

Original Article

GENETIC INVOLVEMENT OF INTERLEUKIN 4 FOR ASTHMA AND IDENTIFICATION OF POTENTIAL PHYTOCHEMICAL SCAFFOLD THROUGH MOLECULAR DOCKING STUDIES

S. SARITHAMOL¹, DIVYA V.³, SUNITHA V. R.³, SUCHITRA SURENDRAN¹, V. L. PUSHPA^{1*}, K. B. MANOJ²

¹P G and Research Department of Chemistry, Sree Narayana College, Kollam, Kerala, India 691001, ²Sree Narayana College, Cherthala, Alapuzha, ³Department of Chemistry, MSM College, Kayamkulam
Email: drpushpavl2017@gmail.com

Received: 21 Oct 2017, Revised and Accepted: 12 Dec 2017

ABSTRACT

Objective: Interleukin 4, an important cytokine, has the major role in the immunomodulatory responses associated with asthma. The present study focused on the involvement of single nucleotide polymorphism variation (SNP) of interleukin 4 (IL4) in the development of disease, asthma and designing small molecules for the inhibition of IL4 through *in silico* strategy.

Methods: Identification of disease causing SNP will be a wise approach towards the phenotype specific treatment. A human origin deleterious non synonymous SNP of IL4 were found out in the chromosome region 5q31-q33 (rs199929962) (T/C). Proteins of the corresponding nucleotide variation were identified and were subjected to characterization studies for selecting the most appropriate one for further mutational analysis and molecular docking studies.

Results: Influence of microbes on SNP variation of IL4 gene leading to asthma was found to be insignificant by metagenomic studies. Gene responsive drugs were identified through environmental factor analysis. The drug candidates including corticosteroids were subjected to protein interaction studies by *in silico* means. The pharmacophoric feature derived from drug receptor interaction was utilized for virtual screening on a dataset of anti-inflammatory phytochemicals. The scaffolds of ellagic acid and quercetin were identified as potential nonsteroidal entities which can shield the asthmatic activities.

Conclusion: Developing small molecules using these scaffolds taking interleukin 4 as a target will be an adequate solution for steroid resistant asthma.

Keywords: Interleukin 4, Single nucleotide polymorphism, Cell signaling pathway, Mutation studies, Protein characterization studies, Pharmacophore, Molecular docking, Phytochemicals

© 2018 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)
DOI: <http://dx.doi.org/10.22159/ijcpr.2018v10i1.24704>

INTRODUCTION

Asthma is an adverse condition engendered by lung inflammatory cells due to their overexpressed inflammatory replications towards sundry agents such as dust, allergic pathogens and chemicals [1]. According to the recent findings by WHO, more than 300 million people including children are the victim of this rigorous condition [2]. Popular medications to overcome this situation are corticosteroids, leukotriene modifiers, mast cell stabilizers and adrenergic receptor [3-4]. But, the most asthmatic conditions are being unresponsive towards these treatments. Most of the treatments are engendering earnest side effects including magnification retardness in children, reduced receptor efficiency and incremented blood pressure [5]. Different phenotypes are exhibiting unique replication towards the treatments due to the variation in their genetic signature [6]. Pharmacogenomics is a study of the relationship between genetic variation and pharmacotherapy. Pharmacogenomic studies are mainly fixated on single nucleotide polymorphisms (SNPs) which occur in the nucleotide sequences of DNA. SNPs are of variant, sometimes they may be considered as unresponsive albeit they occupy the noncoding region of DNA [7]. SNPs occur in the nonsynonymous coding region stand for disease generation [7]. Atopic asthma is occurred by composing immunoglobulin E (IgE) antibody. The main culprit behind the pathophysiological changes associated with asthma is Th2 cytokines including IL4 (interleukin 4), IL5, and IL13 [8]. Genes corresponding to these members are located on the chromosome region 5q31-q33, also, studies revealed that this region is highly prone to asthma [9]. The present study is focusing on the investigation of sundry types of deleterious SNPs in the chromosome region 5q31-q33 associated to IL4. The disease generation occurs through a cell signaling pathway managed by Janus Kinase-1(JAK1) enzyme and signal transducer and activator of transcription 6 (STAT 6) proteins through tyrosine phosphorylation activation [10]. *In silico* bioinformatics implements

and databases nowadays are popular in pharmaceutical field [7]. The external stimuli including pathogens and allergens in the airways have considerable contribution towards the disease pathogenesis. The role of microbes as well as chemicals and drugs in the disease progression and management through interleukin 4 gene were under the scope of the study. Further work included ascertaining of disease-associated SNPs of IL4 gene in the chromosome region 5q31-q33 and detection of the most contributing SNP.

Plant world is a promising repository of compounds with very high therapeutic potential. Most of the today's popular drugs were derived from various potential phytochemicals [11]. The present study is also focusing on searching and discovery of potential anti-inflammatory phytochemicals as a substitute for popular corticosteroids by several *in silico* tools including molecular docking. *In silico* tools nowadays are very popular in the pharmaceutical field for drug design in a speedy curated and economically feasible way [12].

Experimental

MATERIALS AND METHODS

SNP analysis

Single nucleotide polymorphism (SNP) studies are highly recommended for the personalization in drug designing techniques. Since high expression of IL4 is associated with asthma, all the SNPs of IL4 in the chromosome region 5q31-q33 were checked for their deleterious nature [13]. NCBI. National Center for Biotechnology Information Database (NCBI) is an updated repository for the genetic information of various types of genes [14]. All missense SNPs of IL4 of hominids origin was collected. SIFT (sorting intolerant from tolerant) and PolyPhen-2 (Polymorphism Phenotyping v2) analyses were executed to check

whether the SNPs are deleterious or not [15-16]. SIFT analysis can predict the problems associated with amino acid substitution, which is based on the conservation of amino acid residues in sequence alignments derived from closely associated sequences, accumulated via psi-blast. PolyPhen-2 predicts feasible effect of an amino acid substitution at the structure and function of a human protein using straightforward physical and comparative considerations. SIFT has a basic concept that, conserved amino acids in the protein chain will keep as such and the problems created in those regions is due to the miss functioning of proteins. The changes may or may not be deleterious. The cut-off value of SIFT tolerance is 0.05, the score equal to or greater than 0.05 can be treated as tolerating change while the score value less than 0.05 can be treated as deleterious changes [13]. The functional activity of protein has been checked using the PolyPhen-2 tool. Secondary structure information of protein was analyzed using this tool. From the secondary structure information, the algorithm calculated 'position-specific independent count score for all SNPs. The score difference of variations due to different SNPs was recorded. As the score difference increases, the variations generated by those regions can be considered. If the score difference is greater than or equal to 0.9, the variation can be treated as problematic.

Gene association studies

Asthma is a multi-targeted disease. A large variety of genes are working behind the disease generation. Genes having the considerable contribution towards IL4 found out through the online bioinformatics database, STRING (Search Tool for the Retrieval of Interacting Genes/Proteins). STRING database comprised of predicted as well as known protein-protein interactions. The STRING information is based on experimental results, computational predictions and text information.

Metagenomic studies

Metagenomic is the study of influence of microbes on genetic variants of a disease [17]. The identification of involvement of microbes to the atopic diseases asthma is more helpful in disease pathogenesis. The online tool NCBI-BLAST was used to execute metagenomic studies.

Environmental factor analysis

The influence of external factors including chemicals, allergens and other natural materials on the gene IL4 for the disease asthma were analyzed using the online database CTD (Comparative Toxicogenomics Database) [18].

Records of protein from deleterious SNPs

Gene-level information is not enough for drug designing technique. Drug development, treatment associated studies and studies on the side effect of drugs will be possible from protein level studies only. In the present study, all the possible proteins associated to the deleterious SNPs were found out using LS-SNP/PDB online data base [19].

Protein characterization studies

The online tools, ExPASy-Prot Param and SOPMA (self-optimized prediction from multiple alignments) were used to do the primary and secondary structure analysis on all the proteins associated to deleterious SNPs [20-21]. The physicochemical properties such as instability index, aliphatic index, half-life and GRAVY (grand average hydropathy) were analyzed using the ExPASy-Prot Paramtool.

The protein which satisfies all the characterization parameters were selected for checking the thermodynamic stability of SNPs and corresponding X-ray crystal structure has been retrieved from the protein database (PDB). Finally, the protein was subjected to energy minimization and the energy of non-mutated and mutated forms were checked using Swiss-PDB Viewer software at a force field of GROMOS 96.

The protein validation was carried out in Swiss PDB viewer by Ramachandran plot analysis. The stereo chemical aspects were considered for analyzing protein.

Molecular docking studies

The protein obtained as a result of protein characterization studies was subjected to molecular docking studies using all the drug molecules obtained from environmental factor studies on IL4 gene and possible intermolecular interactions furnished by the drug-receptor association were also analyzed. Glide module of the molecular modeling software Schrodinger 2015-1 was used for the molecular docking studies. The protein preparation and energy minimization were done using the protein preparation wizard of the same software. During the preparation process, addition of hydrogen atoms, correction in terms of bond order was done. All the processes were done at a force field of OPLS_2005 at pH 7.4. The potential binding site in the protein crystal structure was assigned by the site map module of Schrodinger 2015-1 software. Site map consider the size, functionality requirement and the solvent exposure to the site and derive site score corresponding the site efficiency towards ligand binding. The site having maximum site score was selected for molecular docking and virtual screening processes.

Energy based pharmacophore generation

A potential pharmacophoric feature has derived from the protein-potential drug association using E-pharmacophoric analyses and executed virtual screening on a set of anti-inflammatory phytochemicals based on the pharmacophore feature thus obtained. All the default settings were kept as such while generating pharmacophoric feature. The phytochemicals showing similarity in terms of functionality, volume and shape were subjected to molecular docking analysis with the IL4 protein and find out the best interacting ligand molecule among them.

Energy minimization calculations of docking conformations

The docked conformations of protein-phytochemical complexes were subjected to energy minimization studies using the prime module of Schrodinger 2015-1 and binding energy changes of the complex and independent systems were calculated by Molecular Mechanics/Generalized Born Surface Area (MM-GBSA) continuum solvent model at a force field of OPLS-2005 incorporating VSGB solvent model and rotamer search algorithms.

RESULTS AND DISCUSSION

SNP analysis

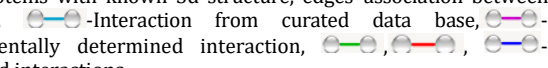


The database dbSNP of NCBI contains 5085 SNPs of the gene, IL4. Out of which 62 human variant missense SNPs were filtered. The SIFT and PolyPhen-2 analyses shortlisted the number of deleterious SNPs into 5. The detailed score values of the 5 deleterious SNPs are displayed in table 1. SIFT and PolyPhen-2 score of those five variations were in the range of deleterious SNPs. These five SNPs were subjected to detailed analysis to check their sustainability.

Table 1: SIFT polyphen-2 score of deleterious SNPs

SNP ids	Codons	Protein sequence change	SIFT score	PolyPhen-2 score
rs149950065	GCG/GGG	ALA 118 GLY	0.02	0.979
	GCG/GGG	ALA102GLY	0.01	0.979
rs199929962	ATG/ACG	MET 144 THR	0.03	0.994
	ATG/ACG	MET128THR	0.04	0.994
rs202231191	AAC/TAC	ASN 113 TYR	0.01	0.992
	AAC/GAC	ASN97ASP	0.05	0.913
	AAC/TAC	ASN97TYR	0.01	0.992
rs376367511	TGT/CGT	CYS 123 ARG	0	0.99
	TGT/CGT	CYS107ARG	0	0.99
rs775753738	CTC/TTC	LEU 110PHE	0.01	0.975
	CTC/TTC	LEU94PHE	0.01	0.975

Gene-gene interaction studies

Involvement of neighboring genes of IL4 associated with asthma has been found out using the online bioinformatics tool, STRING. It has been found out that, there are 10 potential genes having significant interaction with the gene, IL4.

Nodes represents proteins produced by single protein-coding gene locus, empty node-proteins with unknown protein 3D structure, filled node-proteins with known 3d structure, edges-association between proteins, -Interaction from curated data base, -experimentally determined interaction, -predicted interactions.

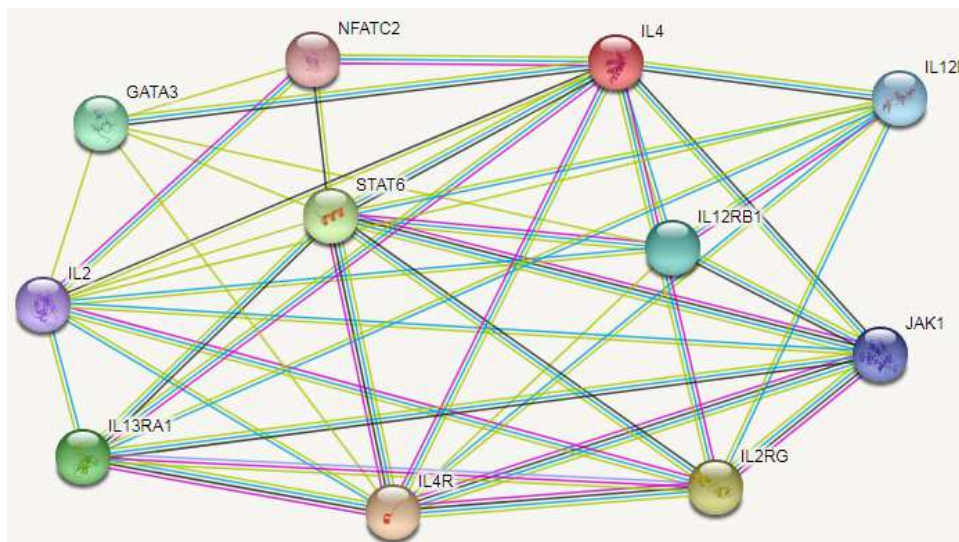


Fig. 1: IL4 gene association with other genes

Table 2: IL4-associated gene interaction score table

S. No.	Associated gene	Interaction score
1	IL4R	0.998
2	IL2RG	0.995
3	STAT6	0.993
4	IL13RA1	0.992
5	GATA3	0.989
6	IL12RB1	0.981
7	IL12B	0.981
8	JAK1	0.980
9	IL2	0.979
10	NFATC2	0.979

The gene-gene interaction score greater than 0.9 are considered as potential interactions. The interacting score of IL4 with some 10 associated genes is shown in table 2. For asthma, IL4 responds through JAK1-STAT6 signaling pathway [22]. The signaling pathway will be active by the entry of allergen and through the association of IL4 with its receptor, IL4R. Due to the interaction of IL4 with IL4R, the non-receptor tyrosine kinase enzyme Janus kinase 1 (JAK1) get activated and which will phosphorylate the receptor, followed by the activation and dimerization of STAT 6 protein. The dimerized STAT6 migrated to the nucleus and there it causes gene transcription associated with asthma. The gene association studies have revealed that the genes of JAK1, IL4R and STAT6 have created an interaction score of >0.9 with the gene, IL4. This signifies the potential of the associated genes and corresponding signaling pathway towards the disease, asthma and hence we choose this pathway to design a potential drug like molecule to prevent it in future works.

Metagenomic analysis

Metagenomic studies were conducted using the online server NCBI/BLAST. IL4 gene is subjected to metagenomic analysis and identified some 9 microbes having >90% similarity with the IL4 in the intron variant region. Neither of the deleterious SNPs has found in this region. These indicate that microbes have no marked effect on IL4 gene which leads to Asthma and is mainly associated with genetic involvements.

Environmental factor analysis

Environmental factor analysis using the CTD software has produced a list of various environmental factors influencing the gene IL4. From the CTD data base, all the popular drug molecules having effect on IL4 were identified, which were subjected to EDR analysis to find out their various biological properties. The drugs include albuterol, beclomethasone, budesonide, dexamethasone, fluticasone, formoterolfumarate, hydrocortisone, methyl prednisolone, prednisolone, prednisone, verlukast and zileuton, all were reported for decreasing IL4 expression. Air pollutants, occupational pollutants, chemicals such as 4, 4'-diphenylmethane diisocyanate, 1, 6-hexamethylene diisocyanate, acetaldehyde, acetaminophen, cyclic AMP, cocaine, curdlan, 2, 4-dichlorophenoxyacetic acid, DDT, heroin, histamine, ovalbumin, and ozone were found to have a tendency to generate Asthma. Chemicals stands for reducing asthma include zinc, vitamin A, theophylline, salicylic acid, resveratrol, phthalic anhydride, Nitric Oxide, Nicardipine, and Emodin. The study revealed the significance of consumption of vitamin rich food, resveratrol containing fruits such as grapes, berry fruits etc, for preventing asthma like atopic diseases.

Effective drug response of drugs of IL4 gene

All the drugs including corticosteroids, adrenergic receptor agonists, bronchodilators, beta2 agonists, and leukotriene antagonist having influence on IL4 gene were analysed for their pharmacological

efficiency against asthma by the use of PASS online server of Way2Drug predictive services. Among the drugs, corticosteroids showed very good anti-inflammatory, antiasthmatic and antiallergic properties. Salbutamol is a better bronchodilator but formoterol fumarate showed greater bronchodilatory effect than salbutamol.

Among the drugs, dexamethasone showed considerable IL4 antagonistic activity. Among corticosteroids, beclomethasone, budesonide and fluticasone showed better antiasthmatic properties. This result highlighted the efficiency of corticosteroid medicines as better treatment for asthma.

Table 3: EDR response of drug molecules

Drug	Antiasthmatic activity	Antiallergic activity	Anti-inflammatory property	Bronchodilator property	IL4 antagonistic activity
salbutamol	0.521	Nil	Nil	0.856	Nil
Beclomethasone	0.952	0.937	0.976	Nil	Nil
Budesonide	0.942	0.955	0.968	Nil	Nil
Dexamethasone	0.561	0.711	0.942	Nil	0.328
Fluticasone	0.996	0.996	0.989	Nil	Nil
Formoterolfumarate	0.613	0.345	Nil	0.954	Nil
Hydrocortisone	0.502	0.791	0.942	Nil	Nil
Methyl prednisolone	0.653	0.872	0.959	Nil	Nil
Prednisolone	0.698	0.879	0.962	Nil	Nil
prednisone	0.556	0.836	0.952	Nil	Nil
Verlukast	0.871	0.872	0.603	Nil	Nil
zileuton	0.837	0.859	0.540	Nil	Nil

Protein characterization studies

All the 5 deleterious SNPs were imported into the online bioinformatics database, LS-SNP/PDB to find out the responsible proteins. A total of 23 proteins were identified from five

deleterious SNPs. All the proteins were subjected to characterization studies using the online tools, ExpASY-Prot Param and SOPMA. The primary and secondary structure analysis has done on all the proteins and analyzed the result to find out better among them for further studies. The result is shown in table 3.

Table 4: protein characterization features

PDB Id	Aliphatic index	GRAVY	Alpha helix (%)	Resolution A °	Random coil (%)	Energy of protein KJ/mol
1BBN	78.57	-0.571	70.68	NMR	18.05	-2639
1BCN	78.57	-0.571	70.68	NMR	18.05	-1859
1CYL	79.46	-0.562	67.44	NMR	20.16	-1101
1HIJ	79.46	-0.554	68.22	3.0	20.16	-5848
1HIK	79.46	-0.554	67.44	2.6	20.16	-6344
1HZI	80.23	-0.521	69.77	2.05	17.83	-6314
1IAR	79.46	-0.562	68.22	2.3	20.16	-12898
1ITI	78.57	-0.571	70.68	NMR	18.05	-3104
1ITL	78.85	-0.543	67.69	NMR	20.00	-4116
1ITM	78.85	-0.543	67.69	NMR	20.00	-4859
1RCB	79.46	-0.562	67.44	2.25	20.16	126
2B8U	79.46	-0.562	67.44	1.8	20.16	17579
2B8X	79.46	-0.661	68.22	1.7	20.16	-6184
2B8Y	79.46	-0.633	70.54	1.8	18.60	-6018
2B8Z	80.23	-0.513	67.44	2.5	20.16	-5512
2B90	80.23	-0.535	69.77	2.1	18.60	-5443
2B91	80.23	-0.562	68.22	2.0	20.16	-5925
2CYK	79.46	-0.562	67.44	NMR	20.16	615050
2D48	79.46	-0.584	69.77	1.65	18.60	-5551
2INT	79.46	-0.562	67.44	2.35	20.16	-3556
3BPL	79.46	-0.562	67.44	2.93	20.16	-23042
3BPN	79.46	-0.562	67.44	3.02	20.16	-17044
3QB7	81.36	-0.461	62.88	3.24	24.24	-31839

The characterization parameters include the aliphatic index, gravity, resolution of protein crystallization, alpha helix, random coil and energy of proteins. The first two features were analyzed from the online database, ExpASY-Prot Param, resolution of protein crystallization has taken from RCSB protein data bank and the alpha helix and random coil parameters were checked using the SOPMA tool. Finally, the energy of all the proteins was calculated using the Swiss-PDBviewer4.0.1 software at a force field of GROMOS 96. The aliphatic index stands for the thermo stability of proteins. All the proteins have very good thermo stability. Among the 23, 3QB7 has exhibited high thermo stability. Grand average hydropathy, GRAVY measures the hydrophilicity of protein. All the listed proteins have the GRAVY value less than zero indicating hydrophilic nature of proteins. Alpha helix and random

coil values furnished information about the secondary structure evolution of the protein. Maximum alpha helix percentage and least random coil percentage indicate that the protein is stable. Alpha helix and random coil values of all the proteins satisfied the condition for stability. On the basis of energy of proteins, X-ray crystal structure resolution, and other primary and secondary structure analysis, the protein, 1IAR has chosen as the best protein among the 23. Also 1IAR is a wild form of IL4 protein. The criteria for a protein to be selected for Insilco studies include high aliphatic index, low X-ray crystal structure resolution, high alpha helix percentage, low random coil and low energy conformation. The protein 3QB7, one of the high-value proteins, is a mutated form. The protein complex 1IAR contains non mutated IL4 chain, which is the wild form of protein. Based on the above-mentioned criteria, the

protein under the PDB id 1IAR has been selected for further mutational and molecular docking studies.

Determination of thermodynamic stability of SNPs

The thermodynamic stability of five deleterious SNPs has been checked using the software, Swiss PDB viewer. The IL4 protein has

been retrieved from the protein data bank (pdb id 1IAR) and subjected to analyze the thermodynamic stability. The energy minimization has been carried out at a force field of GROMOS 96.

The mutations corresponding to the deleterious SNPs were done on the protein and associated energy of protein was recorded in table 4.

Table 5: mutational energy changes of deleterious SNPs

SNP id	Mutations given in amino acid sequence of protein	Initial protein energy (KJ/mol)	Energy after mutation (KJ/mol)	Stability due to mutational change
rs149950065	ALA118 GLY	-12898.637	-12817.289	Decreases
	ALA102 GLY	-12898.637	-12803.396	Decreases
rs199929962	MET 144 THR	-12898.637	-12923.423	Increases
	MET128THR	-12898.637	-12901.033	Increases
rs202231191	ASN 113 TYR	-12898.637	-12140.255	Decreases
	ASN97ASP	-12898.637	-12733.013	Decreases
	ASN97TYR	-12898.637	-12781.694	Decreases
rs376367511	CYS 123 ARG	-12898.637	-11248.243	Decreases
	CYS107ARG	-12898.637	-11248.009	Decreases
rs775753738	LEU 110PHE	-12898.637	-10931.777	Decreases
	LEU94PHE	-12898.637	27316.646	Decreases

On checking the energy changes associated with the mutational effects, it was found that the SNP, rs199929962 is stable among the five deleterious SNPs (for both the positions of amino acids, energy due to mutational change has decreased from -12898.637 to -12923.423 and -12901.033 respectively). All other mutations will make the system unstable. This indicates that mutations correspond to the SNP, rs199929962 will be more sustainable and the proneness of disease due to the mutation of amino acid MET to THR at 144th as well as the 128th position of the protein, IL4 is high. Therefore this SNP has been identified as the most significant nonsynonymous deleterious SNP of IL4 gene in the coding region, for the disease, asthma.

Validation of protein by Ramachandran plot analysis

The quality of protein in terms of stereochemistry was checked by Ramachandran plot analysis in the Swiss PDB viewer software. The result revealed that most of the residues of the protein 1IAR were in the alpha helix conformation. Number of residues analyzed, number of residues in the favoured region, allowed

region and in the outlier region were 98%, 94%, 5% and 2% respectively. Ramachandran outliers identified were SER36 and SER 128; both are in the beta sheet conformation.

Molecular docking analyses

The protein IL4 (pdb id 1IAR) has retrieved from the RCSB protein data bank and subjected to protein preparation studies at a force field of OPLS_2005 using the protein preparation wizard module of the molecular modeling software Schrodinger 2015-1. The ligand binding site was assigned by the site map analysis and the site having Dscore above 0.5 and high volume score was selected for docking studies (table 6). The contour generated at the site map region of IL4 indicates the requirement for a better ligand interaction. The colour code signifies, yellow region indicate the requirement of hydrophobic region, red region stands for H bond acceptor, blue region want the presence of H-bond donor, pink region stands for metal binding site. Here there is no metal binding region. Two sites are generated through site map analysis. Site 1 has produced higher site score, Dscore and volume and hence was chosen for the docking analysis with the protein.

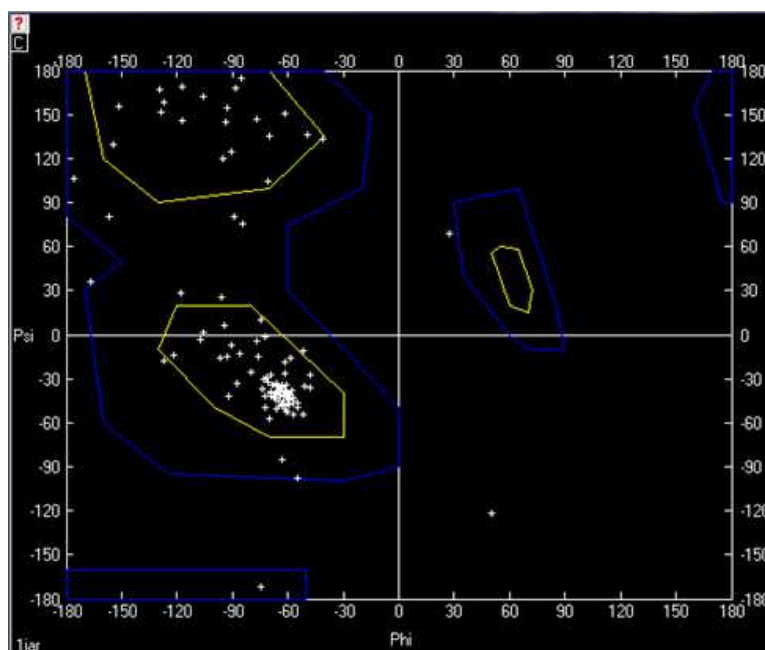


Fig. 2: Ramachandran plot of IL4 (1IAR) energy minimized protein

Table 6: Site map score

Site	Site score	D score	Volume
Site 1	0.679	0.639	83.69
Site 2	0.608	0.603	63.45

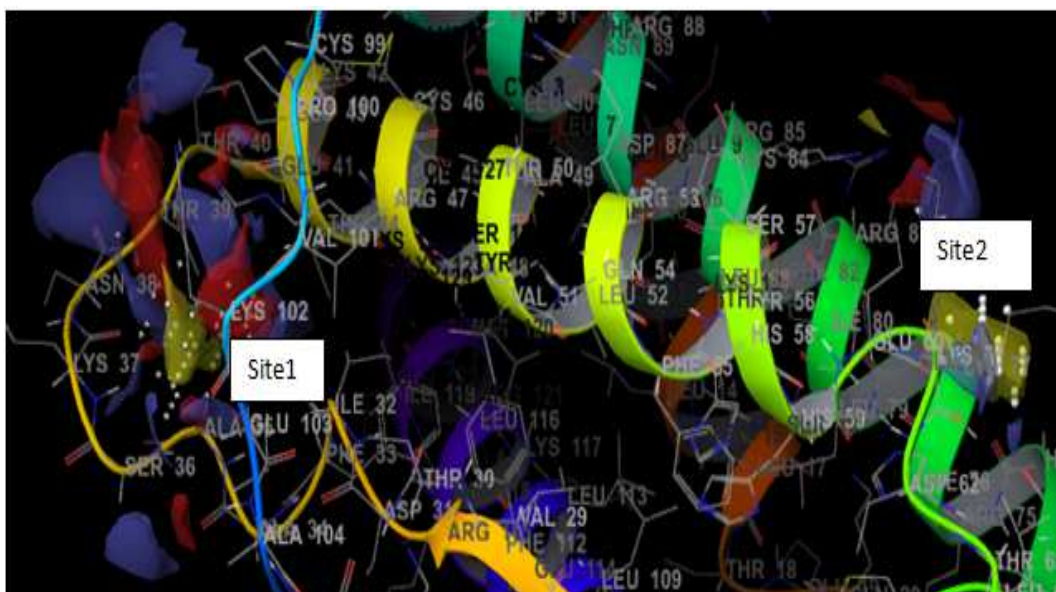


Fig. 3: Site map images of IL4 protein

All the drug molecules showing environmental influence with the IL4 gene were subjected for docking analyses with the protein and the result is shown in the table 6. Dexamethasone is the only drug shows considerable IL4 antagonistic property and both dexamethasone and

budesonide are high value anti-inflammatory antiallergic drugs with very good therapeutic potency (table 3). Therefore pharmacophoric feature has derived from the interaction of both the candidates and were utilized for virtual screening purpose.

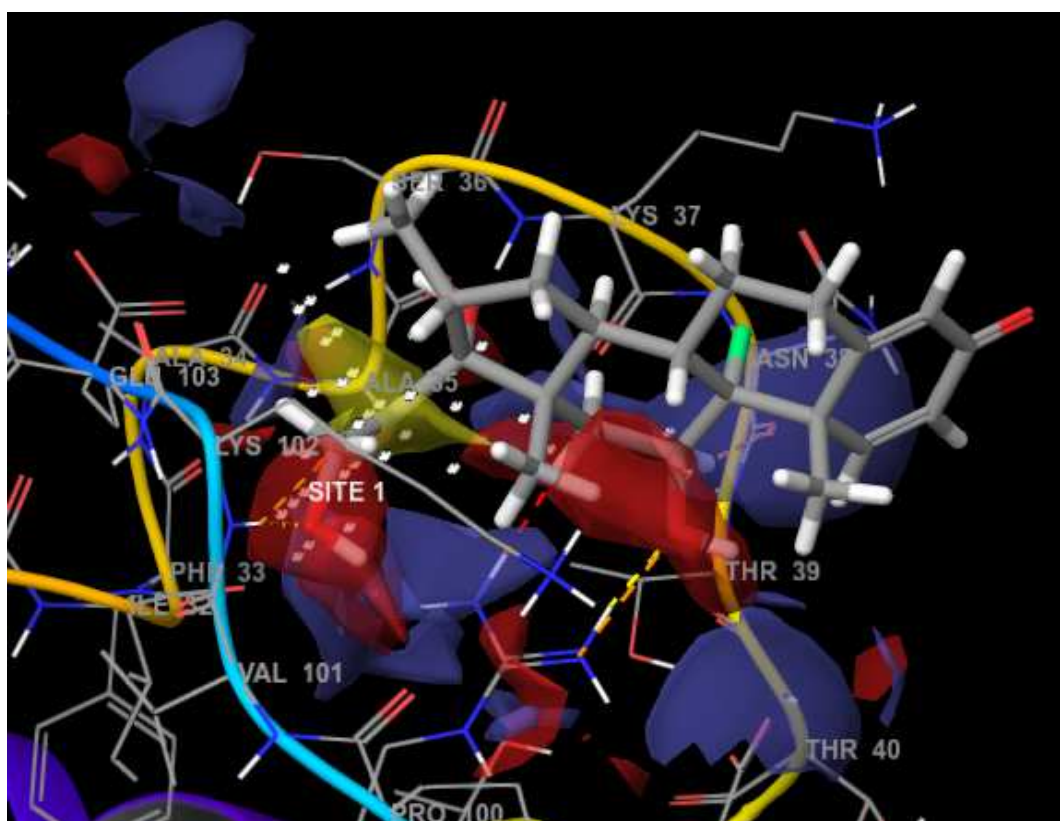


Fig. 4a: docking of dexamethasone with IL4 at the site map region, b: Two dimensional image of docking

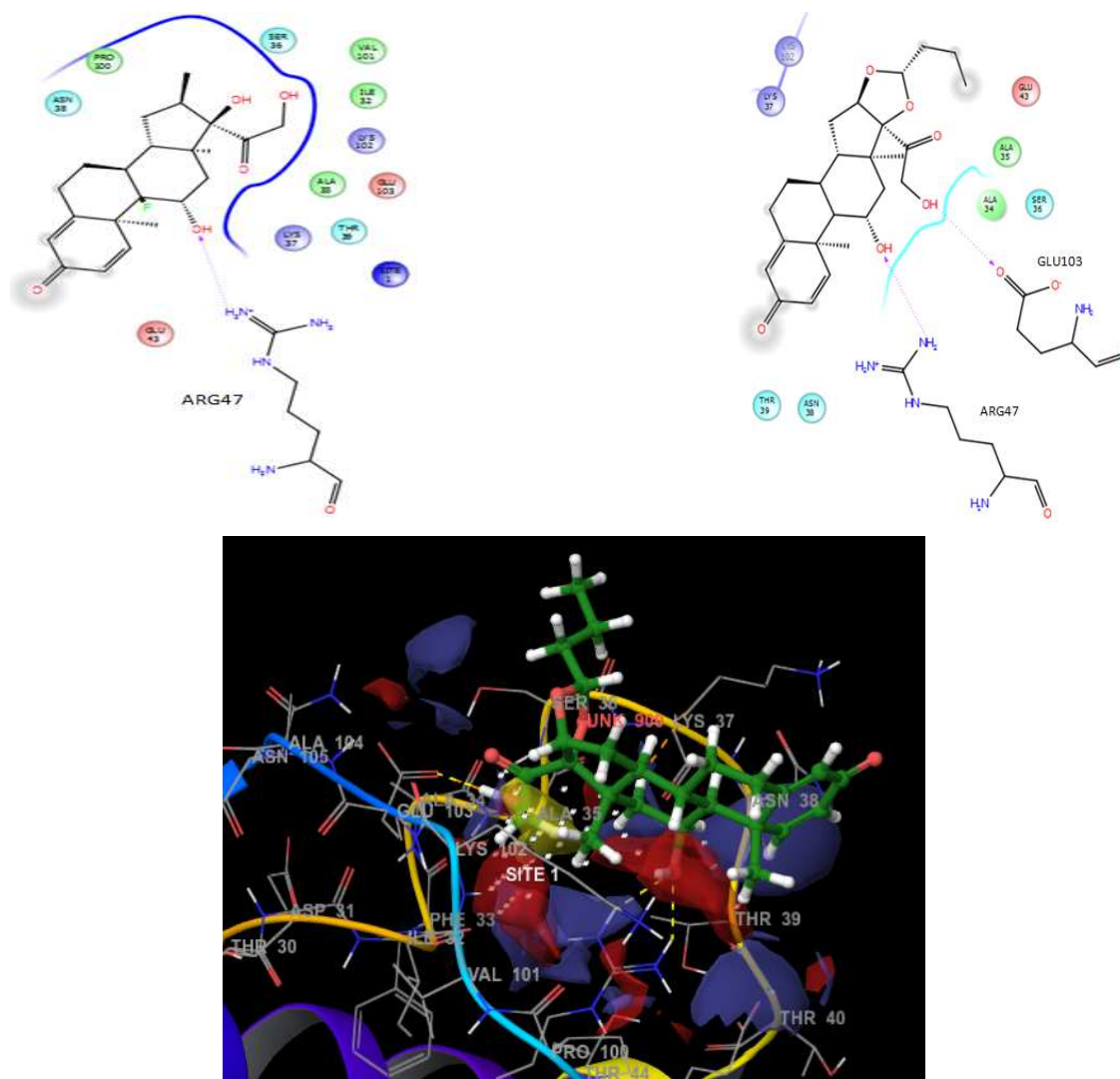


Fig. 5: Budesonide docked with IL4, (2D and 3D interaction images)

Table 7: Docking score table of drug molecules with IL4

Compound	Glide G score	Glide energy Kcal/mol	Interacting residues	Type of secondary bond forces	Binding free energy Kcal/mol
Salbutamol	-4.440	-27.474	GLU103, ILE32, ARG47, LYS37	5 hydrogen bonding interactions	-25.064
Methyl prednisolone	-3.763	-34.055	ARG47, GLU103, ILE32	4 hydrogen bonding interactions	-35.5
Prednisolone	-3.738	-33.978	ARG47, GLU103, ILE32	4 hydrogen bonding interactions	-35.401
Prednisone	-3.703	-33.182	ARG47, GLU103, ILE32, LYS102	5 hydrogen bondings	-31.558
Hydrocortisone	-3.578	-32.110	ARG47, GLU103, ILE32	4 hydrogen bonding interactions	-34.184
Formoterolfumarate	-3.392	-32.165	LYS37, ARG47, GLU103	5 hydrogen bonding interactions	-26.659
Thymoquinone	-2.937	-22.446	LYS37, GLU103, LYS102	3 hydrogen bonding interactions	-24.738
Budesonide	-2.615	-26.287	ARG47, GLU103	2 hydrogen bonding interactions	-13.251
Verlukast	-1.663	-39.294	ARG47, GLU103, LYS102, ALA104	3 hydrogen bonding, two pi-cation interactions	-27.31
Fluticasone	-1.458	-28.587	ARG47, LYS37	2 hydrogen bonding interactions	-29.917
Dexamethasone	-3.028	-27.868	ARG47	1 hydrogen bonding interaction	-23.534

All the drugs have shown different type of intermolecular interactions with the target protein. All the drugs were interacted with the residue ARG47 through hydrogen bonding interactions. The negative binding free energy of the drug-receptor complexes were revealed that all are energy stabilized

The dexamethasone has created hydrogen bonding interaction with the protein at site 1 through NH₂⁺ of the residue ARG 47 while budesonide has created two type of hydrogen bonding interactions with the protein. The C chain of budesonide has generated hydrogen

bond with the lone pair electrons in the N terminal of arginine, while open chain hydroxyl group of budesonide has furnished hydrogen bonding with electron deficient oxygen of glutamic acid residue. Dexamethasone has created two feature pharmacophore AA, while budesonide has created 3 feature pharmacophore ADA. The pharmacophoric feature AA comprised of two hydrogen bond acceptor features, one located on hydroxyl group in ring 3 of steroid frame work and other at the fourth ring of steroid frame work. Feature ADA composed of two acceptor features and one donor feature.

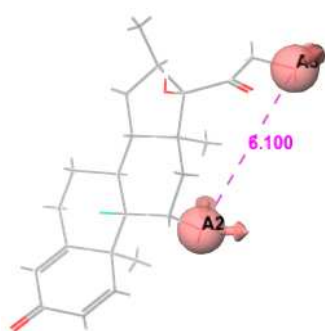


Fig. 6a: AA of dexamethasone

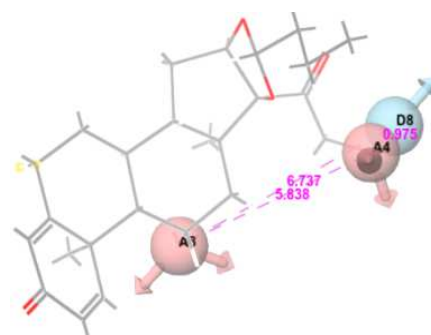


Fig. 6b: ADA of budesonide

Table 8: Pharmacophore feature score

Drug	Feature	score	Score source
Dexamethasone	A3	-1.60	Hydrogen bond acceptor
	A2	-0.35	Hydrogen bond acceptor
Budesonide	A4	-1.01	Hydrogen bond acceptor
	D8	-0.62	Hydrogen bond donar
	A3	-0.32	Hydrogen bond acceptor

Both the pharmacophoric features were utilized for virtual screening on the data set of anti-inflammatory phytochemicals taken from Dr. Duke's ethno botanical data base to find out a pharmacophorically similar scaffold from the plant world. The parameters for virtual screening include shape, functionality and volume similarity. A total of 8 phytomolecule shaving similarity with the pharmacophoric feature was obtained through virtual

screening. All the 8 molecules having fitness score above 1.5 were subjected to docking analyses with the protein, IL4 and found out the best docked compound among them. The pharmacophoric feature of budesonide has a major contribution towards virtual screening. The feature AA could not effectively match with the below mentioned anti-inflammatory phytochemicals while the feature ADA could.

Table 9: Fitness score of phytochemicals based on the features AA and ADA

Phytochemical compound	Align score		Vector score		Volume score		Fitness	
	AA	ADA	AA	ADA	AA	ADA	AA	ADA
Ellagic acid	0.0051	0.144	0.00	0.964	0.00	0.157	0.9956	2.001
Cyanidin	0.0727	0.115	0.00	0.651	0.00	0.212	0.9393	1.768
Quercetin	0.1153	0.355	0.00	0.949	0.00	0.391	0.9038	2.045
Arzanol	0.1738	0.3327	0.00	0.855	0.00	0.174	0.885	1.752
Andrographolide	0.2686	0.725	0.00	0.959	0.00	0.330	0.776	1.685
Ascorbic acid	0.331	0.199	0.00	0.987	0.00	0.083	0.723	1.904
Curcumin	0.451	0.00	0.00	0.00	0.00	0.00	0.623	0.00
apocynin	0.008	0.353	0.00	0.832	0.00	0.099	0.993	1.637

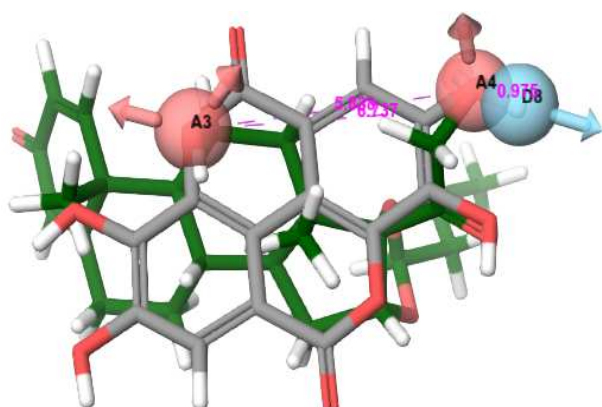
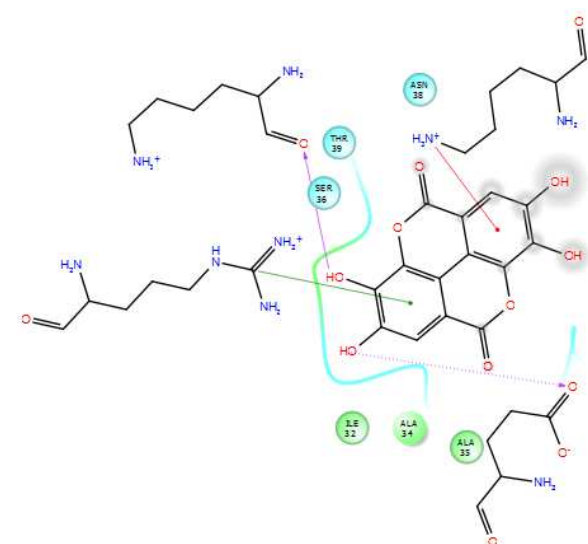


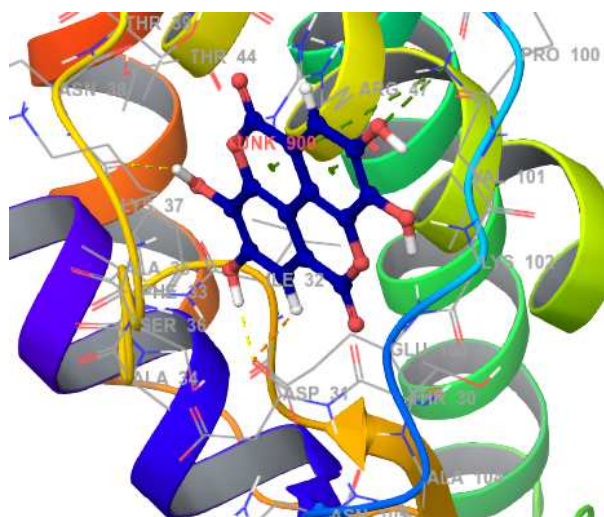
Fig. 7: Alignment of ellagic acid with budesonide, ash colour-ellagic acid, green-budesonide

Among the phytochemicals, ellagic acid has shown high range of similarity with both the drug molecules through alignment, vector and volume matching. Similarity to ellagic acid with the feature of

budesonide is greater than that of dexamethasone. Therefore, the drug likeness of ellagic acid was compared with that of budesonide through docking interaction studies. Glide G score of ellagic acid is -5.986Kcal/mol while that of budesonide is -2.615. Also ellagic acid has furnished two type of hydrogen bonds with C=O of GLU103 at a distance of 1.796 Å and with C=O of LYS37 at a distance of 2.05 Å. An edge to face pi-pi stacking interaction was found with pi electron cloud of one of the phenolic ring of ellagic acid and electron cloud around C=NH₂ bond of ARG47 at a distance of 5.32 Å. Finally, one pi-cation interaction is seen with aromatic electron cloud of second phenolic ring of ellagic acid and NH₂ of LYS102 at 5.09 Å distance. Quercetin is the other compound having all the values very close to that of ellagic acid. It has created 4 type of intermolecular interactions with the protein, in detail, an edge to face pi-pi stacking interaction between the fused phenolic ring of quercetin and C=NH₂ pi electrons of ARG47. A backbone hydrogen bonding interaction is found with phenolic OH of quercetin and C=O of LYS37 at a distance of 1.99 Å, another side chain hydrogen bonding between other phenolic OH of quercetin and C=O of GLU103 at a distance of 1.72 Å finally, a backbone hydrogen bonding between OH of isolated phenolic ring of quercetin and C=O of LYS102 at 1.94 Å distance. All the results furnished through the docking studies pointed out the anti-inflammatory potential of nonsteroidal scaffolds ellagic acid and quercetin over leading anti-inflammatory steroids.

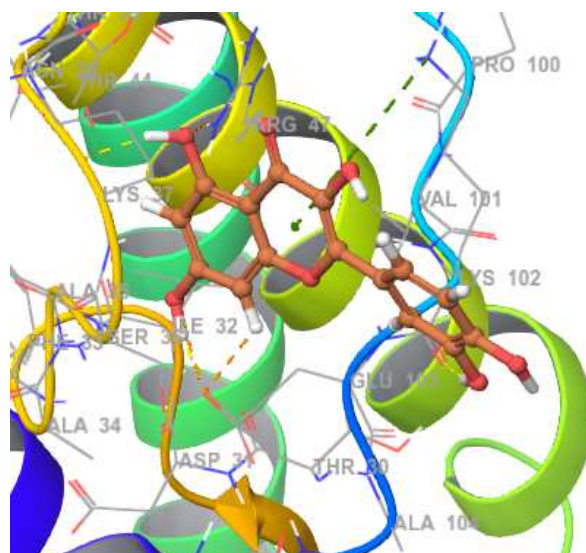


a

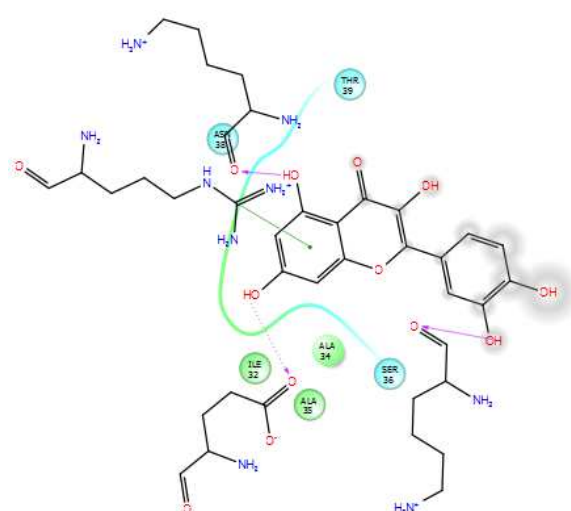


b

Fig. 8:a) 2D interaction fig. of ellagic acid with IL4 b) docking interactions of ellagic acid with IL4



c



d

Fig. 8c: 3D docking image of quercetin with IL4, d: 2D interaction image of quercetin with IL4

Table 10: Docking result and binding energy of phytochemicals

Compound	Glide G score	Glide energy (Kcal/mol)	Interacting residues of protein	Types of bonding interaction	MMGBSA dG Bind (Kcal/mol)
Ellagic acid	-5.986	-29.338	ARG47, LYS37, LYS102, GLU103	2 hydrogen bonding, 1 pi-pi stacking bonding, 1 pi-cation interactions	-27.883
Cyanidin	-3.625	-23.830	LYS102, ARG47, ILE32	2 salt bridges and two hydrogen bonds	-14.066
Quercetin	-5.313	-31.127	LYS37, ARG47, GLU103, LYS102	Three hydrogen bonding, one pi-pi stacking interactions	-26.651
Arzanol	-4.012	-35.613	ARG47, LYS37, GLU103, LYS102	Four hydrogen bonding, one salt bridge, one pi-cation interactions	-32.162
Andrographolide	-3.734	-25.618	LYS37, LYS102	Two hydrogen bonds	-31.073
Ascorbic acid	-5.599	-26.316	LYS102, ARG47, ILE32	Two hydrogen bonds, two salt bridges	-19.107
Curcumin	-3.297	-32.356	LYS37, ARG47, ALA104	Three hydrogen bonds	-26.808
apocynin	-3.689	-24.127	ARG47, LYS37, LYS102	Two hydrogen bonds, one pi-pi stacking, one pi-cation interactions	-18.503

Glide G score includes all the glide docking features such as lipophilic term derived from hydrophobic substitutions of ligand, vander Waals volume parameter, rotatable bond penalties, hydrogen bond acceptor and donor features.

All the phytochemicals mentioned above are highly interactive with the protein than the currently used drugs. Ellagic acid has furnished

high Glide G score with the protein and it could better align with the pharmacophoric feature of budesonide.

The stability in terms of free binding energy of all the phytochemical docked complexes was calculated using MMGBSA studies. It was found that all the complexes were stabilized by releasing a high amount of energy. The final compounds ellagic acid and quercetin

were stabilized by -27.883 and -26.651 Kcal/mol respectively. On comparing the stability of ellagic acid and quercetin with that of reference drugs; it was found that ellagic acid and quercetin were more energy stabilized than the reference drugs, dexamethasone and budesonide. These findings revealed the higher affinity of those phytochemicals towards the target IL4. The protein complexes of arzanol and andrographolide were found to be more energy stabilized by releasing 32.162 and 31.073 Kcal/mol respectively. Among the phytochemicals, cyanidin-protein complex is the least stabilized one and is stabilized by releasing 14.066 Kcal/mol energy.

The phytochemicals were subjected to PASS analysis to make a comparative study with the activity properties of drugs. The findings are included in the table 11. On detailed observation, it is

found that ellagic acid and quercetin has shown anti-allergic, anti-inflammatory and interleukin 4 antagonistic activities. All the phytochemicals listed are not showing bronchodilatory effect. Sum of the mentioned activity properties are high for quercetin, also, the fitness value of quercetin with budesonide pharmacophore is 2.045, slightly greater than that of ellagic acid and both the candidates are showing considerable interleukin 4 antagonistic activity. But the docking score is slightly greater for ellagic acid. All the results obtained through docking studies, pharmacophoric screening and *In silico* biological activity studies revealed the anti-inflammatory potential of both ellagic acid and quercetin. Therefore, both the phytochemical scaffolds can be suggestively introduced as nonsteroidal anti-inflammatory lead for the treatment of asthma.

Table 11: Pass activity of phytochemicals

Compound	Antiasthmatic	Antiallergic	Antiinflammatory	Bronchodilatory	IL4 Antagonist
Ellagic acid	-	0.316	0.748	-	0.550
Quercetin	-	0.443	0.704	-	0.719
Arzanol	0.222	0.369	0.660	-	-
Andrographolide	-	0.471	0.913	-	-
Ascorbic acid	-	-	0.427	-	-
Curcumin	0.242	0.432	0.709	-	0.141
apocynin	0.429	0.517	0.504	-	0.151

Both the ellagic acid and quercetin were subjected to structure activity studies and found the significance of phenolic hydroxyl groups of both the candidates towards anti-inflammatory, anti-allergic and interleukin 4 antagonistic activities. Removal of all the hydroxyl groups of ellagic acid showed a considerable decrease in its anti-inflammatory, and interleukin 4 antagonistic properties and the compound has lost its anti-allergic property. The hydroxyl groups replaced with NH₂ groups and checked for its activity properties and it was found that the compound is no longer showing any anti-allergic, Antiasthmatic and interleukin 4 antagonistic properties while a little anti-inflammatory effect retains there. This signifies the potential of hydroxyl entity towards the concerned activity properties. The removal of phenolic hydroxyls of quercetin also showed a reduction in its anti-allergic, anti-inflammatory

and interleukin 4 antagonistic properties. Here also replacement of OH with NH₂ does not create any positive impact on the activity properties. The study revealed that phenolic hydroxyls and aromatic structural frameworks are the responsible features behind the anti-inflammatory, anti-allergic and interleukin 4 antagonistic properties. In quercetin, phenolic hydroxyls are having more contribution towards the above mentioned properties. Non-phenolic hydroxyl has shown positive effect on anti-allergic and interleukin 4 antagonistic properties. The aromatic nature of rings is highly influenced on the three properties. In the absence of aromatic rings, quercetin is unable to show anti-allergic and interleukin 4 antagonistic properties. The SAR analyses are detailed in table 12 and table 13.

Table 12: Ellagic acid and derivatives and change in activities

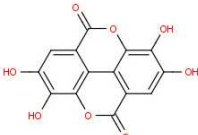
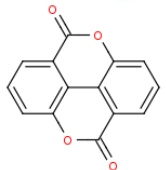
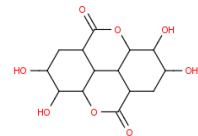
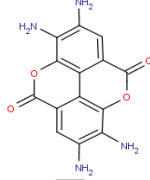

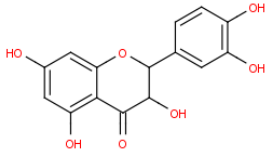
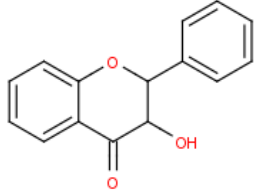
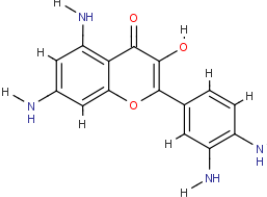
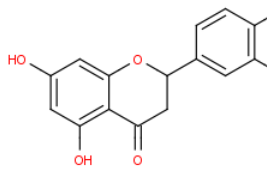
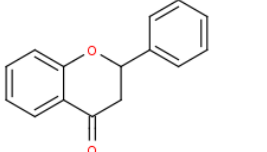
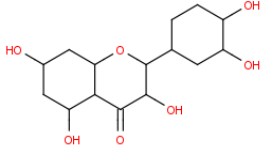
Ellagic acid and derivatives	Antiallergic	Anti-inflammatory	IL4 antagonist
	0.316	0.748	0.550
	-	0.665	0.354
	-	0.307	-
	-	0.436	-
	-	0.246	-

Table 13: Quercetin and derivatives with change in activity values

Quercetin and derivatives	Antiallergic	Anti-inflammatory	IL4 Antagonist
	0.443	0.704	0.719
	0.389	0.602	0.484
	-	0.323	0.330
	0.265	0.691	0.318
	0.203	0.601	-
	-	0.477	-

The information on the structural requirements of both ellagic acid and quercetin for the antiallergic, anti-inflammatory and interleukin 4 antagonistic properties will help in designing more promising candidates for a better treatment for asthma.

CONCLUSION

IL4 is a potential target for the disease asthma. The genetic involvement of the gene IL4 has been investigated through the present study. IL4-JAK1-STAT6 signaling pathway has been confirmed as a potential causative for asthma through gene association studies. This information can be used for designing a druggable molecule in future which will prevent the disease by arresting this pathway. 5000 SNPs of IL4 have retrieved from the dbSNP database of NCBI and 65 human missense SNPs were screened out. The SIFT and PolyPhen-2 score of all the SNPs were utilized to find out their deleterious nature. The proteins of all the five SNPs were identified; primary and secondary structure analysis was done on these 23 proteins. The protein, 1IAR has been selected by primary and secondary structure analysis and subjected to find mutational changes associated with the deleterious SNPs. From the mutational studies, it was found that the SNP happened due to the change of amino acid from methionine to threonine in the 144th as well as the 128th position of the protein; IL4 is a major genetic variation which causes the disease, asthma. Designing a drug, by considering this SNP based on IL4-JAK1-STAT6 signaling pathway will be more beneficial for those patients having the concerned SNP change. Metagenomic studies revealed about the

insignificance of microbes with the IL4 mutational changes. Environmental factor studies and EDR analysis signifies the potential of corticosteroid drugs as better treatment.

Molecular docking studies and pharmacophore based virtual screening revealed the potential of nonsteroidal anti-inflammatory polyphenolic phytochemicals, ellagic acid and quercetin as better lead for designing druggable molecules for asthma. Further SAR studies on the scaffolds of ellagic acid and quercetin produced valuable information on the structural requirements towards anti-inflammatory property. The potential of features such as aromatic nature and presence of phenolic hydroxyl groups were analysed and hence recommended for tuning the scaffolds towards very good anti-inflammatory, antiallergic, antiasthmatic and interleukin 4 antagonistic properties.

ACKNOWLEDGEMENT

The authors are deeply grateful to 'The University of Kerala' for providing junior research fellowship. We convey our authentic gratitude to 'Sree Narayana College, Kollam', Kerala for providing conveniences as a research centre.

AUTHORS CONTRIBUTIONS

All the author have contributed equally

CONFLICT OF INTERESTS

Declared none

REFERENCES

- Ivana V. Yang, David A. Schwartz epigenetic mechanisms and the development of asthma. *J Allergy Clin Immunol* 2013;6:1243-55.
- Global Strategy for Asthma Management and Prevention; 2016.
- RJ Flower. Eleventh gaddum memorial lecture lipocortin and the mechanism of action of the glucocorticoids. *Br J Pharmacol* 1988;94:987-1015.
- J Mullol, A Xaubet, E Lopez, J Roca-ferrer, C Picado. Comparative study of the effects of different glucocorticosteroids on eosinophil survival primed by cultured epithelial cell supernatants obtained from nasal mucosa and nasal polyps. *Thorax* 1995;270:270-4.
- RS Irwin, ND Richardson. Corticosteroids* The Physician's perception. *Chest* 2006;130:41S-53S.
- AV Ignal, DM Ilan. A review on SNP and other types of molecular markers and their use in animal genetics. *Genet Sel Evol* 2002;34:275-305.
- JTL Mah, ESH Low, E Lee. *In silico* SNP analysis and bioinformatics tools : a review of the state of the art to aid drug discovery. *Drug Discovery Today* 2011;16:800-9.
- G Erkel, Hanane Belahmer, Annegret Serwe, Timm Anke, Horst Kunz, Heinz Kolshorn, *et al.* Oxacyclododecandione, a novel inhibitor of IL-4 signaling from *exserohilum rostratum*. *J Antibiot* 2008;61:285-90.
- Michael Kabesch, Iren Tzotcheva, David Carr, Claudia Höfler, Stephan K Weiland, Christian Fritzsche, *et al.* A complete screening of the IL4 gene: novel polymorphisms and their association with asthma and IgE in childhood. *J Allergy Clin Immunol* 2003;112:893-8.
- K Vale. Targeting the JAK-STAT pathway in the treatment of 'Th2-high' severe asthma. *Future Med Chem* 2016;10:4155.
- T Tanaka, R Takahashi. Flavonoids and Asthma. *Nutrients* 2013;5:2128-43.
- Barbara Männel, Mariama Jaiteh, Alexey Zeifman, Alena Randakova, Dorothee Möller, Harald Hübner, *et al.* Structure-guided screening for functionally selective D dopamine receptor ligands from a virtual chemical library structure-guided screening for functionally selective D2 dopamine receptor ligands from a virtual chemical library. *ACS Chem Biol* 2017;12:2652-61.
- Noguchi, Emiko Nukaga-nishio, Yuumi Jian, Zhang Yokouchi, Yukako Kamioka, Masashi Yamakawa-kobayashi, *et al.* Haplotypes of the 5' region of the IL-4 gene and SNPs in the intergene sequence between the IL-4 and IL-13 genes are associated with atopic asthma. *Hum Immunol* 2001;8859:1251-7.
- Sherry S, T Ward, M Kholodov, M Baker, J Phan, L Smigielski, *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29:308-11.
- I Adzhubei, DM Jordan, SR Sunyaev. Predicting functional effect of human missense mutations using polyphen-2 predicting the effect of a single-residue substitution on protein structure and function using the polyphen-2 WEB. *Current Protocols in Human Genetics* 2013;7:1-41.
- PCNg, S Henikoff. Accounting for human polymorphisms predicted to affect protein function. *Genome Res* 2002;12:436-46.
- Simov, Vladimir Deshmukh, Sujal V Dinsmore, Christopher J Elwood, Fiona Fernandez, Rafael B Garcia, *et al.* Lung microbiome analysis in steroid-naïve asthma patients by using whole sputum. *Korean Acad Tuberc Respir Dis* 2016;3536:165-78.
- AB Mukherjee, Z Zhang. Allergic asthma: influence of allergic asthma: a complex syndrome. *J Biol Chem* 2011;286:32883-9.
- M Ryan, M Diekhans, S Lien, Y Liu, R Karchin. LS-SNP/PDB: annotated non-synonymous SNPs mapped to protein data bank structures. *Struct Bioinf* 2009;25:1431-2.
- Gasteiger, Elisabeth Hoogland, Christine Gattiker, Alexandre Duvaud, Séverine Wilkins, Marc R Appel, *et al.* Protein identification and analysis tools on the ExPASy server. *Proteomics Protocols Handbook*; 2005. p. 571-605.
- DG Geourjon. SOPM: a self-optimized method for protein secondary structure prediction. *Protein Eng* 1994;7:157-64.
- AB Pernis, PB Rothman. JAK-STAT signaling in asthma. *J Clin Invest* 2002;109:1279-83.