Modeling of signaling pathways for endocrine disruptors

Tsuguchika Kaminuma *, Takako Takai-Igarashi, Tatsuya Nakano, Kotoko Nakata

Division of Chem-Bio Informatics, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

Abstract

The so called endocrine disruptors have become an important working hypothesis for a wide range of toxicology researchers. This hypothesis has also attracted those who have worked on designer estrogens or selective estrogen receptor modulators. Already numbers of substances have been identified as such chemicals, but there remain a large number of chemicals waiting to be tested for their endocrine modulating capabilities. Because of the time and costs required for wet lab tests, it is unrealistic to apply these kinds of tests to all such suspicious or probable chemicals. Thus some theoretical methods must be developed for this purpose. However the conventional QSAR (quantitative structure activity relationships) approach is of limited relevance to this problem, because these methods do not take detailed mechanisms of molecular interactions in biological systems into account. Thus we have developed a database complex system that enables one to trace molecular interactions triggered by interaction of receptors with xenobiotic chemicals. The main components of this database complex are a potential endocrine disruptor database, a receptor database, a cell signaling networks database, a transcription factor database, and an affinity binding database based on modes of actions. © 2000 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Endocrine disruptor; Database integration; Computer-based screening

1. Introduction

Recently the potential hazardous effects that estrogen-like and androgen-like chemicals and other environmental hormone-like chemicals may have on wildlife and human health have attracted much attention both the scientific community, particularly environmental toxicology, and the general public. A hypothesis has been proposed by several researchers that these chemicals mimic natural hormones and modulate endocrine systems, Colborn et al. (1996). These chemicals are now called endocrine disruptors, endocrine disrupting chemicals, or endocrine modulators.

A large number of projects have been proposed to study this hypothesis, Crisp et al. (1998). These studies are focussed on (1) what is the mechanism by which the endocrine disruptors modulate normal endocrine systems, and (2) among enormous number of existing chemicals how we can screen such endocrine disruptors efficiently. It is obvious that conventional toxicological tests using animals such as fish, rats, or mice are not suitable to screen out endocrine disruptors from the so called
existing chemicals, because the number of the latter is of the order of one hundred thousands.

It is thus 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 tests or the so called high through put screenings (HTPS) can be applicable (http://www.epa.gov/opptintr/opptendo/).

Some feasibility studies have been carried out from this viewpoint, but it was pointed out, that conventional QSAR (quantitative structure activity relationships) methods are not sufficient for such purposes, Tattersfield et al. (1997). The reason is that the well known ligand receptor binding model is not enough to characterize the function of suspected chemicals. For example, two different ligands that bind to a same receptor may not result the same phenomena (endpoints). The effect of a ligand binds to a receptor differs at different tissues. Moreover these approaches do not consider the so called cross talk among different signaling pathways. Such phenomena can only be handled by taking the cellular signaling systems and gene expression mechanisms into explicit consideration.

We had already developed a cell signaling networks database (CSNDB), Takai-Igarashi et al. (1998), Takai-Igarashi and Kaminuma (1999), which stores causal signal transduction networks from receptors to gene expressions in human cells. We have tried to link this database to other cellular molecular interaction databases such as metabolic pathway databases and transcription factor databases like TRANSFAC, Heinemeyer et al. (1999). We have then extended our CSNDB to link it to other databases in order to describe signaling pathways from the binding site with external (exogenous) chemicals to the resultant phenomena of interests (endpoints) such as developmental abnormalities, carcinogenesis, and other pathological effects. These bindings are gate-point(s) for the external signals in the biological systems.

In this paper we present the prototype of an integrated databases and knowledge-bases complex of chemical substances and of biomolecules that can describe internal signaling evoked by endocrine disruptors from gate-points to the endpoints.

2. Conceptual framework of the system

Fig. 1 shows the steps to produce lists of probable endocrine disruptors, related chemicals and reference chemicals. We first did a literature survey of potential endocrine disruptors on different categories of chemicals that include synthetic estrogens for medicine, phytoestrogens, pesticides, industrial chemicals, environmental pollutants, and metals and their compounds. We then made lists of the chemicals per category and add regulatory and other relevant information (http://www.nihs.go.jp/hse/endocrine-e/paradigm/paradigm.html).

From these preliminary lists of endocrine disruptors a database was produced. This database has three dimensional structure data which is important for predicting chemical properties and QSAR.

For each of the potential endocrine disrupting chemicals the mode of action was surveyed, and their receptors were identified if possible. A molecular interaction (affinity binding) database that stores binding data of xenobiotic chemicals (ligands) and their target biomolecules (mainly receptors) has been developed. Those receptors or interacting biomolecules are the gate-points at which some signals will be evoked by these xenobiotic chemicals. These chemical signals will be transmitted or amplified by cellular signaling net-

Fig. 1. Making endocrine disruptor database by literature search.
works, will transcribe mRNA, will produce a protein, and will eventually effect the organism. The effects may be gene mutation, cell death, cancer, abnormal development, organism lethality, reproductive failure, etc. Such effects are called endpoints.

Our eventual goal is to develop a computer-based causal network system that enables one to trace the effect of exogenous endocrine disrupting chemicals on biological signaling systems at the cellular and organism levels. For that purpose we need complete a knowledge base and databases for bio-signaling systems. However at present such a network system is incomplete, for there only exists receptor databases, metabolic pathway databases, cell signaling pathway databases, transcription factor databases, etc independently. Among these we have developed a receptor database, Nakata et al. (1999), and cell signaling database, Takai-Igarashi et al. (1998), Takai-Igarashi and Kaminuma (1999), by ourselves. The emerging WWW technology allows us to link so far independent system into an integrated system complex. We thus started to link existing databases and knowledge bases into a complex system.

The WWW technology has also made excellent software available on the Internet. Molecular viewing tools such as Rasmol, Bradley (1994), Chemscape (http://www.mdli.com/) or Weblab (http://www.msi.com/) were embedded into our systems.

3. Endocrine disruptor database

From the categorical lists of chemicals we have produced a database. This database includes such entries as chemical name, CAS registry numbers, synonyms, physicochemical properties, and two-dimensional and three-dimensional structural data. The three-dimensional data were obtained by ab initio molecular orbital (MO) calculation using a commercial software Gaussian 94 (http://www.gaussian.com/). At present we have obtained optimized coordinates for 149 substances.

With appropriate viewing software we can generate three-dimensional images of chemicals. The three dimensional data are also important for 3D QSAR analysis such as CoMFA (http://tripos.com/).

4. Receptor database

The receptor database (RDB), Nakata et al. (1999), contains biomolecules that are found to be receptors. The definition of receptors changes in time. We have applied a sort of common sense definition. The receptor database stores the biomolecule names, their genes and sequences, spices, ligand etc. An object-oriented database called ACEDB (a Caenorhabditis elegans database), Thierry-Mieg and Durbin (1992), was used as the database management system. ACEDB is a database management system developed for C. elegans genome research. Necessary data were taken from internationally well known biological databases such as PIR, SwissProt, PDB, GenBank, and GDB. The system provides various viewing tools. By such tools one can highlight amino acid sequences for functional regions, such as DNA binding sites, ligand binding sites, and transmembrane regions. Present RDB contains 634 receptors.

5. Binding affinity database

This database stores experimental data for interaction of exogenous chemicals and biomolecules. Endocrine disruptors stimulate target organisms in various modes of actions. We consider the three basic modes of actions in our model; (1) interaction with extracellular binding proteins, (2) interaction with enzyme systems that metabolize hormones, and (3) interaction with hormone receptors, Cheek et al. (1998). Best examples are given by estrogen case.

It is known that availability of internal and external estrogenic signals to target tissues are modulated by extracellular binding proteins in the plasma. Natural hormone estradiol and synthetic estrogen, DES (Diethylstilbestrol) both bind to the estrogen receptors, but they also interact with serum transport proteins such as alpha-feto-protein, sex hormone binding globulin (SHBG),
Steroid hormones are synthesized from cholesterol by a series of lyase and cytochrome P450 reactions. Specific P450 isozymes convert progesterone to 17β-estradiol. Other enzyme systems such as reductases and transferases are also involved in inactivation and elimination of steroids. However, these enzyme systems metabolize estrogenic xenobiotic chemicals. Conversely, xenobiotic estrogenic chemicals up-regulate the P450 isozyme expression, which leads to increased metabolism of these substrates and endogenous steroids and thus modulates the endocrine system.

However, the main mode of action of environmental estrogens, that is, estrogenic xenobiotic chemicals, is interaction with estrogen receptors (ERs). Estrogen receptors have common structure with other hormone receptors such as androgen receptors (ARs) and thyroid hormone receptors (TRs). Together they are classified into nuclear receptors. There also exist at least two types of estrogen receptors α and β. Affinities are different in different combinations of ligands (hormonal chemicals) and their receptors. Another important endocrine disruptor, dioxin, binds to a different receptor called aryl hydrocarbon receptor (AhR).

The fourth probable mode of action is direct interaction with cell signaling pathways without mediation by receptor bindings. The result may increase calcium concentration, affect the sodium channel and calcium homeostasis, and activate some kinases. We excluded this last case and only collected the first three modes of action binding affinity data.

The prototype of this system has been implemented on a stand-alone PC using a database management system ACCESS and is presently available through the Internet (Fig. 2).

6. Ligand-receptor docking studies

With Endocrine Structural Database, Receptor Database, and Binding Affinity Databases one can have a fairly wide view on how xenobiotic endocrine modulating chemicals interact with their receptor molecules. The databases contain three-dimensional structural data of the interacting molecules. Structures of the ligands are mostly obtained by theoretical (ab initio molecular orbital) calculation, while receptor structures were obtained from X-ray crystallography analyses. These data are relevant for the so-called ‘docking study’ and QSAR studies.

At present we are studying three typical cases: estrogen-like chemicals and the estrogen receptors, dioxins which interact with Ah receptors, and the interaction of hormone transport proteins. For these studies we use various computational tools such as ab initio MO methods, Discover/Insight II, etc. Fig. 3 shows the relation between binding affinity database and modeling studies.

7. Cell signaling networks database

We have developed a database, CSNDB, for modeling and analyzing the signaling pathways, Takai-Igarashi et al. (1998), Takai-Igarashi and Kaminuma (1999). Using CSNDB one can retrieve pathways connecting two arbitrary molecules. In CSNDB, no pathways are pre-encoded but only binary relationships are recorded in the database. So it is the user’s request that creates pathways. Pathways are subsets filtered from the whole networks. This filtration is based on our consideration that any cellular response can be subsets of networks selected in response to various external and internal conditions of the cells. This pathway representation enabled one to overcome the difficult problem that there are no consensus grouping of signaling pathways. In the case of metabolic pathways, we can easily refer to the consensus grouping, such as chlorophyll synthesis, fatty acid desaturation, and tyrosine synthesis. This is one of major distinctions between cell signaling and metabolic pathways.

The pathway finding system in CSNDB enables queries about pathways. The users can retrieve pathways around arbitrary molecules or between two arbitrary molecules. The retrieved pathways are represented as graphs, which are connected or disconnected and cyclic or acyclic directed graph
with labels. Particularly from biological viewpoint, the cyclic graph corresponds to the feedback regulation that is an important mechanism for balancing a whole system. Since the graph is produced as the hypertext, a clicking on a node or an arrow calls a database entry. The arrows are liked to entries of molecular interactions and the nodes are linked to entries of biological molecules. The entries of biological molecules contain structural and functional information and linkages to the external resources such as GenBank and SwissProt through the Internet. In CSNDB, all the recorded information is abstracted from published articles that are annotated in every entry.

CSNDB consists of three modules; the database management system, the inference engine, and the automatic graph drawer. We use object-oriented database system ACEDB, Thierry-Mieg and Durbin (1992), for the database management system, rule-based production system CLIPS, Riley et al. (1988), for the inference engine, and the
Letovsky’s algorithm, Letovsky (1995) for the automatic graph drawing. All the modules are integrated with the user interface developed in the programming language C. From the standpoint of the database management system, CSNDB is a novel application of object-oriented technology to biological modeling. From the standpoint of production system, CSNDB works as real biologist in inferring possible pathways using biological principles as rules. The inference engine of CSNDB is programmed in LISP syntax. It executes data-driven forwarded-chaining inference. It finds connections between binary relationships when a transmitter of a relationship and a receiver of another relationship are identical. Automatic graph drawing made us free from manually drawing of graphs and enabled frequent updating (Fig. 4).

A great advance of CSNDB is the data-exchange with the transcription database, TRANSFAC, Heinemeyer et al. (1999). Obviously, many cellular signals induce some gene expressions that evoke second-phase cellular responses. The integration of pathways in cellular signaling and transcription regulations will cover whole the regulatory pathways occurred in cells. For the linkage of two databases, we developed a translation table linking both entries. Through mediation of the table, one can obtain possibly up- or down-regulated gene expressions from cell signaling pathways.

The present CSNDB contains 1968 biomolecules, 1060 molecular interactions, 968 references, and 117 connections with TRANSFAC entries. We are planning to link some metabolic pathway databases to CSNDB.
8. From gatepoint(s) to endpoints

Endpoints are the biological phenomena or simply makers by which researchers identify positive effects of chemical agents. Since the mechanisms of endocrine modulation due to xenobiotic chemicals have not been totally understood, we do not exactly know proper endpoints. Nevertheless wide varieties of endpoints were proposed by experimental researchers.

From purely theoretical viewpoint, it would be possible some day to identify signal pathways to the reactions of biological systems at molecular base. Of course cell–cell interactions also play very important roles here. However our current knowledge is far from such viewpoints. We simply take these pictures as conceptual frame for the future studies (Fig. 5).

9. Linkage of databases

All of our component databases share at least two common entries; a chemical substance name and a biological molecular name. We assume that chemical substances which we take into the databases have CAS registry numbers. We then created the index database which is nothing but a big relational table composed of name of chemicals and the databases contain them. Indexing with CAS registry numbers tends to be a standard method in integrating chemical databases. However we need indexing of relationships of chemicals and biomolecules although such an index does not yet exist. This is the reason that we created this index database. We consider the index database is important, because it can bridge between toxic chemicals and biological responses through the mediation of biological pathways.

Technically, each database has interface to WWW, and the interface (Common Gateway Interface) program converts queries to WWW to queries to the database. This enables one to cross-search through the hyper-linked database by his client machine browser. Fig. 6 shows a schematic diagram of our hyper-linked database complex.

10. Discussion

The system we have implemented 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. The good examples of such drugs, designer estrogens, are two new anti-cancer drugs tamoxifen and raloxifene, Jordan (1998).
Since expertise of cellular and organism signaling pathways and various assay data for endocrine disrupting chemicals will rapidly be accumulated in the near future, we can easily increase and refine data and knowledge contents of our databases. Our complex system has enough framework and flexibility to add rapidly growing new experimental facts. Our complex system is also very flexible for adding new molecular databases either on xenobiotic chemicals or biomolecules.

Current lack of knowledge on signal pathway at organism level and the fact that many assay experiments are still on going make it difficult to develop reliable computer system to predict organism reaction to endocrine disrupting chemicals at organism level. Our current system is only relevant to predict cellular response to exogenous hormones. Predicting higher level responses that are characteristic to multi-cellular organisms such as cancer or reproductive abnormalities are our next targets. The so called endpoints relate to these levels of responses. However we need some models which explicitly take cell–cell interactions into account. As a new approach we have proposed a more realistic modeling of bio-system called three-world model, Kanimuma et al. (1998). To combine the complex system introduced in this paper with the three-world model system is the next target of our studies.

11. Conclusion

The so called endocrine disruptors or endocrine disrupting (modulating) chemicals have become a very important working hypothesis for wide range of toxicological research. They are also important for drug design particularly designer estrogens such as tamoxifen and raloxifene.

We have developed data and knowledge base complex for endocrine receptors in order to support experimental endocrine disruptor research. The main components of this database complex are a potential endocrine disruptor database, a receptor database, a cell signaling networks database, a transcription factor database, and affinity binding database based on modes of actions. The index database and WWW technology link these databases. This complex system provides infrastructure for describing cellular and organism responds to xenobiotic hormone-like chemicals in terms of signaling pathways. Such an infrastructure is complementary to conventional QSAR approach for theoretically predicting biological responses to endocrine disruptors. Only component (modular) systems are now available on the Internet (http://impact.nih.go.jp/RDB.html, http://geo.nih.go.jp/csndb.html, http://hse.nih.go.jp/eddb/afdb.html). The full prototype system will soon be put on the Internet.

Acknowledgements

We thank Drs Miyako Yamamoto and Michiko Yamamoto, C. Ohtake, S. Hasegawa, H. Taki, and M. Yukawa for making heavy literature survey and listing potential endocrine disruptors. We also thank to Dr K. Koyano who carried out molecular modeling calculations for these chemicals.

References