RJC
http://www.rasayanjournal.com

RASĀYAN *J. Chem.*
Vol.2, No.2 (2009), 475-487
ISSN: 0974-1496
CODEN: RJCABP

CLASSIFICATION OF FUNCTIONAL GROUPS OF A PHYTOCHEMICAL: GYMNEMIC ACID USING FUZZY LOGIC

S. Palaniswamy[1], M.Mayilvanan[2] and Antony Selvadass Dharamani[3]
[1] Department of Physics, PSG College of Arts and Science, Coimbatore-641 014.
[2] Department of Computer Science, PSG College of Arts and Science, Coimbatore-641 014.
[3] Department of Computer Science, NGM College, Pollachi -642 001
*E-mail: palazhamni_5003@yahoo.com

ABSTRACT

This paper deals with simplest method to identify the active functional groups in phytochemicals. A photochemical is a natural bioactive compound found in plant foods that work with nutrient the dietary fiber to protect against disease. In organic chemistry, functional groups specific groups of atoms within molecules that are responsible for the characteristic chemical reactions of those molecules. The same functional groups will undergo the same or similar chemical reactions. A chemical database is one of the database that exclusively designed to store phytochemical information such as molecular surface properties number atom , position of the atom and bond relation between atom are derived from histogram. In this study we developed a method to find the functional group of clustered phytochemicals which will lead to new millstone in drug discovery.

Keywords: Phytochemical, Data ware house, Homogenous, Federated Database, Data mart.

INTRODUCTION

Carbohydrates are a main source of energy for the body and are made of carbon, hydrogen, and oxygen. Humans and other animals obtain carbohydrates by eating foods that contain them. In order to use the energy contained in the carbohydrates, humans must metabolize, or break down, structure of the molecule in a process that is opposite that of photosynthesis. It starts with the carbohydrate and oxygen and produces carbon dioxide, water, and energy. The body utilizes the energy and water and rids itself of the carbon dioxide. Molecular structure of the organic compounds contains carbon, hydrogen and oxygen atoms as major constituents. Among the above three atoms, carbon atom plays a major role in the skeleton of the compounds and also the carbon atom chain is considered in the naming of the compound as well as the physical and chemical properties of the compound . Along with hydrogen and oxygen atom constitutes normal functional groups that is organic compounds .To identify those organic compounds is a challenging problem.

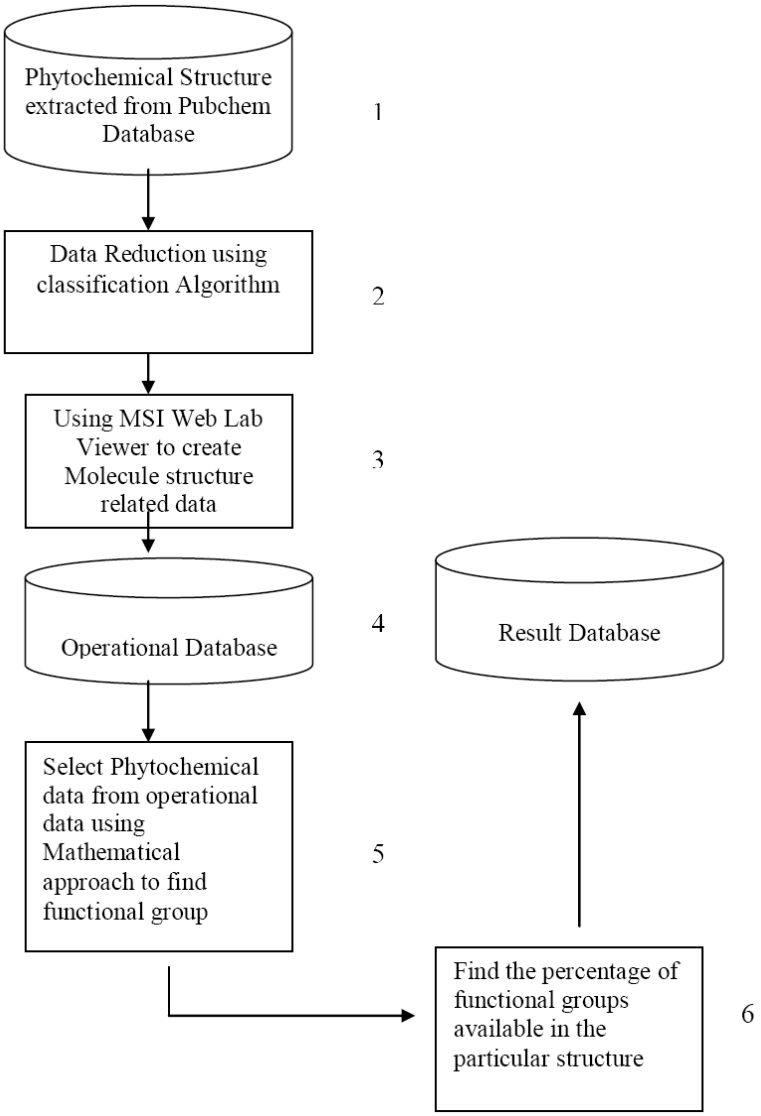
Related work:

The earliest research work involved the application of textual representation of the molecular structures vector representation and Smiles representation. To use the Chemical Search Engine, Structure of the molecule to be searched smiles which represents the molecule into a text field, or open a file which stores a smile value in the corresponding field. The engine converts the query into an International Chemical Identifier standard value and returns of a listing of molecules and their similarity values[1,3]. Our research involves, instead of textual representation use numerical values which represent the number of atom, position and distance between atom .The above attribute values are derived from MSI Web Lab Viewer. Phytochmical structure is taken from Pub Chem. free internet software toll which is connected to authorized plant phytochemical database. This approach is faster then the string search due to the less memory space and also minimizes the unnecessary data search.

EXPERIMENTAL

The Materials are collected from near and surrounding area of Dindugal and Plants propagation centre in Top Slip belonging to Tamil Nadu State , INDIA. There are about three hundred and forty eight medicinal plants and certain endemic plants were propagated. Few importance medicinal plants from this were procured, therein binomial, family, vernacular names, identifying features medicinal uses and chemical components were recorded in detail in the present investigation. The periodical visits were carried out and the first hand information's were collected from the siddha medicinal practitioner in the surrounding area of Dindugal and Nelagri Districts and the herbal medicines practiced by them. The information about collected phytochemicals are taken from various books and internet also present here.

Research Methodology:



Field	Description	Details
P_Id	Plant Identification	Gymne_Dia_10
P_Name	Botanical Name	Gymneme Sylvester R.B
P_V_Name	Vernacular name	Asdepiadaceae in English , Sakkaraikolli and Srukurinjan in Tamil

P_Id_F	Identifying Features	A large woody , much branches , leaves simple opposite ecliptic or ovate , Flowers small and yellow in colour , Fruits slender follicles
P_P_U	Parts Used	Leaf
P_Med_U	Medicinal Uses	The leaf of Gymnema Sylvester and Flowers of cocosnucifera are taken in the ratio 1:2 and pounded together to a baste, dried in shade. This is made into small tablets and given twice a day with hot water to control diabetes. The powdered leaves are taken twice daily with water or honey to control diabetes.
P_Act_Pri	Active Principles	Gymnemic Acid
P_P_Family	Plant Family	Asclepiadaceae

ACTIVE PRINCIPLES: GYMNEMIC ACID

Active Principles are the bioactive substances or Phytochmicals responsible for the Curative property of disease. The figure of Gymnemic acid is-

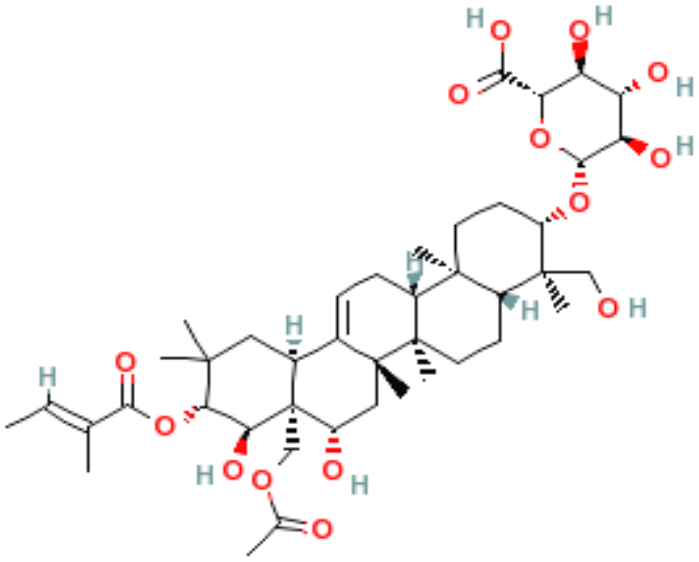


Table-1

Atom	X	Y	Atom Type
1	9.1130	-4.2910	O
2	14.4650	1.9226	O
3	6.9333	-4.2610	O
4	5.4641	-.2741	O
5	7.5507	-4.5312	O
6	15.0976	-0.5375	O
7	13.5932	3.4192	O

8	16.1932	2.9293	O
9	4.6038	-1.7708	O
10	8.8690	-5.6548	O
11	16.1855	4.9293	O
12	14.4496	5.9226	O
13	12.7176	5.9159	O
14	11.8574	4.4126	O
15	10.8572	-1.2218	C
16	10.8572	-0.2218	C
17	11.7511	0.3129	C
18	9.9911	-1.7218	C
19	12.6572	-0.2010	C
20	11.7511	-1.7564	C
21	9.1251	-1.2218	C
22	8.2151	-1.7286	C
23	8.2071	-2.7702	C
24	12.6572	-1.2426	C
25	13.6010	0.3343	C
26	10.0072	-2.7633	C
27	9.9911	0.2782	C
28	11.7260	1.3976	C
29	9.1092	-3.2910	C
30	11.1993	-2.1614	C
31	13.6010	1.4193	C
32	9.1251	-0.2218	C
33	10.8812	0.8062	C
34	7.2883	-1.1645	C
35	10.6339	-2.4878	C
36	12.6572	1.9545	C
37	7.2717	-3.3200	C
38	6.3404	-1.6925	C
39	6.3320	-2.7775	C
40	8.1994	-3.7702	C
41	14.6009	0.3304	C
42	14.0976	-0.5337	C
43	5.8504	-0.8207	C
44	5.3405	-1.6809	C
45	14.4612	2.9226	C
46	15.3253	3.4259	C
47	4.6000	-2.7708	C
48	7.8855	-5.4736	C
49	15.3214	4.4259	C
50	13.5894	4.4192	C
51	14.4535	4.9226	C
52	3.7320	-3.2674	C
53	7.2368	-6.2346	C
54	12.7214	4.9159	C
55	3.7282	-4.2674	C
56	2.8679	-2.7641	C
57	2.0000	-3.2608	C

58	10.8510	0.5782	H
59	13.3484	-0.6036	H
60	12.1447	-2.2355	H
61	11.3465	-2.2262	H
62	8.2305	-0.9288	H
63	12.8662	-1.8263	H
64	13.2680	-1.1365	H
65	10.6163	-2.6478	H
66	10.2252	-3.3437	H
67	9.5926	0.7532	H
68	10.3897	0.7532	H
69	11.5035	1.9763	H
70	11.1178	1.2774	H
71	8.5728	-3.6021	H
72	10.6167	-2.3736	H
73	11.4114	-2.7440	H
74	11.7819	-1.9493	H
75	14.1385	1.1103	H
76	8.5882	0.0882	H
77	10.5754	0.2669	H
78	10.3419	1.1120	H
79	11.1871	1.3455	H
80	7.6965	-0.6979	H
81	6.8984	-0.6825	H
82	10.1590	-2.8864	H
83	11.0325	-2.9628	H
84	11.1089	-2.0893	H
85	12.2526	2.4243	H
86	13.0508	2.4335	H
87	7.7489	-3.7157	H
88	5.7969	-2.4644	H
89	8.7947	-3.5970	H
90	8.5360	-4.2908	H
91	15.1844	0.5402	H
92	14.4956	0.9414	H
93	13.5595	-0.8416	H
94	14.4055	-1.0718	H
95	14.6357	-0.2257	H
96	9.6511	-4.5990	H
97	6.3909	-0.5170	H
98	5.5467	-0.2803	H
99	5.3099	-1.1245	H
100	5.3476	-1.0610	H
101	4.7205	-1.6738	H
102	5.3333	-2.3009	H
103	6.3231	-4.3710	H
104	13.9255	2.6105	H
105	15.7176	-0.5399	H
106	15.3277	2.8059	H
107	15.8596	4.1180	H

108	13.0537	4.1072	H
109	14.9892	5.2346	H
110	6.7649	-5.8324	H
111	6.8346	-6.7065	H
112	7.7086	-6.6368	H
113	16.7290	3.2413	H
114	16.7236	4.6213	H
115	4.3482	-4.2698	H
116	3.7258	-4.8874	H
117	3.1082	-4.2650	H
118	2.8703	-2.1441	H
119	14.9854	6.2346	H
120	1.6921	-2.7226	H
121	1.4619	-3.5687	H
122	2.3079	-3.7989	H
123	12.1795	6.2238	H

where **Atom** – Atom serial number, Total number of atoms in Gymnemic Acid is 123, **O** – Oxygen **C**-Carbon **H** – Hydrogen, and X and Y are Atom possession with respect to X and Y Axis.

Table-2

A1	**A2**	**B**
29	1	1
1	96	1
31	2	1
45	2	1
37	3	1
3	103	1
39	4	1
4	47	1
5	40	1
5	48	1
6	41	1
6	105	1
7	45	1
7	50	1
46	8	1
8	113	1
9	47	2
10	48	2
49	11	1
11	114	1
51	12	1
12	119	1
13	54	1
13	123	1
14	54	2
15	16	1
15	18	1
15	20	1

15	30	1
16	17	1
16	27	1
16	58	1
17	19	1
17	28	1
17	33	1
18	21	1
18	26	1
18	35	1
19	24	1
19	25	1
19	59	1
20	24	1
20	60	1
20	61	1
21	22	1
21	32	2
22	23	1
22	34	1
22	62	1
23	29	1
23	37	1
23	40	1
24	63	1
24	64	1
25	31	1
25	41	1
25	42	1
26	29	1
26	65	1
26	66	1
27	32	1
27	67	1
27	68	1
28	36	1
28	69	1
28	70	1
29	71	1
30	72	1
30	73	1
30	74	1
31	36	1
31	75	1
32	76	1
33	77	1
33	78	1
33	79	1
34	38	1
34	80	1

34	81	1
35	82	1
35	83	1
35	84	1
36	85	1
36	86	1
37	39	1
37	87	1
38	39	1
38	43	1
38	44	1
39	88	1
40	89	1
40	90	1
41	91	1
41	92	1
42	93	1
42	94	1
42	95	1
43	97	1
43	98	1
43	99	1
44	100	1
44	101	1
44	102	1
45	46	1
45	104	1
46	49	1
46	106	1
47	52	1
48	53	1
49	51	1
49	107	1
50	51	1
50	54	1
50	108	1
51	109	1
52	55	1
52	56	2
53	110	1
53	111	1
53	112	1
55	115	1
55	116	1
55	117	1
56	57	1
56	118	1
57	120	1
57	121	1
57	122	1

where **A1** – Atom Number, **A2** – Connected Atom Number, **B** -Bond

FUZZY SET WITH RESPECT TO ATOMS CONNECTED TO PARTICULAR ATOM
Grouping with respect to connected atom
Using Table-2.

Where; An = ATOM NUMBER
An = {(Atom, Connected Atom, Bond)}

A29={(29,1,1}
A1 = {(1,96,1)}
A31={(31,2,1)}
A45={(45,2,1)}
A37={(37,3,1)}
A3={(3,103,1)}
A39={(39,4,1)}
A4={(4,47,1)}
A5={(5,40,1),(5,48,1)}
A6={(6,41,1),(6,105,1)}
A7={(7,45,1),(7,50,1)}
A46={(46,8,1)}
A81={(81,13,1)}
A9={(9,47,2)}
A10={(10,48,2)}
A49={(49,11,1)}
A11={(11,114,1)}
A51={(51,12,1)}
A12={(12,119)}
A13={(13,54,1),(13,123,1)}
A14={14,54,2)}
A15={(15,16,1),(15,18,1),(15,20,1),(15,30,1)}
A16={(16,17,1),(16,27,1),(16,58,1)}
A17={(17,19),(17,28),(17,33,1)}
A18={(18,21,1),(18,26,1),(18,35,1)}
A19={(19,24,1),(19,25,1),(19,59,1)}
A20={(20,24,1),(20,60,1),(20,61,1)}
A21={(21,22,1),(21,32,2)}
A22={(22,23,1),(22,34,1),(22,62,1)}
A23={(23,29,1),(23,29,1),(23,40,1)}
A24={(24,63,1),(24,64,1)}
A25={(25,31,1),(25,41,1),(25,42,1)}
A26={(26,29,1),(26,65,1),(26,66,1)}
A27={(27,32,1),(27,67,1),(27,68,1)}
A28={(28,36,1),(28,69,1)(29,70,1)}
A29={(29,71,1),(29,72,1),(29,73,1)}
A30={(30,73,1),(30,74,1)}
A31={(31,36,1),(31,75,1)}
A32={(32,76,1)}
A33={(33,77,1),(33,78,1),(33,79,1)}
A34={(34,38,1),(34,80,1),(34,81,1)}

A35={(35,82,1),(35,83,1),(35,84,1)}
A36={(36,85,1),(36,86,1)}
A37={(37,39,1),(37,87,1)}
A38={(38,39,1),(38,43,1),(38,44,1)}
A39={(39,88,1)
A40={(40,89,1),(40,90,1)}
A41={(41,91,1),(41,92,1)}
A42={(42,93,1),(42,94,1),(42,95,1)}
A43={(43,97,1),(43,98,1),(43,991)}
A44={(44,100,1),(44,101,1),(44,102,1)}
A45={(45,46,1),(45,104,1)}
A46={(46,49,1),(46,106,1)}
A47={(47,52,1)}
A48={(48,53,1)}
A49={(49,51,1),(49,107,1)}
A50={(50,108,1),(50,54,1)}
A51={(51,109,1)}
A52={(52,55,1),(52,56,2)}
A53={(53,110,1),(53,111,1),(53,112,1)}
A55={(55,115,1),(55,116,1),(55,117,1)}
A56={(56,57,1),(56,118,1)}
A57={(57,120,1),(57,121,1),(57,122,1)}

Using Table-1
Assign Atom Type (C-CARBON, O-OXGEN, H-HYDROGEN)
An = ATOM NUMBER
An = {(Atom, Connected Atom, Bond)}

A29= {(C-O)} = CO
A1 = {(O-H)} = OH
A31= {(C-C)} = C2
A45= {(C-O)} = CO
A37= {(C-O)} = CO
A3= {(O-H)} = OH
A39= {(C-O)} = CO
A4= {(O-C)} = CO
A5= {(O-C), (O-C)} = CO
A6= {(O-C), (O-H)} = COH
A7= {(O-C), (O-C)} = CO
A46= {(C-O)} = CO
A81= {(H-O)} = OH
A9= {(O=H)} = OH
A10= {(O=H)} = OH
A49= {(C-O)} = CO
A11= {(O-H)} = OH
A51= {(C-O)} = CO
A12= {(O-H)} = OH
A13= {(O-C), (O-H)} = COH
A14={O=C)} = CO
A15={(C-C),(C-C),(C-C),(C-C)} = C4

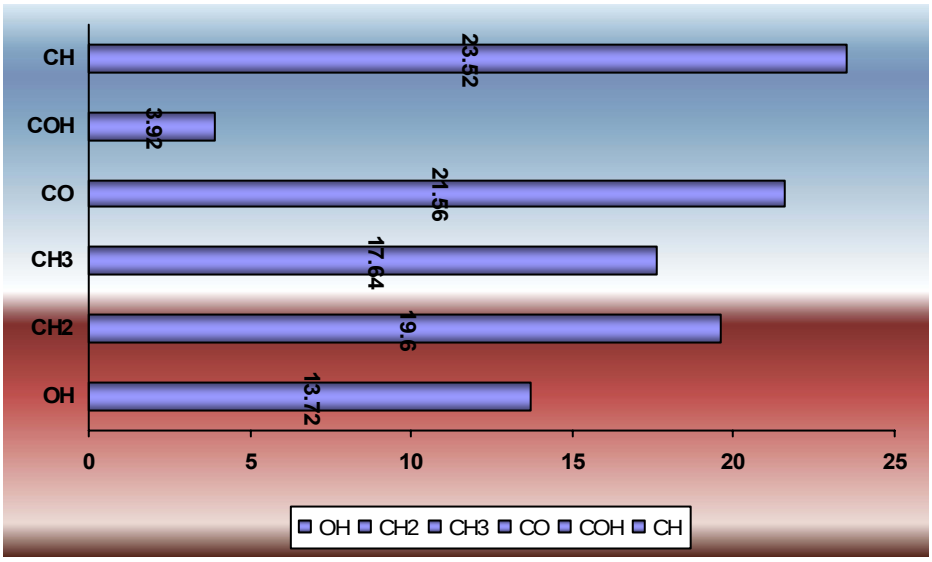
A16={(C-C),(C-C),(C-H)} = CH
A17={(C-C),(C-C),(C-C)} = C3
A18={(C-C),(C-C),(C-C)} = C3
A19={(C-C),(C-C),(C-H)} =CH
A20={(C-C),(C-H),(C-H)} = CH2
A21={(C-C),(C=C)} = C2
A22={(C-C),(C-C),(C-H)} = CH
A23={(C-C),(C-C),(C-C)} = C3
A24={(C-H),(C-H)} = CH2
A25={(C-C),(C-C),(C-C)} = C3
A26={(C-C),(C-H),(C-H)} = CH2
A27={(C-C),(C-H),(C-H)} = CH2
A28={(C-C),(C-H)(C-H)} = CH2
A29={(C-H),(C-H),(C-H)} =CH3
A30={(C-H),(C-H)} = CH2
A31={(C-C),(C-H)} = CH
A32={(C-H)} = CH
A33={(C-H),(C-H),(C-H)} = CH3
A34={(C-C),(C-H),(C-H)} = CH2
A35={(C-H),(C-H),(C-H)} = CH3
A36={(C-H),(C-H)} =CH2
A37={(C-C),(C-H)} = CH
A38={(C-C),(C-C),(C-C)} = C3
A39={(C-H) = CH
A40={(C-H),(C-H)} = CH2
A41={(C-H),(C-H)} = CH2
A42={(C-H),(C-H),(C-H)} = CH3
A43={(C-H),(C-H),(C-H)} = CH3
A44={(C-H),(C-H),(C-H)} = CH3
A45={(C-C),(C-H)} = CH
A46={(C-C),(C-H)} = CH
A47={(C-C)} = C2
A48={(C-C)} = C2
A49={(C-C),(C-H)} = CH
A50={(C-H),(C-C)} = CH
A51={(C-H)} = CH
A52={(C-C),(C-C)} = C2
A53={(C-H),(C-H),(C-H)} = CH3
A55={(C-H),(C-H),(C-H)} = CH3
A56={(C-C),(C-H)} = CH
A57={(C-H),(C-H),(C-H)} = CH3

A = { CO, OH , C2 , CO ,CO ,OH , CO ,CO, CO , COH, CO ,CO , OH ,OH,
OH ,CO , OH,CO , OH ,COH,CO,CH,CH,CH2,CH,CH2,CH2, CH2,CH2,CH3,
CH2,CH,CH3,CH,CH3,CH2,CH3,CH2,CH,CH,CH2,CH2,CH3,CH3,CH3,CH,CH,CH,
CH, CH, CH3, CH3, CH, CH3}

Functional Group Found: {OH, CH2, CH3, CO, COH, CH}
Number of times functional Groups found:

OH= {7 TIMES}
CH2= {10 TIMES}
CH3= {9 TIMES}
CO= {11 TIMES}
COH = {2 TIMES}
CH = {13 TIMES}

The percentage of OH, CH2, CH3, CO, COH, CH found:



Advantages of the Proposed Algorithm
The advantage of the proposed method is that, it not required checking the main dataset more than once. In the first scan of the data set the Algorithm will create the clusters at the time of checking the molecular related Fuzzy set. This way it will goon grouping the data set to find the functional group. This method is also useful to find the percentage of functional groups available in phytochemical structure.

CONCLUSIONS

In this paper the molecular structure of the Phytochemical compounds and the role of constructing the compounds based on carbon, hydrogen and oxygen atoms are discussed briefly. Finally, this paper has made following concluding remarks.

1. A new algorithm which is completely new in its design is proposed to find the active functional groups in molecular structure.
2. Algorithm is used to find percentage of functional groups which are available in the molecule structure.

It is believed that this research is a necessary to first step to representing the imprecision found in phytochemical molecule structure and will, hopefully, lead to a better understanding of functional and surface properties. In this study we extract any cluster of phytochemicals from database and compare the

surface properties and find the functional group of phytochemicals Active principles which will lead to new millstone in drug discovery.

REFRENCES

1. A.F. Kalib Zurinahni, 'Application of external string matching algorithms towards smiles representation of chemical structure', IJCISE (2007).
2. A. Mehta C., 'Database Management Systems for Efficient Management of Biological Data', 459 (2007).
3. C.T. Bhunia, 'A Handbook of Information Technology', Vitasta Publications Pvt. Ltd. 459 (2007).
4. C.G. Li, 'String matching and the Kruth – Morris Pratt Algorithms', Carleton University, Canada, pp.1-8 (2006).
5. C. Charrs and T. Lecroq, 'Exact string Matching Algorithms', De Rouen University, France (2006).
6. A. Dalby, J.G. Nourse and W.D .Hournshell, *J. Chem. Inf. Comp. Sci*., **32**, 244 (1992).
7. D. Cantone and S. Faro, 'Forward –fast-search: Another fast variant of the Boyer – Moore string matching algorithm', Department of Computer Science, University of Texas at Austin (2005).
8. D. Flower, *J. Chem. Inf. Comp. Sci.*, **38**, 379 (1998).
9. Inc. Dylight Chemical Information Systems. Daylight Theory; Fingerprints, 2005. Http: //WWW. Daylight .com /dayhtml/doc /theory /theory .finger.html.
10. Inc.Daylight chemical information systems. 2006 Day light Cheminformatics SMILES, 2006.http://daylight.com/ smiles.
11. J. Rhodes and S. Boyer, Pacific Symposium on Computing, **12**, 304 (2007).
12. J. Han, and M. Kamber, 'Data Mining Concepts and Techniques', Morgan Kaufmann Publications, pp. 338-341 (2006).
13. M. Crochemore and T. Lecroq, 'Pattern matching and text compression algorithms', Chapter 2, pp, 12-14 (2006).
14. SMILS –A 2007 Simplified Chemical language .Available:http://WWW.daylight .com
15. Stein's, S.R. Heller, and D. Techkhovskoi, 'An Open Standard for Chemical Structure Representation', In proceedings of the 2003 International Chemical Information Conference (Nimes).
16. V.V. Poroikov and L. Dmitri, *J. Chem Info. Comp. Sci.*, **43**, 228 (2003).
17. P. Willett, J.M. Barnard and G.M. Downs, *J. Chem. Info. Comp. Sci.*, **38**, 983 (1998).

(Received: 16 April 2009 Accepted: 7 May 2009 RJC-377)