Bioinformatics Approach for Designing Biomolecule based Therapy

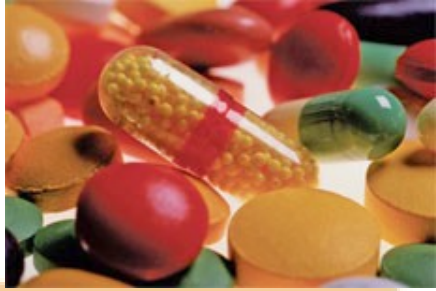
Dr G P S Raghava, Head Bioinformatics Centre

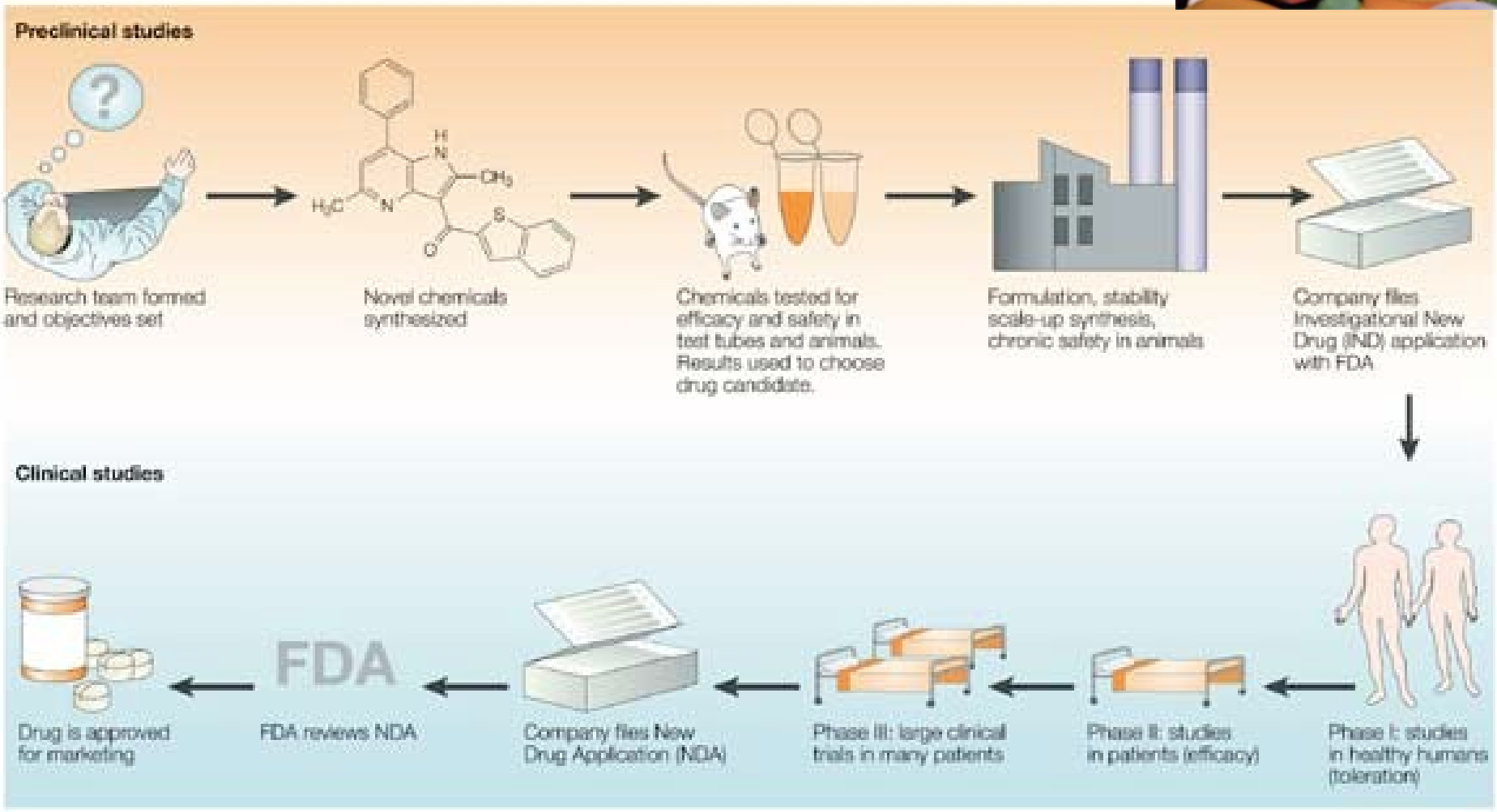
Institute of Microbial Technology, Chandigarh, India

Email: **raghava@imtech.res.in** **http://crdd.osdd.net/**

http://www.imtech.res.in/raghava/

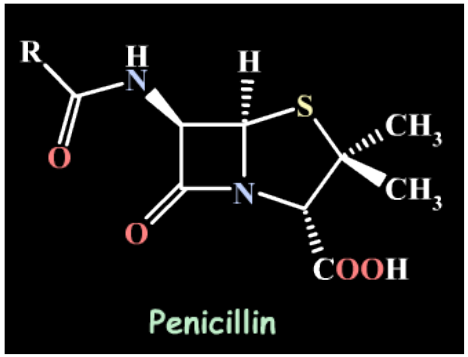
Drug discovery is a long process
(Computer-aided drug discovery)



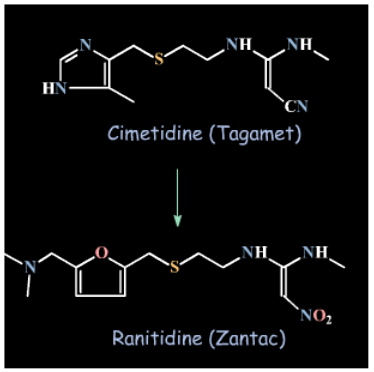


History of Drug Discovery

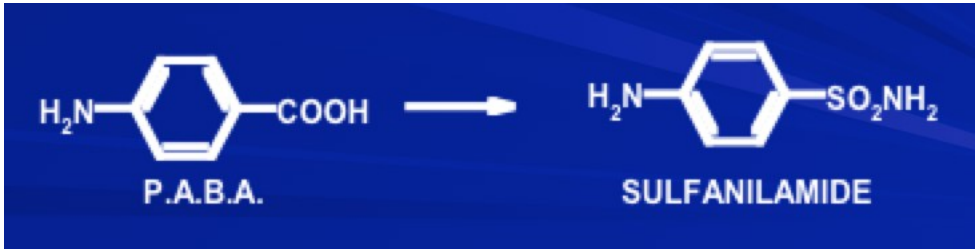
Plants



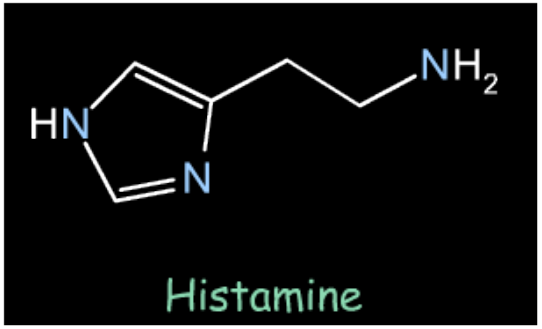
Serendipity

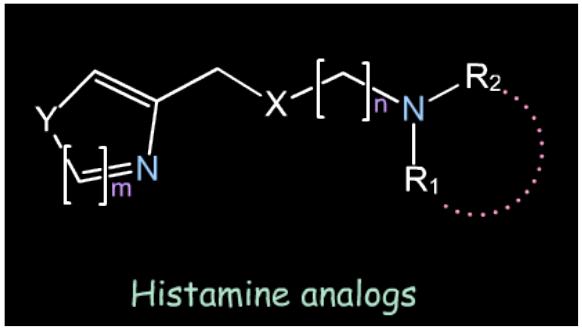


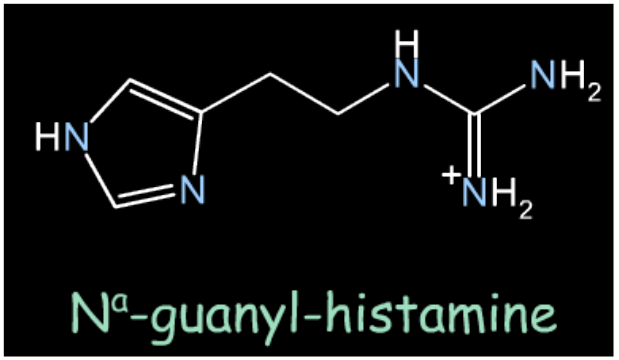
Chemical modifications



Chemical modifications

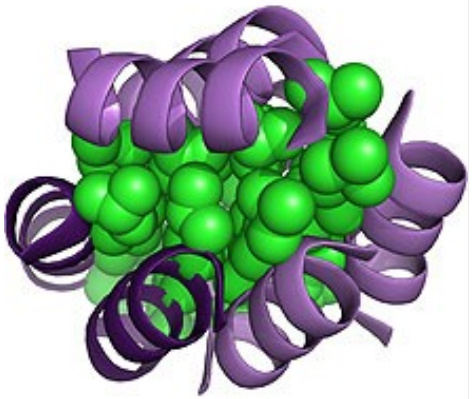




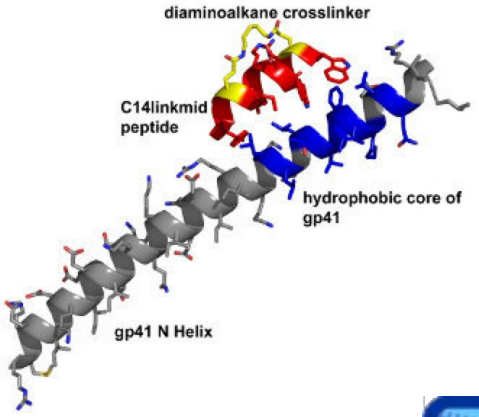


Chemical analogs (rational drug design)

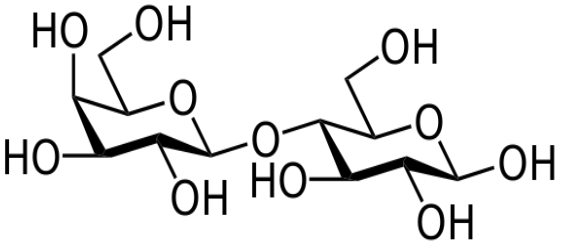
Biomolecules Based Drugs



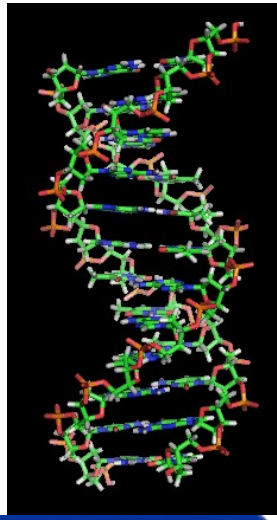
Protein

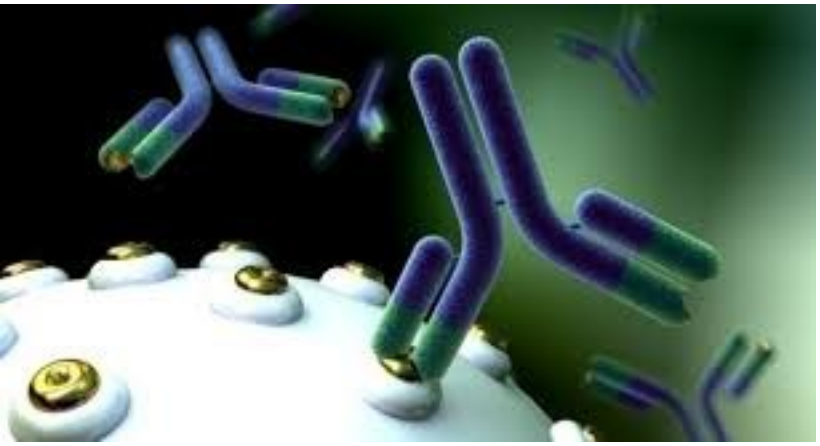


Peptide

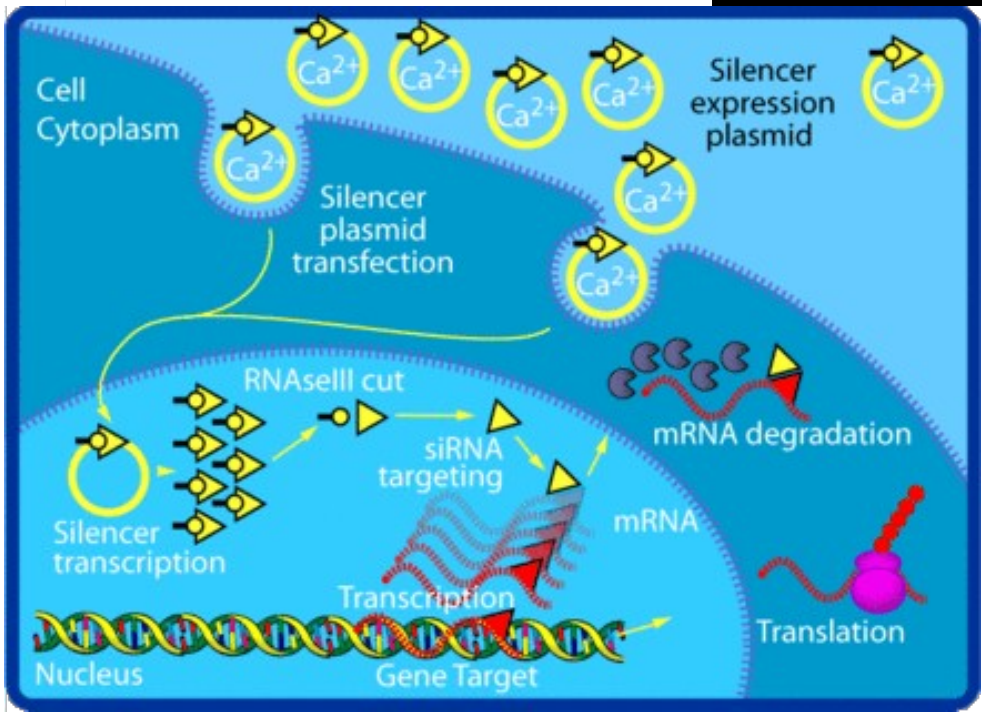


Carbohydrate



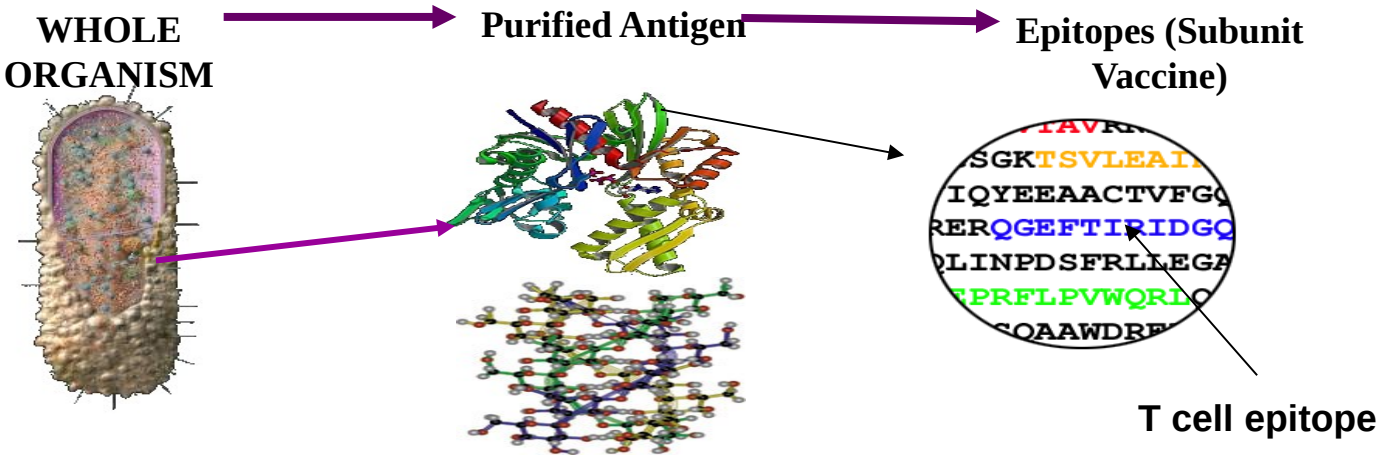


Herceptin (Antibody)

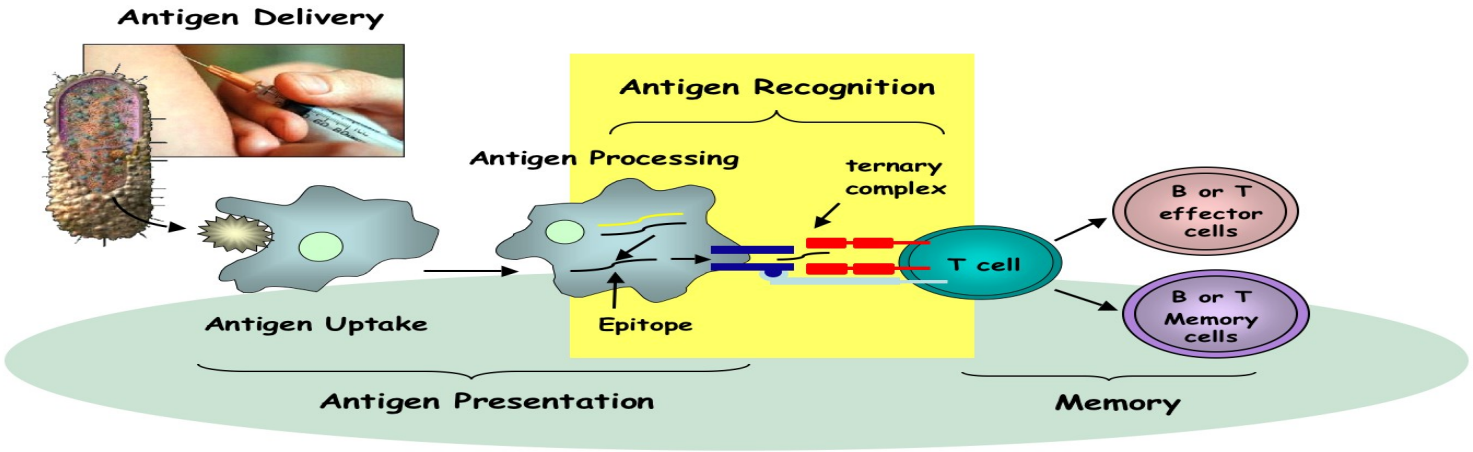


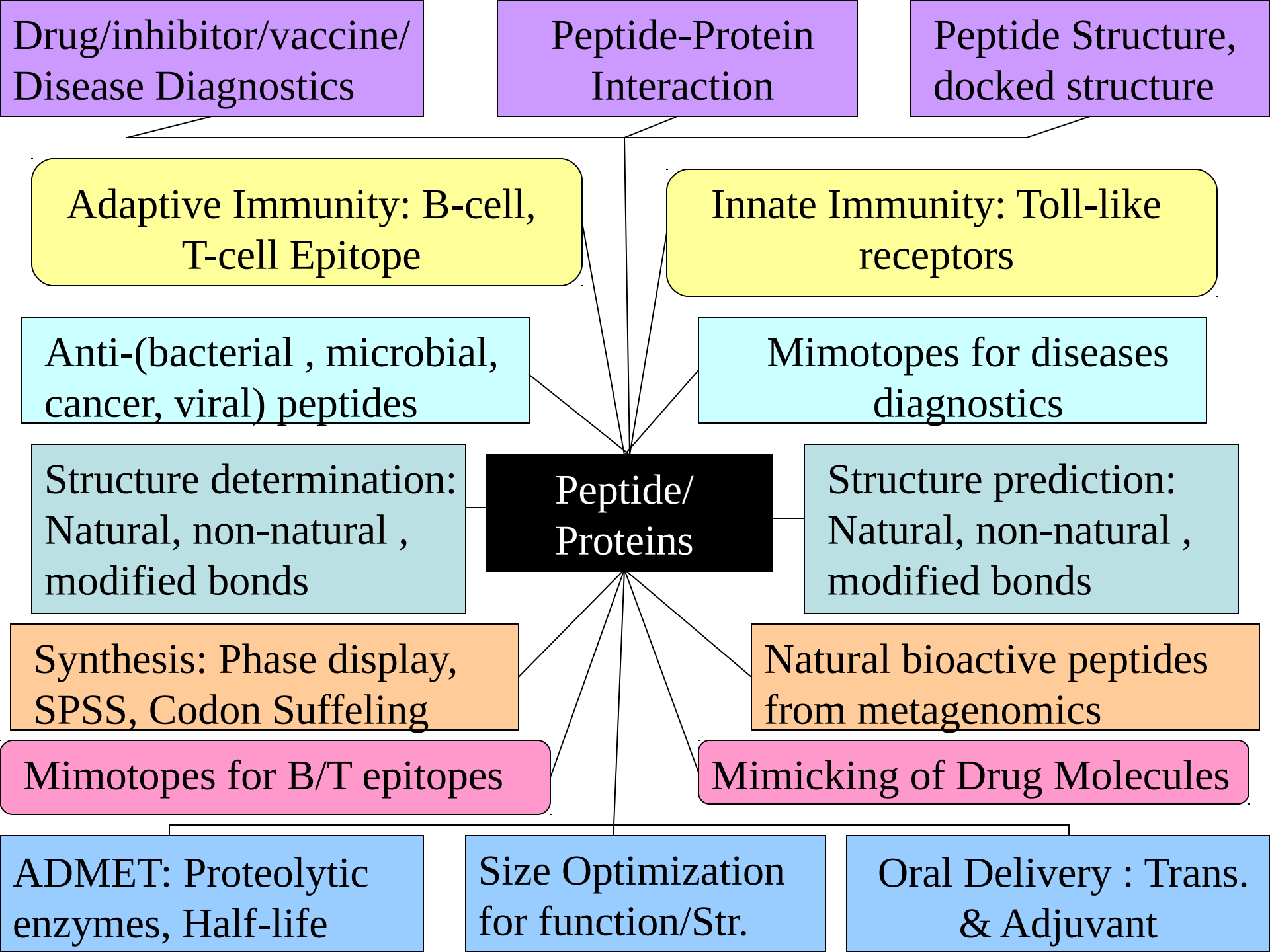
siRNA

Biomolecules Based Vaccines



Attenuated




Drug/inhibitor/vaccine/ Disease Diagnostics
Peptide-Protein Interaction
Peptide Structure, docked structure
Adaptive Immunity: B-cell, T-cell Epitope
Innate Immunity: Toll-like receptors
Anti-(bacterial , microbial, cancer, viral) peptides
Mimotopes for diseases diagnostics
Structure determination: Natural, non-natural , modified bonds
Peptide/ Proteins
Structure prediction: Natural, non-natural , modified bonds
Synthesis: Phase display, SPSS, Codon Suffeling
Natural bioactive peptides from metagenomics
Mimotopes for B/T epitopes
Mimicking of Drug Molecules
ADMET: Proteolytic enzymes, Half-life
Size Optimization for function/Str.
Oral Delivery : Trans. & Adjuvant

Concept of Drug and Vaccine

- Concept of Drug
  - Kill invaders of foreign pathogens
  - Inhibit the growth of pathogens
- Concept of Vaccine
  - Generate memory cells
  - Trained immune system to face various existing disease agents

Human Vaccines against pathogens

Organism	Type	Vaccine Type	Year
Variola virus	Virus	Live	1798
Rabies virus	Virus	Inactivated	1885
Salmonella typhi	Bacteria	Live	1896
Vibrio cholerae	Bacteria	Inactivated	1896
Yersinia pestis	Bacteria	Inactivated	1897
Corynebacterium diphtheriae	Bacteria	Toxoid	1923
Bordetella pertussis	Bacteria	Acellular	1926
Clostridium tetani	Bacteria	Toxoid	1927
Mycobacterium tuberculosis	Bacteria	Live	1927
Yellow fever virus	Virus	Live	1935
Influenza virus type A	Virus	Inactivated	1936
Influenza virus type B	Virus	Inactivated	1936
Coxiella burnetii	Bacteria	Inactivated	1938
Rickettsia prowazekii	Bacteria	Inactivated	1938
Rickettsia rickettsii	Bacteria	Inactivated	1938
Central European encephalitis virus	Virus	Inactivated	1939
Poliovirus types 1, 2, and 3	Virus	Inactivated/Live	1962
Measles virus	Virus	Live	1963
Mumps virus	Virus	Live	1967
Rubivirus	Virus	Live	1969
Staphylococcus aureus	Bacteria	Staphage lysate	1976
Streptococcus pneumoniae	Bacteria	Polysaccharide	1977
Human adenovirus types 4 and 7	Virus	Live	1980
Neisseria meningitidis	Bacteria	Polysaccharide	1981
Hepatitis B	Virus	Recombinant	1986
Haemophilus influenzae	Bacteria	Conjugate	1987
Hantaan virus	Virus	Inactivated	1989
Japanese encephalitis virus	Virus	Inactivated	1992
Varicella-zoster virus	Virus	Live	1994
Hepatitis A	Virus	Inactivated	1995
Escherichia coli	Bacteria	Inactivated	1995
Junin virus	Virus	Live	1996
Bacillus anthracis	Bacteria	Adsorbed	1998
Borrelia burgdorferi	Bacteria	Recombinant	1998

Immunological Bioinformatics, The MIT press.

History of Immunization

- **Children protected who recovered from smallpox**
- **Immunity induce, a process known as variolation**
- **Variolation spread to England and America**
- **Stopped due to the risk of death**
- **Edward Jenner found that protection against smallpox**
- **Inoculation with material from an individual infected with cowpox**
- **This process was called vaccination (cowpox is vaccina)**
- **Inoculum was termed a vaccine**
- **Protective antibodies was developed**

Vaccination

Vaccination: a substance to a person for preventing a disease

- Traditionally composed of a killed or weakened microorganism
- Enables memory cells to respond to an organism before it can cause disease

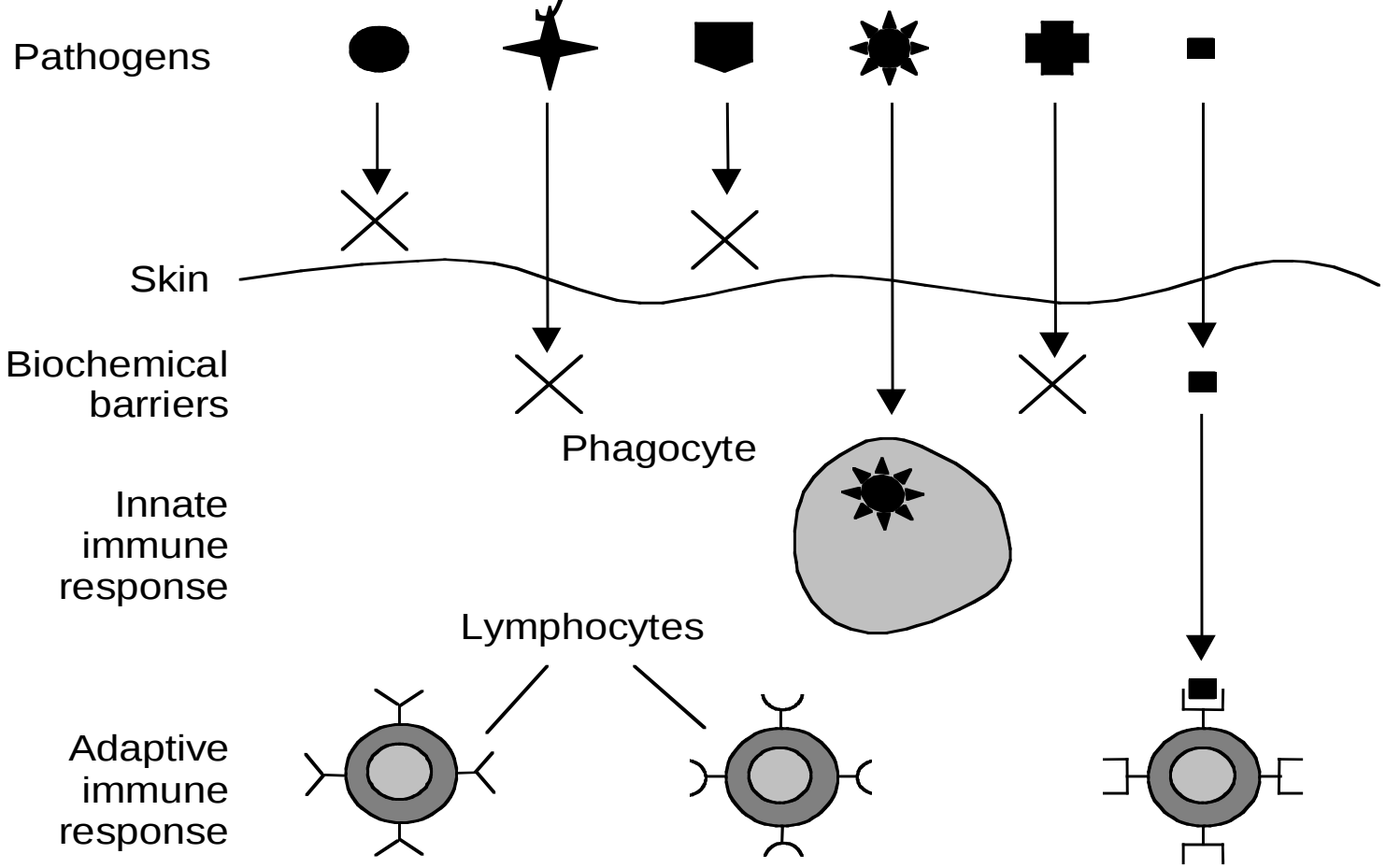
Importance

- **Saves more than 3 million children each year**
- **More than 2 million lives could be saved if existing vaccines were applied on a full-scale worldwide**
- **Complete eradication of Smallpox**

Need of hours

- **Vaccines have been made for only 34 of the more than 400 known pathogens.**
- **Searching of effective vaccines for AIDS, Malaria and Tuberculosis**
- **Development of low cost vaccines**

Multiple layers of the immune system



Adaptive Immunity

Innate Immunity

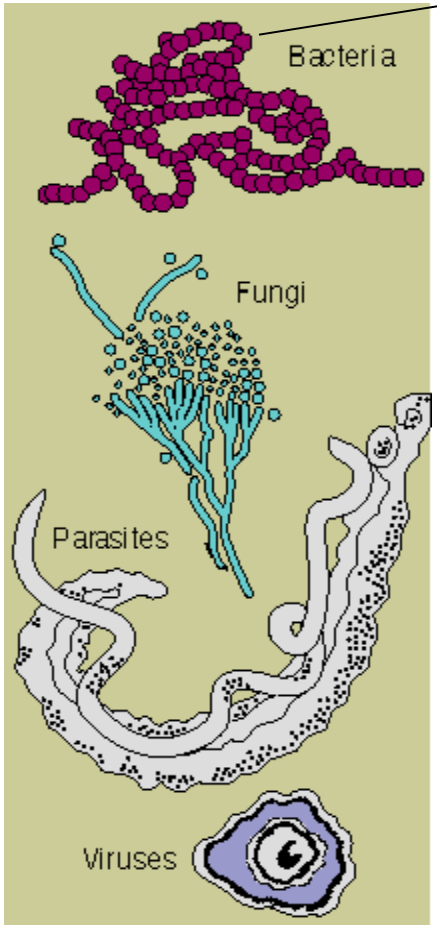
Vaccine Informatics

Bioinformatics Centre
IMTECH, Chandigarh

Protective Antigens

Vaccine Delivery

Disease Causing Agents



Pathogens/Invaders

Host Immune Defense System
For Microbial (Pathogens) Attack

Non-Specific Defense
(Innate Immunity)

Specific Defense
Adaptive Immunity
Responsible for Generating
Memory B & T Cells

TOLL LIKE RECEPTORS
PATHOGEN ASSOCIATED
MOLECULAR PATTERNS

Humoral Immunity
(B-Lymphocytes)

Cell-Mediated Immunity
(T-Lymphocytes)

B-Cell Epitopes
Antibody Prod.
Neutralization

Endogenous Antigen Processing

Exogenous Antigen Processing

CTL Epitopes
MHC I binders
TAP Binders
P. Cleavage Sites

T-Helper Epitopes
MHC II binders

Adaptive Immunity

Innate Immunity

Vaccine Informatics

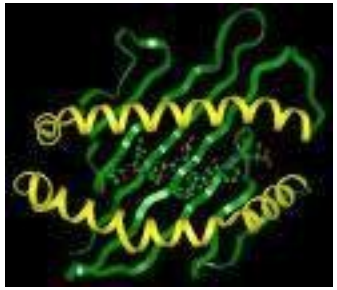
Bioinformatics Centre
IMTECH, Chandigarh

Protective Antigens

Vaccine Delivery

MHCBN: A database of MHC/TAP binders and T-cell epitopes

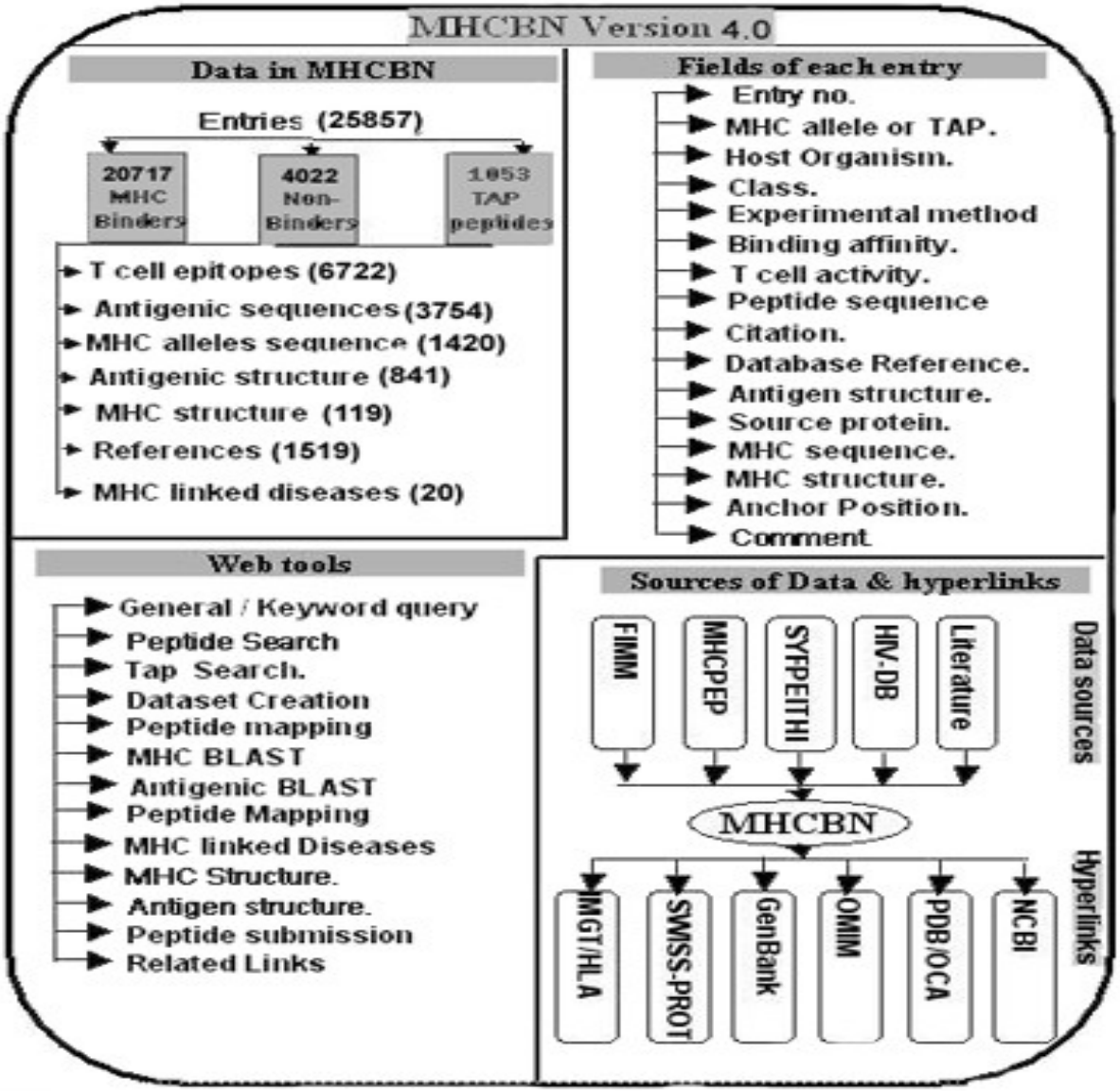
Distributed by EBI, UK

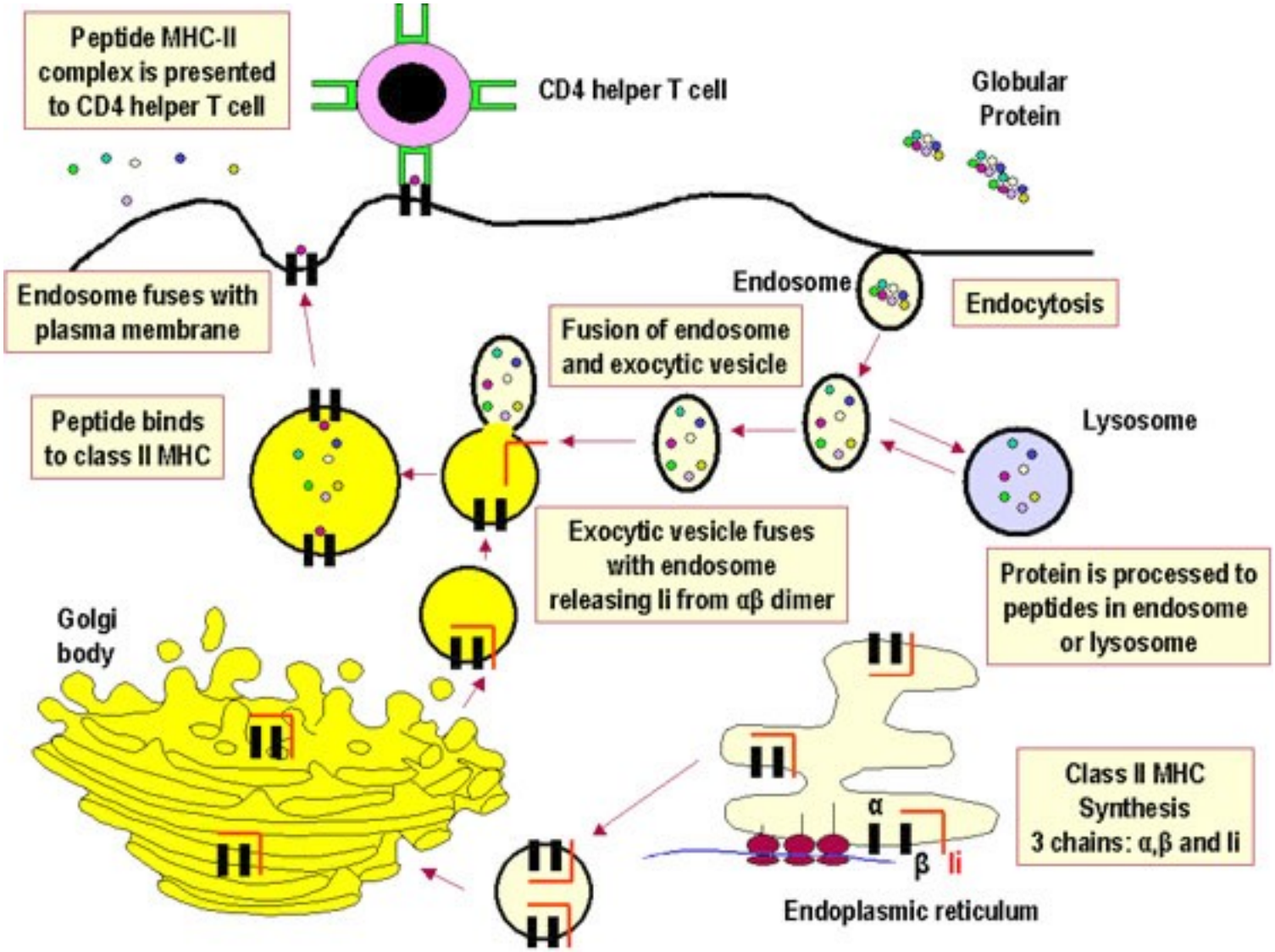


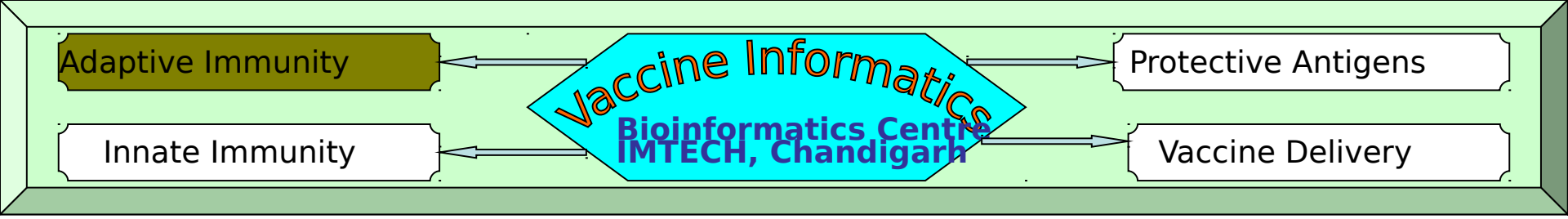
Reference database in T-cell epitopes
Highly Cited (~ 70 citations)

Bhasin et al. (2003) Bioinformatics 19: 665

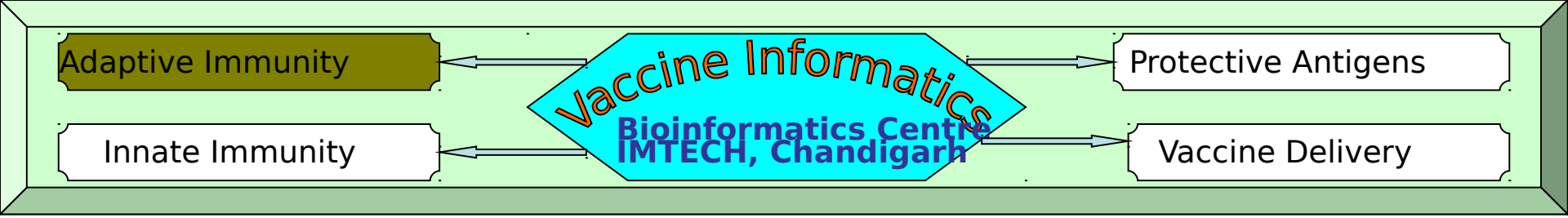
Bhasin et al. (2004) NAR (Online)




Peptide MHC-II complex is presented to CD4 helper T cell
CD4 helper T cell
Globular Protein
Endosome fuses with plasma membrane
Endosome
Endocytosis
Fusion of endosome and exocytic vesicle
Peptide binds to class II MHC
Lysosome
Exocytic vesicle fuses with endosome releasing Ii from αβ dimer
Protein is processed to peptides in endosome or lysosome
Golgi body
Class II MHC Synthesis 3 chains: α,β and Ii
α
β
Ii
Endoplasmic reticulum



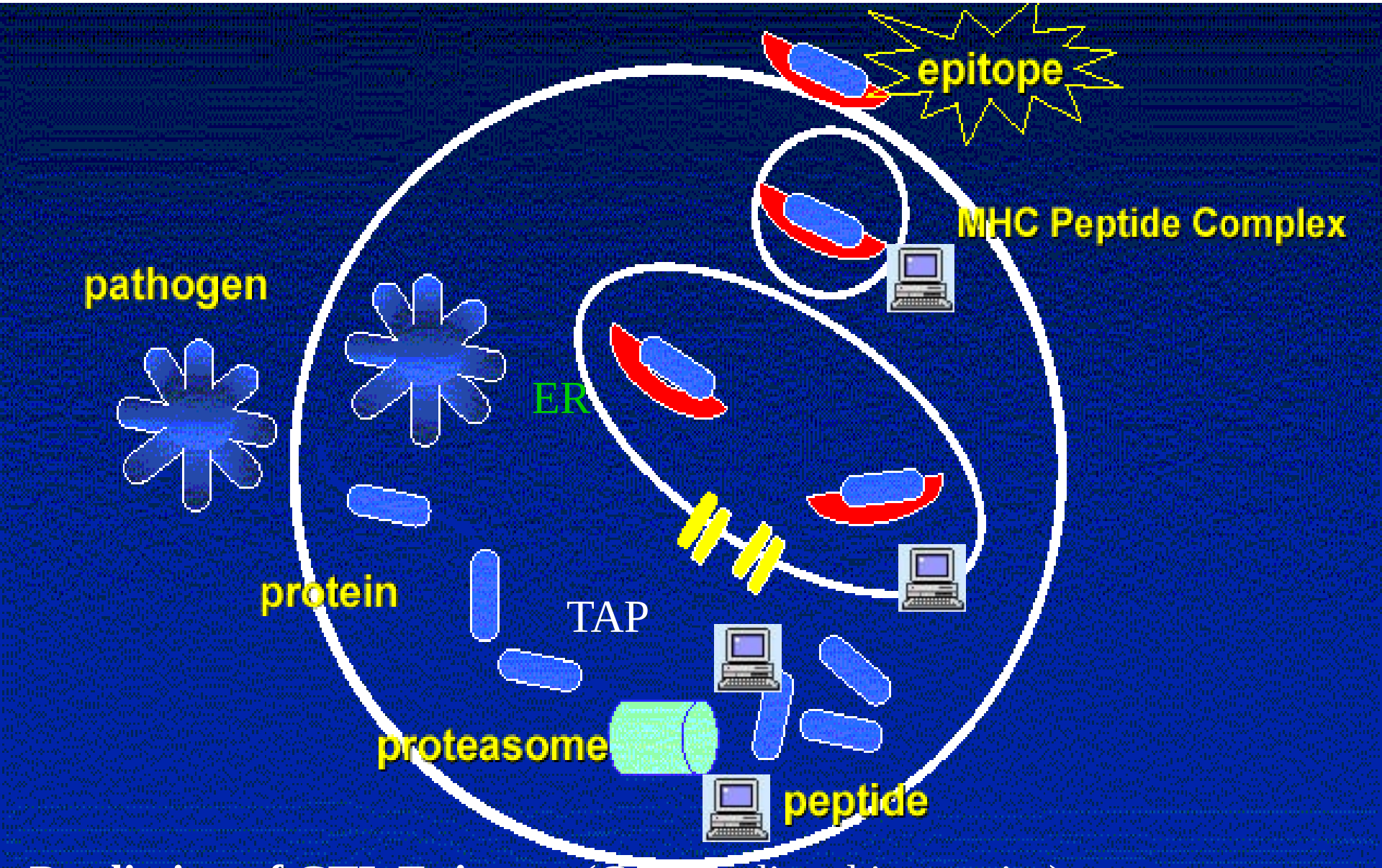

```
                ---------10---------20---------30---------40---------50---------60--
DRB1_0101:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0102:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0301:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0305:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0306:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0307:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0308:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0309:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0311:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0401:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0402:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0404:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0405:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0408:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0410:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0421:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0423:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0426:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0701:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0703:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0801:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0802:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0804:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0806:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0813:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_0817:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_1101:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
DRB1_1102:  MKVKYALLSAGALQLLVVGCGSSHHETHYGYATLSYADYWAGELGQSRDVLLAGNAEADRAGI
```



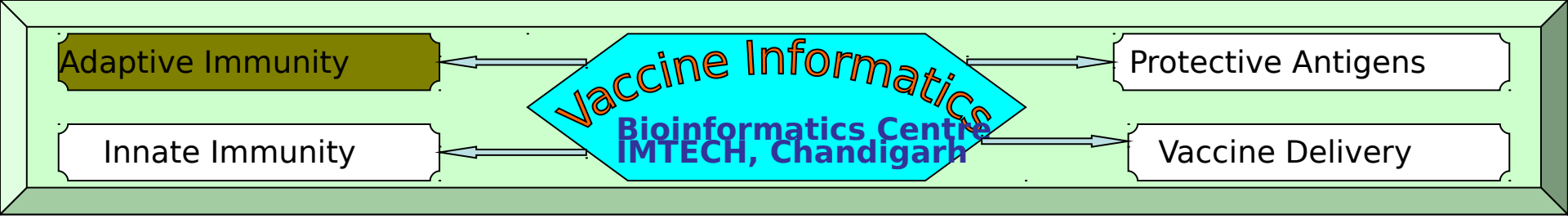
Prediction of MHC II Epitopes (T_{helper} Epitopes)

- **Propred: Promiscuous of binders for 51 MHC Class II binders**
  - **Virtual matrices**
  - **Singh and Raghava (2001) Bioinformatics 17:1236**
- **HLADR4pred: Prediction of HLA-DRB1*0401 binding peptides**
  - **Dominating MHC class II allele**
  - **ANN and SVM techniques**
  - Bhasin and Raghava (2004) Bioinformatics 12:421.
- **MHC2Pred: Prediction of MHC class II binders for 41 alleles**
  - **Human and mouse**
  - **Support vector machine (SVM) technique**
  - **Extension of HLADR4pred**
- **MMBpred: Prediction pf Mutated MHC Binder**
  - Mutations required to increase affinity
  - Mutation required for make a binder promiscuous
  - **Bhasin and Raghava (2003) Hybrid Hybridomics, 22:229**
- **MOT : Matrix optimization technique for binding core**
- **MHCBench: Benchmarting of methods for MHC binders**

Endogenous Antigen Processing



Prediction of CTL Epitopes (Cell-mediated immunity)



Prediction of MHC I binders and CTL Epitopes

Propred1: Promiscuous binders for 47 MHC class I alleles

- **Cleavage site at C-terminal**
- ***Singh and Raghava (2003) Bioinformatics 19:1109***

nHLApred: Promiscuous binders for 67 alleles using ANN and QM

- **Bhasin and Raghava (2007) J. Biosci. 32:31-42**

TAPpred: Analysis and prediction of TAP binders

- ***Bhasin and Raghava (2004) Protein Science 13:596***

Pcleavage: Proteasome and Immuno-proteasome cleavage site.

- **Trained and test on in vitro and in vivo data**
- ***Bhasin and Raghava (2005) Nucleic Acids Research 33: W202-7***

CTLpred: Direct method for Predicting CTL Epitopes

- ***Bhasin and Raghava (2004) Vaccine 22:3195***

Adaptive Immunity

Innate Immunity

Vaccine Informatics

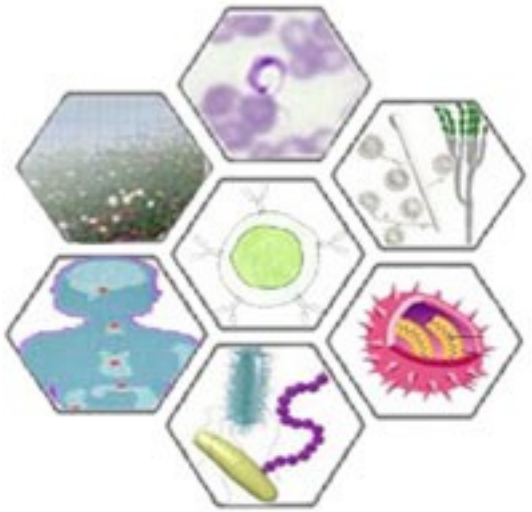
Bioinformatics Centre

IMTECH, Chandigarh

Protective Antigens

Vaccine Delivery

BCIPEP: A database of B-cell epitopes.



SWISS-PROT | PDB | MHCBN | PUBMED | GenBank

Structure of Each Entry

- Entry No
- Peptide Sequence
- Source Protein
- Experimental Method
- Immunogenicity
- Neutralization
- Model Organism
- Antibody
- Pathogen Group
- Publication reference
- Database Reference
- Antigen Structure
- Comment

Hyperlinks

Bcipep

Data structure

Interactive Tools

- Keyword Search
- Peptide Search
- Peptide Mapping
- Antigenic BLAST
- Data Submission

Web Resources

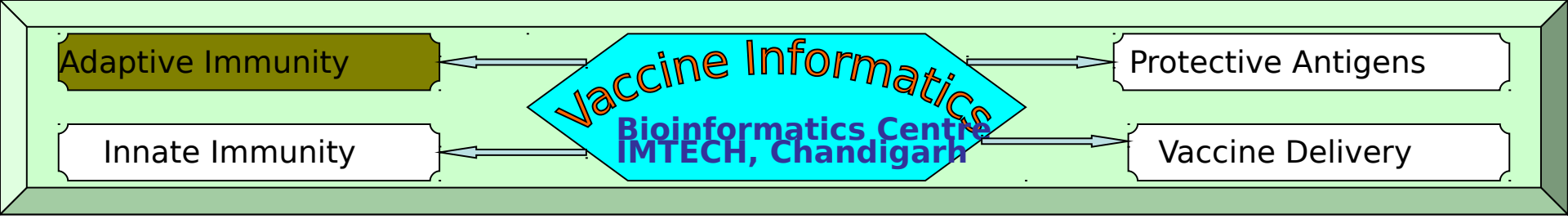
- Ab Structure
- Db. Information
- Related Links
- SRS @ IMTECH
- SRS @ EBI

Total Entries (3031)

Virus (2046) | Bacteria (539) | Others (157) | Fungi (53) | Protozoa (236)

Saha et al.(2005) BMC Genomics 6:79.

Saha et al. (2006) NAR (Online)



Prediction of B-Cell Epitopes

- **BCEpred: Prediction of Continuous B-cell epitopes**
  - **Benchmarking of existing methods**
  - **Poor performance slightly better than random**
  - **Combine all properties and achieve accuracy around 58%**
  - **Saha and Raghava (2004) ICARIS 197-204.**
- **ABCpred: ANN based method for B-cell epitope prediction**
  - **Extract all epitopes from BCIPEP (around 2400)**
  - **700 non-redundant epitopes used for testing and training**
  - **Recurrent neural network**
  - **Accuracy 66% achieved**
  - **Saha and Raghava (2006) Proteins,65:40-48**
- **ALGpred: Mapping and Prediction of Allergenic Epitopes**
  - **Allergenic proteins**
  - **IgE epitope and mapping**
  - **Saha and Raghava (2006) Nucleic Acids Research 34:W202-W209**
- **CBTOPE: Prediction of conformational epitopes**

Adaptive Immunity

Innate Immunity

Vaccine Informatics

Bioinformatics Centre
IMTECH, Chandigarh

Protective Antigens

Vaccine Delivery

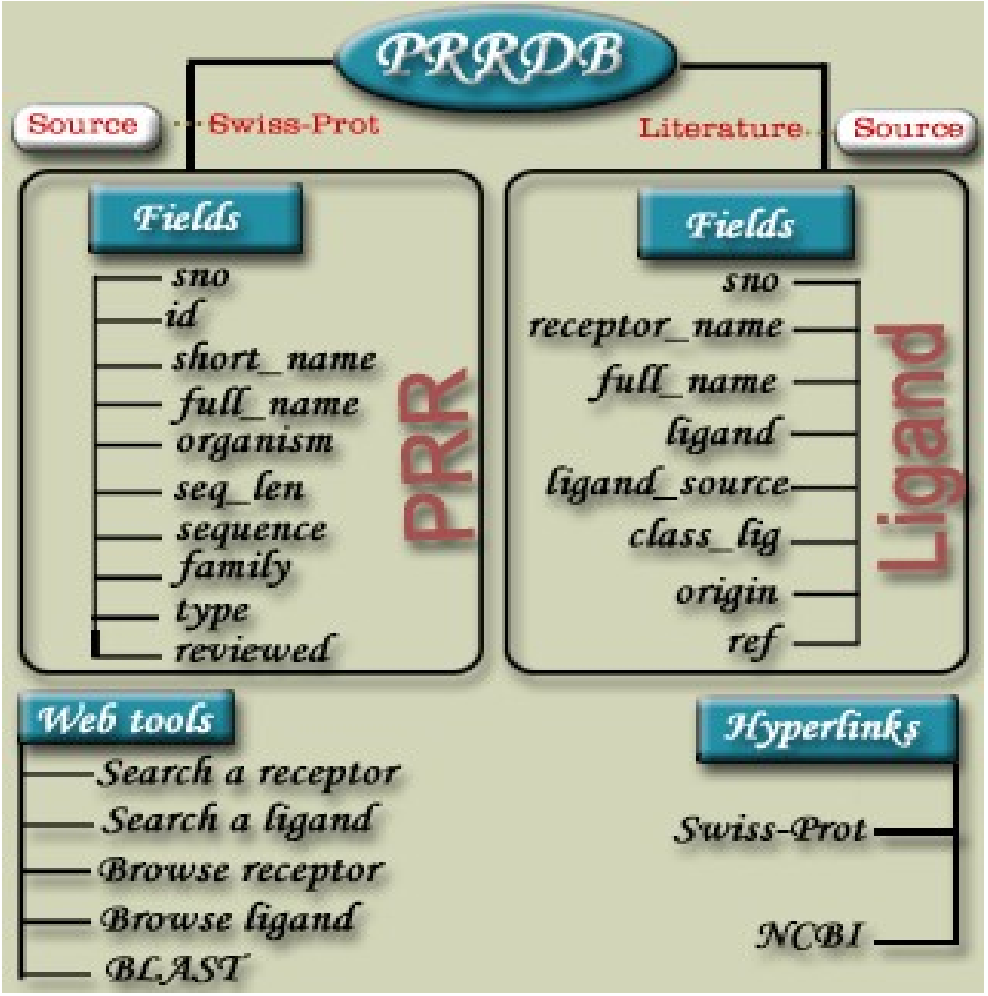
PRRDB is a database of pattern recognition receptors and their ligands

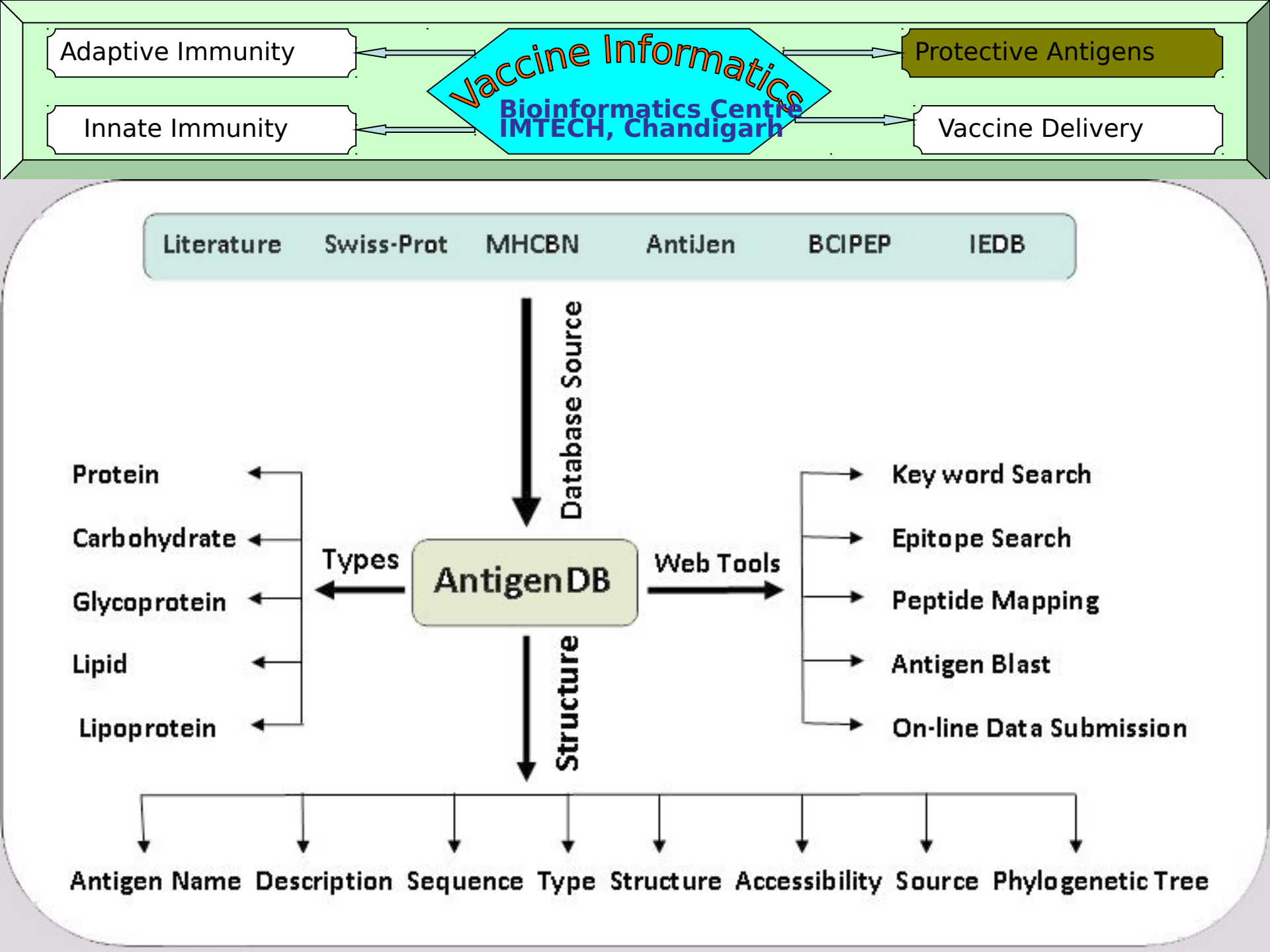
~500 Pattern-recognition Receptors

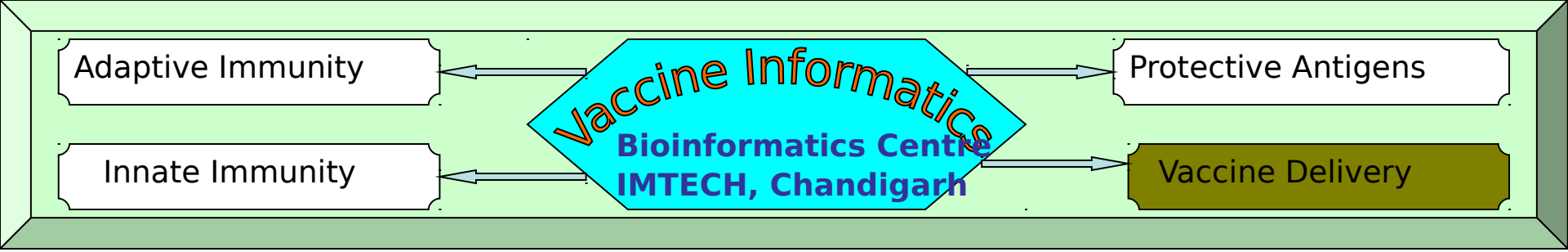
228 ligands (PAMPs)

77 distinct organisms

720 entries




Adaptive Immunity
Innate Immunity
Vaccine Informatics
Bioinformatics Centre
IMTECH, Chandigarh
Protective Antigens
Vaccine Delivery
Literature
Swiss-Prot
MHCBN
AntiJen
BCIPEP
IEDB
Database Source
Protein
Carbohydrate
Glycoprotein
Lipid
Lipoprotein
Types
AntigenDB
Web Tools
Key word Search
Epitope Search
Peptide Mapping
Antigen Blast
On-line Data Submission
Structure
Antigen Name
Description
Sequence
Type
Structure
Accessibility
Source
Phylogenetic Tree



Major Challenges in Vaccine Design

- **ADMET of peptides and proteins**
- **Activate innate and adaptive immunity**
- **Prediction of carrier molecules**
- **Avoid cross reactivity (autoimmunity)**
- **Prediction of allergic epitopes**
- **Solubility and degradability**
- **Absorption and distribution**
- **Glycocylated epitopes**

Modelling of Immune System for Designing Epitope-based Vaccines

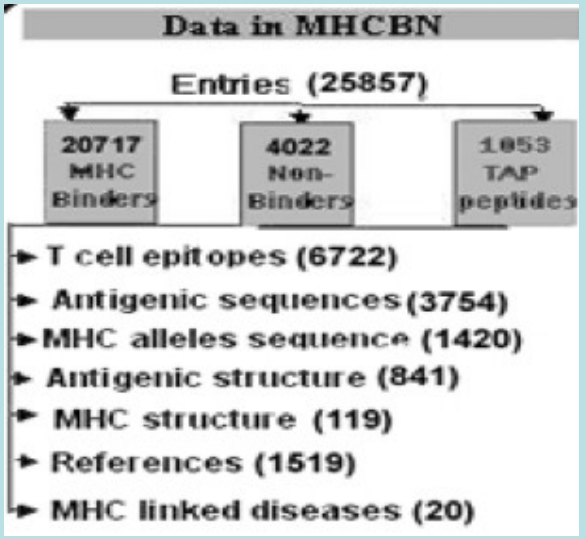
Adaptive Immunity (Cellular Response) : T_{helper} Epitopes

Propred: for promiscuous MHC II binders

MMBpred:for high affinity mutated binders

MHC2pred: SVM based method

MHCBN: A database of MHC/TAP binders and non-binders



Adaptive Immunity (Cellular Response) : CTL Epitopes

Pcleavage: for proteome cleavage sites

TAPpred: for predicting TAP binders

Propred1: for promiscuous MHC I binders

CTLpred: Prediction of CTL epitopes

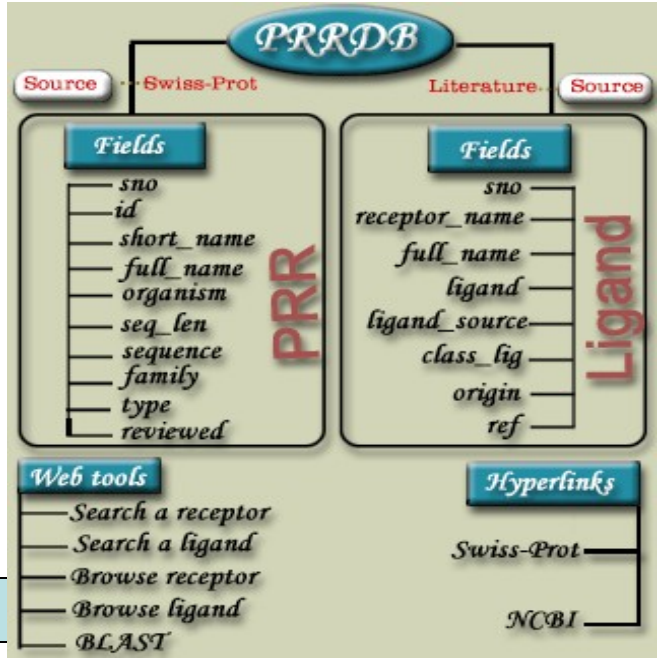
Adaptive Immunity (Humoral Response) :B-cell Epitopes

BCIpep: A database of B-cell eptioes;

ABCpred: for predicting B-cell epitopes

ALGpred: for allergens and IgE eptopes

HaptenDB: A datbase of haptens



Innate Immunity : Pathogen Recognizing Receptors and ligands

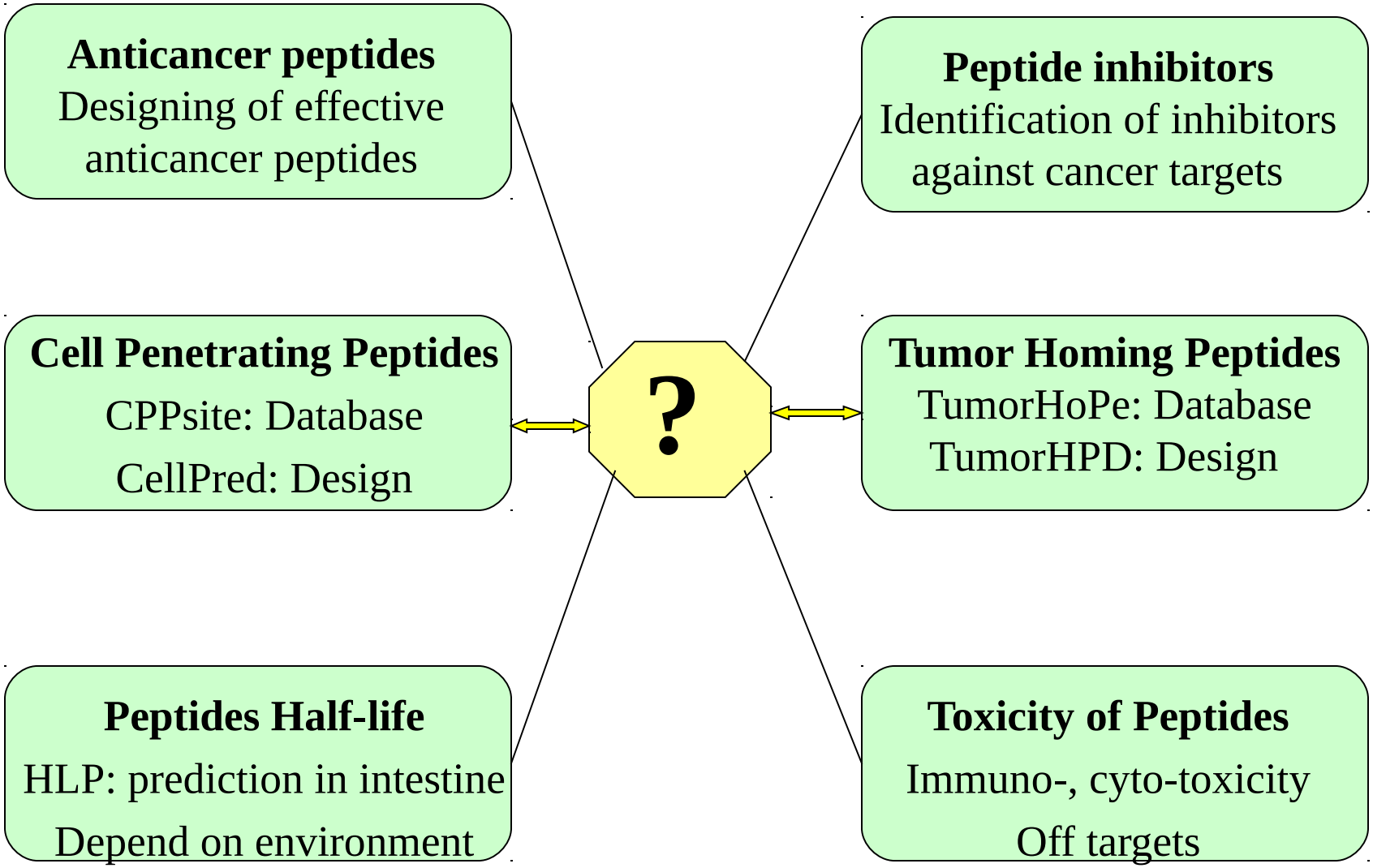
PRRDB: A database of PRRs & ligands

Antibp: for anti-bacterial peptides

Signal transduction in Immune System

Cytopred: for classification of Cytokines

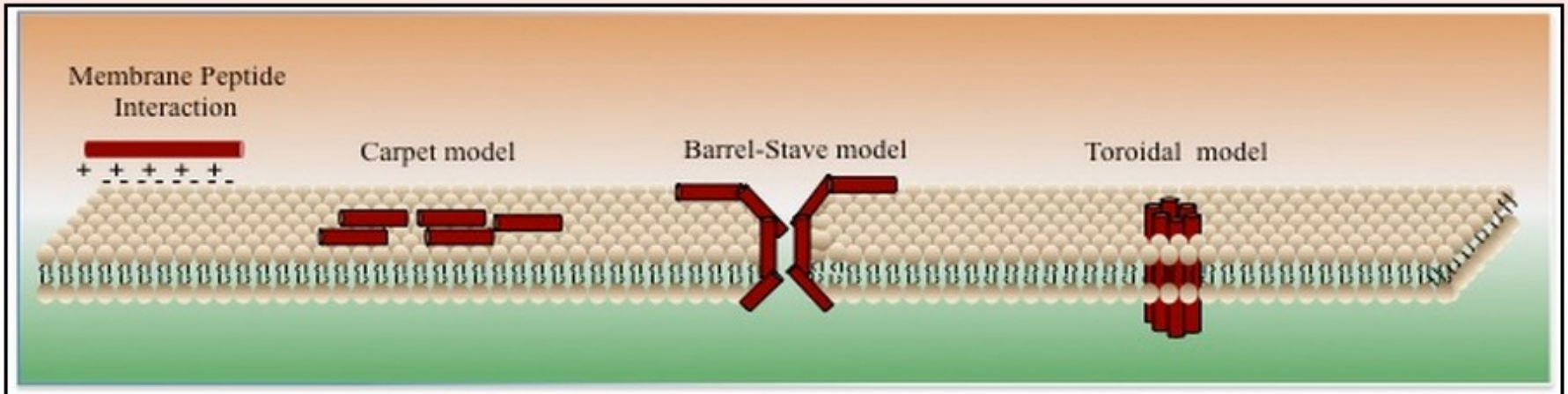
Designing of therapeutic peptides against cancer



AntiCP:- Designing of Anticancer Peptides

Institute of Microbial Technology, Chandigarh India

Welcome to AntiCP



Various models of membrane permeation by anti-cancer peptides

AntiCP is web based prediction server for Anticancer peptides. SVM models developed are based on amino acid composition and binary profile features. Positive dataset consists of 225 antimicrobial peptides with anticancer properties. This server is extremely useful for the researchers working in the field of Anticancer peptides. This server allows the users to design ACPs and their mutants with

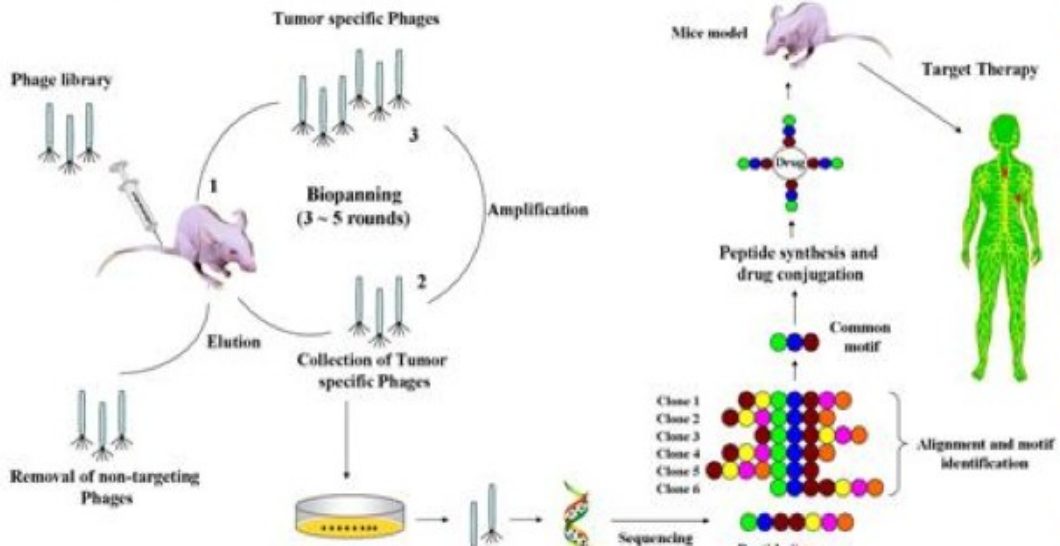
Peptide Resources/Databases

TumorHope - Tumor Homing Peptide Database

Home | Search | Browse | Structure | Tools | Important | Help | About Us | Contact Us

Welcome to TumorHope - A comprehensive database of Tumor Homing Peptides

TumorHope is a manually curated comprehensive database of experimentally characterized tumor homing peptides. These peptides recogninze tumor tissues and tumor associated micro environment, including tumor metastasis. Thus, they can be used to deliver drugs selectively in tumors.



Importance of Peptides: Poor selectivity of chemotherapeutic drugs for cancer is a major challenge for successful clinical outcome. Conjugation of drug with homing peptide may enhance the selectivity and efficacy of the therapy. Current efforts are being focused on tumor homing peptides that may target tumor tissues.

Information about Peptides: Tumor Homing Peptide Database has been developed using extensive literature search. It contains detailed information about the tumor targeting/homing peptides. Each entry contains following type of information about a peptide; its sequence, source, target tumor, target cell, biomarker, applications and clones. Experimental details like phage display libraries used, cell lines, *in*

TumorHPD: Designing of Tumor Homing Peptides

(Institute of Microbial Technology, Chandigarh, India)

Welcome to TumorHPD

Tumor homing peptides are the short peptides having average length between 7 to 12 residues. These peptides h
bind to tumor cells or tissues. These peptides can be used to deliver target specific drugs and as imaging agents for t
Thus prediction of tumor homing peptide is important for managing cancer treatment effectively.

TumorHPD is a web server for predicting and designing tumor homing peptides. This server is extremely useful fo
the field of therapeutic peptides. This server allows the users to design tumor homing peptides and their mutants
and physicochemical properties.

ference: Sharma, A. et al. Computational approach for designing tumor homing peptides. Sci. Rep. 3, 1607; DOI:10.

CellPPD: Designing of Cell Penetrating Peptides

Welcome to CellPPD

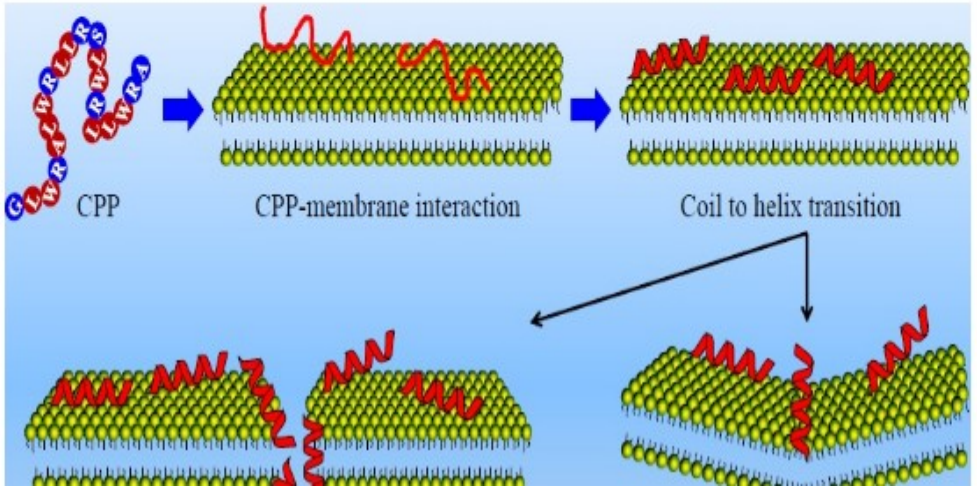
utam *et al.*: *In silico* approaches for designing highly effective cell penetrating peptides. *Journal of Translational Medicine* 13 11:74.Link

CellPPD is an *in silico* method, which is developed to predict and design efficient cell penetrating peptides (CPPs). The main dataset used in this method consists of 708 experimentally validated CPPs.

Major Features include:

(1) **Desing Peptide:** This module allows user to generate all possible single mutant analogues of their peptides and predict whether the analogue is cell penetrating or not.

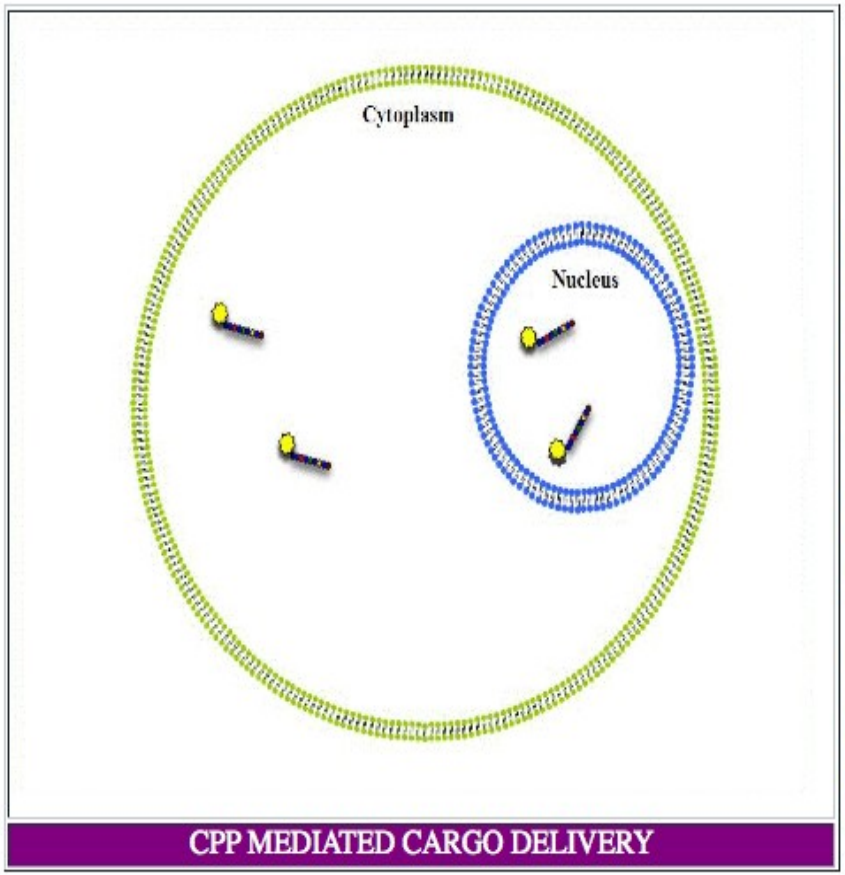
(2) **Multiple Peptides:** This module of CellPPD allows user to predict number of CPPs in peptides submitted by the user



CPPsite: a webSite for Cell Penetrating Peptides

Navigation

- Home
- Search
- Browse
- Structure
- Tools
- Important
- Help
- About us
- Contact us



CPP MEDIATED CARGO DELIVERY

CPPsite: CPPsite is a database of experimentally validated Cell Penetrating Peptides (10-30 amino acids).

Importance of CPPsite: CPPs have tremendous therapeutic applications. These are widely used to promote intracellular uptake of conjugated cargos (nucleic acids, peptide nucleic acids, proteins, drugs, liposomes etc.) and thus play role to overcome the problem of poor delivery and low bioavailability of therapeutic molecules. CPP conjugated drugs when delivered *in vivo* have s hown promising results with high efficacy. Many CPP-conjugated compounds are under clinical trials. CPPsite database provides comprehensive information on CPPs, which may be helpful to scientific community working in the area of peptide based drug discovery.

What type of information it has: CPPsite database's current version contains comprehensive information of 843 CPPs with multiple entries in terms of peptide sequence, source/origin, localization, uptake efficiency, uptake mechanism, hydrophobicity, charge etc.

Is it a manually curated database: Yes, we have collected and compiled all the information from published literature. In addition, we have also generated structural information of CPPs. We predicted tertiary and secondary structure of these peptides using PepStr and DSSP.

Work in Progress

1. Prediction of CPP 2. Designing CPP 3. Scanning in proteins

CellPPD

Designing of Cell Penetrating Peptides

Your job id is 2149

Go Back

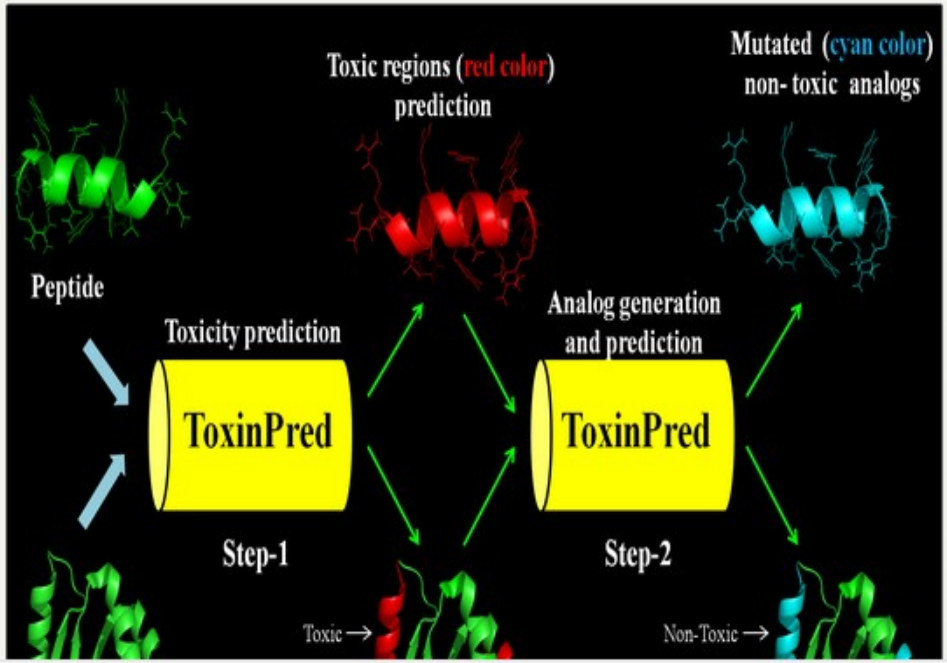
Original Peptide								
Peptide Sequence ▲	Mutation Position	SVM score	Prediction	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge	Mol wt
KMPQACEERTDSLALLA	No	-0.42	CPP	-0.20	-0.25	0.35	-1.00	1876.41
Mutant Peptides								
AMPQACEERTDSLALLA	1	-0.57	CPP	-0.12	0.08	0.14	-2.00	1819.31
CMPQACEERTDSLALLA	1	-0.51	CPP	-0.13	0.12	0.11	-2.00	1851.37
DMPQACEERTDSLALLA	1	-0.55	CPP	-0.18	-0.23	0.35	-3.00	1863.32
EMPQACEERTDSLALLA	1	-0.61	CPP	-0.17	-0.23	0.35	-3.00	1877.35
FMPQACEERTDSLALLA	1	-0.51	CPP	-0.10	0.14	0.02	-2.00	1895.41
GMPQACEERTDSLALLA	1	-0.46	CPP	-0.12	-0.05	0.17	-2.00	1805.29
HMPQACEERTDSLALLA	1	-0.51	CPP	-0.16	-0.21	0.14	-1.50	1885.38
IMPQACEERTDSLALLA	1	-0.59	CPP	-0.09	0.24	0.06	-2.00	1861.40

ToxinPred

Designing and prediction of toxic peptides

Welcome to ToxinPred

ToxinPred is an *in silico* method, which is developed to predict and design toxic/non-toxic peptides. The main dataset used in this method consists of 1805 toxic peptides (<=35 residues).



Major Features include:

(1) **Desing Peptide:** This module allows user to generate all possible single mutant analogs of their peptides and predict whether the analog is toxic or not.

(2) **Batch Submission:** This module of ToxinPred allows user to predict number of toxic peptides submitted by the user.

(3) **Protein Scanning:** This module generates all possible overlapping peptides and their single mutant analogs of protein submitted by the user. It also predicts whether overlapping peptide/analog is toxic or not.

(4) **QMS Calculator:** This tool allows the users to submit query peptide in FASTA format and to optimize the peptide sequence to get maximum/minimum/desired toxicity based upon the Quantitative Matrix based position specific scores. It will help the user to tweak any residue from the predecessor peptide to attain the analog with desired property (highest/lowest toxicity).

ParaPep -A Database of Anti-parasitic peptides

Home Page of ParaPep

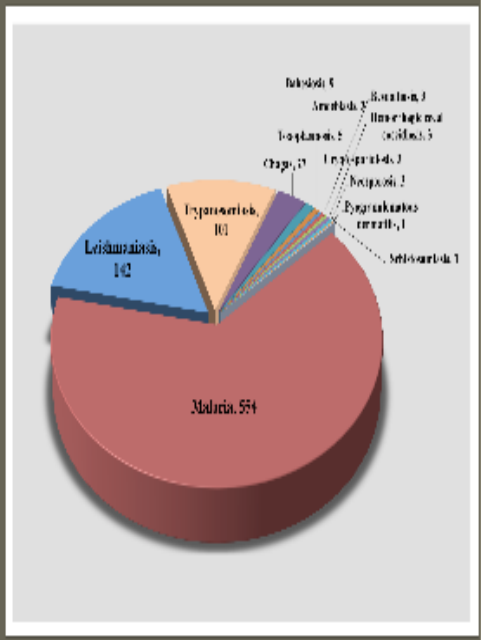
ParaPep: It is a manually curated repository of experientially validated anti-parasitic peptides and their structures. Data have been collected from research papers, published patents and other databases.

Peptide sequences: The current release of ParaPep contains 863 anti-parasite peptide entries, which have been tested against 12 different types of parasites. Most of the entries have been compiled for Malaria followed by Leishmaniasis and Trypanosomiasis.

Type of Peptides: ParaPep consists of various types of peptides, which includes linear peptides, cyclic peptides and peptides having L-amino acids, non-natural amino acids (e.g., D-amino acid, ornithine, etc.) and chemically modified residues.

Structure of Peptides: We determined secondary and tertiary structure of each peptide in ParaPep. using PepStr software. First, we scan PDB to identify all identical peptides to assign their tertiary structure. Structure of remaining peptides were predicted using PEPstr. Secondary structure of peptides were assigned using DSSP from their tertiary structure.

Parasitic Disease Covered



Quick View of ParaPep

Peptide Web Servers

AntiBP2 : Server for antibacterial peptide prediction

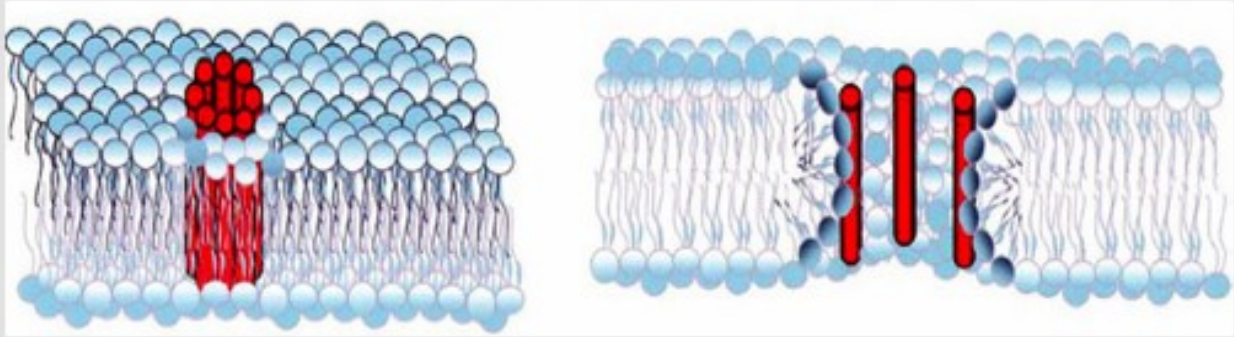
- Home
- Algorithm
- Submission
- Help
- Team
- Contact

Antibp : Our previous version for the prediction of antibacterial peptides for given protein sequence.

Sneh Lata, B K Sharma, GPS Raghava. Analysis and prediction of antibacterial peptides. *BMC Bioinformatics* 2007,8:263

About AntiBP2

Antibacterial peptides are important components of innate immune system, used by the host to protect itself from different types of pathogenic bacteria. Antimicrobial peptides have broad spectrum of activity against bacteria, fungai, viruses and even cancer cells.



AntiBP2 server predicts the antibacterial peptides in a protein sequence. Prediction can be done by using Support Vector Machine (SVM) based method using coposition of peptide sequences and overall accuracy of this server is ~92.14%. This server can also predict the source of these antibacterial peptides with ~98.52% accuracy. If the source of these antibacterial peptides are insect, frog or mammal then it gives the information of its family also. This server can help in finding and designing of peptides based antibiotics.

If You are using this server, please site:
Lata, S., Mishra, N.K. and Raghava, G. P. S. (2009) AntiBP2: Improved version of antibacterial peptide prediction. **BMC Bioinformatics 11:S19** .

HEMOLYTIK: A Database of Hemolytic and Non-hemolytic Peptides

SEARCH	CATEGORIZATION	SIMILARITY	DOWNLOAD	IMPORTANT	GENERAL
Basic	Source	BLAST	Sequence	Submit Form	Acknowledgment
Conditional	Peptide Type	Smith-Waterman	Structure	Statistics	Important Links
Peptide	Function	Mapping	Refrences	Guide/Help	Developers
SMILES	Length	Alignment	Datasets	Recent Papers	Contact

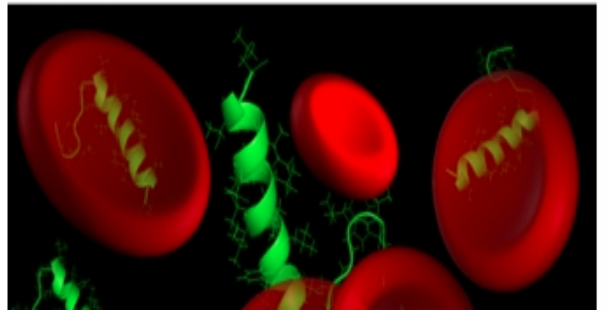
Welcome to Homepage

experimentally determined hemolytic and non-hemolytic peptides, Nucl. Acids Res. (2013) doi: 10.1093/nar/gkt1008.

Hemolytik is a manually curated database of experimentally validated Hemolytic and Non-hemolytic peptides. In this database, peptides have been collected from both published articles as well as from other repositories like CAMP, DAMPD, APD2 and Swiss-Prot. In addition, tertiary structure of peptides have been predicted using PEPstr and secondary structure states are assigned using DSSP. First time, structure of modified peptides (containing Non-natural, D-amino acids, Modified-amino acid like Ornithine, Terminal modifications like Acetylation/Amidation) have also been predicted. In order to provide comprehensive information, peptides were searched and linked with important peptide and protein databases such as IEDB, PDB, Swiss-Prot and TrEMBL.

Major Features

(1) **Resource:** It provides comprehensive information about hemolytic peptides that include their Sequence, Name, Origin, Type (Linear/Cyclic), Chirality, End modification, Chemical modification, Source of RBCs, Hemolytic activity and Function. Data is collected from wide sources like literature and various other databases. Basic and Conditional Search facility enables the users to search a specific peptide/query in



Peptide Web Servers

PEPstr: PEPTIDE TERTIARY STRUCTURE PREDICTION SERVER

Bioinformatics Centre, Institute of Microbial Technology, Chandigarh

---

The Pepstr server predicts the tertiary structure of small peptides with sequence length varying between 7 to 25 residues. The prediction strategy is based on the realization that β-turn is an important and consistent feature of small peptides in addition to regular structures. Thus, the methods uses both the regular secondary structure information predicted from PSIPRED and β-turns information predicted from BetaTurns. The side-chain abgles are placed using standard backbone-dependent rotamer library. The structure is further refined with energy minimization and molecular dynamic simulations using Amber *version6*.

Usage: Paste your one-letter amino acid sequence in the textarea provided below..

Sequence name :

Choose the peptide environment: Vacuum

Paste the peptide sequence below : Help

Enter your e-mail address:

SUBMIT CLEAR

DESTAMP: Designing of stable antibacterial mutant peptides

S.No.	Peptides	Prediction	Mutation Position	Half-life(s)	Antibacterial Activity (%)	HPLC parameter	Hydrophobicity (KJ/mol)	pKa	pKb	Residue volume	Molecular weight
Original peptide sequence											
0	DKADSFGPLMNCERT	Antibacterial Peptide	NO	0.393	71.329	45.100	193.000	31.720	141.480	1953.500	1936.060
Mutant peptide sequences predicted to be antibacterial. Mutant residues are colored RED.											
Sorting											
1	AKADSFGPLMNCERT	Antibacterial Peptide	1	0.316	74.192	37.200	258.000	32.180	141.570	1931.000	1892.050
2	CKADSFGPLMNCERT	Antibacterial Peptide	1	0.345	73.814	36.500	263.000	31.800	142.160	1955.100	1924.110
3	EKADSFGPLMNCERT	Antibacterial Peptide	1	0.322	71.539	42.900	219.000	32.030	141.010	1983.200	1950.090
4	FKADSFGPLMNCERT	Antibacterial Peptide	1	0.383	81.156	25.900	303.000	31.670	141.010	2031.700	1968.150
5	GKADSFGPLMNCERT	Antibacterial Peptide	1	0.698	83.114	40.800	211.000	32.180	141.480	1902.700	1878.030
6	HKADSFGPLMNCERT	Antibacterial Peptide	1	0.327	73.580	37.200	169.000	31.660	141.050	1995.300	1958.120
7	IKADSFGPLMNCERT	Antibacterial Peptide	1	0.463	71.569	27.100	311.000	32.200	141.480	2011.200	1934.130
8	KKADSFGPLMNCERT	Antibacterial Peptide								2018.300	1949.150
9	LKADSFGPLMNCERT	Antibacterial Peptide								2011.200	1934.130
10	MKADSFGPLMNCERT	Antibacterial Peptide								2004.900	1952.170

Designing of antibacterial stable peptides

Scanning of peptides in a protein

Submition of multiple peptides

al Conference on Open Source for Computer Aided Drug Discovery (March 22-26, 2009

Computational Resources for Drug Discovery

OSDD Forum is an initiative with a vision to provide affordable healthcare to the developing world. The OSDD concept aims to synergize the power of genomics, computational technologies and facilitate the participation of young and brilliant talent from Universities and industry. It seeks to provide a global platform where the best brains can collaborate and collectively endeavor to solve the complex problems associated with discovering novel therapies for neglected diseases like Tuberculosis.

CRDD (Computational Resources for Drug Discovery) is an important module of the *in silico* module of OSDD. The CRDD web portal provides computer resources related to drug discovery on a single platform. Following are major features of CRDD:

- **CRDD provides computational resources for researchers in the field of computer-aided drug design.**
- **CRDD allows users to discuss their problem with other members.**
- **CRDD gives equal opportunity to those willing to solve these problems.**
- **CRDD Wiki maintain wikipedia related to drug discovery.**
- **Contributors may host their database or web server on CRDD portal.**

Thus, CRDD provides a platform for researchers having limited resources.

Software name	Description
AntiCP	Prediction and design of anticancer peptides.
Toxinpred	Prediction and designing of toxic/non-toxic peptides.
AntiBP	Mapping of antibacterial peptides in a protein sequence.
AntiBP2	Advanced server for predicting antibacterial peptides with high precision.
CPPsite	CPPsite is a database of experimentally validated cell penetrating Peptides
CellPPD	Computer-aided Designing of efficient cell penetrating peptides.
TumorHoPe	A comprehensive database experimentally characterized tumor homing peptides.
TumorHPD	Server dedicated for designing tumor homing peptides.
Destamp	Designing of stable antibacterial peptides
HEMOLYTIK	A repository of experimentally validated hemolytic and non-hemolytic peptides.
ParaPep	Compilation of experimentally validated anti-parasitic peptides and their structure

Chemoinformatics and Pharmacoinformatics

Web Server	Description
DrugMint	A Server for Identification of Drug-like Molecules
ABMPred	Prediction of AntiBacterial Compounds against MurA Enzyme
MDRIpred	Prediction of Inhibitor against Drug Resistant M.Tuberculosis
DMKpred	Prediction of Drug molecules for kinase protein
KiDoQ	Prediction of inhibition constant of a molecule against Dihydrodipicolinate synthase enzyme
TOXIpred	Prediction of aqueous toxicity of small chemical molecules in T. pyriformis.
MetaPred	Prediction of cytochrome P450 isoform responsible for metabolizing a drug molecule.
GDoQ	Model for prediction of GLMU inhibitors using QSAR and docking apprach.
KetoDrug	Binding affinity prediction of ketoxazole derivatives against fatty acid amide hydrolase.
WebCDK	Web Interface for CDK libraries
TLR4HI	SVM based model for computing inhbitors against human TLR4 (Toll like receptor).
DMKPred	A webserver for the prediction of binding of chemical molecules with specific kinases.
ntEGFR	Predicting and designing imidazothiazoles/pyrazolopyrimidines based inhibitors against wild/mutant EGFR.
CancerIn	Classification and designing of anti-cancer inhibitors.
EGFRpred	Prediction of inhibitor of anti-EGFR molecules of diverse class.
DiPCell	Designing of inhibitors against pancreatic cancer cell lines.
HIVfin	Prediction of fusion protein inhibitors against HIV.

Molecular Interactions

Software name	Description
ADPint	Prediction of ADP interacting residues in a protein.
ATPint	Identification of ATP binding sites in ATP-binding proteins.
DOMprint	SVM based model for predicting domain-domain interaction (DDI).
GlycoEP	Prediction of C-, N- and O-glycosylation site in eukaryotic proteins.
GlycoPP	Prediction of potential N-and O-glycosites in prokaryotic proteins.
GTPbinder	Identification of GTP binding residue in protein sequences.
MYCOprint	A tool fort exploration of the interactome of Mycobacterium tuberculosis.
NADbinder	Prediction of NAD binding proteins and their interacting residues.
Pprint	ANN based method for identification of RNA-interacting residues in a protein.
PreMieR	Identification of mannose interacting residues (MIRs) in protein sequences.
PROprint	Prediction of physical/functional interaction between two protein molecules.
RNApin	A server for the prediction of protein interacting nucleotides in RNA sequences.
tRNAmod	Prediction of post transcriptional modifications in transfer-RNA (tRNA) sequence.
VitaPred	Identification of different class of vitamin interacting residues in a protein.

Biological Databases

Database name	Description
MHCBN	A curated database of MHC-binding, Non-binding peptides and T-cell epitopes.
Bcipep	A database of B-cell epitopes.
HaptenDB	A database of hapten molecules that can not activate immune system.
PolysacDB	Compilation of antigenic polysaccharides found on surface of microbial organism.
TumorHope	A database of experimentally characterized tumor homing peptides.
AntigenDB	Information about a wide range of experimentally-validated antigens.
OXDBase	Compilation of oxygenases involved in the biodegradation on xenobiotic compounds.
HMRBase	A manually curated database of hormones and their Receptors.
CPPsite	Compilation of experimentally validated Cell Penetrating Peptides (10-30 amino acids).
BIAdb	Information about Benzylisoquinoline Alkaloid molecules
HIVsir	A manually curated database of anti-HIV siRNAs.
CCDB	Catalog of genes involved in the different stages of cervical carcinogenesis.
ProGlycProt	Repository of experimentally characterized eubacterial and archaeal glycoproteins.
NPACT	A database of plant derived natural compounds that exhibit anti-cancerous activity.
CancerDR	Compilation of anticancer drugs and their effectiveness against various cancer cell lines.
ccPDB	Compilation and creation of datasets from PDB for structural/functional annoation of proteins.
ParaPep	HIPdb is a manually curated database of experimentally validated antiparasite peptides.
EGFRindb	Collection of EGFR inhibitors from literature.
CancerPPD	Collection and compilation of experimentally validated anticancer peptides
PCMdb	Pancreatic cancer methylation database provides large scale collection of metylated genes.
HerceptinR	Information about assays performend to test sensitivity/resistance of Herceptin Antibody.
HemolytiK	A resource of experimentally tested hemolytic peptides.
CancerTope	A database of epitopes found in protein involved in cancer.
AHTPDB	AHTPDB is an ideal platform for complete & relevant information for large number of antihypertensive peptides

Genome Annotations

Server Name	Description
FTG	Locating probable protein coding region in nucleotide sequence using FFT based algorithm.
GWBLAST	Genome wide similarity search using BLAST
GWFASTA	Genome Wise Sequence Similarity Search using FASTA.
EGPred	Prediction of gene (protein coding regions) in eukaryote genomes that includes introns/exons.
SVMgene	SVM based approach to identify the protein coding regions in human genomic DNA.
SRF	Find repeats through an analysis of the power spectrum of a given DNA sequence.
MyPattern	A program for detection of a 'motif' in DNA sequence using an exact search method.
GeneBench	A suite of datasets and tools for evaluating gene prediction methods.
FTGPred	A web server for predicting genes in a DNAsequence.
PHDcleav	Prediction of Human Dicer cleavage sites.
PolyApred	Prediction of polyadenylation signal (PAS) in human DNA sequence.
siRNAPred	Predicting actual efficacy of both 21mer and 19mer siRNAs with high accuracy.
ECGPred	Analsis of expresion data and correlation between gene expression and nucleotides composition of genes.
desiRam	Designing of highly efficient siRNA with minimum mutation approach
MARSpred	Discriminating between Mitochondrial and Cytosolic Aminoacyl tRNA Synthetases
Icaars	Identification & Classification of Aminoacyl tRNA Synthetases.
LGEpred	Prediction of correlation between amino acid residue and gene expression level.

Immunoinformatics or Vaccine Informatics

Software name	Description
T-Helper Epitopes or MHC/HLA Class II binders (Adaptive Immunity, Exogenous Antigen)	
MHCBN	A database of MHC-Binding, Non-binding peptides and T-cell epitopes.
ProPred	Identification of promiscuous MHC Class-II binding regions in an antigen sequence
HLA-DR4Pred	Identification of HLA-DRB1*0401(MHC class II alleles) binding peptides.
MHC	Matrix Optimization Technique for identification of binding core in MHC II binding peptides
MHC2pred	The MHC2Pred is an SVM based method for prediction of promiscuous MHC class II binding peptides.
MHCBENCH	Benchmarking of MHC binding peptide prediction algorithms.
FDR4	Prediction of binding affinity of HLA-DRB*0401 binders in an antigenic sequence.
IL4pred	In silico platform for designing and disovering of interleukin-4 inducing peptides.
IFnepitope	Designing of interferon-gamma inducing epitopes.
CTL Epitopes or MHC/HLA Class I binders (Adaptive Immunity, Endogenous Antigens)	
PROPRED1	Prediction of promiscuous binders for 47 MHC/HLA class I alleles using quantitative matrices;
Pcleavage	Identification of protesosomal cleavage sites in a protein sequence.
TPPred	Prediction of TAP binding peptides for understanding of peptide internalization to endoplasmic reticulum
CTLPred	A direct method for prediction of CTL epitopes.
nHLApred	This is a comprehensive method for prediction of MHC binding peptides or CTL epitopes of 67 MHC class alleles.
MMBPred	Prediction of mutated MHC class I binders in an antigen, having high affinity and promiscuousity.
HLAPRED	The method can identify and predict HLA (both class I & II) binding regions in an antigen sequence.
Linear & Conformational B-cell Epitopes	
BCIPEP	Collection & compilation of B-cell epitopes from literature
BCEPRED	Prediction of linear B-cell epitopes, using Physico-chemical properties
ABCPred	Mapping of B-cell epitope(s) in an antigen sequence, using artificial neural network.

Functional Annotation of Proteins

Server name	Description
NRpred	Prediction and classification of nuclear receptors, SVM models based on composition.
GPCRpred	Prediction of families and superfamilies of G-protein coupled receptors (GPCR)
ESLpred	Subcellular localization of the eukaryotic proteins using dipeptide compostion and PSI-BLAST.
PSLPred	Prediction of subcellular localization of bacterial proteins
BTXPred	It predicts bacterial toxins and their function from primary amino acid sequence.
GPCRsclass	This webserver predicts amine type of G-protein coupled receptors
Mitpred	Specifically trained to predict mitochondrial proteins with high accuracy
Oxypred	Classification and prediction of oxygen binding proteins.
VGIchan	Classification and prediction of proteins involved in voltage gated ion channels.
HSLpred	Subcellular localization of human proteins with high accuracy
DNAsize	Compute length of DNA or protein fragments from gel using a graphical method.
GSTpred	SVM-based method for predicting Glutathione S-transferase protein.
Mango	A server for predicting functional class of a protein.
LGEpred	Calculate correlation coefficient between amino acid residue and gene expression level.
NTXpred	Identification of neurotoxins their source and function from primary amino acid sequence.
VICMpred	Classification of bacterila proteins particularly virulent proteins
ALGpred	Prediction of allergenic proteins and mapping of IgE epitopes in antigens.
PseaPred	Prediction of proteins secreted by Malarial Parasite P. falciparum into infected-erythrocyte.
RSLPred	SVM based method for subcellular localizaton of rice proteins.
COPid	Composition based identification and classification of proteins.
ESLPred2	Advanced method for subcellular localization of eukaryotic proteins.
ISSpred	Identification of Inteins hiding in their protein sequences.
CyclinPred	CyclinPred is a SVM based prediction method to identify novel cyclins.

Proteins Structure Prediction

Web Server	Description
AlphaPred	A neural network based method for predicting alpha-turn in a protein.
APSSP2	Prediction of secondary structure of proteins from their amino acid sequence.
AR_NHPred	Identification of aromatic-backbone NH interaction in protein residues.
BetatPred	Statistical-based method for predicting Beta Turns in a protein.
Betatpred2	Prediction of Beta-turns with high accuracy using multiple sequence alignment.
BetaTurns	It predict different types of beta-turns (e.g., Types I/II/IV/VIII) in a protein.
BhairPred	Prediction of beta hairpins in proteins using ANN and SVM techniques.
CHpredicts	Prediction of CH-O, CH-PI interactions in backbone residues of a protein
GammaPred	Identification of gamma-turn containing residues in a given protein sequence.
PEPstr	Prediction of tertiary structure of small peptides (7 to 25 residues).
Proclass	Classification of proteins based on secondary structure contents.
PSA	Analyze the amino acid sequence and multiple sequence alignment of proteins.
RPFOLD	A fold recognition server for searching protein fold in PDB.
SARpred	ANN-model for redicting real-value of surface acessibility of protein residues.
TBBpred	This server predict Transmembrane Beta Barrel regions in a protein.
PEP2D	This server allows you to predict secondary structure of peptides.

HOME | ABOUT | FEEDBACK | FAQ | MIRROR SITE | LICENSE | REGISTER | THANKS | TEAM | RAGH

OSDD LINUX

A Custo
fo

Operating System for Drug Discovery

- General Inoformation
- Software Packages
- Install/Download
- Service to Community
- OSDDlinux Online
- Important Resources

Install/Download

- Live DVD/USB
- Full Installation
- Virtual Box
- On existing machine
- Package Repository
- Upgrade/New Packages

General Inoformation

- Major Features
- Installation Guide
- Users Guide
- Drug Discovery manual
- GPSR manual
- Required Package
- List of Packages

Service to Community

- Command Mode
- Web Services
- Galaxy Portal
- GUI-based Software

Software Packages

- Bioinformatics
- Vaccine Informatics
- Drug Informatics
- Biotherapeutics
- Analysis of NGS data
- Education & Research
- Basic Scripts

Thankyou