Databases and ontologies

HaptenDB: a comprehensive database of haptens, carrier proteins and anti-hapten antibodies

Mahender Kumar Singh[†], Shilpy Srivastava[†], G.P.S. Raghava and Grish C. Varshney[*]

Institute of Microbial Technology, Chandigarh, India

ABSTRACT

The key requirement for successful immunochemical assay is the availability of antibodies with high specificity and desired affinity. Small molecules, when used as haptens, are not immunogenic. However, on conjugating with carrier molecule they elicit antibody response. The production of anti-hapten antibodies of desired specificity largely depends on the hapten design (preserving greatly the chemical structure and spatial conformation of target compound), selection of the appropriate carrier protein and the conjugation method. This manuscript describes a curated database HaptenDB, where information is collected from published literature and web resources. The current version of the database has 2021 entries for 1087 haptens and 25 carrier proteins, where each entry provides comprehensive details about (1) nature of the hapten, (2) 2D and 3D structures of haptens, (3) carrier proteins, (4) coupling method, (5) method of anti-hapten antibody production, (6) assay method (used for characterization) and (7) specificities of antibodies. The current version of HaptenDB covers a wide array of haptens including pesticides, herbicides, insecticides, drugs, vitamins, steroids, hormones, toxins, dyes, explosives, etc. It provides internal and external links to various databases/resources to obtain further information about the nature of haptens, carriers and respective antibodies. For structure similarity comparison of haptens, the database also integrates tools like JME Editor and JMOL for sketching, displaying and manipulating hapten 2D/3D structures online. So the database would be of great help in identifying functional group(s) in smaller molecules using antibodies as well as for the development of immunodiagnostics/therapeutics by providing data and procedures available so far for the generation of specific or cross-reactive antibodies.

Availability: HaptenDB is available on http://www.imtech.res.in/ raghava/haptendb/ and http://bioinformatics.uams.edu/raghava/ haptendb/ (Mirror site).

Contact: grish@imtech.res.in

There is an exponential growth in data and databases in biology particularly in sequence databases (see Database Issue of NAR 2005). These databases comprise one of the important components in biological research, specifically in bioinformatics, where one can extract useful information from these well-organized databases (Brusic *et al*., 2000). Immunology is an important area where databases assist the immunologists in understanding the immune system and in searching potential drug/vaccine candidates (Yau *et al*., 2003). Though there are a number of immunological databases on protein sequences and peptides (epitopes) (KABAT, IMGT, FIMM, MHCBN, BCIPEP and AntiJen 2.0) (Saha *et al*., 2005; Bhasin *et al*., 2003; McSparron *et al*., 2003; Johnson and Wu, 2000; Ruiz *et al*., 2000; Schönbach *et al*., 2000), there is no database on haptens. The haptens are low molecular weight molecules, which do not elicit immune response until and unless conjugated with an immunogenic carrier, such as protein (Goodrow *et al*., 1990; Singh *et al*., 2004). Once the antibody is formed, it can bind to the hapten. Since these molecules are so small (e.g. pesticides, fungicides, drugs, hormones, toxins, synthetic peptides, etc.) they do not directly induce antibody response. Biologists have found ways to raise antibodies against these haptens by conjugating them with antigenic macromolecular proteins (carrier in this context). Antibodies thus generated can prove useful in serology, drug delivery, development of immunodiagnostic kits or biosensors (Abad *et al*., 2001; Moreno *et al*., 2001; Choi *et al*., 1999). A large number of research articles have been published on haptens, which describe different aspects of the hapten molecules (Stix, 2001; Goodrow and Hammock, 1998; Hall *et al*., 1990), but this information has not been compiled suitably in the literature. Moreover, it is difficult to draw any conclusion with regard to hapten properties, efficacy of different carriers towards antibody generation and the specificity of antibodies. In order to overcome these problems, for first time we have made an attempt to collect information about important haptens that are found in biological systems and the environment around us.

In this manuscript we describe a database called HaptenDB developed to provide comprehensive information about the hapten molecules, ways to raise antibodies against particular group of haptens, specificity and cross-reactivity of raised antibody with related haptens, use of antibodies in constructing cost-effective and simple detection kits (Fig. 1). The first version of the database contains 2021 entries for antibodies either raised against haptens or cross-reactivity of an antibody raised against one hapten and checked with the other related haptens. The database also contains 1087 entries for physical, chemical and biological aspects of different hapten molecules used in this study. All the data for HaptenDB are collected from published literature. A simple search in the HaptenDB database against a hapten molecule will display a summary of all records where either an antibody is raised against that hapten or some other antibody, raised against somewhat different hapten but having significant cross-reactivity with the searched hapten. Individual records of interest can then be viewed in greater detail. Alternatively, a more specific advanced search

[*]To whom the correspondence should be addressed.
[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.
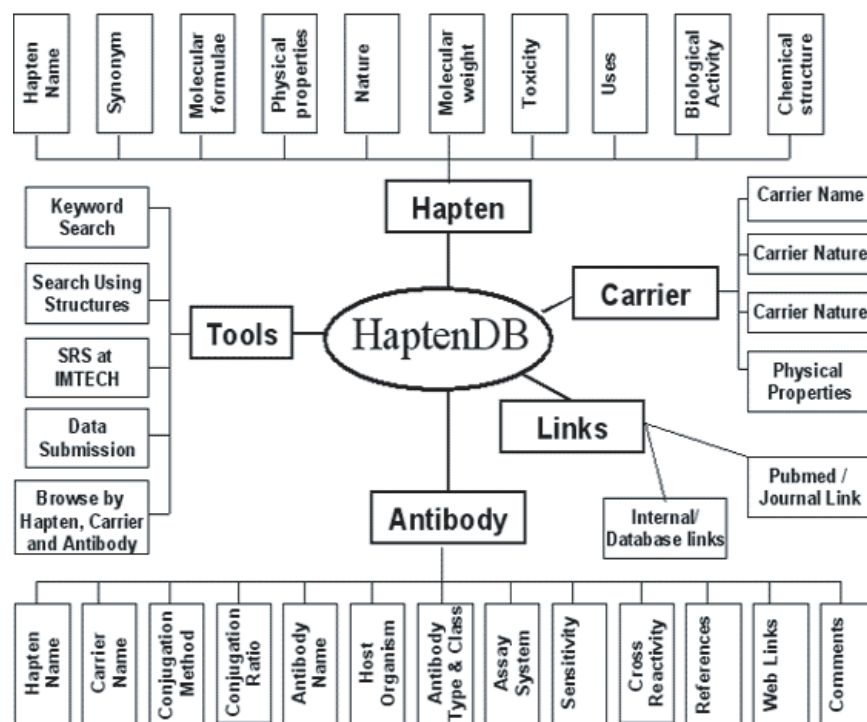
Fig. 1. Architecture of database HaptenDB.

will give effective search on database according to the user's requirement.

Every entry in the database contains detailed information about the hapten, the carrier (used to prepare the immunizing conjugate) and the antibodies (elicited in response to the immunization) along with the assay methods and their sensitivity towards hapten detection. The information about the hapten includes (1) its common and chemical name, (2) molecular mass, physical and chemical properties and (3) biological importance. In addition, we have generated 2D structures of most of the haptens in the database, based on the information obtained from the literature and web resources. These 2D structures were further converted to 3D structures using CORINA web server (http://www2.chemie.uni-erlangen.de/software/corina/index.html). Thus, 2D and 3D structures of most of the haptens are also available in the database. The details of the carrier molecule include the nature of the carrier, its molecular mass, sequence, origin and physical properties. The details of the antibody production included are (1) hapten-carrier conjugation protocol (many times the original hapten is modified by introduction of spacer arms so that it is better recognized), (2) conjugation ratio of the hapten and the carrier molecules, (3) the organism and the protocol used for the immunization, (4) properties of the antibodies/antiserum, (5) assay method and its merits, (6) sensitivity and the specificity of the system, (7) cross-reactivity (with related compounds) and (8) reference along with web links.

HaptenDB provides online web tools that allow users to retrieve and analyze the data that include (1) tools for searching database using keywords with many options and (2) browsing tool to browse the database on hapten name, carrier protein and antibody. In addition, it provides an internal hyperlink to display detailed information and an external hyperlink to other database like PubMed. One of the important uses of the HaptenDB is to compare the structure similarities. Therefore, it allows the user to search similar structures. In order to provide structure similarity search against haptens in our database, we generated 2D structures of large number of haptens (>95%) in standard format based on information in the literature. We integrated the jcsearch program in HaptenDB for searching similar structures. The jcsearch program (http://www.chemaxon.com/jchem/doc/user/Jcsearch.html) is a command-line interface of the JChem chemical structure search (http://www.jchem.com/). It allows to perform substructure, exact and perfect searches on the specified query. We integrated JMOL (http://jmol.sourceforge.net/) in HaptenDB that allows one to display and manipulate 3D structures of the haptens. We also incorporated JME molecular editor in HaptenDB, which allows the user to draw/edit molecules and to depict molecules directly in the HTML page. This editor allows one to sketch structures online and to generate SMILES or MDL mol file of created structures (http://www.molinspiration.com/jme/index.html). HaptenDB has integrated the BABEL program that is designed to interconvert a number of file formats currently used in molecular modeling. This allows the user to submit query structures in any standard format (MOL, MOL2, PDB, SMI, etc.) for performing similarity search using the jcsearch program.

The HaptenDB allows the users to submit their haptens and related antibodies online. In order to maintain the quality, we will examine the submitted data before integrating it in the database. Our team is also searching and adding more entries in the database from published literature. In order to maintain the consistency we will release HaptenDB database quarterly. The HaptenDB web server was developed in a UNIX environment on the SUN server 420R in Solaris 7.0. This server is designed to provide easy access to the user, based upon a set of simple graphical user interface forms.

Methods, for searching the database and displaying the selected objects, were built with a combination of JavaScripts and CGI-scripts in PERL 5.4. In order to provide search on any field of database and to maintain standards we have also launched the SRS version of HaptenDB (http://www.imtech.res.in/srs/).

In conclusion, this is the first attempt to provide comprehensive information about haptens and anti-hapten antibodies from a single source. On one hand the database can be useful in dissecting the functional region(s) of hapten while on the other the applied front would also provide baseline information for the development of immunodiagnostics.

ACKNOWLEDGEMENTS

REFERENCES

Abad,A. *et al.* (2001) Monoclonal enzyme immunoassay for the analysis of carbaryl in fruits and vegetables without sample cleanup. *J. Agric. Food. Chem.*, **49**, 1707–1712.

Bhasin,M. *et al.* (2003) MHCBN: a comprehensive database of MHC binding and non-binding peptides. *Bioinformatics*, **19**, 665–666.

Brusic,V. *et al.* (2000) Molecular immunology databases and data repositories. *J. Immunol. Methods*, **238**, 17–28.

Choi,M.J. *et al.* (1999) Production and characterization of monoclonal antibodies specific to atrazine group compounds: effects of coating ligand structure on the variation of sensitivity and specificity. *J. Immunoassay*, **20**, 57–77.

Goodrow,M.H. and Hammock,B.D. (1998) Hapten design for compound selective antibodies: ELISAS for environmentally deleterious small molecules. *Anal. Chim. Acta*, **376**, 83–91.

Goodrow,M.H. *et al.* (1990) Hapten synthesis, antibody development and competitive inhibition enzyme immunoassay for s-triazine herbicides. *J. Agric. Food Chem.*, **38**, 990–996.

Hall,J.C. *et al.* (1990) Immunoassays to detect and quantitate herbicides in the environment. *Weed Technol.*, **4**, 226–234.

Johnson,G. and Wu,T.T. (2000) Kabat database and its applications: 30 years after the first variability plot. *Nucleic Acids Res.*, **28**, 214–218.

Moreno,M.J. *et al.* (2001) Production of monoclonal antibodies to the *N*-methylcarbamate pesticide propoxur. *J Agric. Food Chem.*, **49**, 72–78.

McSparron,H. *et al.* (2003) JenPep: a novel computational information resource for immunobiology and vaccinology. *J. Chem. Inf. Comput. Sci.*, **43**, 1276–1287.

Ruiz,M. *et al.* (2000) IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res.*, **28**, 219–221.

Saha,S. *et al.* (2005) BCIPEP: a database of B-cell epitopes. *BMC Genomics*, **6**, 79.

Schönbach,C. *et al.* (2000) FIMM, a database of functional molecular immunology. *Nucleic Acids Res.*, **28**, 222–224.

Singh,K.V. *et al.* (2004) Synthesis and characterisation of hapten–protein conjugates for antibody production against small molecules. *Bioconjug. Chem.*, **15**, 168–173.

Stix,G. (2001) The mice that warred. *Sci. Am. (NY)*, **284**, 34–35.

Yau,K.Y.F. *et al.* (2003) Emerging trends in the synthesis and improvement of hapten-specific recombinant antibodies. *Biotechnol. Adv.*, **21**, 599–637.