**DEVELOPMENT OF A RECEPTOR DATABASE.**

[KOTOKO NAKATA](http://www.bionet.nsc.ru/meeting/bgrs/author/nakata.html), [TAKAKO TAKAI](http://www.bionet.nsc.ru/meeting/bgrs/author/takai.html), [TSUGUCHIKA KAMINUMA](http://www.bionet.nsc.ru/meeting/bgrs/author/tsuguchika.html)

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

[Keywords](http://www.bionet.nsc.ru/meeting/bgrs/keywords_index.html): receptor, database, protein, DNA binding sites, three dimensional structure, ligand binding sites, signal transdiction

A cell responds to a particular extracellular signaling molecule depends on its having specific proteins, called receptors, that bind the signal molecule. Each receptor recognize an unique ligand or class of ligands, and is realized as the site of action of endogenous regulators whether called transmitters, hormones, growth controllers. These information regarding the structure of receptors and the site of action on the receptor are highly interesting in biologically, medically and pharmacologically. We collected such information as much as possible, and represent hierarchically using a database management system ACEDB (A Caenorhabditis elegans Data Base). The system was implemented on the UNIX workstation (IRIS, INDIGO 2).

ACEDB is an object oriented database management system, which has been developed as part of the Caenorhabditis elegans genome research. The system is a generalized genome database system, and can be used to create new database without the need for any reprogramming or in fact any sophisticated computer skills. We made perl programs which were combined to ACEDB so that the system can collect data items such as attributes of proteins from distributed data sources on the Internet and that the system provides various viewing tools effectively depend on different types of receptor data. Such sources include internationally standard biological databases such as the updated genetic database of PIR, Swiss Prot, PDB, GenBank, EMBL and GDB.

Providing information include the receptor protein entry name which links to PIR or Swiss Prot database, the sequence data, the secondary structure prediction results, the entry name for PDB database and the three dimensional structure images of molecules. In the sequence data representation, DNA binding sites, ligand binding sites, transmembrane positions and PDB matching positions are maked out by different colors. A perl program, which generate an input data to another analysis program and send it to outer Web sites, was added to the system to display the secondary structure prediction results. For the three dimensional structure image, we used a freeware program RASMOL. DNA information whose sequences are translated to the receptor protein are also providing, as for the entry name and the sequence data. As for a gene, it links to GDB database. For a cellular signal transduction of human, it links to CSNDB (Cell Signaling Networks Database).

The system has a browser interface so that the receptor database can be accessed via World Wide Web. The database may be useful for quick reference for ligand - membrane receptors, the secondary structure prediction of the receptor protein and the signal transduction in the cell. In April ‘98 version, one thousand proteins were included. The number of proteins which have three dimensional structural data was 20, and that have DNA binding sites or ligand binding site were 120. At present, we have insufficient data for DNA binding site, ligand binding site and three dimensional structure in the database. However, we expect that the receptor database become more useful for the basic research of drug design, with the increase of the experimental results.

As a particular application, we are interesting in modeling of opioid receptor structure, based on rhodopsin crystallographic structure. Using our receptor database, we get easily the secondary structure prediction results for receptor proteins and correspond them to multiple sequence alignments. Through a modeling of the ligand and receptor complex, we are trying to research leading a drug design which has no addiction liability nor other deleterious effects.

**References**

1. Richard Durbin and Jean Thierry Mieg: (1991-). A C. elegans Database. Documentation, code and data available from anonymous FTP servers at <ftp://lirmm.lirmm.fr/>, <ftp://cele.mrc-lmb.cam.ac.uk/> and <ftp://ncbi.nlm.nih.gov/>
2. Kotoko Nakata, MasatakaHayakawa, Takako Igarashi and Tsuguchika Kaminuma: Building A Receptor Database. Genome Informatics (1996) <http://impact.nihs.go.jp/RDB.html>
3. Takako Igarashi and Tsuguchika Kaminuma: Development of Cell Signaling Networks Database, Pacific Symposium on Biocomputing '97, pp.187-197, (1997), World Scientific.<http://geo.nihs.go.jp/csndb.html>