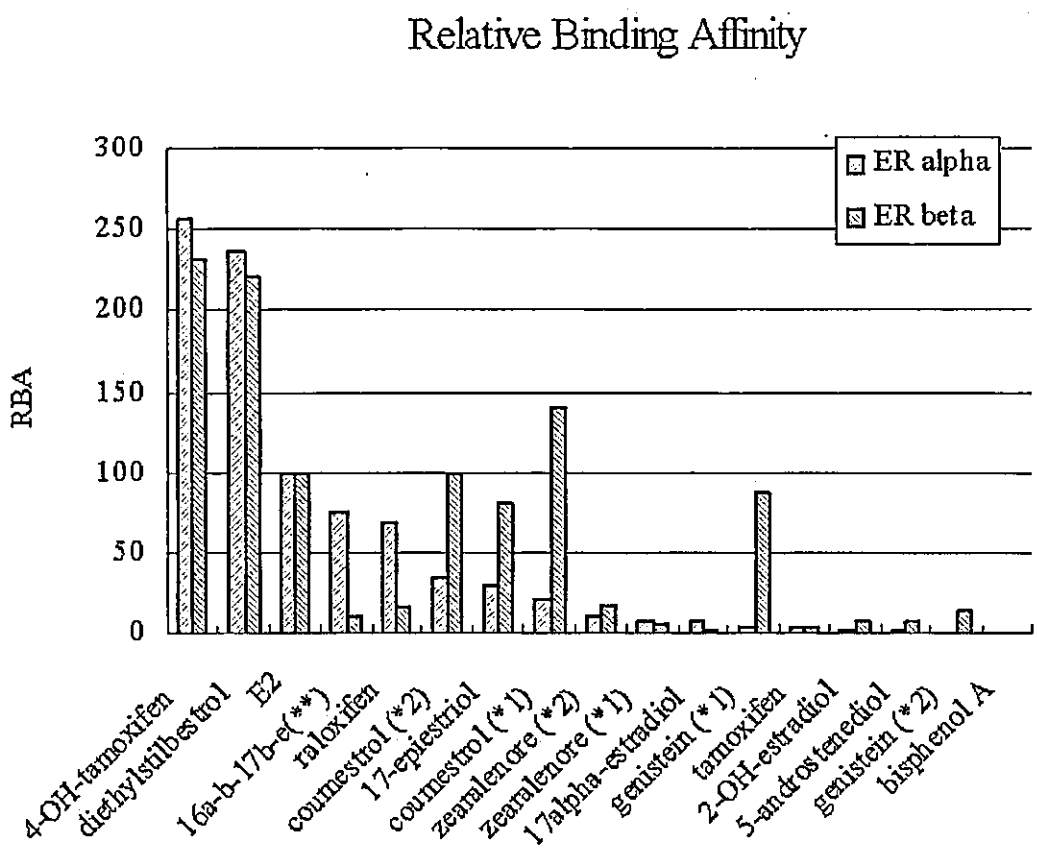

Figure 1. RBA between Estrogen Receptor alpha/beta and ligands.

molecular interaction database (BADB, http://moldb.nihs.go.jp/eddb/afdb/) that stores binding data of xenobiotic chemicals (ligands) and their target biomolecules has been developed (27). BADB stores experimental data for interactions of exogenous chemicals and biomolecules, which included 376 enzyme induction experiments and 742 competition binding experiments. These values were recalculated according to the definition of the relative binding affinity (RBA) proposed by Bolger *et al.* (30). Figure 1 shows the RBA between the estrogen receptor (ER) and ligands in BADB.

3.6. Receptor Database (RDB)

The Receptor Database (RDB, http://impact.nihs.go.jp/RDB.html) has been developed, which retrieves various receptor related data and provides hierarchical and graphical representation (31-32). RDB provides the receptor protein structure (amino acid sequences, secondary structure, and three-dimensional structure), DNA/ligand binding sites and binding affinity information, SNPs and cell signaling information, etc. The RDB aims to support structural biologists, not only for the examination of receptor-ligand binding but also for elucidation of the post- binding signal transduction pathway. At present RDB contains 1772 receptors, including endocrine-related receptors. The potential endocrine disrupting chemicals were stored in RDB and are included in the table "Steroid hormone/Aryl hydrocarbon receptors and possible endocrine disruptors" (see Figure 2).

3.7. Ligand-Receptor docking simulation

Structures of ligands are mostly obtained by theoretical calculation, for instance MM2; whereas target receptor structures of EDs, such as the nuclear receptor of estrogens, androgens, and thyroid hormone, were obtained from X-ray crystallography analyses (Protein Databank, http://www.rcsb.org/pdb/). When the steric structure of a target protein is available, one of the most effective methods to predict the binding strengths of ligands is docking simulation.

Itai *et al.* (33-36) developed a program called "ADAM" originally as a tool for docking simulation of a protein and ligand. The program has been used for rational drug design and investigations of biochemical reaction mechanism. Then they applied it for endocrine disruptors (37). In their method, empirical parameters were used for the computer calculation. Nakano *et al.* applied an approximation method of ab initio fragment molecular orbital calculation, which was developed by Kitaura *et al.* (38-39) for estimating the binding energies between the estrogen receptor and ligands. They got good correlation between the calculated binding energies and the RBA values. The advantage of this method is that empirical parameters need not be taken into account.

3.8. Cell signaling networks database (CSNDB)

Takai-Igarashi *et al.* (40-41) have developed the Cell Signaling Networks Database (CSNDB,

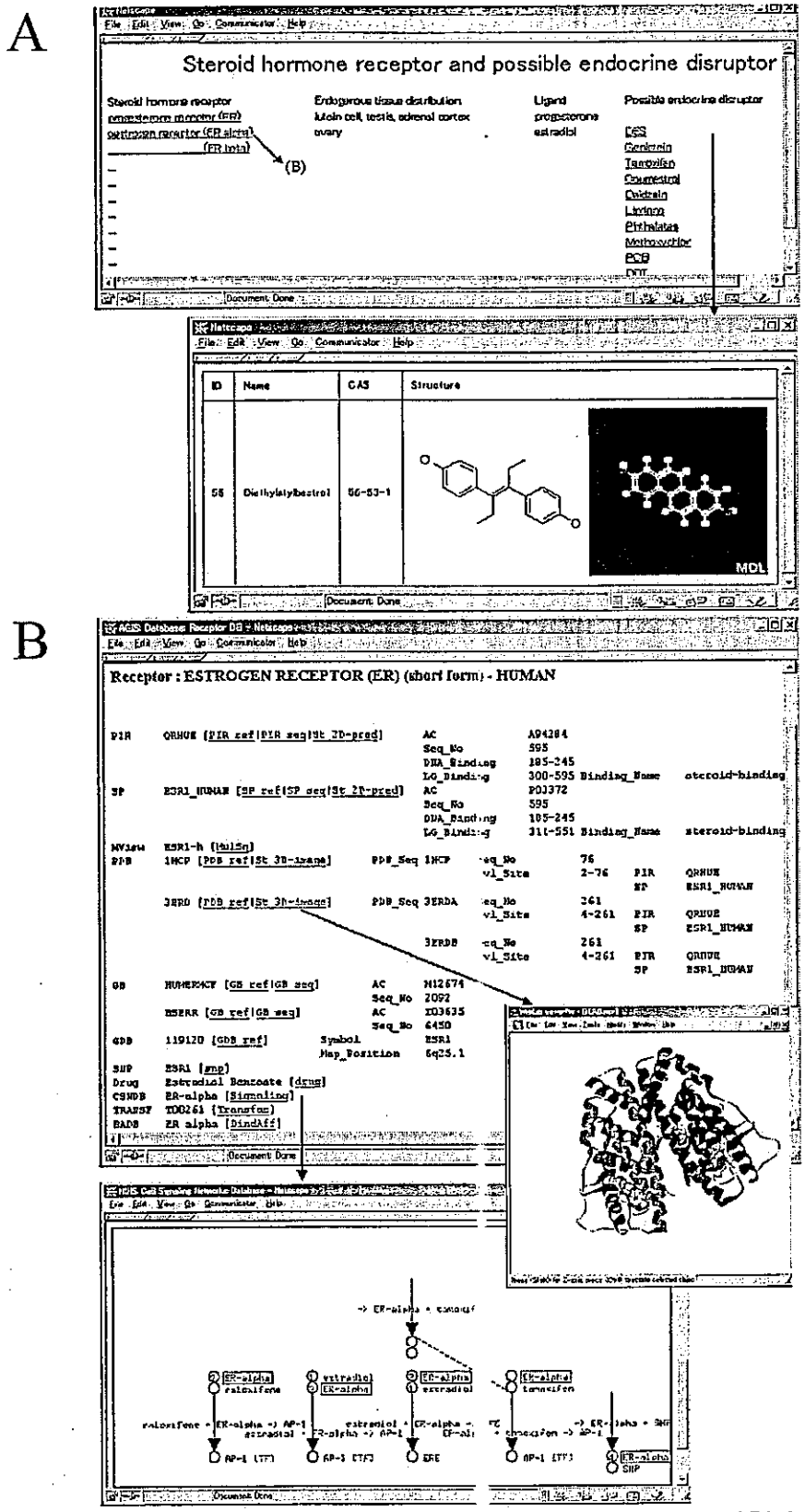

Figure 2. (A) Steroid hormone receptor and possible endocrine disrupt s. Example of ER alpha and Diethyl-stylbestrol (DES). (B) ER alpha. Three dimensional image and cell signaling.

http://geo.nihs.go.jp/csndb.html), for modeling and analyzing the signaling pathways. In CSNDB, only binary relationships between two arbitrary molecules are recorded. Then pathways connecting such molecules are retrieved by automatic graph drawing. A great advantage of CSNDB is the provision for data exchange with the transcription factor database TRANSFAC (42). As many cellular signals induce gene expression that evokes a second-phase cellular response, the integration of pathways for cellular signaling and transcription regulation will eventually cover all of the regulatory pathways occurring in cells. At present the CSNDB contains 1968 biomolecules and 1060 molecular interactions.

3.9. Statistical Approach

Haseman *et al.* examined data supporting the presence or absence of low-dose effects of endocrine disruptors in specific studies and then evaluated the likelihood and significance of these and/or other potential low dose effects for humans (43) (http://ntp-server.niehs.nih.gov/htdocs/liason/LowDosePeerFinalRpt.pdf). They re-evaluated 38 studies from 12 different investigators. The reevaluation focused primarily on the experimental design, data analysis, and interpretation of experimental results for each individual study within the context of its own experimental conditions rather than by comparisons of results across studies.

4. PERSPECTIVE

Endocrine disruptors are now a world-wide concern, and many domestic and international projects related to them are ongoing. Although there had been only a limited contribution of theoretical works until now, the needs and advanced anticipation of them are growing. The endocrine-related complex database system might not only be relevant for predicting cellular responses to exogenous hormonal chemicals but also be useful for designing drugs that interact or control endocrine systems. As examples of such drugs, tamoxifen and raloxifene are designer estrogens, now being used as anti-cancer drugs (44-45). It is expected that future comprehensive investigations will include docking simulations and database complex systems that cover cell signaling information and gene expression (46).

5. ACKNOWLEDGEMENT

The author thanks Dr. T. Kaminuma and DCBI members in NIHS for useful discussions and encouragement.

6. REFERENCES

1. Colborn, T., F. S. vom Saal & A. M. Soto: Developmental effects of endocrine-disrupting chemicals in wildlife and humans. *Environ Health Perspec* 101, 378-384 (1993)
2. Colborn, T.: Environmental estrogens: health implications for humans and wildlife. *Environ Health Perspect* 103, Suppl. 7, 135-136 (1995)
3. Colborn, T., D. Dumanoski & J. P. Myers: Our Stolen Future. A Dutton Book, New York, (1996)
4. Danzo, B. J., The effects of environmental estrogens and reproduction. *Cell Mol Life Sci* 54, 1249-1264 (1998)
5. Daston, G. P., J. W. Gooch, W. J. Breslin, D. L. Shuey, A. I. Nikiforov, T. A. Fico, & J. W. Gorsuch: Environmental estrogens and reproductive health: A discussion of the human and environmental data. *Reprod Toxicol* 11, 465-481 (1997)
6. Haseman, J. K., A. J. Bailer, R. I. Kodell, R. Morris & K. Portiew: Statistical issues in the analysis of low-dose endocrine disruptor data. *Toxicol Sci* 61, 201-210 (2001)
7. Hutchinson, T. H., R. Brown, K. E. Brugger, P. M. Campbell, M. Hol, R. Lange, P. McCahon, L. J. Tattersfield & R. van Egmond: Ecological risk assessment of endocrine disruptor. *Environ Health Perspect* 108, 1007-1014 (2000)
8. Kavlock , R. J., G. P. Daston, C. DeRosa, P. Fenner-Crisp, L. E. Gray, S. Kaattari, G. Lucier, M. Luster, M. J. Mac, C. Maczka, R. Miller, J. Moore, R. Rolland, G. Scott, D. M. Sheehan, T. Sinks & H. A. Tilson: Research needs for the risk assessment of health and environmental effects of endocrine disruptors: A report of the U.S. EPA-sponsored workshop. *Environ Health Perspect* 104, *Suppl* 4, 715-740 (1996)
9. Kavlock, R. J.: Overview of endocrine disruptor research activity in the United States. *Chemosphere* 39, 1227-1236 (1999)
10. Nilsson, R.: Endocrine modulators in the food chain and environment. *Toxicol Phthol* 28, 420-431 (2000)
11. Olea, N., P. Pazos & J. Exposito: Inadvertent exposure to xenoestrogens. *Eur J Cancer Prev Suppl* 1, S17-23 (1998)
12. Spearow, J. L. & M. Barkley: Reassessment of models used to test xenobiotics for oestrogenic potency is overdue. *Hum Reprod* 16, 1027-1029 (2001)
13. Sonnenschein, C. & Soto, A. M., An updated review of environmental estrogen and androgen mimics and antagonists. J. Steroid Biochem. *Mol Biol* 65, 143-150 (1998)
14. Toppari, J., J. C. Larsen, P. Chrisriansen, A. Giwercman, P. Grandjean, L. J. Guillette, Jr., B. Jegou, T. K. Jensen, P. Jouannet, N. Keiding, H. Leffers, J. A. McLachlan, O.Meyer, J. Muller, E. R.-D. Rajpert-De Meyts, T. Scheike, R. Sharpe, J. Sumpter & N. E. Skakkeboek: Male reproductive health and environmental xenoestrogens. *Environ Health Perspect* 104, *Suppl* 4, 741-803 (1996)
15. Crisp, T. M., E. D. Clegg, R. L. Cooper, W. P. Wood, D. G. Anderson, K. P. Baetcke, J. L. Hoffmann, M. S. Morrow,
D. J. Rodiew, J. E. Schaeffer, L. W. Touart, M. G. Zeeman & Y. M. Patel: Environmental Endocrine Disruption: An Effects Assessment and Analysis. *Environ Health Perspect* 106, *Suppl* 1, 11-56 (1998)
16. Hansch, C. & A. Leo: Exploring QSAR-Fundamentals and applications in chemistry and biology. American Chemical Society, Washington, DC, (1995)
17. Shi, L. M. , W. Tong, H. Fang, R. Perkins, J. Wu, M. Tu, R. Blair, W. Branham, C. Waller & D. Sheehan: An integrated "Four-Phase" approach for priority setting of endocrine disruptors-Phase I and II for prediction of potential estrogenic endocrine disruptor. *SAR QSAR Envir Res.* 2000, submitted for publication.

18. Lipinski, C. A, F. Lombardo, B. W. Dominy & P. J. Feeney: Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. *Advanced Drug Delivery Reviews* 23, 3-25 (1997)
19. Tong, W., R. Perkins, R. Strelitz, E. R. Collantes, S. Keenan, W. J. Welsh, W. S. Branham & D. M. Sheehan: Quantitative structure-activity relationships (QSAR) for estrogen binding to the estrogen receptor: predictions across species. *Environ Health Perspect* 105, 1116-1124 (1997)
20. Tong, W., R. Perkins, L. Xing, W. J. Welsh & D. M. Sheehan: QSAR models for binding of estrogenic compound to estrogen receptor alpha and beta subtypes. *Endocrinology* 138, 4022-4025 (1997)
21. Tong, W., D. R. Lowis, R. Perkins, Y. Chen, W. J. Welsh, D. W. Goddette, T. W. Heritage & D. M. Sheehan: Evaluation of quatitative structure-activity relationship methods for large-scale prediction of chemicals binding to the estrogen receptor. *J Chem Inf Comput Sci* 38, 669-677 (1998)
22. Koyano, K., T. Nakano & T. Kaminuma: Electrostatic Potentials and CoMFA Analysis of Toxicity of Dioxins. *Chem-Bio Informatics Journal* 1, 60-72 (2001)
23. Waller, C. L., T.I. Oprea, K. Chae, H. K. Park, K.S. Korach, S. C. Laws, T. E. Wiese, W. R. Kelce & L.E. Gray, Jr.: Ligand-based identification of environmental estrogens. *Chem. Res. Toxicol* 9, 1240-1248 (1996)
24. Xing, L., W. J. Welsh, W. Tong, R. Perkins & D. M. Sheehan: Comparison of estrogen receptor alpha and beta subtypes based on comparative molecular field analysis (CoMFA). *SAR QSAR Envir Res* 10,215-237 (1999)
25. Sadler, B. R., S. J. Cho, K. S. Ishaq, K. Chae & K. S. Korach: Three-dimensional quantitative structure-activity relationship study of nonsteroidal estrogen receptor ligands using the comparative molecular field analysis cross - validated r2-guided region selection approach. *J Med Chem* 41, 2261-2267 (1998)
26. Shi, L. M., H. Fang, W. Tong, J.Wu, R. Perkins, R. M. Blair, W. S., Branham, S. L. Dial C. L. Moland & D. M. Sheehan, QSAR Models Using a Large Diverse Set of Estrogens. *J Chem Inf Comput Sci* 41, 186-195 (2001)
27. Kaminuma, T., T. Takai-Igarashi, T. Nakano & K. Nakata: Modeling of signal pathways for endocrine disruptors. *BioSystems* 55, 23-31 (2000)
28. Allinger, N. L.: Conformational Analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *J Am Chem Soc* 99, 8127-8134 (1977)
29. Cheek, A. O., P. M. Vonier, E. Oberdoerster, B. C. Burow & J. A. McLachlan: Environmental Signaling: A Biological Context for Endocrine Disruption. *Environ Health Perspect* 106, Suppl. 1, 5-10 (1998)
30. Bolger, R., T. E. Wiese, K. Ervin, S. Nestich & W. Checovich: *Environ Health Perspect* 106, 551-557 (1998)
31. Nakata, K., T. Takai, & T. Kaminuma: Development of A Receptor Database (RDB): Application to the endocrine disruptor problem. *Bioinformatics* 15, 544-552 (1999)
32. Nakata, K., T. Takai-Igarashi, T. Nakano & T. Kaminuma: Extension of the Receptor Database (RDB). *Chem-Bio Informatics Journal* 3, 115-119 (2001)
33. Mizutani, M. Y., N. Tomioka & A. Itai: Rational automatic search method for stable docking models of protein and ligand. *J Mol Biol* 243, 310-326 (1994)
34. Tomioka, N. & A. Itai: GREEN: a program package for docking studies in rational drug design. *J Comput Aided Mol Des* 8, 347-366 (1994)
35. Toyoda, T., R.K. Brobey, G. Sano, T. Horii, N. Tomioka & A. Itai: Lead discovery of inhibitors of the dihydrofolate reductase domain of Plasmodium falciparum dihydrofolate reductase-thymidylate synthase. *Bioche Biophys Res Commun* 235, 515-519 (1997)
36. Iwata, Y., M. Arisawa, R. Hamada, Y. Kita, M. Y. Mizutani, N. Tomioka, A. Itai & S. Miyamoto: Discovery of novel aldose reductase inhibitors using a protein structure-based approach: 3D-database search followed by design and synthesis. *J Med Chem* 44, 1718-1728 (2001)
37. Endo, Y., T. Iijima, Y. Yamakoshi, H. Fukasawa, C. Miyaura, M. Inada, A. Kubo & A. Itai: Potent estrogen agonists based on carborane as a hydrophobic skeletal structure. A new medicinal applicationof boron clusters. *Chem Biol* 8, 341-355 (2001)
38. Kitaura, K., E. Ikeo, T. Asada, T. Nakano & M. Uebayasi: Fragment molecular orbital method: An approximate computational method for large molecules. *Chem Phys Lett* 313, 701-706 (1999)
39. Nakano, T., T. Kaminuma, T. Sato, Y. Akiyama, M. Uebayasi & K. Kitaura: Fragment molecular orbital: Application to polypeptides. *Chemical Physics Letters* 318, 614-618 (2000)
1. Takai-Igarashi T, Y. Nadaoka & T. Kaminuma: (1998) A database for cell signaling networks. *J Comp Biol* 5, 747-754.
41. Takai-Igarashi, T. & T. Kaminuma: A Pathway Finding System for the Cell Signaling Networks Database, *In Silico Biology* 1, 129-146 (1998)
42. Wingender, E., X. Chen, E. Fricke, R. Geffers, R. Hehl, I. Liebich, M. Krull, V. Matys, H. Michael, R. Ohnhauser, M. Prus, F. Schachere, S. Thiele & S. Urbach: The TRANSFAC system on gene expression regulation. *Nucleic Acids Res* 29, 281-283 (2001)
43. Haseman, J. K., A. J. Bailer, R. L. Kodell, R. Morris & K. Portier: FORUM: Statistical Issues in the Analysis of Low-Dose Endocrine Disruptor Data. *Toxicological Sciences* 61, 201-210 (2001)
44. Jordan V. C.: Designer Estrogens. *Scientific American* 36-40 (1998)
45. Jeng, M. H., S. Y. Jiang & V. C. Jordan: Paradoxical regulation of estrogen-dependent growth factor gene expression in estrogen receptor (ER)-negative human breast cancer cells stably expressing ER. *Cancer Lett* 82, 123-128 (1994)
46. T. Kaminuma, K. Nakata, T. Nakano & T. Takai-Igarashi: Pharmacoinformatics Infrastructure for Genome-based Personalized Medicine. *Chem-Bio Informatics Journal* 1, 1-17 (2001)

Key Words: Endocrine Disruptor, Receptor, Signal Pathway, Binding Affinity, Docking Analysis, Review

Send correspondence to: Dr Kotoko Nakata, Division of Chem-Bio Informatics, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan, Tel: 81-3-3700-9572, Fax: 81-3-5717-7180, E-mail: nakata@nihs.go.jp