June 2009 *Volume 3 Issue 1*

Trade Science Inc.

Research & Reviews in

BioSciences

Regular Paper

RRBS, 3(1), 2009 [1-10]

Insilico analyis of potential inhibitors for glutathione S-transferase of *Wuchereria Bancroft causing* lymphatic filariasis

M.Abhilash[1]*, S.Uma Maheshwari[2]

[1]Department of Biotechnology, Oxford college of Engineering, Bangalore, (INDIA)

[2]Department of Nanotechnology, Karunya University, Coimbatore, Tamil Nadu, (INDIA)

E-mail : abhibiotek@gmail.com

Received: 7th January, 2009 ; Accepted: 12th January, 2009

ABSTRACT

Lymphatic filariasis is caused by the infection with *Wuchereria bancrofti*, *Brugia malayi* and *B.timori*, parasitic filarial nematodes transmitted by mosquito vectors. Currently the drug available for the treatment of filariasis, Diethylcarbamazine (DEC) is effective only against microfilariae. Three dimensional structure of the filarial drug target Glutathione S-transferase of *W.bancrofti* (*wb*GST) was retrieved from the Protein data bank and used as the biological receptor for macrofilaricidal drug development based on SBDD. Active sites of target protein *wb*GST were mapped using a computational tool PASS and the binding pockets using program CASTp. Plumbagin, macrofilaricidal lead molecule identified by VCRC was selected as the seed structure for NCI database search. Preliminary database search resulted in 100 hits. The hits from NCI database were docked with 1SFM, the target enzyme using the program Autodock 3.0.5 and from the resulting conformations, top 21 conformations were selected based on their Binding Energy values. The binding energies of the top 21 ligands varies from -8.36 to -6.02 Kcal/mol. The physicochemical properties influencing the pharmacokinetic properties of the drug molecules namely log P, molecular volume, polar surface area and hydrogen bond donor/acceptor properties were calculated for the top 21 hits and their ADME violation has been generated from their 2D structures using TSAR 3.3.

KEYWORDS

Lymphatic filariasis;
Macrofilaricidal drug development based on SBDD;
The physicochemical properties calculations using TSAR 3.3.

1. INTRODUCTION

Filariasis contributes the highest morbidity of human population of many tropical and subtropical countries of the world[1,2]. Lymphatic filariasis is a complex disease caused by parasitic nematodes. Three species of filariae, *Wuchereria bancrofti*, *Brugia malayi*, and *Brugia timori* are responsible for human lymphatic filariasis, which can lead to such conditions as elephantiasis and tropical pulmonary eosinophilia.

The successful treatment of filariasis is not possible because of the non-availability of macrofilaricidal drugs[1,3]. The age-old drug diethyl carbamazine (DEC) continues to be the mainstay of clinical practice despite its well-known deficiencies[4,5]. Ivermectin, a semisynthetic macrocyclic lactone antibiotic, may take an impact as microfilaricide for onchocerciasis but it did not irreversibly damage the adult filarial worms[6]. Although organic arsenical compounds have long been known as good macrofilaricides[7], their potential toxicity to the

Regular Paper

host has prevented their development as useful antifilarial drugs. Besides these antifilarials, a number of phenoxycyclohexane derivatives[8], 2,4,6-substituted triazines[9], 5-amino and 5,8-diaminoisoquinolines[10] aplysinoposin derivatives[11] and 1,10-dicyano-2- substituted ethylenes[12] were identified as potential filaricides but most of the compounds exhibited very poor adulticidal response. Benzimidazole group of anthelmintics exhibit high order of activity against intestinal helminths but have not found application for the treatment of tissue dwelling helminthes[13,14]. Therefore, the need arose to identify structural prototypes associated with macrofilaricidal activity.

Hence it is imperative to identify the key enzymes and biochemical pathways which are pivotal to the parasite's survival in the host's hostile environment, including their oxidative stresses and immune responses[15]. These enzymes should provide excellent biochemical targets for developing effective chemotherapies and vaccines[16,17]. One such enzyme is Glutathione S transferase[18], which is involved in xenobiotic metabolism, intracellular binding, and biosynthesis of endogenous substrates such as prostaglandins and leukotrienes[19-21]. GSTs may potentially favour parasite survival by neutralizing the toxins acting against them and may repair host-induced damage[22]. Because GSTs are major detoxifying enzymes in helminthes, they may be able to scavenge the products of lipid peroxidation and to metabolize toxic products, including anthelmintics. These biological functions make GSTs molecular targets for new antifilarial drugs[18].

The application of computational methods to study the formation of intermolecular complexes has been the subject of intensive research during the last decade. It is widely accepted that drug activity is obtained through the molecular binding of one molecule (the ligand) to the pocket of another, usually larger, molecule (the receptor), which is commonly a protein. In their binding conformations, the molecules exhibit geometric and chemical complementarity, both of which are essential for successful drug activity. The computational process of searching for a ligand that is able to fit both geometrically and energetically the binding site of a protein is called molecular docking.

Three-dimensional structure (3D) of the target enzyme is essential for defining the active site and also for designing, improving and docking of small ligands to the complex target protein. In the absence of a crystal structure of the target enzyme filarial GST, the 3D-structures of GSTs of *W.bancrofti* and *B.malayi viz.*, *wb*GST and *bm*GST have been modelled at Vector Control Research Centre, Pondicherry by comparative protein modelling for which the PDB and RCSB IDs are 1SFM and RCSB021668 and 1SJO and RCSB021767 respectively. These 3D structures have been used for docking small ligands[23] and the results encouraged to work further in this area for the development of a macrofilaricidal molecule that can destroy the adult filarial parasites which causes tremendous concern in combating filariasis. Earlier work from this Centre has identified a macrofilaricidal lead molecule plumbagin[24] isolated from the medicinal plant *Plumbago rosea* which also binds to the filarial GST.

Therefore, for the identification of more potent molecules, a structure based approach for macro filaricidal drug development has been proposed in this work with following objectives;

(i). To retrieve the structure of the filarial drug target enzyme Glutathione S transferase (*wb*GST) from the Protein database and to identify the active sites in the target using bioinformatics tools

(ii) To apply molecular docking to identify compounds appropriate for the biological receptor (*wb*GST)

2. MATERIALS AND METHODS

Structure based drug design (SBDD)

The structure based drug design of macrofilaricidal molecules followed a standard stepwise procedure[26] as given in **scheme** 1.

2.1. Selection of receptor structure

The enzyme Glutathione S Transferase of the filarial parasite *Wuchereria bancrofti* (*wb*) was selected as the molecular target for drug design. Since the crystal structure of the *wb*GST is not available currently, the comparatively modeled three-dimensional structure was used as the biological receptor for this study. The 3D structure of the target protein *wb*GST was retrieved from the protein data bank with PDB ID: 1SFM which is available in the following URL: http://www.rcsb.org.

2.2. Mapping of active sites, cavities and binding pockets

Active sites present in the target protein *wb*GST

were determined using a computational tool PASS (Putative Active Sites with Spheres)[27] based on size, shape and buried volumes in the protein. Cavities and binding pockets of *wb*GST were determined using CASTp, an online server[28].

2.3. Selection of ligands for docking

Plumbagin, the reported[23,24] lead molecule for macrofilaricidal drug development, was selected as the primary ligand in docking experiments. The plumbagin was drawn in the Java Molecular Editor provided in the browser page and it was transferred to query form. A search was conducted in NCI database (National Cancer Institute, USA) (http://129.43.27.140\ncidb2) using plumbagin as substructure and the hits were selected for further analysis. All the structures were downloaded in SDF (Structure Data File) format and converted to standard PDB format using the program MarvinView (http://www.chemaxom.com/marvin).

2.4. Virtual screening

The 3D structure of wbGST retrieved from PDB database was used as the target structure for the molecular docking. The 100 small molecules retrieved from NCI by substructure search were virtually screened to give small molecules with apt geometric and chemical complementarities using Autodock 3.0.5[29].

2.4.1. Generation of receptor-ligand complexes

To calculate the binding energy using AutoDock, polar hydrogens were added to the receptor 1SFM coordinates with the PROTONATE utility from AMBER[30]. AMBER united atom force field charges were assigned, and solvation parameters were added using the ADDSOL utility. The 3D affinity grid fields were created using the auxiliary program AutoGrid. The residue TYR 116 was chosen as the grid center. In this stage, the protein was embedded in the 3D grid and a probe atom was placed at each grid point. The affinity and electrostatic potential grid was calculated for each type of atom in the ligand molecule. The number of grid points in x, y, z-axis was 60×60×60 with grid points separated by 0.375 Å.

Ligands that had a peptide-like N- or C-terminal end were assigned a charge. Hydrogen atoms were added to fill all empty valences, and Kollman united-atom charges[31] were also assigned to the ligands. Rotatable dihedrals in the ligands were assigned using the program AutoTors and were allowed to rotate freely. The nonpolar hydrogens were removed and the partial charges from these were added to the carbon that held the hydrogen. The atom type for the aromatic carbons was reassigned to be handled by the aromatic carbon grid map. These preparations were done for each ligand using the AutoTors module.

2.4.2. Automated docking

Docking calculations were carried out using AutoDock, version 3.0.5[29]. Three binding energy terms were taken into account in the docking step: the van der Waals interaction represented as a Lennard-Jones 12-6 dispersion/repulsion term, the hydrogen bonding represented as a directional 12-10 term, and the Coulombic electrostatic potential.

Docking runs were performed using the Larmarckian genetic algorithm (LGA)[29] with default docking parameters. The LGA describes the relationship between the protein and the ligand by the translation, orientation, and conformation of the ligand. These "state variables" are the ligand's genotype, and the resulting atomic coordinates together with the interaction and intermolecular energies are the ligand's phenotype. The environmental adaptation of the ligand's phenotype was reverse transcribed into its genotype and became heritable traits.

Docking began with a population of random ligand conformations in random orientations and at random translations. Each docking experiment was derived from 100 different runs that was set to terminate after a maximum of 2500,000 energy evaluations or 27,000 generations, yielding 100 docked conformations. The population size was set to 50. The elitism number, the rate of gene mutation and the rate of gene crossover were 1, 0.02 and 0.8 respectively. A pseudo-Solis and Wets local search was then used to minimize energy of the population. The probability that docking solution in the population would undergo a local search was set to 0.06 and the constraint was set to a maximum of 300 iterations per search. The maximum number of successes or failures before changing the size of local search space (rho) were both set to 4. The starting conformations of the ligand were set to random positions. Translations were set to have a maximum limit of 2 Å/step and the orientation step size for the angular component and the torsions had a maximum limit at 50 degrees/step.

At the end of a docking job with multiple runs,

Regular Paper

>1SFM :Gluta thione S Transferase:A: Wuchereria Bancrofti:Theoretical Model::::
MSYKLTYFPIRGLAEPIRLVLVDQGIKFTDDRINASDWPSMKSHFHFGQL
PCLYDGDHQIVQSGAILRHLARKHNLNGGNELETTHIDMFCEGIRDLHTK
YAKMIYQAYDTEKDSYIKDILPVELAKFEKLLATRDDGKNFILGEKISYV
DFVLFEELDIHQILDPHCLDKFPLLKAYHQRMEDRPGLKEYCKQRNRAKI
PVNGNGKQ

Figure 1(a): The single letter code aminoacid sequence of 1SFM retrieved from PDB

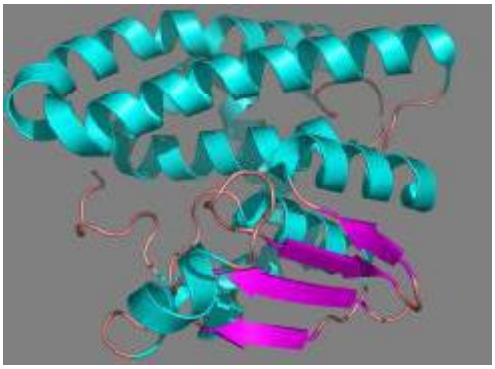

Figure 1(b): *wb*GST in cartoon form

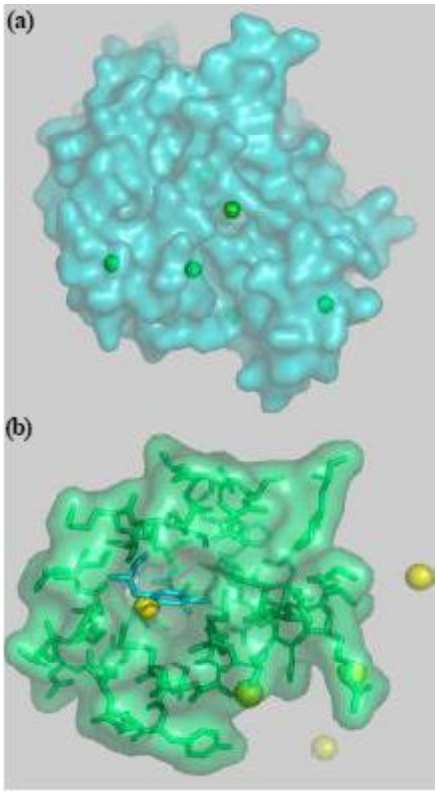

Figure 2: (a) PASS predicted active sites of *wb*GST, (b) residues around TYR116 (upto10Å)

AutoDock performed cluster analysis. Docking solutions with ligand all-atom RMSDs within 1.0 Å of each other were clustered together and ranked by the lowest energy representative. The lowest-energy solution of the lowest ligand all-atom RMSD cluster was accepted as the calculated binding energy.

2.5. ADME properties of top scored hits

The important pharmacokinetic properties like Absorption, Distribution, Metabolization, Excretion are all involve in passage across cell membranes. It is essential to consider the mechanisms by which drugs cross cell membranes and physico chemical properties of molecules and membrane that influence this transfer. Important characteristics include molecular size, shapes, solubility at the site of its absorption, degree of ionization and relative lipid solubility of its ionized and non-ionized forms[32].

ADME properties were calculated for the top scored 20 small molecules from *virtual* screening using TSAR 3.3 (http://www.accelrys.com).

3. RESULTS AND DISCUSSION

3.1. Receptor structure retrieval

The computationally modelled 3D structure of enzyme Glutathione S Transferase of the filarial parasite *Wuchereria bancrofti* (*wb*) with PDB ID: 1SFM was selected as the molecular target for macrofilaricidal drug design. The sequence was retrieved from the Protein Data Bank and its single letter code amino acid sequence is given in figure 1a and its 3D structure in 'cartoon' form as visualized by PYMOL is given in figure 1b.

3.2. Active sites and binding pockets

The most probable active sites in 1SFM were determined using computational tool PASS and the output of the PASS is shown in figure 2a. There are six putative active sites with spheres in the receptor model 1SFM. The residues around TYR116 (upto10Å) are shown in figure 2b. The actives sites predicted by the tool PASS are in agreement with earlier reports[23,33].

Binding sites and active sites of proteins are associated with structural pockets and cavities. A cavity (or

POC:	Molecule	ID	N_mth	Area_sa	Area_ms	Vol_sa	Vol_ms	Lenth	cnr
POC:	/cast/A	1	0	0.000	24.76	0.000	11.58	0.05	4
POC:	/cast/A	2	1	0.071	15.10	0.001	7.45	0.90	3
POC:	/cast/A	3	1	2.768	3.98	0.891	5.25	2.42	2
POC:	/cast/A	4	0	0.030	26.82	0.000	12.97	0.82	6
POC:	/cast/A	5	0	0.051	28.52	0.000	13.92	1.33	6
POC:	/cast/A	6	0	0.067	28.78	0.001	13.98	1.34	6
POC:	/cast/A	7	1	0.440	12.59	0.009	5.87	2.42	3
POC:	/cast/A	8	1	1.268	12.44	0.110	8.38	3.91	5
POC:	/cast/A	9	1	0.247	20.70	0.004	11.58	3.22	7
POC:	/cast/A	10	1	3.198	38.95	0.233	23.55	9.07	8
POC:	/cast/A	11	1	1.864	20.47	1.028	15.90	10.08	7
POC:	/cast/A	12	0	1.049	46.36	0.030	23.19	6.18	12
POC:	/cast/A	13	1	8.714	26.99	2.590	27.79	14.22	10
POC:	/cast/A	14	1	3.316	28.67	0.384	20.14	8.88	12
POC:	/cast/A	15	1	22.918	98.50	4.148	71.42	30.57	19
POC:	/cast/A	16	0	25.046	111.79	5.908	95.14	37.99	28
POC:	/cast/A	17	1	24.094	55.69	12.108	67.62	34.60	22
POC:	/cast/A	18	1	48.904	122.48	16.438	130.17	51.72	28
POC:	/cast/A	19	0	22.675	142.48	3.420	97.65	50.51	36
POC:	/cast/A	20	2	86.521	165.23	40.196	207.83	74.22	35
POC:	/cast/A	21	1	60.463	148.63	18.997	161.70	70.37	44
POC:	/cast/A	22	0	62.780	249.40	16.468	210.86	92.99	64
POC:	/cast/A	23	2	69.756	287.55	11.095	243.51	118.39	69
POC:	/cast/A	24	1	63.385	235.48	20.033	209.54	96.33	67
POC:	/cast/A	25	0	92.219	249.72	42.555	274.41	110.12	60
POC:	/cast/A	26	2	179.870	366.87	115.942	477.06	181.60	93
POC:	/cast/A	27	4	436.085	837.11	248.830	1086.11	406.351	87

N_mth: number of mouth openings for the pocket, Area_sa and Area_ms: The SA and MS area of the pocket or cavity, Vol_sa and Vol_ms: The SA and MS volume of the pocket or cavity, Length: Length sums the arcs of the pocket, cnr: cnr is the total count of the corner points.

Figure 3 : 1SFM Binding pocket informations

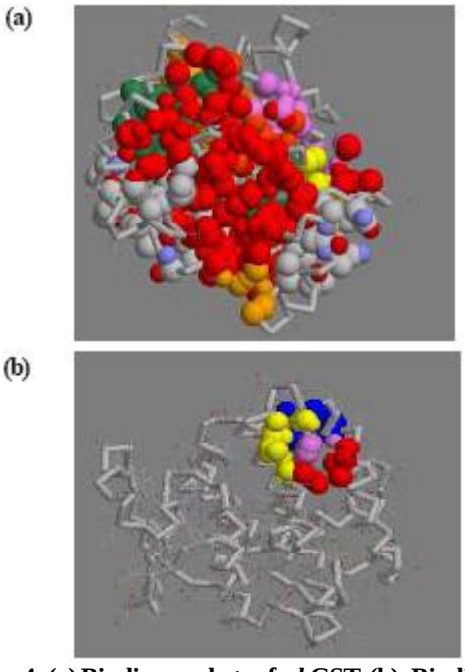

Figure 4: (a) Binding pockets of *wb*GST, (b): Binding pockets around TYR 116 of *wb*GST

void) is an interior empty space that is not accessible to the solvent probe. It has no mouth openings to the outside bulk solution.

CASTp provides identification and measurements of surface accessible pockets as well as interior inaccessible cavities, for proteins and other molecules. It measures analytically the area and volume of each pocket and cavity, both in solvent accessible surface (SA, Richards' surface) and molecular surface (MS, Connolly's surface). It also measures the number of mouth openings, area of the openings, circumference of mouth lips, in both SA and MS surfaces for each pocket.

CASTp identified pockets and cavity, along with the measurements for 1SFM is given in figure 3. All the binding pockets in 1SFM predicted using CASTp are given in figure 4a as visualized by Rasmol[34]. Different binding pockets are shown in spacefill model with grey coloured backbone. The binding pockets around TYR 116 of wbGST are given in figure 4b.

Pockets are empty concavities on a protein surface into which solvent (probe sphere of 1.4 A) can gain access, i.e., these concavities have mouth openings connecting their interior with the outside bulk solution. A total of 27 pockets were identified in 1SFM. The binding pockets predicted by the CASTp server revealed that the TYR 116 is involved in the formation of three binding pockets namely 15, 17 and 21. The largest pocket 21 holds 27 atoms and the residues involved are 109, 113, 117, 163, 164 and 208 besides 116. Pocket 17 holds 16 atoms with residues 100, 103, 104, 108 and 112 excluding residue 116. Similarly, the smallest pocket 15 contains 13 atoms formed from residues 97, 100, 119, 120 and 123 other than 116.

3.3 Ligand generation

Plumbagin, the lead molecule for macrofilaricidal drug development was selected as the seed structure and is shown in figure 5. NCI database search with plumbagin as substructure resulted in 100 hits. The list of hits with NSC IDs and chemical names are given in TABLE 1.

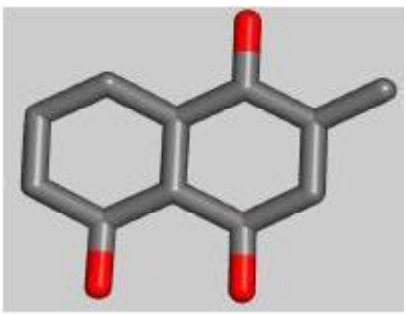

Figure 5 : Plumbagin molecule

Regular Paper

TABLE 1: Hits from NCI database with NSC IDs and chemical names

S.no.	NSC ID	Name
1	NSC8	No name
2	NSC504	1-chloro-2-methylanthra-9,10-quinone
3	NSC607	2-methylanthra-9,10-quinone
4	NSC3871	7,8,9,10-tetrahydro-5,12-naphthacenedione
5	NSC4722	5-O-(3-(hydroxymethyl)-9,10-dioxo-9,10-dihydro-1-anthracenyl)pentose
6	NSC5001	9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
7	NSC6196	7-(hexopyranosyloxy)-3,5,6,8-tetrahydroxy-1-methyl-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
8	NSC7216	2-ethylanthra-9,10-quinone
9	NSC7230	Dichinyl
10	NSC7578	1-anilino-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
11	NSC7581	1-((4'-((2-carboxy-9,10-dioxo-9,10-dihydro-1-anthracenyl)amino)[1,1'-biphenyl]-4-yl)amino)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
12	NSC7824	1-chloro-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
13	NSC7961	benzo[a]anthracene-7,12-dione
14	NSC9026	1-amino-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
15	NSC9027	2-methyl-9,10-dioxo-9,10-dihydro-1-anthracenesulfonyl chloride
16	NSC9608	1-(2-naphthyl)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
17	NSC15908	5,6,11,12,17,18-trinaphthylenehexone
18	NSC16228	4-((9,10-dioxo-8-((5,8,14-trioxo-5,8,13,14-tetrahydronaphtho[2,3-c] acridin-4-yl)amino)-9,10-dihydro-1-anthracenyl)amino)naphtho[2,3-c] acridine-5,8,14(13H)-trione
19	NSC17351	5a,11a-dihydro-5,6,11,12-naphthacenetetrone
20	NSC18334	Methyl-2-ethyl-2,5,7,10-tetrahydroxy-6,11-dioxo-4-((2,3,6-trideoxy-4-O-(2,6-dideoxy-4-O-(6-methyl-5-oxotetrahydro-2H-pyran-2-yl) hexopyranosyl)-3-(dimethylamino) hexopyranosyl)oxy)-1,2,3,4,6,11-hexahydro-1-naphthacenecarboxylate
21	NSC18335	Cinerubin B
22	NSC21205	Pluramycin A
23	NSC27034	1,4-dihydroxy-9,10-dioxo-9,10-dihydro-2,3-anthracenedicarbonitrile
24	NSC30546	1,3-dihydroxy-2-(hydroxymethyl)anthra-9,10-quinone
25	NSC30548	2-tert-butylanthra-9,10-quinone
26	NSC30555	7-acetyl-6-ethyl-3,5,8-trihydroxy-9,10-dioxo-9,10-dihydro-1,2-anthracenedicarboxylic acid
27	NSC30839	benzo[b]triphenylene-9,14-dione
28	NSC30869	10-methylbenzo[a]anthracene-7,12-dione
29	NSC31000	dibenzo[a,h]anthracene-7,14-dione
30	NSC37130	1-hydroxy-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
31	NSC37131	1-hydroxy-2-methylanthra-9,10-quinone
32	NSC37132	1,8-dihydroxy-3-methylanthra-9,10-quinone
33	NSC37223	7,7'-Bializarin
34	NSC37598	1-chloro-9,10-dioxo-9,10-dihydro-2-anthracenecarbonyl chloride
35	NSC37599	1-chloro-2-(dibromomethyl)anthra-9,10-quinone
36	NSC37600	2-(dibromomethyl)-1-(hydroxy(oxido)amino)anthra-9,10-quinone
37	NSC37601	1-mercapto-2-methylanthra-9,10-quinone
38	NSC37615	1,3,5,7-tetramethylanthra-9,10-quinone
39	NSC37621	2-benzoyl-1-chloroanthra-9,10-quinone
40	NSC37623	1-((2-carboxyphenyl)thio)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
41	NSC38628	1,8-dihydroxy-3-(hydroxymethyl)anthra-9,10-quinone
42	NSC38629	4,5-dihydroxy-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
43	NSC39889	1-amino-3-bromo-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
44	NSC39896	6-chloro-11-hydroxy-5,12-naphthacenedione
45	NSC39897	6-hydroxy-5,12-naphthacenedione
46	NSC39898	6-amino-11-hydroxy-5,12-naphthacenedione
47	NSC39911	1-chloro-N-(9,10-dioxo-9,10-dihydro-2-anthracenyl)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxamide
48	NSC39913	1-amino-N-(9,10-dioxo-9,10-dihydro-1-anthracenyl)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxamide
49	NSC39914	1-amino-N-(9,10-dioxo-9,10-dihydro-2-anthracenyl)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxamide
50	NSC39918	N-(3-methyl-2,7-dioxo-2,7-dihydro-3H-naphtho[1,2,3-de]quinolin-6-yl)-5,8,14-trioxo-5,8,13,14-tetrahydronaphtho[2,3-c]acridine-10-carboxamide
51	NSC39939	5-amino-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
52	NSC39940	1-amino-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
53	NSC39943	1-amino-2-methylanthra-9,10-quinone
54	NSC39944	4-amino-1-hydroxy-2-methylanthra-9,10-quinone

To be countinue next page

S.no.	NSC ID	Name
55	NSC39950	1-amino-2-ethylanthra-9,10-quinone
56	NSC39958	1-amino-2-benzoylanthra-9,10-quinone
57	NSC39964	2-(1-chloro-9,10-dioxo-9,10-dihydro-2-anthracenyl)anthra[2,1-d][1,3]thiazole-6,11-dione
58	NSC39965	2-(1-amino-9,10-dioxo-9,10-dihydro-2-anthracenyl)anthra[2,1-d][1,3]thiazole-6,11-dione
59	NSC47720	6,10,12-trichloronaphtho[2,3-c]acridine-5,8,14(13H)-trione
60	NSC49891	No Name
61	NSC51543	N-(9-(benzoylamino)-5,10,15,17-tetraoxo-10,15,16,17-tetrahydro-5H-dinaphtho[2,3-a:2,3-i]carbazol-6-yl)benzamide
62	NSC53156	9,10-dihydroxy-2,3,5,6,7,8-hexahydro-1,4-anthracenedione
63	NSC59063	3-hydroxy-1-methoxy-2-methylanthra-9,10-quinone
64	NSC62490	Cinerubin B
65	NSC70845	Nogalamycin
66	NSC70929	Hedamycin
67	NSC76511	5H-naphtho[2,3-a]carbazole-5,13(12H)-dione
68	NSC76612	2,3-dimethylanthra-9,10-quinone
69	NSC79232	7-methyl-5H-naphtho[2,3-a]carbazole-5,13(12H)-dione
70	NSC82151	3-acetyl-3,5,12-trihydroxy-10-methoxy-6,11-dioxo-1,2,3,4,6,11-hexahydro-1-naphthacenyl 3-amino-2,3,6-trideoxyhexopyranoside
71	NSC82289	1-amino-4-chloro-2-methylanthra-9,10-quinone
72	NSC82322	1-(hydroxy(oxido)amino)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid
73	NSC82892	Nogalamycin compound B
74	NSC83142	3-acetyl-3,5,12-trihydroxy-10-methoxy-6,11-dioxo-1,2,3,4,6,11-hexahydro-1-naphthacenyl 3-amino-2,3,6-trideoxyhexopyranoside
75	NSC86005	NOGALAROL
76	NSC86262	12-hydroxy-6,11-dioxo-6,11-dihydro-5-naphthacenesulfonic acid
77	NSC90571	9,10-dimethyl-5H-naphtho[2,3-a]carbazole-5,13(12H)-dione
78	NSC90986	9,10-dichloro-5H-naphtho[2,3-a]carbazole-5,13(12H)-dione
79	NSC92938	5H-naphtho[2,3-a]carbazole-5,13(12H)-dione
80	NSC92941	12-methyl-5H-naphtho[2,3-a]carbazole-5,13(12H)-dione
81	NSC93419	3,10,12-trihydroxy-2,8-dimethoxy-3-methyl-4,6,11-trioxo-1,2,3,4,6,11-hexahydro-1-naphthacenyl 6-deoxy-2-O-methylhexopyranoside
82	NSC95407	6,11-dihydroxy-5,12-dioxo-5,12-dihydro-2-naphthacenecarboxylic acid
83	NSC98908	2-((2-(1,3-dioxolan-2-yl)-1lambda~5~-pyridin-1-yl)methyl)anthra-9,10-quinone
84	NSC100290	methyl 2-ethyl-2,5,7-trihydroxy-6,11-dioxo-4-((2,3,6-trideoxy-3-(dimethylamino)hexopyranosyl)oxy)-1,2,3,4,6,11-hexahydro-1-naphthacenecarboxylate hydrochloride
85	NSC102813	.beta.-Rhodomycin
86	NSC102815	Daunomycin compound D
87	NSC106763	5,12-dioxo-1,4,4a,5,12,12a-hexahydro-1-naphthacenyl acetate
88	NSC109351	8-acetyl-6,8,10,11-tetrahydroxy-1-methoxy-7,8,9,10-tetrahydro-5,12-naphthacenedione
89	NSC110428	3-acetyl-1-((4-O-benzoyl-3-(benzoylamino)-2,3,6-trideoxyhexopyranosyl)oxy)-3,12-bis(benzoyloxy)-10-methoxy-6,11-dioxo-1,2,3,4,6,11-hexahydro-5-naphthacenyl benzoate
90	NSC111708	12-methoxy-6,11-dioxo-1,2,3,4,6,11-hexahydro-5-naphthacenyl acetate
91	NSC112758	3-acetyl-3,5,12-trihydroxy-10-methoxy-6,11-dioxo-1,2,3,4,6,11-hexahydro-1-naphthacenyl 3-amino-2,3,6-trideoxyhexopyranoside
92	NSC112857	12-(acetyloxy)-6,11-dioxo-1,2,3,4,6,11-hexahydro-5-naphthacenyl acetate
93	NSC112877	2,4-dimethylpentyl 1-(hydroxy(oxido)amino)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylate
94	NSC112894	3-ethylpentyl 1-(hydroxy(oxido)amino)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylate
95	NSC112923	8-(acetyloxy)-3-methyl-9,10-dioxo-9,10-dihydro-1-anthracenyl acetate
96	NSC113465	1,12-bis(acetyloxy)-6,11-dioxo-1,2,3,4,6,11-hexahydro-5-naphthacenyl acetate
97	NSC113467	6,11-dihydroxy-7,8,9,10-tetrahydro-5,12-naphthacenedione
98	NSC114019	6,7,11-trihydroxy-7,8,9,10-tetrahydro-5,12-naphthacenedione
99	NSC114343	1,4,5,8-tetrahydroxy-2-methylanthra-9,10-quinone
100	NSC114345	7-Deoxynogalarol

These hits were used as ligands for virtual screening with the filarial drug target glutathione S transferase of *W.bancrofti* (1SFM) as the biological receptor.

3.4. Virtual screening

The one hundred hits from NCI database were docked with 1SFM, the target enzyme using the program Autodock and from the resulting conformations, top 21 conformations were selected based on their Binding Energy values. The results are shown in TABLE 2.

The accurate prediction of enzyme-substrate interaction energies is one of the major challenges in com-

Regular Paper

TABLE 2 : Binding energies and inhibition constants of top 21 hits

NSC ID	B-ENERGY kcal/mol	K_i
NSC39965	-8.36	+7.47e -07
NSC39964	-8.30	+8.30e -07
NSC37223	-7.66	+2.43e -06
NSC9608	-6.98	+7.62e -06
NSC95407	-6.49	+1.76e -05
NSC47720	-6.45	+1.87e -05
NSC92941	-6.42	+1.98e -05
NSC7824	-6.35	+2.21e -05
NSC30839	-6.35	+2.21e -05
NSC79232	-6.31	+2.38e -05
NSC39914	-6.31	+2.36e -05
NSC90571	-6.30	+2.40e -05
NSC7230	-6.29	+2.47e -05
NSC86262	-6.29	+2.44e -05
NSC90986	-6.20	+2.87e -05
NSC39913	-6.18	+2.93e -05
NSC76511	-6.14	+3.17e -05
NSC92938	-6.14	+3.15e -05
NSC106763	-6.04	+3.72e -05
NSC37621	-6.02	+3.89e -05
NSC38629	-6.02	+3.90e -05

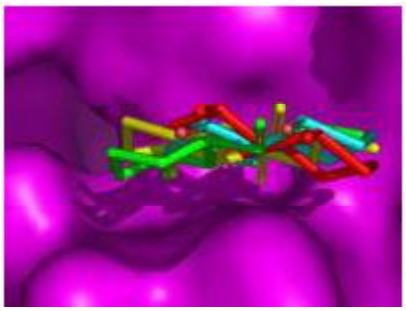

Figure 6: Plumbagin binding with receptor in various conformations

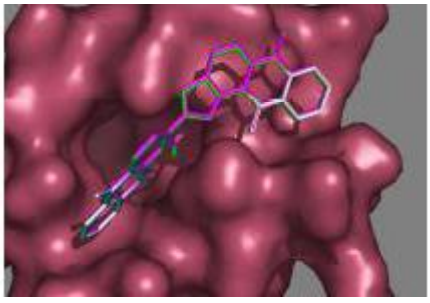

Figure 7 : Compound with highest affinity among NCI hits binding to receptor

putational biology. Protein-ligand docking methods aim to predict the binding energy of the protein-ligand complex given the atomic coordinates. In such calculations, both the protein and ligand can be treated as rigid bodies[35,36]; alternately, the ligand, the protein, or both molecules, can be completely or partially flexible[37,38].

One advantage of incorporating flexibility is that it enables a search without bias introduced by the initial model. This bias normally influences both the orientation and conformation of the ligand in the active site, which usually represents a local minimum conformation[39]. More importantly, the lock and key concepts used to evaluate enzyme-substrate binding, in reality, refer to flexible locks and keys that are both in constant dynamic (thermal) motion[40].

Binding energy is the sum of final intermolecular energy and torsional free energy. The scoring method used here is based on the binding energies of receptor-ligand complexes. A more negative binding energy represents a more favourable binding interaction. The binding energies of the top 21 ligands varies from –8.36 to –6.02 Kcal/mol. The five best scored ligands were found to be compounds with NSC Ids as follows; NSC39965 (2-(1-amino-9,10-dioxo-9,10-dihydro-2-anthracenyl)anthra[2,1-d] [1,3]thiazole-6,11-dione), NSC39964 (2-(1-chloro-9,10-dioxo-9,10-dihydro-2-anthra cenyl)anthra [2,1-d][1,3]thiazole-6,11-dione), NSC37223 (7,7'-Bializarin), NSC9608 (1-(2-naphthyl)-9,10-dioxo-9,10-dihydro-2-anthracenecarboxylic acid), NSC95407 (6,11-dihydroxy-5,12-dioxo-5,12-dihydro-2-naphthacenecarboxylic acid).

Plumbagin docked to the receptor 1SFM is given in figure 6. Top scored hit (NSC39965) among the set of 21 compounds is given in figure 7. The top scoring 5 compounds docking conformations is given in **figure** 8.

Database search with plumbagin as substructure yielded 100 compounds which on subsequent virtual screening using Autodock with filarial GST as the receptor resulted in 21 hits containing anthraquinone type molecules. These molecules may play a vital role in the future development of macrofilaricidal molecules which may be useful in eliminating the human lymphatic filariasis. The importance or the validity of this work is confirmed by the recent report on the antifilarial activity of anthraquinone type molecules[41].

3.5. Pharmacokinetic properties of top scored hits

The pharmacokinetic properties of the molecules mainly depend on their physicochemical properties. The physico-chemical properties of the top 21 hits from virtual screening for macrofilaricidal activity were calculated and the results are given in TABLE 3.

The physical and chemical properties of a com-

TABLE 3 : Physicochemical properties of top 21 NCI hits

NSC ID	Physicochemical properties mol Wt	H B acc	H B do	log p	ADME vio	Sur area	Mol vol
7230	442.48	4	0	5.4578	1	235.416	148.994
7824	286.67	4	1	2.6592	0	166.703	104.193
9608	378.39	4	1	4.8278	0	208.765	134.224
30839	308.34	2	0	4.4472	0	185.687	117.03
37223	478.42	8	4	3.3858	0	271.947	174.014
37621	346.77	3	0	4.1822	0	193.81	126.483
38629	284.23	6	3	1.5724	0	172.132	105.723
39913	472.47	5	2	2.8176	0	269.243	173.771
39914	472.47	5	2	2.8176	0	275.881	174.901
39964	505.93	5	0	4.9448	1	289.906	177.576
39965	486.51	5	1	3.6436	0	285.058	174.131
47720	428.65	3	1	4.2839	0	218.078	151.955
76511	297.32	2	1	2.9824	0	180.404	113.995
79232	311.35	2	1	3.4496	0	187.109	118.357
86262	354.34	6	2		Invalid operand	185.711	114.696
90571	325.38	2	1	3.9168	0	200.524	124.11
90986	366.2	2	1	4.0184	0	199.456	133.245
92938	297.32	2	1	2.9824	0	180.404	113.995
92941	311.35	2	0	3.2289	0	185.241	116.782
95407	334.29	6	3	2.5746	0	198.258	123.187
106763	320.36	4	0	2.1391	0	194.288	119.876

pound are a function of its molecular structure. Structure-property relationships are developed by finding one or more molecular descriptors explaining variations in the physical or chemical properties of a group of congeners/analogs. While some descriptors can be determined experimentally, deriving them from either the 2D or 3D molecular structure is generally more convenient and practical. A relationship, once quantified, can be used to estimate the properties of other molecules simply from their structures and without the need for experimental determination or synthesis.

Here the physicochemical properties such as log P, molecular volume, polar surface area and hydrogen bond donor/acceptor properties were calculated for the top 21 hits and their ADME violation has been generated from their 2D structures using TSAR 3.3. The molecular weight ranged from 284.23 to 505.93, hydrogen bond acceptors from 2 to 8, hydrogen bond donors from zero to 3, log P from 1.5724 to 5.4578, polar surface area from 166.703 to 289.906 and molecular volume from 105.723 to 177.576. ADME violation was observed only for two compounds bearing NSC ids 7230 and 39964. The violations were seen with respect to logP value in the case of compound NSC 7230 and with respect to molecular weight in the case of NSC39964. An invalid parameterization problem was observed with compound NSC86262. All the other hits agree with Lipinski's "rule-of-five"[25].

4. CONCLUSION

Database search with plumbagin as substructure yielded 100 compounds which on subsequent virtual screening using Autodock with filarial GST as the receptor resulted in 21 hits containing anthraquinone type molecules. From the work done it is clear that the lead compounds generated using virtual screening are of very valuable for further work in drug development, and also the principles of computational biology gives us better results in short duration and results are in agreement with reported experimental results. Lead compounds selected may be tested *in vitro* and *in vivo* against adult filarial parasites. These molecules may play a vital role in the future development of macrofilaricidal molecules which may be useful in eliminating the human lymphatic filariasis.

5. REFERENCES

[1] W.Cao, C.P.B.Ploeg, Z.Ren, K.D.F.Habbema; World Health Forum, **18**, 17-20 **(1997)**.

[2] S.K.Srivastava, A.Agarwal, P.M.Chauhan, S.K. Agarwal, A.P.Bhaduri, S.N.Singh, N.Fatima, R.K. Chatterjee; J.Med.Chem., **42(9)**, 1667-72 **(1999)**.

[3] E.A.Ottesen, C.P.Ramachandran; Parasitol Today, **2**, 129-131 **(1995)**.

[4] S.Sharma; Prog Drug Res., **35**, 365-485 **(1990)**.

[5] P.E.Simonsen, D.W.Meyrowitsch, W.H.Makunde; Trans Royal Soc., Trop Med. Hyg., **91**, 290-93 **(1997)**.

[6] J.L.Bennett, J.F.Williams, V.Dave; Parasitol Today, **4**, 226-228 **(1988)**.

[7] C.D.Ginger; Parasitol Today, **2**, 38-40 **(1986)**.

[8] P.M.Loiseau, P.Depreux; Ann.Trop Med.Parasitol., **87**, 469-476 **(1993)**.

[9] P.M.S.Chauhan, R.K.Chatterjee; Indian J.Chem., **33B**, 32-37 **(1994)**.

[10] S.K.Srivastava, P.M.S.Chauhan, S.K.Agarwal, A.P.Bhaduri, S.N.Singh, N.Fatma, R.K.Chatterjee, C.Bose, V.M.L.Srivastava; Bioorg.Med.Chem. Lett., **6**, 2623-2628 **(1996)**.

[11] S.N.Singh, S.Bhatnagar, N.Fatma, P.M.S.Chauhan, R.K.Chatterjee; Trop Med Inter Health., **2**, 535-543 **(1997)**.

[12] S.Tewari, P.M.S.Chauhan, A.P.Bhaduri, S.N.Singh, N.Fatma, R.K.Chatterjee, V.M.L.Srivastava; Bioorg.Med.Chem.Lett., **7**, 1891-1896 **(1997)**.

[13] S.Sharma; Adv.Drug Res., **25**, 103-172 **(1994)**.

[14] L.B.Townsend, D.S.Wise; Parasitol Today, **6**, 107-

Regular Paper

112 **(1990)**.
[15] U.R.Rao, G.Salinas, K.Mehta, T.R.Klei; Parasitol. Res., **86(11)**, 908-15 **(2000)**.
[16] R.B.Grieve, N.Wisnewski, G.R.Frank, C.A.Tripp, M.F.Powell, M.J.Newman; 'Vaccine Research and Development for the Prevention of Filarial Nematode Infections', Plenum Press, NewYork, 737-768 **(1995)**.
[17] J.Lazdins, M.Kron; Parasitol Today, **15(8)**, 305-306 **(1999)**.
[18] P.M.Brophy, D.I.Pritchard; Exp.Parasitol., **79**, 89-96 **(1994)**.
[19] T.D.Boyer; Hepatol., **9**, 486-496 **(1989)**.
[20] R.N.Armstrong; Chem.Res.Toxicol., **4**, 131-140 **(1991)**.
[21] P.J.Van Bladeen, B.Van Oommen; Pharmacol Ther., **51**, 35-46 **(1991)**.
[22] G.F.Mitchell; Parasitol Today, **5**, 34-37 **(1989)**.
[23] S.T.Nathan, M.Nisha, M.Kalyanasundaram, K.Balaraman; J.Mol.Model, Published online, DOI, 10.1007/s00894-005-0234-0).
[24] M.Nisha, K.P.Paily, P.Vanamail, Abidha, M. Kalyanasundaram, K.Balaraman; Drug.Dev Res., **56(1)**, 33-39 **(2000)**.
[25] C.A.Lipinski, F.Lombardo, B.W.Dominy, P.J. Feeney; Adv.Drug Del.Res., **23**, 3-25 **(1997)**.
[26] A.V.Veselovsky, A.S.Ivanov; Curr Drug Targets-Infectious Disorders, **3**, 33-40 33 **(2003)**.
[27] G.P.Brady, P.F.Stouten; J.Comput Aided Mol.Des., **14**, 383-401 **(2000)**.
[28] J.Liang, H.Edelsbrunner, C.Woodward; Protein Science, **7**, 1884-1897 **(1998)**.
[29] G.M.Morris, D.S.Goodsell, R.S.Halliday, R.Huey, W.E.Hart, R.K.Belew, A.J.Olson; J.Comput. Chem., **19**, 1639-1662 **(1998)**.
[30] D.A.Pearlman, D.A.Case, J.W.Caldwell, W.S.Ross, T.E.Cheatham, S.I.DeBolt, D.Ferguson, G.Seibel, P.Kollman; Comp.Phys.Commun., **91**, 1-41 **(1995)**.
[31] S.J.Weiner, P.A.Kollman, D.A.Case, U.C.Singh, C.Ghio, G.Alagona, S.Profeta, P.Weiner; J.Am. Chem.Soc., **106**, 765-784 **(1984)**.
[32] J.G.Hardman, L.E.Limbird, P.B.Molinoff, R.W. Ruddon, A.G.Gilman; The Pharmacological Basis of Therapeutics., McGraw-Hill, 9th Edition, 1905 **(1996)**.
[33] P.M.Brophy, A.M.Campbell, A.J.Van Eldik, P.H. Teesdale-Spittle, E.Liebau, M.F.Wang; Bioorg. Med.Chem.Lett., **10(9)**, 979-81 **(2000)**.
[34] Roger Sayle, E.James Milner-White; Trends Biochem.Sci., (TIBS), **20(9)**, 374 **(1995)**.
[35] I.A.Vakser; Proteins, Suppl., **1**, 226-230 **(1997)**.
[36] I.A.Vakser, O.G.Matar, C.F.Lam; Proc.Natl.Acad Sci., USA, **96**, 8477-8482 **(1999)**.
[37] L.Schaffer, G.M.Verkhivker; Proteins, **33**, 295-310 **(1998)**.
[38] G.M.Verkhivker, P.A.Rejto; Proc.Natl.Acad Sci., USA, **93**, 60-64 **(1996)**.
[39] N.Ota, D.A.Agard; J.Mol.Biol., **30**, 607-617 **(2001)**.
[40] M.J.Betts, M.J.Sternberg; Protein Eng, **12**, 271-283 **(1999)**.
[41] M.R.Dhananjeyan, Y.P.Milev, M.A.Kron, M.G. Nair; J.Med.Chem., **48(8)**, 2822-2830 **(2005)**.