Vol 2 | Issue 1 | Jan –Jun 2011 | 52-62. **Print ISSN – 2229 7502**

Research Article

TOXICITY EVALUATION USING NOVEL INSILICO SCREENING SYSTEM FOR PREDICTING MUTAGENICITY AND CARCINOGENICITY

A.Padmanabha Rao1, R.Ravinder[2*], K.Surendranath[2], A.N.Singh[2], Kaiser Jamil[1], Prakash V Diwan[3]

[1]School of Biotechnology (MGNIRSA) University of Mysore, India- 500 029.
[2] Celestial Biolabs, Banjarahills, Hyderabad- 500 038.
[3]Department of Pharmacology, Lalitha College of Pharmacy, Ghatkesar, Hyderabad-501 301.

ABSTRACT

Drug Discovery involves synthesis of new chemical entities which are devoid of toxicity. The animal experiments are not economical and also time consuming. Hence the study was aimed to develop an authenticated software program for predicting the carcinogenicity and mutagenicity. The study summarizes the evidence on the structural alerts for mutagenicity and carcinogenicity. The structural alerts are the molecular substructures or fragments that are related to the toxic properties of the chemicals, and represent a codification of a long series of studies aimed at highlighting the mechanism of toxicity. The identification of the structural alerts has a great potential in terms of both understanding mechanisms, and also in assessing the toxicity possed by the new chemical entities. The program is based on the structural facts and was examined by taking the available database and confirmed the results. Based on the available and established tools the rules were used to construct a robust and predictable system to categorize the toxic end points. The predictability of the system developed was monitored using 150 molecules consisting of toxic and non toxic entities from CPDB database. The test results are encouraging with maximum predictability of > 85% of selectivity with higher rate of concordance. The present work is aimed at expanding and refining the knowledge on the structural alerts by using expert system.

Key words: Expert System, Carcinogenicity, Mutagenicity, Structural Alerts,Insilico Screening System, Toxicophores.

INTRODUCTION

The pharmaceutical industry is involved in the drug discovery programs for the cure of deadly diseases. In the process the chemists have used the modern tools to synthesize thousands of compounds. It is mandatory to evaluate the toxicity before subjecting to the preclinical studies. It is very difficult to screen large number of chemicals by using animal experimentation. Hence a theoretical full proof method is proffered to identify the toxic molecules. The present scenario of the artificial Intelligent systems is becoming popular day by day with the advent of high end computer based intensive methods for various stages of drug discovery e.g., insilico toxicity prediction. Computer aided drug designing tools are playing a vital role in generating a large numbers of new chemical entities based on analogy and/or given the target structure. The human dependence on the intelligent systems is paving the way for errors, which might incur huge loss and they must be addressed strongly by utilizing existing knowledge. The objective is to provide efficient toxicity screening system with accuracy and speed (batch screening) to simulate today's industrial requirements.

The expert systems are being popularly exploited at different stages of screening new chemical entities for their toxicity profiles. Expert system is a knowledge-based computer aided screening system, which is embedded with chemical profile in the form of rules to guide the chemists in drug discovery [1]. It mainly involves rules derived from the existing knowledge, by minimizing the time to identify the lead molecule. The literature survey reveals that toxicity predicting expert systems viz., DEREK [2], Hazard Expert [3, 4] COMPACT [5] are limited to few end points.

Corresponding Author :- **A.Padmanabha Rao Email:-** padmanabha.rao.a@gmail.com

However, the experimental tests for predicting the toxicity profiles of chemicals is available, they are expensive and require the use of animals, which needs animal ethical approvals. Recently many artificial intelligence methods have been used in conjunction with diverse combinations of statistical protocols for better accuracy in predicting the toxicity [6]. The presently available guidelines to predict the toxicity are based on the rules drawn by scientific experts, using statistical methods or mechanistic processes [7].

Our work is aimed at exploiting the available up-to-date knowledge and utilizing the same with effective toxicophores matching search engine for predicting the toxic profiles of huge sized databases consisting of new chemical entities. The batch screening is the need of the hour as the size of experimental data is increasing in exponential proportions. The newly developed insilico tool APEXPERT is a knowledge driven, and the data collected plays a key role in its predictability.

MATERIALS AND METHODS

APEXPERT is an in-house developed insilico toxicity prediction module, which predicts toxicity of NCE's based on well defined rules. The rules considered for prediction are in the form of structural patterns/Toxicophores. The batch processing of the molecules is enabled as it runs on Linux clusters. The toxic fragments database was developed by using literature survey and salient features of toxicophore were identified [8, 9 and 10]. A vast literature was scanned for mutagenicity [11, 12, 13, 14, 5, 15 and 16] and carcinogenicity [5, 15, 17, 18 and 19] structural class alerts and they are used in our database.

Expert system requires experimental data stored in the form of rules viz., toxicophores falling in different structural classes' viz., aromatic amines, aromatic nitro groups, epoxides etc. [15] and a fragment matching/search engine which helps in identifying respective toxicophores was developed.

Based on the database information the present module was designed and the applications were subjected for evaluating the probable toxic moieties in new chemical entities. The method developed searches for the identical substructures in the new chemical moieties and if identified will be ear marked as toxic substances. The study was conducted by using toxic and non toxic experimentally proven mutagenic and carcinogenic compounds. To understand the predictability of the new developed method was validated by using statistical parameters.

ITMP (Insilico Toxicophores Matching Program) Working Module

ITMP is s software program to find identical matches in a given chemical structure. The strength of our program lies in considering the atom types, bond types, connectivities and above all the stereochemistry while finding the identical matches. The logic behind the working of ITMP is presented in the form of a flow diagram (Figure 1). The current version of ITMP supports only sybyl mol2 file format. The NCE's and toxic fragment must be converted to sybyl mol2 file format before subjecting them to ITMP. ITMP intelligently decodes the information stored in the sybyl mol2 file formats and extracts information relating to atom types, bond types and connectivities for toxic fragments and as well NCE's. The matching/search engines associated with expert systems mainly try to find the toxicophores in a given set of NCE's (New Chemical Entities). The similarity/diversity of the match is normally calculated using tanimoto co-efficient and euclidean distance respectively [20]. Keeping in view a robust search engines was developed using Shape and stereochemistry for similarity search.

The information in the mol2 files will be taken as basis and n X n' (where n or n'= no of atoms in a molecule) matrix is built for each and every toxic fragments and NCE's. The population of values in respective matrices will be on the basis of atom type, bond type and the connectivity between the atoms in a particular chemical structure. Thus a Connection Table Matrix (CTM) is built and matching of chemical structures will be done at matrices level.

The nxn' matrix for a particular chemical structure contains encoded values on the diagonal of the matrix basing on atom type. The non-diagonal values are filled basing on atom-atom connectivity, for example the atoms present in a (i_1j_1) and a' (i_3j_3) positions in the matrix are two bonds for the corresponding value will be two and if two respective atom under consideration are not connected, then the corresponding value will be zero. As the matrix comparison is computer intensive the populated values will be rearranged and higher values will be brought to left top end of matrix and therefore ranking of CTM is done. The toxicophores to NCE's matching will be done and the predicted toxicity profiles will be determined. The results were analyzed and the predictions are grouped into four categories viz. True positive, True negative, False positive and False negatives. The criteria for categorizing the results are presented in Table 1.

RESULTS AND DISCUSSION

The two-dimensional structures from CPDB with mutagenicity and carcinogenicity profiles were evaluated and the respective results are presented in Table 2 and Table 3. The concordance of APEXPERT results with that

of CPDB was encouraging and found to be 85% accurate when compared with commercially available systems. The categorizations of results are presented in Figure 2. The study was validated with CPDB database and the results are depicted in table-4.

Rule-based expert systems are suitable when some of the toxic information is uncertain or even unknown. The routine procedure utilizes the database related rules and confirms the toxicity by directly matching the available similar data [15, 17]. Today the insilico systems are available to synthesize thousands of compounds by using automated synthesis for identifying a lead molecule [21].

This may not be a true picture of toxicity of the compound. Hence more parameters have to be included to avoid false positive and false negative results. Based on the lead molecule obtained using the software tool the results can be confirmed by performing experiments on animals. The expert systems for toxicity screening must be thoroughly evaluated with laboratory experiments to authenticate the predicted results. The sensitivity of any knowledge-based system directly depends on the quality and quantity of the data it harbors.

Figure 1: A FLOW DIAGRAM OF ITMP ALGORITHM.

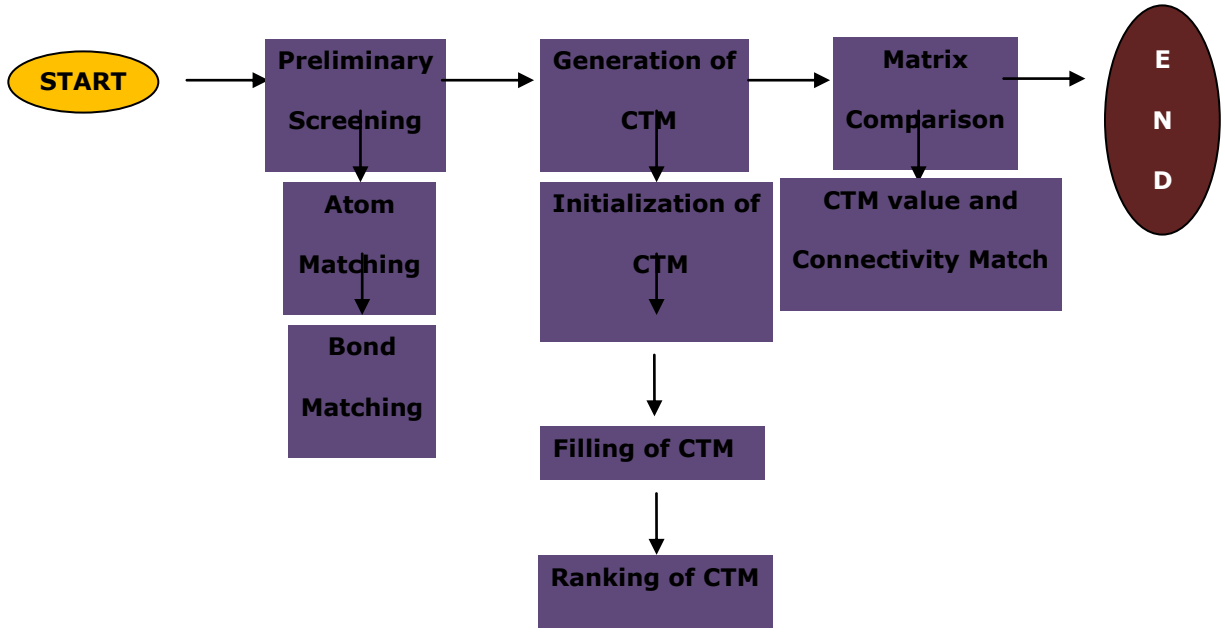

Figure 2: Categorization of APEXPERT results for Mutagenicity and Carcinogenicity

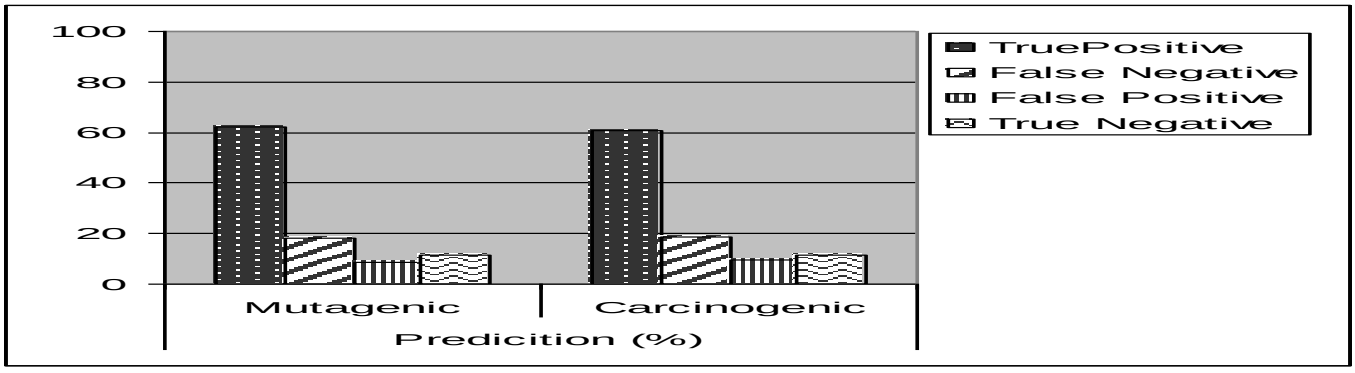

Table: 1 Categorization of APEXPERT results for calculating sensitivity and selectivity.

Experimental Result	Predicted Result	Categorization of result
Mutagenic/Carcinogenic	Mutagenic/Carcinogenic	True Positive (TP)
NonMutagenic/Non Carcinogenic	Mutagenic/Carcinogenic	False Positive (FP)
Mutagenic/Carcinogenic	NonMutagenic/Non Carcinogenic	False Negative (FN)
NonMutagenic/Non Carcinogenic	NonMutagenic/Non Carcinogenic	True Negative (TN)

Table 2: APEXPERT Predictions for Mutagenicity of 150 selected compounds from CPDB database.

Compound Number	CAS Number of Test Molecules	Toxicity Profile		Structural Alert Classes
		CPDB*	APEXPERT+	
1	1728-95-6	(+)	(-)	
2	91-08-7	(+)	(-)	
3	120314-14-9	(+)	(-)	
4	146795-42-8	(+)	(+)	Aromatic Halo compounds
5	614-00-6	(+)	(+)	Aromatic amide
6	54378-38-9	(+)	(+)	Aromatic Amines
7	1836-75-5	(+)	(+)	Aromatic Nitro Compounds
8	146177-59-5	(+)	(-)	
9	106264-79-3	(+)	(+)	Aromatic Amines
10	61203-01-8	(+)	(+)	Aliphatic Aldehydes
11	112-24-3	(+)	(+)	Aromatic Amines
12	50438-75-0	(+)	(+)	Aromatic Amides
13	126983-60-6	(+)	(+)	Aliphatic Aldehydes
14	97606-15-0	(+)	(+)	Aromatic Amines
15	952-23-8	(+)	(+)	Hetero Atomic Aromatic Nitrogen
16	3807-77-0	(+)	(+)	Aromatic Nitro Compounds
17	71628-96-1	(+)	(+)	Aromatic amines
18	4629-58-7	(+)	(+)	Aromatic Nitro compounds
19	103942-97-8	(+)	(+)	Aromatic Amines
20	95-64-7	(+)	(-)	
21	21655-84-5	(+)	(+)	Hetero Atomic Aromatic Nitrogen
22	67587-52-4	(+)	(+)	Hydrazine Compound & Hetero Atomic Aromatic Nitrogen
23	610-14-0	(+)	(+)	Aromatic Nitro
24	140-08-9	(+)	(+)	Aliphatic Halo
25	10605-21-7	(+)	(-)	
26	68743-68-0	(+)	(+)	Aromatic Amide
27	505-29-3	(+)	(-)	
28	135-68-2	(+)	(+)	Aromatic Amine
29	484-23-1	(+)	(-)	
30	6149-33-3	(+)	(+)	Aromatic Nitro
31	91-23-6	(+)	(+)	Aromatic Nitro
32	218-19-9	(+)	(+)	Metal Acids, Aromatic Polycyclic
33	6449-35-0	(+)	(+)	Aromatic Azo
34	51640-90-5	(+)	(+)	Aromatic Nitro
35	2541-69-7	(+)	(+)	Metal Acids, Aromatic Polycyclic
36	4450-68-4	(+)	(+)	Aromatic Nitro
37	77121-90-5	(+)	(+)	Aliphatic Aldehydes
38	138580-48-0	(+)	(+)	Aromatic Amines
39	6268-09-3	(+)	(+)	Aromatic Ketones
40	96563-10-9	(+)	(+)	Aromatic Halogens
41	2508-20-5	(+)	(+)	Aromatic Nitroso
42	96-99-1	(+)	(+)	Aromatic Nitro, Aromatic Halogens
43	74-31-7	(+)	(+)	Aromatic amine
44	142044-37-9	(+)	(+)	Aliphatic Epoxide
45	615-05-4	(+)	(+)	Aromatic Amine, Aromatic Ethers
46	146177-65-3	(+)	(-)	
47	29916-56-1	(+)	(-)	
48	125239-64-7	(+)	(+)	Aromatic Ethers
49	156215-58-6	(+)	(-)	
50	578-06-3	(+)	(-)	

Compound Number	CAS Number of Test Molecules	Toxicity Profile		Structural Alert Classes
		CPDB*	APEXPERT+	
51	3442-62-4	(+)	(+)	Aromatic Nitro
52	10147-37-2	(+)	(-)	
53	62450-07-1	(+)	(+)	Hetero Atomic Aromatic Nitrogen
54	161697-03-6	(+)	(-)	
55	99-09-2	(+)	(+)	Aromatic Nitro, Aromatic Amines
56	41533-75-9	(+)	(+)	Aromatic Nitro
57	113698-18-3	(+)	(+)	Enols
58	123-73-9	(+)	(+)	Aliphatic Aldehydes
59	3930-19-6	(+)	(+)	Enols
60	75586-69-5	(+)	(-)	
61	69177-41-9	(+)	(+)	Aromatic Amines
62	116505-02-3	(+)	(+)	Mutliple Fuction Group
63	96910-73-5	(+)	(+)	Aliphatic Aldehydes
64	90-13-1	(+)	(+)	Aromatic Halo compound
65	212-54-4	(+)	(+)	Metal Acids/Salts, Aromatic Polycyclic
66	95-70-5	(+)	(+)	Aromatic Amines
67	512-56-1	(+)	(+)	Metal Acids/Salts
68	3018-12-0	(+)	(+)	Aliphatic Halo compounds
69	146795-37-1	(+)	(+)	Aromatic Halo
70	104445-28-5	(+)	(+)	Aromatic Amines
71	106063-42-7	(+)	(+)	Aromatic Ethers
72	119-36-8	(+)	(-)	
73	15131-84-7	(+)	(+)	Aromatic Epoxides
74	610-39-9	(+)	(+)	Aromatic Nitro
75	112575-91-4	(+)	(+)	Aromatic Aldehydes
76	1672-88-4	(+)	(-)	
77	105650-23-5	(+)	(-)	
78	14925-39-4	(+)	(+)	Aliphatic Aldehydes
79	61740-00-9	(+)	(+)	Enols
80	85-33-6	(+)	(-)	
81	67730-11-4	(+)	(+)	Hetero Atomic Aromatic Nitrogen
82	68162-37-8	(+)	(+)	Aliphatic Carboxylic Acid
83	64188-64-3	(+)	(+)	Aromatic Amines
84	114454-59-0	(+)	(+)	Metal Acids/Salts
85	2426-08-6	(+)	(+)	Aliphatic Epoxide
86	127-07-1	(+)	(-)	
87	2226-96-2	(+)	(-)	
88	54818-87-0	(+)	(-)	
89	80115-73-7	(+)	(+)	N-Methylol Compounds
90	501-30-4	(+)	(-)	
91	2224-15-9	(+)	(+)	Aliphatic Epoxide
92	151-56-4	(+)	(+)	Aromatic Amines
93	52898-98-3	(+)	(+)	Aliphatic Aldehydes
94	139953-77-8	(+)	(-)	
95	21229-99-2	(+)	(-)	
96	73341-53-4	(+)	(+)	Hetero Atomic Aromatic Nitrogen.
97	109-09-1	(+)	(-)	
98	396-32-7	(+)	(+)	Aromatic Halo Compounds
99	112022-06-7	(+)	(+)	Aromatic Ketones, Amines
100	2562-38-1	(+)	(-)	

Compound Number	CAS Number of Test Molecules	Toxicity Profile		Structural Alert Classes
		CPDB*	APEXPERT+	
101	106-50-3	(+)	(+)	Aromatic Amines
102	33611-88-0	(+)	(+)	Aromatic Nitro
103	100-13-0	(+)	(+)	Aromatic Nitro
104	5907-38-0	(+)	(+)	Aromatic Ketones
105	3688-53-7	(+)	(+)	Aliphatic Aldehydes
106	146795-43-9	(+)	(+)	Aromatic Halo
107	89-80-5	(+)	(+)	Aliphatic Aldehydes
108	62-50-0	(+)	(+)	Metal Acids/Salts
109	159394-73-7	(+)	(+)	Aromatic Nitro
110	18829-55-5	(+)	(+)	Aliphatic Aldehydes
111	820-75-7	(+)	(+)	Hydroxyl Amines
112	28166-06-5	(+)	(+)	Aromatic Nitro, Halo
113	96-09-3	(+)	(+)	Aromatic Epoxides
114	16430-32-3	(+)	(+)	Aromatic Polycyclic
115	1516-21-8	(+)	(+)	Aromatic Nitroso
116	2185-92-4	(+)	(+)	Aromatic amines
117	4812-40-2	(+)	(+)	Hetero Atomic Aromatic Nitrogen
118	644-21-3	(+)	(+)	Hydrazine
119	102059-18-7	(+)	(+)	Aliphatic Aldehydes
120	96-45-7	(+)	(+)	Thio Compounds
121	58-08-2	(-)	(-)	
122	58-93-5	(-)	(+)	Aromatic Halo Compound
123	58-95-7	(-)	(-)	
124	59-67-6	(-)	(-)	
125	60-35-5	(-)	(-)	
126	60-56-0	(-)	(-)	
127	60-80-0	(-)	(-)	
128	61-82-5	(-)	(-)	
129	62-38-4	(-)	(-)	
130	62-53-3	(-)	(-)	
131	62-55-5	(-)	(-)	
132	62-56-6	(-)	(-)	
133	64-17-5	(-)	(-)	
134	64-75-5	(-)	(+)	Aliphatic Aldehydes, Enols
135	64-77-7	(-)	(-)	
136	66-22-8	(-)	(-)	
137	67-21-0	(-)	(+)	Aliphatic Aldehydes
138	67-48-1	(-)	(+)	Aromatic Amines
139	67-66-3	(-)	(+)	Aliphatic Halo
140	69-65-8	(-)	(-)	
141	71-43-2	(-)	(-)	
142	72-33-3	(-)	(-)	
143	72-55-9	(-)	(+)	Aromatic Halo
144	73-22-3	(-)	(+)	Aliphatic Aldehydes
145	77-65-6	(-)	(+)	Aliphatic Halo Compounds
146	597-25-1	(-)	(+)	Metal Acids/Salts
147	634-93-5	(-)	(+)	Aromatic Amines/Halo
148	671-16-9	(-)	(+)	Hydrazine
149	756-79-6	(-)	(+)	Metal Acids/Salts
150	7722-84-1	(-)	(+)	Peroxy Compounds

* CPDB: Carcinogenic Potency Data Base
+ APEXPERT: AP Toxicophore Identification Program.
(+) Toxic (-) Nontoxic

Table 3: APEXPERT Predictions for Carcinogenicity of 150 selected Compounds from CPDB database.

Compound Number	CAS Number of Test Molecules	Toxicity Profile		Structural Alert Classes
		CPDB*	APEXPERT+	
1	60-11-7	(+)	(+)	Aromatic Polycyclic
2	67-66-3	(+)	(+)	Aliphatic Halogens
3	56-53-1	(+)	(+)	Aromatic Polycyclic
4	92-87-5	(+)	(+)	Aromatic Amide
5	53-96-3	(+)	(+)	Aromatic Aldehydes
6	50-32-8	(+)	(+)	Aromatic Polycyclic
7	10034-93-2	(+)	(+)	Hydrazine
8	86-30-6	(+)	(+)	Aromatic Nitroso
9	96-45-7	(+)	(+)	Aromatic Nitro
10	3068-88-0	(+)	(+)	Aliphatic Aldehydes
11	94-59-7	(+)	(+)	Aromatic Alcohols
12	106-89-8	(+)	(+)	Aliphatic Saturated Hydrocarbons
13	59-89-2	(+)	(+)	Aliphatic Nitroso
14	57-57-8	(+)	(+)	Aliphatic Aldehydes
15	51-79-6	(+)	(+)	Aliphatic Amides
16	61-82-5	(+)	(-)	
17	101-14-4	(+)	(+)	Aromatic Amines
18	636-21-5	(+)	(+)	Aromatic Amines
19	57-06-7	(+)	(-)	
20	106-99-0	(+)	(-)	
21	2832-40-8	(+)	(+)	Aliphatic Amides
22	106-93-4	(+)	(+)	Aliphatic halogens
23	101-90-6	(+)	(+)	Aliphatic Saturated Hydrocarbons
24	140-88-5	(+)	(-)	
25	108-78-1	(+)	(-)	
26	150-68-5	(+)	(+)	Aliphatic Amides
27	75-56-9	(+)	(+)	Aliphatic Saturated Hydrocarbons
28	542-75-6	(+)	(+)	Aliphatic unsaturated Hydrocarbons
29	121-66-4	(+)	(-)	
30	20265-96-7	(+)	(+)	Aromatic amines
31	563-47-3	(+)	(-)	
32	3-(Chloromethyl) Pyridine Hydrochloride	(+)	(-)	
33	4-Chloro-metaphenylenediamine	(+)	(+)	Aromatic amines
34	meta-Cresidine	(+)	(+)	Aromatic Alcohols
35	105-55-5	(+)	(-)	
36	91-93-0	(+)	(+)	Aromatic Alcohols
37	Dimethyl morpholinophosphoramidate	(+)	(+)	Aliphatic Amines
38	123-91-1	(+)	(+)	Aromatic Ether
39	67-72-1	(+)	(+)	Aliphatic Halogens
40	78-59-1	(+)	(+)	Aliphatic Aldehydes
41	139-94-6	(+)	(+)	Aliphatic Halogens
42	139-13-9	(+)	(+)	Aliphatic Aldehydes
43	1955-45-9	(+)	(+)	Aliphatic Aldehydes
44	105-11-3	(+)	(-)	
45	1596-84-5	(+)	(+)	Aliphatic Aldehydes
46	95-06-7	(+)	(+)	Aliphatic Unsaturated Hydrocarbons
47	127-18-4	(+)	(+)	Aliphatic Unsaturated Hydrocarbons
48	88-06-2	(+)	(+)	Aromatic Halogens
49	137-17-7	(+)	(+)	Aromatic Polycyclic
50	1,1,3-trimethyl-2-thiourea	(+)	(-)	

Compound Number	CAS Number of Test Molecules	Toxicity Profile		Structural Alert Classes
		CPDB*	APEXPERT+	
51	126-72-7	(+)	(+)	Aliphatic Halogens
52	107-13-1	(+)	(-)	
53	542-88-1	(+)	(+)	Aliphatic Ethers
54	107-30-2	(+)	(+)	Aliphatic Ethers
55	75-21-8	(+)	(-)	
56	56-23-5	(+)	(+)	Aliphatic Halogens
57	107-06-2	(+)	(-)	
58	75-01-4	(+)	(+)	Aliphatic Unsaturated Hydrocarbons
59	75-27-4	(+)	(+)	Aliphatic Halogens
60	75-09-2	(+)	(+)	Aliphatic Halogens
61	513-37-1	(+)	(+)	Aliphatic Halogens
62	106-88-7	(+)	(+)	Aliphatic Saturated Hydrocarbons
63	5989-27-5	(+)	(+)	Aliphatic Polycyclic
64	24382-04-5	(+)	(-)	
65	598-55-0	(+)	(+)	Aliphatic Amides
66	67-20-9	(+)	(+)	Multiple functional groups
67	59-87-0	(+)	(+)	Multiple functional groups
68	75-25-2	(+)	(+)	Aliphatic Halogens
69	79-06-1	(+)	(+)	Aliphatic Halogens
70	759-73-9	(+)	(+)	Aliphatic Halogens
71	52-24-4	(+)	(+)	Aliphatic Halogens
72	68-76-8	(+)	(+)	Aliphatic Halogens
73	55-86-7	(+)	(+)	Aliphatic Halogens
74	126-85-2	(+)	(-)	
75	119-84-6	(+)	(+)	Aromatic Carboxylic Acid
76	1825-21-4	(+)	(+)	Aromatic Halogens
77	115-96-8	(+)	(-)	
78	91-64-5	(+)	(-)	
79	75-07-2	(+)	(+)	Aliphatic Aldehydes
80	60-35-5	(+)	(+)	Aromatic Esters
81	Acetone [4-(5-nitro-2-furyl)-2-thiazolyl] hydrazone	(+)	(-)	
82	Acetoxime	(+)	(-)	
83	1'-Acetoxysafrole	(+)	(+)	Aliphatic Aldehydes
84	4-Acetylaminobiphenyl	(+)	(+)	Aliphatic Amides
85	Aflatoxin B1	(+)	(+)	Aromatic Aldehydes
86	67730-10-3	(+)	(-)	
87	301-04-2	(+)	(+)	Aliphatic Aldehydes
88	56-49-5	(+)	(+)	Aromatic Aldehydes
89	19408-74-3	(+)	(+)	Aromatic halogens
90	3778-73-2	(+)	(+)	Aliphatic Amines
91	54-85-3	(+)	(+)	Aliphatic Amides
92	4-(4-(N-methyl-N-nitrosoamino)styryl) quinoline	(+)	(+)	Aromatic Nitroso
93	7758-01-2	(+)	(-)	
94	70-25-7	(+)	(+)	Mutiple Functinal Groups
95	N-methyl-N-nitrosobenzamide	(+)	(+)	Aliphatic Halogens
96	Dinitrosohomopiperazine	(+)	(+)	Aliphatic Nitroso
97	Nifurdazil	(+)	(+)	Mutiple functional Groups
98	25013-16-5	(+)	(+)	Aromatic Alcohols
99	3570-75-0	(+)	(+)	Hetero Atomic Aromatic (Oxygen)
100	3775-55-1	(+)	(-)	

Compound Number	CAS Number of Test Molecules	Toxicity Profile		Structural Alert Classes
		CPDB*	APEXPERT+	
101	N-butyl-N-(4-hydroxybutyl)nitrosamine	(+)	(+)	Aliphatic Nitroso
102	1-Amyl-1-nitrosourea	(+)	(+)	Aliphatic Amides
103	4342-03-4	(+)	(+)	Aliphatic Amides
104	Formic acid 2-(4-methyl-2-thiazolyl) hydrazide	(+)	(+)	Aliphatic Amides
105	50-00-0	(+)	(-)	
106	1,3-Dibutyl-1-nitrosourea	(+)	(+)	Aliphatic Amides
107	Hexa(hydroxymethyl)melamine	(+)	(-)	
108	2-(2,2-Dimethylhydrazino)-4-(5-nitro-2-furyl) thiazole	(+)	(+)	Hetero Atomic Aromatic (Nitrogen)
109	60391-92-6	(+)	(+)	Aliphatic Amides
110	25843-45-2	(+)	(-)	
111	N-(4'-fluorobiphenyl-4yl)acetamide	(+)	(+)	Aromatic Amides
112	Dimethylnitramine	(+)	(-)	
113	305-03-3	(+)	(+)	Aromatic Ester
114	319-84-6	(+)	(+)	Aliphatic Saturated Hydrocarbon
115	123-73-9	(+)	(-)	
116	N-(2-fluorenyl)-2,2,2-trifluoroacetamide	(+)	(+)	Aliphatic Amides
117	87-68-3	(+)	(+)	Aliphatic Unsaturated Hydrocarbons
118	4-Methyl-1-[(5-nitrofurfurylidene)amino]-2-imidazolidinone	(+)	(+)	Multiple Functional Groups
119	3761-53-3	(+)	(+)	Aromatic Polycyclic
120	3761-53-3	(+)	(+)	Aromatic Polycyclic
121	68-12-2	(-)	(+)	Aliphatic amides
122	50-81-7	(-)	(+)	Aliphatic Aldehydes
123	140-11-4	(-)	(+)	Aromatic Aldehydes
124	19408-74-3	(-)	(+)	Aromatic Halogens
125	119-53-9	(-)	(+)	Aromatic Aldehydes
126	107-07-3	(-)	(-)	
127	1936-15-8	(-)	(-)	
128	105-87-3	(-)	(+)	Aliphatic Aldehydes
129	3567-69-9	(-)	(-)	
130	95-50-1	(-)	(-)	
131	33229-34-4	(-)	(-)	
132	148-24-3	(-)	(-)	
133	69-65-8	(-)	(+)	Aliphatic Saturated Hydrocarbon
134	108-95-2	(-)	(-)	
135	115-07-1	(-)	(-)	
136	75-35-4	(-)	(+)	Aliphatic Unsaturated Hydrocarbon
137	140-49-8	(-)	(+)	Aromatic Amide
138	999-81-5	(-)	(-)	
139	6959-47-3	(-)	(-)	
140	61702-44-1	(-)	(-)	
141	56-72-4	(-)	(-)	
142	74-83-9	(-)	(+)	Aliphatic Halogens
143	75-69-4	(-)	(+)	Aliphatic Halogens
144	79-10-7	(-)	(-)	
145	3844-45-9	(-)	(-)	
146	97-00-7	(-)	(+)	Aromatic Halogens
147	75-45-6	(-)	(-)	
148	75-71-8	(-)	(+)	Aliphatic Halogens
149	81-21-0	(-)	(+)	Aliphatic Saturated Hydrocarbon
150	74-31-7	(-)	(-)	

* CPDB: Carcinogenic Potency Data Base
+ APEXPERT: APToxicophore Identification Program.
(+) Toxic (-) Nontoxic

Table: 4 Validation of results using APEXPERT predictions with experimental results of Mutagenicity and Carcinogenicity of CPDB.

Toxic End Point	Sensitivity	Selectivity	Concordance
MUTAGENICITY	77.5%	87.7%	73.3%
CARCINOGENICITY	76.4%	86.7%	72%

CONCLUSION:

Insilico toxicity prediction tool has not solved all the problems related to expert system. The development of a hybrid based module will aim to solve the problem(s) from all possible angles i.e., considering the metabolic decay, the structural alerts in residues, and the molecule as a whole or as a set of different substructures is desirable. We are in the process of developing a new hybrid module which can quantify the toxicity associated with new chemical entity.

ACKNOWLEDGEMENTS:

The authors are grateful to Director and Registrar MGNIRSA for the encouragement during the study. The author A. Padmanabha Rao is grateful to Dr.A.N.Singh Director, Celestial Biolabs for encouraging and providing facilities for carrying out research work.

REFERENCES:

1. Benfennati, Emilio, Gini, Giusepppina. Computational predictive programs (expert systems) in toxicology. *Toxicology*, 119, 1997, 213-225.
2. Ridlings JE, Baratt, MD, Cary R, et al. Computer prediction of possible toxic action from chemical structure; an update of DEREK system. *Toxicology*, 106, 1996, 267-279.
3. Lewis DFV. Computer Assisted methods in the evaluation of chemical toxicity. In: K.B. Lipkowitz and D.B. Boyd (Eds), Reviews in Computational Chemistry, VCH Publishers, Inc., New York. (3), 1992, 172-221.
4. Smithing MP, Darvas F. Hazard Expert: an expert system for predicting chemical toxicity, In: Finlay, J.W. & S.F. Robinson, D.J. Armstrong (Eds.), Food Safety Assessment, American Chemical Society, Washington DC, 1992, 191-200.
5. Lewis DFV, Costas Ioannides Dennis V Parke. A Combined COMPACT and Hazard Expert study of 40 chemicals for which information on mutagenicity and carcinogenicity is known, including the results of human epidemiological studies. *Human and Experimental Toxicology*, 17, 1998, 577-586.
6. Basak SC, Grunwald GD, Gute BD, et al. Use of Statistical and Neural Net Approaches in Predicting Toxicity of Chemicals. *J. Chem. Inf. Comput. Sci*, 40(4), 2000, 885-890.
7. Nakadate M, Hayashi M, Sofuni T, et al. The Expert System for Toxicity prediction of chemicals based on Structure Activity Relationship. *Environmental Health Perspectives*, 96, 1991, 77-89.
8. Aliuska Morales Helguera, Miguel Angel Cabrera Perez, Robert D. Combes et al. The Prediction of Carcinogenicity from Molecular Structure. *Current Computer-Aided Drug Design*, 1, 2005, 65-72.
9. Ashby J, Paton D. The Influence of Chemical Structure on the Extent and Sites of Carcinogenesis for 522 Rodent Carcinogens and 55 Different Human Carcinogen Exposures. *Mutation Research*, 286, 1993, 3-74.
10. Pierre Crettaz, Romualdo Benigni. Prediction of the Rodent Carcinogenicity of 60 pesticides by the DEREKfW Expert System, *J. Chem. Inf. Model*, 45, 2005, 1864-1873.
11. Ernesto Estrado, Enrique Molina. Automatic Extraction of Structural Alerts for Predicting Chromosomal aberrations of Organic Compounds. *Journal of Molecular Graphics and Modeling*, 25, 2006, 275-288.
12. Hakimelahi GH, Kodarahmi GA. The Identification of Toxicophores for the Prediction of Mutagenicity, Hepatotoxicity and Cardiotoxicity. *Journal of Iranian Chemical Society*, 2(4), 2005, 244-267.
13. Helma C, Cremer T, Kramer S, et al. Data mining and Machine Learning Techniques for the identification of Mutagenic Inducing substructures and Structure activity relationships of Noncongeneric compounds. *J.Chem. Inf. Comput. Sci*, 44(4), 2004, 1402-1411.
14. Katsuhiko Sawatari, Yoshifumi, Nakanishi, et al. Relationships between Chemical Structures and Mutagenicity: A Preliminary Survey for a Database of Mutagenicity Test Results of New Work Place Chemicals. *Industrial Health*, 39, 2001, 341-345.

15. Wang Jiansuo, Luhua Lai, Youqi Tang. Structural Features of Toxic Chemicals for Specific Toxicity *J. Chem. Inf. Comput. Sci*, 39, 1999, 1173-1189.
16. http://www.epa.gov/oppt/newchems/chemcat.htm
17. Cunningham, Albert R, Gilles Klopman, Herbert S Rosenkranz. Identification of structural features and associated mechanisms of action for carcinogens in rats, *Mutation Research*, 405(a), 1998, 9-28.
18. Cunningham, Albert R, Herbert S Rosenkranz, Ying Ping Zhang, et al. Identification of 'genotoxic' and 'non-genotoxic' alerts for cancer in mice: the carcinogenic potency database. *Mutation Research*, 398, 1998(b),1-17.
19. Ross D King, Ashwin Srinivasan. Prediction of Rodent Carcinogenicity Bioassays from Molecular Structure Using Inductive Logic Programming *Environmental Health Perspectives Supplements*, 104: 1996, supplement 5.
20. Jun Xu, Arnold Hagler. Review: Cheminformatics and Drug Discovery. *Molecules*, 7, 2002, 566-600.
21. Lim Wilby M, Jiang J, Waldman M, et al. Virtual high throughput screening using Ligand Fit as an accurate and very fast tool for docking, scoring, and ranking," in: *Wendy Warr & Associates' Chemical Information* and Computation 224th ACS National Meeting and Exposition, Boston, Massachusetts, August 18-22, 2002, 87-91.