J Mol Model (2012) 18:2513–2525
DOI 10.1007/s00894-011-1265-3

ORIGINAL PAPER

Pharmacophore modeling, molecular docking, QSAR, and in silico ADMET studies of gallic acid derivatives for immunomodulatory activity

Dharmendra Kumar Yadav · Feroz Khan · Arvind Singh Negi

Received: 2 June 2011 / Accepted: 29 September 2011 / Published online: 27 October 2011

Abstract Immunomodulation refers to an alteration in the immune response due to the intrusion of foreign molecules into the body. In the present communication, QSAR and docking studies of gallic acid derivatives were performed in relation to their immunomodulatory activities. Screening through the use of a QSAR model suggested that the compounds G-4, G-7, G-9, G-10, G-12, and G-13 possess immunomodulatory activity. Activity was predicted using a statistical model developed by the forward stepwise multiple linear regression method. The correlation coefficient (r^2) and the prediction accuracy ($r\mathrm{CV}^2$) of the QSAR model were 0.99 and 0.96, respectively. The QSAR study indicated that chemical descriptors—dipole moment, steric energy, amide group count, λ_{max} (UV-visible) and molar refractivity—are well correlated with activity, while decreases in the dipole moment, steric energy, and molar refractivity were negatively correlated. A molecular docking study showed that the compounds had high binding affinities for the INFα-2, IL-6, and IL-4 receptors. Binding site residues formed H-bonds with the designed gallic acid derivatives G-3, G-4, G-5, G-6, G-7, and G-10. Moreover, based on screening for oral bioavailability, in silico ADME, and toxicity risk assessment, we concluded that compound G-7 exhibits marked immunomodulatory activity, comparable to levamisole.

Keywords Gallic acid · Immunomodulatory · Molecular docking · QSAR · ADME · Toxicity · Druglikeness

Electronic supplementary material The online version of this article (doi:10.1007/s00894-011-1265-3) contains supplementary material, which is available to authorized users.

D. K. Yadav · F. Khan (✉)
Metabolic and Structural Biology Department,
Council of Scientific and Industrial Research,
Central Institute of Medicinal and Aromatic Plants,
Lucknow 226015, UP, India
e-mail: f.khan@cimap.res.in

A. S. Negi
Analytical Chemistry Department,
Council of Scientific and Industrial Research,
Central Institute of Medicinal and Aromatic Plants,
Lucknow 226015, UP, India

Introduction

Immunomodulation refers to an alteration in the immune response caused by the intrusion of foreign molecules into the body. It can be either immunostimulation or immunosuppression. A large number of herbal drugs are mentioned in Ayurveda (a traditional system of Indian medicine) due to their immunomodulating activities [1, 2]. In the past, autologous and heterologous proteins from living and attenuated microorganisms as well as injections of animal organ preparations have been used to restore an impaired defense mechanism. Thymus peptides and other biological response modifiers (BRM) (e.g., interferon, interleukins), synthetic low molecular weight compounds (e.g., levamisole), chemically modified nucleotides, polysaccharides from fungi (e.g., lentinan), and some plant extracts are also being used for this purpose, especially in Europe and Asia. Many medicinal plant products have been reported to exhibit immunomodulatory effects, such as berberine, boswellic acid, aristolochic acid, cichoric acid, and plumbagin [2]. Gallic acid is also one of the myriad of herbal biochemicals whose activities have been largely unexplored. Gallic acid and its derivatives are polyphenolic compounds found mostly in gallnuts, grapes, tea, hops, oak

bark [1, 3, 4], as well as in processed beverages such as red wine [2]. It has been known to exhibit a wide range of biological activities, including antioxidant, anti-inflammatory, antimutagenic, antimicrobial, and anticancer activities [5]. Gallic acid has been a building block of choice for various pharmaceutical leads due to the presence of this moiety in several bioactive natural molecules, such as combretastatin A-4 and podophyllotoxin [6–8]. A number of gallic acid derivatives have already been synthesized and found to act as protease inhibitors, antimalarials, and anticancer compounds [9, 10]. Gallic acid appears to have antifungal and antiviral properties, and it acts as an antioxidant, thus helping to protect our cells from oxidative damage [11]. It has also been found to show cytotoxicity towards cancer cells but not healthy cells [12]. Apart from this, it has been used as a remote astringent in cases of internal hemorrhage [5]. Recently, an anti-inflammatory action of gallic acid arising from its inhibitory action on histamine release and proinflammatory cytokine production in mast cells was reported [13, 14].

In the work described in this paper, we screened a library of gallic acid derivatives for potential immunomodulatory compounds using quantitative structure–activity relationship (QSAR), molecular docking, and in silico ADME/Tox studies. Based on the binding affinity energy, possible immunomodulatory receptors were identified. A multiple linear regression QSAR mathematical model was developed for activity prediction that successfully and accurately (noting the corresponding experimental activities) predicted the immunomodulatory activities of some newly designed gallic acid derivatives (G-4, G-7, G-9, G-10, G-12, and G-13) that had the basic naphthophenone pharmacophore [13, 14]. The QSAR model also quantified the activity-dependent chemical descriptors and predicted the lethal dose (log LD_{50}) of each derivative, thus indicating it potential range of toxicity. In the QSAR model, the regression coefficient (r^2), which indicates the relationship correlation, was 0.99, while the cross-validation coefficient (rCV^2), which indicates the prediction accuracy, was 0.96. The activity of each derivative was assessed using the standard computational pharmacokinetic parameters (ADMET) of druglikeness and bioavailability. QSAR studies indicated that dipole moment, steric energy, amide group count, λ_{max} (UV-visible), and molar refractivity correlated well with immunomodulatory activity. Moreover, based on oral bioavailability, in silico ADME, and toxicity risk assessment screening, we concluded that compound G-7 has greater immunomodulatory activity then G-4, G-9, G-10, G-12, and G-13. These results offer useful references for understanding the molecular mechanism and directing the molecular design of pharmacophore-based lead compounds with improved immunomodulatory activity.

Materials and methods

Structure cleaning and molecular docking

The chemical structures of the gallic acid derivatives were constructed using the Scigress Explorer v.7.7.0.47 (formerly CaChe) software package (Fujitsu Ltd., Tokyo, Japan). Energy minimization of the compounds with "cleaned" geometries was achieved through the MO-G application in Scigress, which computes and minimizes an energy related to the heat of formation. MO-G solves the Schrödinger equation for the best molecular orbital and geometry of the ligand molecule. The augmented molecular mechanics (MM2/MM3) parameter was used to optimize the energy of each molecule up to its lowest stable energy state. This energy minimization process was performed until the energy change was less than 0.001 kcal mol^{-1} or the molecules had been updated almost 300 times. The 3D chemical structures of known drugs were retrieved from the PubChem compound database at NCBI (http://www.pubchem.ncbi.nlm.nih.gov). Crystallographic 3D structures of target proteins were retrieved from the Brookhaven Protein Databank (http://www.pdb.org). The valency and hydrogen bonding of each ligand as well as each target protein were subsequently checked using the Workspace module of the Scigress Explorer software. Hydrogen atoms were added to the protein targets to achieve the correct ionization and tautomeric states of amino acid residues such as His, Asp, Ser, and Glu. Molecular docking of the drugs and the gallic acid derivatives with the immunomodulatory receptors was achieved using the FastDock Manager and FastDock Compute engines that are available with the Scigress Explorer software. To perform the automated docking of ligands into the active sites, we used a genetic algorithm with a fast and simplified potential of mean force (PMF) scoring scheme [3, 15]. PMF uses atom types that are similar to the empirical force fields used in mechanics and dynamics. A minimization is performed by the FastDock engine, which uses a Lamarkian genetic algorithm (LGA) so that individuals adapt to the surrounding environment. The best fits are sustained by analyzing the PMF scores of all chromosomes and assigning more reproductive opportunities to those with lower scores. This process was repeated for 3000 generations with 500 individuals and 100,000 energy evaluations. Other parameters were left as their default values. Structure-based screening involves docking candidate ligands into protein targets and then applying a PMF scoring function to estimate the likelihood that the ligand will bind to the protein with high affinity [15–17].

Parameters for QSAR model development

Initially, a total of 61 immunomodulatory compounds/drugs were used for QSAR modeling against 50 chemical

descriptors. Out of these 61, only 22 compounds/drugs were selected to provide a training data set for QSAR model development. Selection was made on the basis of structural/pharmacophore or chemical class similarity, to ensure that a diverse set of data was used rather than only data from compounds of the same family. Similarly, when selecting the best subset of descriptors, highly correlated descriptors were excluded through covariance analysis using a correlation matrix. Finally, out of the 50 chemical descriptors investigated initially, only 28 were used for model development based on the forward stepwise multiple linear regression method. The resulting QSAR model exhibited a high regression coefficient, was successfully validated using random test set compounds, and was evaluated for the robustness of its predictions via the cross-validation coefficient (Table 1, Fig. 8).

Statistical calculations used in QSAR modeling

Selecting a statistical method: stepwise multiple linear regression

The stepwise multiple linear regression method calculates QSAR equations by adding one variable at a time and testing each addition for significance. Only variables that are found to be significant are used in the QSAR equation. This regression method is especially useful when the number of variables is large and when the key descriptors are not known. In the forward mode, the calculation begins with no variables and builds a model by entering one variable at a time into the equation. In backward mode, the calculation begins with all variables included and drops variables one at a time until the calculation is complete; however, backward regression calculations can lead to overfitting.

Multiple correlation coefficient (r)

Variation in the data is quantified by the correlation coefficient (r), which measures how closely the observed data tracks the fitted regression line. This is a measure of how well the equation fits the data (i.e., it measures how good the correlation is). A perfect relation has $r=+1$ (positively correlated) or -1 (negatively correlated); no correlation has $r=0$. The regression coefficient (r^2) is sometimes quoted, and this gives the fraction of the variance (in %) that is explained by the regression line. The more scattered the data points, the lower the value of r. A satisfactory explanation of the data is usually indicated by an r^2 of at least 0.9; compare $r=0.9$ ($r^2=0.81$; 81% of the variance is explained) with $r=0.7$ ($r^2=0.49$; 49% of the variance is explained; 51% is unexplained). Errors in either the model or in the data will lead to a bad fit. This indicator of fit to the regression line is calculated as

$$r^2 = (\text{sum of the squares of the deviations from the regression line}) / (\text{sum of the squares of the deviations from the mean}) \quad (1)$$

$$r^2 = (\text{regression variance}) / (\text{original variance}), \quad (2)$$

Table 1 Comparison of experimental and predicted *in vivo* activity calculated by using the derived QSAR model equation

Compound	Exp. log LD_{50} (mg/kg)	Chemical descriptors of the QSAR model equation					Predicted log LD_{50}
		Dipole moment (debye)	Steric energy (kcal/mol)	Group count (amide)	λ_{max} (UV-visible) (nm)	Molar refractivity	
Levamisole *	2.255	4.12	27.812	0	218.536	60.744	2.258
G-10#	-	4.157	7.616	1	226.93	141.911	2.682
G-13#	-	4.257	14.924	1	224.692	141.911	2.615
G-9#	-	6.083	6.187	1	223.253	117.237	2.513
G-12#	-	5.56	10.522	1	226.218	136.007	2.494
G-4#	-	1.562	22.139	0	223.963	114.538	2.328
G-7#	-	1.479	24.453	0	223.38	119.978	2.281
G-3	-	2.753	−27.813	0	225.584	116.577	2.217
G-1	-	2.662	11.588	0	213.185	94.167	2.116
G-6	-	3.699	4.93	0	225.643	119.238	2.009
G-8	-	2.278	47.049	0	213.123	115.254	1.957
G-2	-	3.839	18.516	0	224.884	124.932	1.908
G-5	-	5.015	19.903	0	223.576	105.021	1.862

* Standard immunomodulatory compound used as control, # predicted active gallic acid derivatives

where the regression variance is defined as the original variance minus the variance around the regression line. The original variance is the sum of the squares of the distances of the original data from the mean.

Validating QSAR equations and data: cross-validation coefficient (rCV^2)

The cross-validation coefficient is a squared correlation coefficient generated during the validation procedure.

When the predictor variables are fixed,

$$rCV^2 = 1 - (N - 1/N)(N + k + 1/N - k - 1)\left(1 - r^2\right). \quad (3)$$

When the predictor variables are random,

$$rCV^2 = 1 - (N - 1/N - k - 1)(N - 2/N - k - 2) \times (N + 1/N)\left(1 - r^2\right), \quad (4)$$

where rCV^2 refers to the cross-validation regression coefficient, r^2 refers to the regression coefficient, N refers to the number of observations (compounds), and k refers to the number of variables (descriptors).

Screening via pharmacokinetic properties

The ideal oral drug is one that is rapidly and completely absorbed from the gastrointestinal tract, distributed specifically to its site of action in the body, metabolized in a way that does not instantly remove its activity, and eliminated in a suitable manner without causing any harm. It has been reported that around half of all drugs in development fail to make it to the market because of poor pharmacokinetics (PK) [18]. The PK properties depend on the chemical properties of the molecule. PK properties such as absorption, distribution, metabolism, excretion, and toxicity (ADMET) are important determinants of the success of the compound for human therapeutic use [18–20]. Some important chemical descriptors correlate well with PK properties, such as the polar surface area (PSA; a primary determinant of fractional absorption) and low molecular weight (MW; for oral absorption) [21]. The distribution of the compound in the human body depends on factors such as the blood–brain barrier (log BB), permeability (such as the apparent Caco-2 permeability, apparent MDCK permeability, $\log K_p$ for skin permeability), the volume of distribution, and plasma protein binding ($\log K_{hsa}$ for serum protein binding) [21], so these parameters were calculated and checked for compliance with their standard ranges. The octanol–water partition coefficient ($\log P$) has been implicated in BBB penetration and permeability prediction, as has PSA. It has been reported that the process of excreting the compound from the human body depends on the MW and $\log P$. Likewise, rapid renal clearance is associated with small and hydrophilic compounds. On the other hand, the metabolism of most drugs, which takes place in the liver, is associated with large and hydrophobic compounds [22]. Higher compound lipophilicity leads to increased metabolism and poor absorption, along with an increased probability of binding to unwanted hydrophobic macromolecules, thereby increasing the potential for toxicity. In spite of some observed exceptions to Lipinski's rule, the property values of the vast majority (90%) of orally active compounds are within their cut-off limits [23]. Molecules that violate more than one of these rules may not be sufficiently bioavailable. When studying PK properties, screening based on Lipinski's rule of five (which is used to assess druglikeness) was applied to the gallic acid derivatives. In addition, the oral bioavailability of each gallic acid derivative was assessed through its topological polar surface area (TPSA) using ChemAxon's MarvinView 5.2.6:PSA plugin software [24]. This descriptor has been shown to correlate well with passive molecular transport through membranes, thus allowing the prediction of drug transport properties, and it has been linked to drug bioavailability (the percentage of the dose of the drug that reaches the blood circulation). Also, the number of rotatable bonds is a simple topological parameter used by researchers as part of an extended Lipinski's rule of five as a measure of molecular flexibility. This is a very good chemical descriptor for oral bioavailability [25]. A rotatable bond is defined as any single nonring bond bound to a nonterminal heavy (i.e., nonhydrogen) atom. Amide C–N bonds are not considered in this context because of their high rotational energy barrier. Moreover, some researchers have also included the sum of H-bond donors and H-bond acceptors as a secondary determinant of fractional absorption. The primary determinant of fractional absorption is PSA [26]. According to the extended Lipinski's rule of five, the sum of H-bond donors and acceptors should be ≤12 or the PSA should be ≤140 Å^2 [26], and the number of rotatable bonds should be ≤10 [25]. ADMET properties were calculated using QikProp v.3.2 software (Schrödinger, Portland, OR, USA, 2009).

Results and discussion

Chemical structure–activity relationship

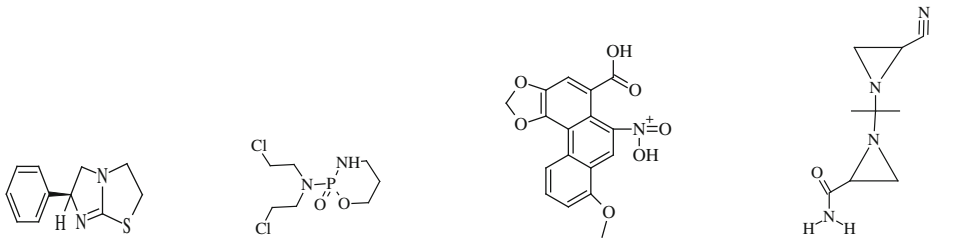
In the present work, derivatives of gallic acid were evaluated for their immunomodulatory activity through QSAR and docking studies. The QSAR results indicated that compounds G-10, G-13, G-9, G-12, G-4, and G-7 showed activity levels similar to or higher than that of levamisole. Gallic acid and its derivatives have been reported to possess immunomodulatory activity [11, 30].

Thus, we designed a prototype in which the gallic acid part was used as one of the naphthophenone rings. A fatty acid chain was also used to add some flexibility to the molecule. Thus, we designed and virtually optimized a number of gallic acid derivatives based on the conformationally restricted naphthophenone moiety as a basic unit along with different linear side chains at the 2-*O*-position. In the present work, we report the immunomodulatory activities of these newly designed gallic acid derivatives with the basic naphthophenone pharmacophore, which were found to be comparable to potent immunomodulatory and anti-inflammatory compounds (Fig. 1). Figure 2 shows the gallic acid based pharmacophore and its derivatives that were predicted to be active immunomodulatory compounds through QSAR and docking studies.

Docking-based detection of immunomodulatory targets

The aim of the molecular docking study was to elucidate whether gallic acid and its derivatives modulate the anti-inflammatory and immunomodulatory receptors, and to study their possible mechanisms of action. The results of the molecular docking study are comparable to those obtained from experimental studies of the activity of gallic acid in humans; they suggest that gallic acid inhibits histamine release and proinflammatory cytokine production in human mast cells [14]. It was reported that the inhibitory effect of gallic acid on histamine release was mediated by the modulation of cAMP and intracellular calcium, and gallic acid decreased proinflammatory cytokine gene expression and production (e.g., of TNF-α and IL-6). The inhibitory effect of gallic acid on proinflammatory cytokines was found to be dependent on nuclear factor κB and p38 mitogen-activated protein kinase [14]. However, reports also suggest that gallic acid directly suppressed the *in vitro* anti-sheep red blood cell (SRBC) antibody response at noncytotoxic doses when several chemicals such as azathioprine (Imuran) (0.5 μg/culture), gallic acid (7 μg/culture), dextran sulfate (100 μg/culture), methylparaben (100 μg/culture), and vanillin (200 μg/culture) were examined for immunomodulatory effects using the Mishell–Dutton *in vitro* antibody producing assay. All of these chemicals were reported to interrupt an early phase of the immune response, and had no effect on the actual release of the specific anti-SRBC antibody [27]. In vitro experiments relating to the anti-inflammatory and immunomodulatory activities of gallic acid derivatives showed a significant decrease in the expression of proinflammatory mediators such as IL-6, TNF-α, and nitric oxide. Also, the expression of immunomodulatory mediator IL-4 was found to increase with gallic acid administration [14].

In the work presented here, we explored the orientations and binding affinities (in terms of the docking energy in kcal mol^{-1}) of gallic acid derivatives towards proinflammatory targets. It is well known that innate immune recognition is mediated by a structurally diverse set of receptors that belong to several distinct protein families. Among them are humoral proteins circulating in the plasma, endocytic receptors expressed on the cell surface, and signaling receptors that can be expressed either on the cell surface or intracellularly [28]. Proinflammatory cytokines such as interleukin-1 (IL-1), interleukin-6 (IL-6), or tumor necrosis factor alpha (TNF-α) have been known to contribute to a variety of inflammatory conditions, such as ischemic tolerance [29], rheumatoid arthritis [30], nephritis [31], and liver diseases [32]. Nitric oxide generated through inducible NO synthase (iNOS) enzymatic activity has been found to participate in various immune and inflammatory reactions, while immunomodulatory cytokines such as interleukin 4 (IL-4), interleukin 10 (IL-10), and interleukin 13 (IL-13) are responsible for inhibiting proinflammatory signaling and hence reduce inflammation. Recent advances made in studies of innate immunity have yielded a better understanding of inflammatory mechanisms. Toll-like receptors (TLRs) have been found to recognize and respond to the moieties related to tissue injury and microbial infections [33]. TLRs are mediators of various cell-mediated and humoral immune responses caused by different agents or TLR-specific ligands. Different TLRs have been known to respond to a variety of pathogen-associated molecular patterns (PAMPs), such as microbial agents, viral proteins, RNA, CpG DNA, bacterial lipopolysaccharides (LPSs), and peptidoglycan. Signaling through TLRs results in inflammatory reactions mediated



(LD_{50} = 180 mg/kg) Levamisole

(LD_{50} = 94 mg/kg) Cyclophosphamide

LD_{50} = 81mg/kg) Aristolochic acid

(LD_{50} = 170 mg/kg) Azimexon

Fig. 1 Some potent immunomodulatory compounds along with their experimental activities

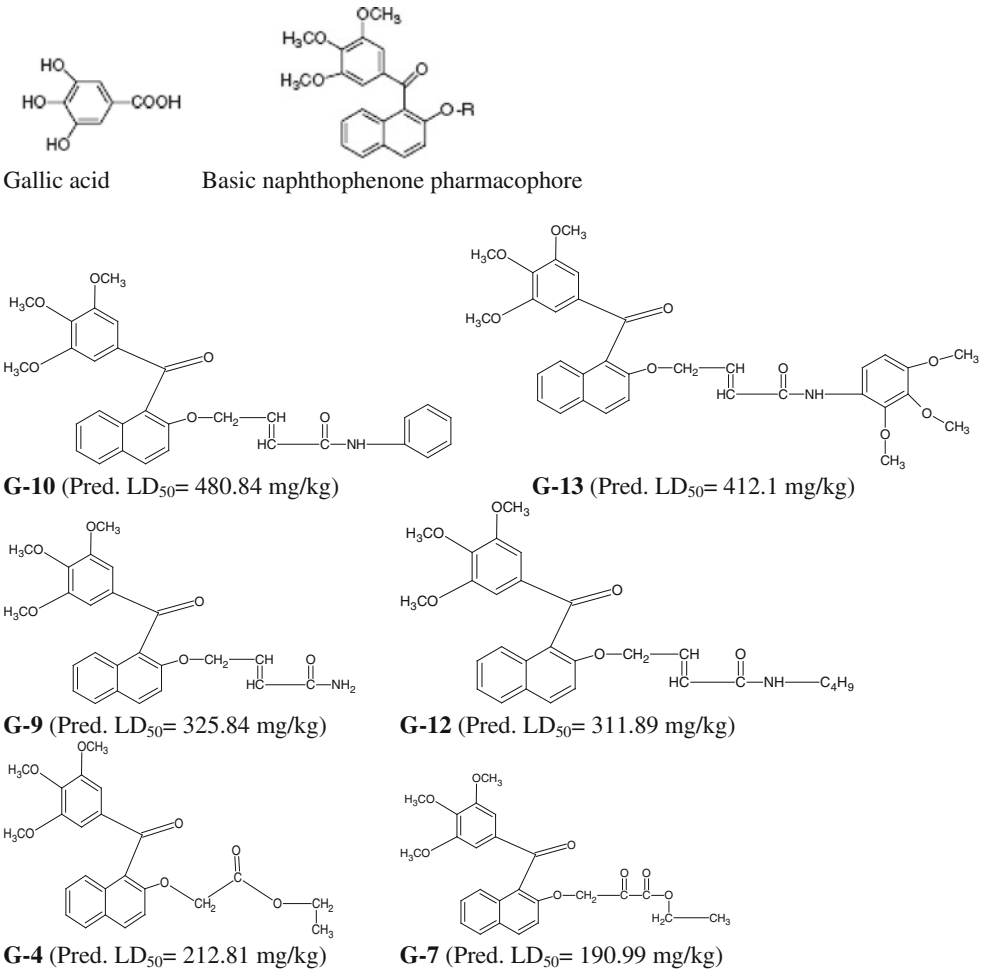


Fig. 2 Gallic acid based pharmacophore and its derivatives that were predicted to be active immunomodulatory compounds using the derived QSAR model

by various cytokines such as TNF-α, IL-6, IL-8, and IL-1β. The inhibitors of the TLR-mediated signaling of inflammatory reactions are the decoy receptors, signaling inhibitors, and immunomodulatory cytokines IL-4, IL-10, and IL-13. [34]. Cluster of differentiation (CD) molecules also play a very important role in the various immunological cascades of reactions, and act as co-stimulatory signaling molecules for the activation of several lymphocytes. Their activities are responsible for producing numerous immune responses, such as the production of T helper cells, cytotoxic T cells, macrophage activation, and antibody production [35, 36]. The results obtained from molecular docking are comparable to corresponding reported experimental data, which suggests that the decreased proinflammatory mediator expression is due to gallic acid derivatives. This in turn indicates that the oral administration of gallic acid derivatives will inhibit proinflammatory mediators and thus enhance the production of immunomodulatory mediators. Gallic acid derivatives showed high binding affinities (in terms of the docking energy in kcal mol^{-1}) with the immunomodulatory receptors INFα-2, IL-4, and IL-6 (Table 2). Moreover, high binding affinities (lower docking energies) with the INFα-2 receptor along with hydrogen (H) bond formation were noted for the active gallic acid derivatives G-4 (Fig. 3), G-7 (Fig. 4), and G-10 (Fig. 5). Similarly, the active gallic acid derivatives G-7 and G-10 showed high binding affinities for the IL-4 (Fig. 6) and IL-6 (Fig. 7) receptors, respectively, along with H-bonding.

Comparison of the binding pocket residues for interferon α-2 (INFα-2)

The activities of the gallic acid derivatives were analyzed by performing molecular docking experiments with the immunomodulatory receptors (INFα-2 and interleukins). IL-4 and IL-6 are known proinflammatory cytokines that play an important role in the immunomodulatory pathway. The binding affinities obtained in the docking study allowed the activities of the gallic acid derivatives to be compared to that of the standard immunomodulatory compound levamisole. All of the derivatives showed high binding affinities (low docking energies) for INFα-2. When we compared how the binding pocket residues of INFα-2 interacted with the gallic acid derivatives, we found that

Table 2 Docking scores (kcal mol^{-1}) of gallic acid derivatives with respect to the immunomodulatory targets INFα-2, IL-4, and IL-6

Compound	Docking energy (kcal/mol) with Immunomodulatory receptors		
	INFα-2	IL-4	IL-6
Levamisole	−41.53	−47.78	−66.76
G-1	−58.66	−72.80	−92.96
G-2	−62.09	−90.25	−90.71
G-3	−66.50	−70.01	−87.73
G-4#	−59.87	−69.39	−94.76
G-5	−62.80	−86.25	−105.20
G-6	−57.32	−75.40	−99.09
G-7#	−60.59	−87.55	−88.97
G-8	−57.67	−83.26	−97.41
G-9#	−62.17	−81.32	−88.04
G-10#	−70.26	−89.11	−94.44
G-11	−75.10	−93.24	−78.05
G-12#	−60.93	−80.79	−95.73
G-13#	−80.38	−108.22	−89.71

Numeric values in boldface indicate H-bond formation, and # indicates a predicted active gallic acid derivative

only compounds G-4 (Fig. 3), G-5, G-7 (Fig. 4), and G-10 (Fig. 5) form H-bonds, leading to more stability and potency in these cases (Table 2). The docking results for the active derivatives showed that compound G-4 docked onto INFα-2 with a low interaction energy (−59.87 kcal mol^{-1}) and formed an H-bond of length 2.045 Å to the basic amino acid residue Lys-15. In this complex, the binding pocket residues within a radius of 3 Å were Pro-109, Ile-194 (hydrophobic), Glu-108 (acidic), Ser-17 (nucleophilic), and Lys-15 (basic) (Fig. 3). On the other hand, the docking results for levamisole with INFα-2 showed a docking energy of −41.53 kcal mol^{-1} and the formation of an H-bond of length 1.938 Å to the basic residue Arg-144. Other residues within a radius of 3 Å were Arg-22, Arg-144, and Arg-149, which are basic in nature. Similarly, compound G-5 (predicted to be inactive) had a docking energy of −62.80 kcal mol^{-1} and formed an H-bond of length 1.994 Å to the acidic residue Glu-108. In this complex, the binding pocket residues within 3 Å were Pro-109, Pro-197 (hydrophobic), Glu-108 (acidic), Ser-26, Ser-17 (nucleophilic), and Lys-15 (basic). Likewise, compound G-7 docked with an interaction energy of −60.59 kcal mol^{-1} and formed an H-bond of length 2.167 Å to the basic residue Lys-15. The binding pocket residues within 3 Å were Lys-15 (basic) and Glu-108 (acidic) (Fig. 4). Compound G-10 docked with a docking energy of −70.26 kcal mol^{-1} and formed an H-bond of length 2.029 Å to the acidic residue

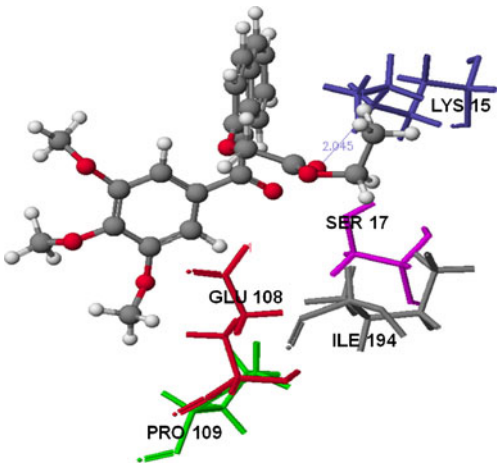


Fig. 3 Compound G-4 was docked onto immunomodulatory receptor INFα-2 with a docking energy of −59.87 kcal mol^{-1}, and an H-bond of length 2.045 Å to the binding pocket residue Lys-15 was observed

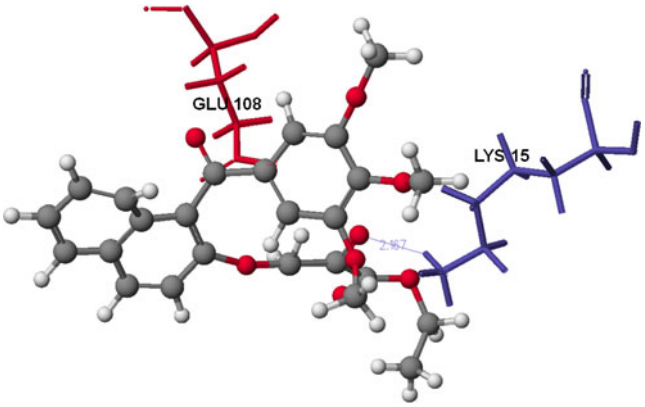


Fig. 4 Compound G-7 was docked onto immunomodulatory receptor INFα-2 with a docking energy of −60.59 kcal mol^{-1}, and an H-bond of length 2.167 Å to the binding pocket residue Lys-15 was observed

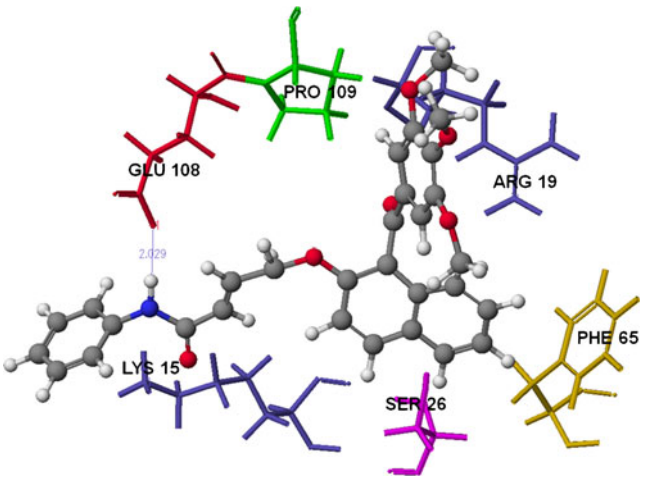


Fig. 5 Compound G-10 was docked onto immunomodulatory receptor INFα-2 with a docking energy of −70.26 kcal mol^{-1}, and an H-bond of length 2.029 Å to the binding pocket residue Glu-108 was observed

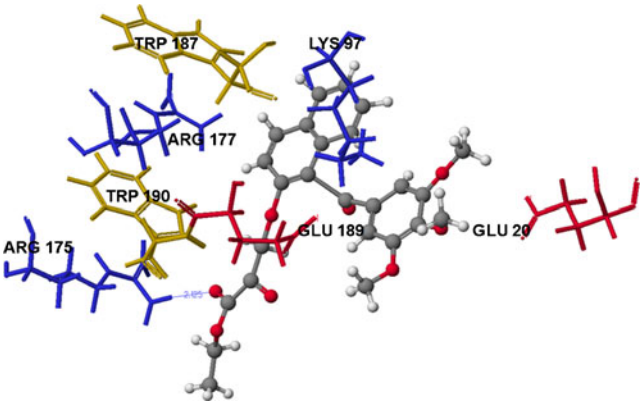


Fig. 6 Compound G-7 was docked onto immunomodulatory receptor IL-4 with a docking energy of −87.55 kcal mol^{-1}, and an H-bond of length 2.85 Å to the binding pocket residue Arg-175 was observed

Glu-108. The binding pocket residues within 3 Å were Arg-19 (basic), Glu-108 (acidic), Pro-109 (hydrophobic), Lys-15 (basic), Phe-65 (aromatic), and Ser-26 (nucleophilic) (Fig. 5).

Comparison of the binding pocket residues for interleukin-4 (IL-4)

The results of molecular docking showed high binding affinities of the gallic acid derivatives for the immunomodulatory receptor IL-4, comparable to that of levamisole (Table 2). When we compared the binding pocket residues that interacted with the active conformations of the derivatives within the IL-4 complex, we concluded that only compounds G-3, G-5, G-6, and G-7 form H-bonds, so these were considered the most stable and potent compounds. The docking results showed that compound G-3 docked with a docking energy of −70.01 kcal mol^{-1} and formed an H-bond of length 1.739 Å to the binding pocket basic residue Arg-53. Other binding pocket residues within a radius of 3 Å were Arg-53, Arg-89 (basic), Asp-87 (acidic), Tyr-56 (aromatic), Lys-61, Lys-84 (basic), Ser-57 (nucleophilic), and Glu-60 (acidic). Similarly, compound G-5 docked with a docking energy of −86.25 kcal mol^{-1} and formed an H-bond of length 2.085 Å to the basic residue His-62. The binding pocket residues of the complex within 3 Å were Ile-16 (hydrophobic), His-62 (basic), Gln-52, Gln-189 (amide), Thr-18 (nucleophilic), Val-10 (hydrophobic), and Lys-97 (basic). Likewise, compound G-6 docked with a docking energy of −75.40 kcal mol^{-1} and formed an H-bond of length 2.104 Å to the basic residue Lys-97. The binding pocket residues within 3 Å were Ile-16 (hydrophobic), Glu-20 (acidic), His-62, Lys-97, Arg-91 (basic), Thr-18 (nucleophilic), and Val-60 (hydrophobic). Lastly, active compound G-7 docked with a docking energy of −87.55 kcal mol^{-1} and formed an H-bond of length 2.85 Å to the basic residue Arg-175. The binding pocket residues (within a radius of 3 Å) were Glu-20, Glu-189 (acidic), Arg-97, Arg-175, Arg-177 (basic), Trp-187, and Trp-190 (both aromatic) (Fig. 6).

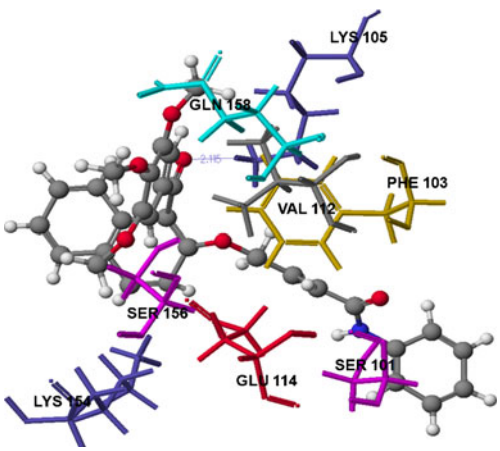


Fig. 7 Compound G-10 was docked onto immunomodulatory receptor IL-6 with a docking energy of −94.44 kcal mol^{-1}, and an H-bond of length 2.115 Å to the binding pocket residue Lys-105 was observed

Comparison of the binding pocket residues for interleukin-6 (IL-6)

The results of molecular docking showed high binding affinities of the gallic acid derivatives for the immunomodulatory receptor IL-6, comparable to that of levamisole (Table 2). When we compared the binding pocket residues that interacted with the active conformations of the gallic acid derivatives within the IL-6 complex, only compounds G-3, G-6, and G-10 were observed to form H-bonds, so they were considered the most stable and potent. The docking results showed that compound G-3 docked with a docking energy of −87.73 kcal mol^{-1} and formed an H-bond of length 2.075 Å to the basic residue Lys-105. Other binding pocket residues (within a radius of 3 Å) were Gln-196 (amide), Glu-114, Glu-286 (acidic), Lys-105, Lys-154 (basic), and Phe-103 (aromatic). Similarly, compound G-6 docked with a docking energy of −99.09 kcal mol^{-1} and formed an H-bond of length 2.132 Å to the aromatic residue Phe-103. Other conserved binding pocket residues of the complex within 3 Å were Lys-105, Lys-154 (basic), Trp-103, Trp-115, Phe-103 (aromatic), Gln-99, Gln-196 (amide), and Asp-198 (acidic). Likewise, the active compound G-10 docked onto IL-6 with a docking energy of −94.44 kcal mol^{-1} and formed an H-bond of length 2.135 Å to the basic residue Lys-105. Other conserved binding pocket residues of the complex within 3 Å were Gln-158 (amide), Ser-101, Ser-156 (nucleophilic), Lys-105, Lys-154 (basic), Glu-114 (acidic), and Phe-103 (aromatic) (Fig. 7).

Predicting activity with the QSAR model

The structure–activity relationship denoted by the QSAR model yielded a very high activity–descriptors relationship accuracy of 99% (r^2=0.99) and a high activity prediction accuracy of 96% ($r\mathrm{CV}^2$=0.96) (Fig. 9). Five chemical descriptors were found to be applicable to the immunomodulatory activity. The QSAR equation indicated that dipole moment, steric energy, amide group count, λ_{max} (UV-visible), and molar refractivity correlated well with activity. The QSAR model equation is given below, showing the relationship between experimental activity *in vivo* [i.e., the dose that is lethal to 50% of the population (LD_{50})] as the dependent variable and five independent variables (descriptors):

$$\begin{aligned}\text{predicted log } LD_{50}(\text{mg/kg}) = {} & -0.156436 \times \text{dipole moment(debye)} \\ & -0.00118794 \times \text{steric energy}\left(\text{kcal mol}^{-1}\right) \\ & +0.910351 \times \text{amide group count} \\ & +0.0206362 \times \lambda\,\text{max(UV} - \text{visible)(nm)} \\ & -0.00834447 \times \text{molar refractivity} \\ & -1.06753\end{aligned}$$

Here, $r\mathrm{CV}^2$ (the cross-validation regression coefficient) =0.96, which indicates that the newly derived QSAR model has a prediction accuracy of 96%, and r^2 (regression coefficient)=0.99, which indicates that the correlation between the activity (dependent variable) and the descriptors (independent variables) for the training data set compounds was 99% (Table 1, Fig. 8). Thus, we successfully developed a QSAR model for immunomodulatory activity. Results showed that the predicted activities were comparable with those obtained experimentally (Fig. 9), and that compounds G-4, G-7, G-9, G-10, G-12, and G-13 have higher immunomodulatory activities than the standard compound levamisole.

Assessing the pharmacokinetic parameters

We considered several physiochemical properties related to PK when screening for active gallic acid derivatives. The results revealed that, except for the compounds G-10 and G-13, all of the gallic acid derivatives followed Lipinski's rule of five (Table 3). Compound G-10 violated Lipinski's rule as log$P$ > 5, so it was likely to be poorly soluble in aqueous solution and hence unable to gain access to membrane surfaces. Lipophilicity (ratio of a molecule's solubility in octanol to solubility in water) is measured through log$P$. Log$P$ has been linked to blood–brain barrier penetration and utilized to predict permeability. The process of excretion, which eliminates the compound from the human body, depends on its molecular weight and log$P$ [21]. Molecules with intermediate lipophilicities have a better chance of arriving at the receptor site [22]. Similarly, compound G-13 violated Lipinski's rule as MW > 500 Da, making it likely to have low solubility and to pass through cell membranes with difficulty. The other active derivatives followed Lipinski's rule and had polarities that enabled better permeation and absorption, as revealed by the number of H-bond donors and H-bond acceptors. Similarly, the ADME parameters were calculated for the active gallic acid derivatives G-3, G-4, G-5, G-6, G-7, and G-10; the values of these parameters also showed close correspondence with those of levamisole and fell within the standard range of values exhibited by 95% of all known drugs. Calculations related to aqueous solubility, serum protein binding, the blood–brain barrier (log BB and apparent MDCK cell permeability), gut–blood barrier (Caco-2 cell permeability), predicted central nervous system activity, number of likely metabolic reactions, log IC_{50} for hERG K^+ channel blockage, transdermal

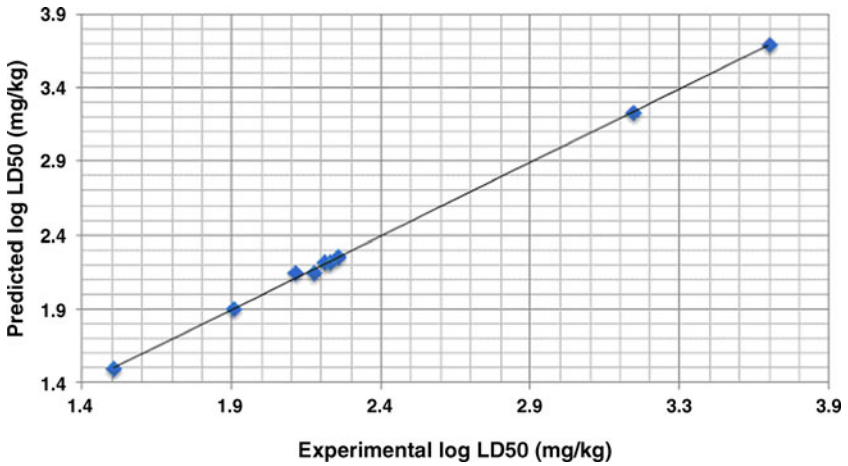


Fig. 8 Multiple linear regression analysis indicates a linear relationship between the experimental and predicted log LD_{50} (mg/kg) values for the training set

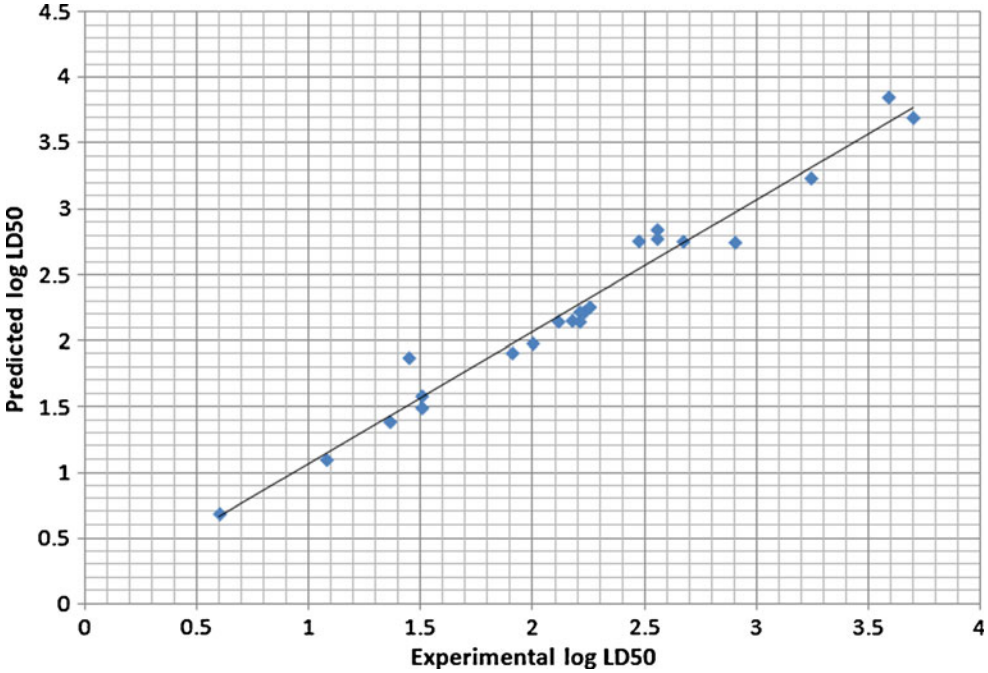


Fig. 9 Multiple linear regression analysis indicates a linear relationship between the experimental and predicted log LD_{50} (mg/kg) values for the test set

transport rate (J_m), skin permeability (K_p), and human oral absorption in the gastrointestinal tract showed that these values for the active gallic acid derivatives fell within the standard ranges generally observed for drugs (Table 4).

Toxicity risk assessment

It is now possible to predict the activities and toxicity risks of compounds using reliable bioinformatics tools. In the

Table 3 Compliance of gallic acid derivatives with standard ranges of computational parameters of druglikeness and ADME properties. Compounds G-10 and G-13 were found to violate Lipinski's rule of five

Compound	Pharmacokinetic property (ADME) dependent on chemical descriptors								Rule of five violation
	ADM	AE	ADME	AD					
	Oral bioavailability: TPSA (Å^2)	MW	log*P*	H-bond donor			H-bond acceptor		
				Amine group count	*sec*-Amine group count	Hydroxyl group count	Nitrogen atom count	Oxygen atom count	
Levamisol	40.9	204.29	3.259	0	0	0	2	0	0
G-1	64.99	338.35	3.228	0	0	1	0	5	0
G-2	53.99	450.48	3.832	0	0	0	0	7	0
G-3	91.29	424.44	3.243	0	0	2	0	7	0
G-4#	80.29	424.44	3.175	0	0	0	0	7	0
G-5	91.29	396.39	2.801	0	0	0	0	7	0
G-6	80.29	438.47	3.357	0	0	0	0	7	0
G-7#	97.36	452.46	2.712	0	0	0	0	8	0
G-8	80.29	460.43	4.377	0	0	0	0	7	0
G-9#	97.08	421.44	2.593	0	0	0	1	6	0
G-10#	83.09	497.54	5.23	0	1	0	1	6	1
G-12#	83.09	477.55	4.047	0	1	0	1	6	0
G-13#	110.78	587.62	3.763	0	1	0	1	9	1

*A*absorption, *D*distribution, *M*metabolism, *E* excretion, *TPSA*topological polar surface area, *MW*molecular weight, log*P*octanol/water partition coefficient

# indicates a QSAR-based predicted active gallic acid derivative

Table 4 Compliance of active gallic acid derivatives with the standard ranges of computational pharmacokinetic parameters (ADME)

Principal descriptors	Levamisole	G-3	G-4#	G-5	G-6	G-7#	G-10#	Stand. range*
log*S* (aqueous solubility)	−3.476	−5.549	−5.378	−4.425	−5.598	−5.297	−7.594	−6.5 / 0.5
logK_{hsa} (serum protein binding)	0.112	0.266	0.319	−0.004	0.394	−0.020	0.964	−1.5 / 1.5
log BB for brain/blood	0.462	−1.546	−0.924	−1.240	−0.923	−1.526	−1.023	−3.0 / 1.2
No. of metabolic reactions	2	5	5	5	5	5	6	1.0 / 8.0
Predicted CNS activity	+2	−2	−1	−2	−1	−2	−2	−2 (inactive), +2 (active)
log IC_{50} for hERG K^+channel blockage	−4.198	−4.306	−6.116	−3.717	−6.193	−6.721	−7.702	Concern below −5
Apparent Caco-2 permeability (nm/s)	5589	99	1448	131	1682	597	1580	<25 poor, >500 great
Apparent MDCK permeability (nm/s)	5839	51	738	70	867 M	283 M	811 M	<25 poor, >500 great
logK_p for skin permeability	−3.392	−2.469	−1.425	−2.377	−1.210	−1.971	−0.482	−8.0 to −1.0, K_p in cm/h
J_m (max. transdermal transport rate)	0.028	0.004	0.067	0.063	0.068	0.024	0.004	μg/cm^2 h
Jorgensen rule of three violations	0	0	0	0	0	0	1	Maximum is 3
% human oral absorption in GI (±20%)	100	89	100	87	100	100	89	<25% is poor
Qual. model for human oral absorption	High	High	High	High	High	High	Low	>80% is high

* For 95% of known drugs, based on –Qikprop v.3.2 (Schrödinger, USA, 2009) software results

# indicates a QSAR-based predicted active gallic acid derivative

present study, we calculated toxicity risk parameters such as mutagenicity, tumorogenicity, irritation, and reproduction of the gallic acid derivatives (G3–G13) (Table 5). The toxicity risk predictor locates fragments within a molecule that indicate a potential toxicity risk. Toxicity screening results showed that none of the compounds presented a risk of tumorogenicity or reproductive toxicity, although there was a partial mutagenicity risk. On the other hand, compounds G-3, G-9, G-10, G-12, and G-13 presented a high risk of irritation and were thus rejected, while compounds G-4, G-5, and G-7 presented no risk of irritation. Compound G-6 yielded a medium risk of irritation. The hydrophilicity of each compound was measured through its log*P* value. Low hydrophilicity and therefore a high log*P* value may lead to poor absorption or permeation. For compounds to have a reasonable probability of being well absorbed, it has been found that their log*P* values must not be >5. This study suggests that, except for compounds G-10 and G-12, all of the compounds conformed to this limit. Typically, low solubility is associated with bad absorption, so the general aim is to avoid poorly soluble compounds. The aqueous solubility (log*S*) of a compound significantly affects its absorption and distribution characteristics. The calculated log*S* values of the studied compounds were within the

Table 5 Compliance of the active gallic acid derivatives with the standard intervals for computational toxicity risk parameters

Compound	Toxicity risk parameters				Druglikeness parameters (Osiris)				
	MUT	TUMO	IRRI	REP	MW	CLP	S	DL	DS
Levamisole	No risk	No risk	No risk	No risk	206	1.38	−1.52	3.73	0.95
G-3	Medium risk	No risk	High risk	No risk	422	3.67	−5.48	4.81	0.25
G-4#	Medium risk	No risk	No risk	No risk	424	3.99	−5.6	0.24	0.31
G-5	Medium risk	No risk	No risk	No risk	396	3.1	−5.17	4.21	0.48
G-6	Medium risk	No risk	Medium risk	No risk	438	4.45	−5.87	−6.57	0.14
G-7#	Medium risk	No risk	No risk	No risk	452	3.35	−5.52	−10.7	0.21
G-9#	Medium risk	No risk	High risk	No risk	421	3.13	−5.56	3.35	0.26
G-10#	Medium risk	No risk	High risk	No risk	497	5.23	−6.98	2.55	0.14
G-12#	Medium risk	No risk	High risk	No risk	477	5.01	−6.31	2.42	0.16
G-13#	Medium risk	No risk	High risk	No risk	587	4.92	−7.03	4.23	0.13

MUT mutagenicity, *TUMO* tumorogenicity, *IRRI* irritation, *REP* reproduction, *MW* molecular weight, *CLP* C log*P*, *S* solubility, *DL* druglikeness, *DS* drug score

# indicates a QSAR-based predicted active gallic acid derivative

acceptable interval. To judge the compound's overall potential to act as a drug, we calculated its overall drug score, which combines its druglikeness, *C*log*P*, log*S*, MW, and toxicity risk parameter values. Generally speaking, the calculated parameters for the active compounds were within the acceptable interval. Results revealed that the overall drug scores of compounds G-5, G-4, G-7, and G-6 were good to moderate compared to the standard immunomodulatory compound levamisole.

Conclusions

Molecular docking and QSAR studies were performed on gallic acid derivatives in order to predict the potential immunomodulatory compounds. During the molecular docking studies, all of the derivatives showed high binding affinities with INFα-2, IL-4, and IL-6. The binding site residues of INFα-2 exhibited H-bond formation with compounds G-4, G-5, G-7, and G-10. Similarly, compounds G-3, G-6, and G-10 formed H-bonds with IL-6 binding site residues. On the other hand, the binding site residues of IL-4 exhibited H-bond formation with compounds G-3, G-5, G-6, and G-7, which were thus considered to be the most stable and potent of the compounds. Moreover, virtual screening performed using the derived QSAR model suggested that compounds G-4, G-7, G-9, G-10, G-12, and G-13 possess immunomodulatory activity. However, compounds G-10 and G-13 violate Lipinski's rule, indicating low oral bioavailability. Based on bioavailability, *in silico* ADME, and toxicity risk assessments for mutagenicity, tumorogenicity, irritation, and reproduction, we concluded that compound G-7 possesses greater immunomodulatory activity then G-4, G-9, G-10, G-12, and G-13.

Acknowledgments We acknowledge the Council of Scientific and Industrial Research, New Delhi, for financial support through networking project (NWP-09) at the Central Institute of Medicinal and Aromatic Plants, Lucknow, India.

References

1. Patwardhan B, Dnyaneshwar W, Pushpangadan P, Bhatt N (2005) Ayurveda and traditional Chinese medicine: a comparative overview. Evid-Based Compl Alt 2:465–473
2. Wagner H (1990) Search for plant derived natural products with immunostimulatory activity (recent advances). Pure Appl Chem 62:1217–1222
3. Muegge I (2000) A knowledge-based scoring function for protein-ligand interactions: probing the reference state. Perspect Drug Discov 20:99–114
4. Kashiwada Y, Nonaka NG, Chang IJ, Lee KHI (1992) Antitumor agents, 129. Tannins and related compounds as selective cytotoxic agents. J Nat Prod 55:1033–1043
5. Yoshida F, Topliss JG (2000) QSAR model for drug human oral bioavailability. J Med Chem 43:2575–2585
6. Cho J, Tropsha SA, Lee KH (1996) Antitumor agents, 163. Three-dimensional QSAR study of 4′-*O*-demethylepipodophyllotoxin analogs using the CoMFA/q2-GRS approach. J Med Chem 39:1383–1395
7. Leung D, Abbenante G, Fairlie DP (2000) Protease inhibitors: current status and future prospects. J Med Chem 43:305–341
8. Bibby MC (2002) Combretastatin anticancer drugs. Drugs Future 27:475–480
9. Srivastava V, Saxena HO, Shanker K, Kumar JK, Luqman S, Gupta MM, Khanuja SPS, Negi AS (2006) Synthesis of gallic acid based naphthophenone fatty acid amides as cathepsin D inhibitors. Bioorg Med Chem Lett 16:4603–4608
10. Negi AS, Darokar MP, Chattopadhyay SK, Garg A, Bhattacharya AK, Srivastava V, Khanuja SPS (2005) Synthesis of a novel plant growth promoter from gallic acid. Bioorg Med Chem Lett 15:1243–1247
11. Sohi KK, Mittal N, Hundal MK, Khanduja KL (2003) Gallic acid, an antioxidant, exhibits antiapoptotic potential in normal human lymphocytes: a Bcl-2 independent mechanism. J Nutr Sci Vitaminol 49:221–227
12. Maurya DK, Nandakumar N, Devasagayam TP (2010) Anticancer property of gallic acid in A549, a human lung adenocarcinoma cell line and possible mechanisms. J Clin Biochem Nutr 48:85–90
13. Kayser O, Kolodziej H, Kiderlen AF (2001) Immunomodulatory principles of *Pelargonium sidoides*. Phytother Res 15:122–126
14. Kim SH, Jun CD, Suk K, Choi BJ, Lim H, Park S, Lee SH, Shin HY, Kim DK, Shin TY (2006) Gallic acid inhibits histamine release and proinflammatory cytokine production in mast cells. Toxicol Sci 91:123–131
15. Martin C (1999) A general and fast scoring function for protein–ligand interactions: a simplified potential approach. J Med Chem 42:791–804
16. Sanda PD, Silva LD, Juranic ZD, Markovic BD, Vladimirov SM, Juranic IO (2008) Docking studies and immuno-modulatory activity of β-hydroxy-β-arylpropanoic acids. Molecules 13:603–615
17. Meena A, Yadav DK, Srivastava A, Khan F, Chanda D, Chattopadhyay SK (2011) In silico exploration of anti-inflammatory activity of natural coumarinolignoids. Chem Biol Drug Des 78:567–579
18. Hodgson J (2001) ADMET—turning chemicals into drugs. Nat Biotechnol 19:722–726
19. Sean E, Sergey A, Andy R, Eugene K, Eugene A, Rakhmatulin SS, Bugrim A, Nikolskaya T (2006) A combined approach to drug metabolism and toxicity assessment. Drug Metab Dispos 34:495–503
20. Norinder U, Bergström CAS (2006) Prediction of ADMET properties. Chem Med Chem 1:920–937
21. Reichel A, Begley DJ (1998) Potential of immobilized artificial membranes for predicting drug penetration across the blood–brain barrier. Pharm Res 15:1270–1274
22. Lombardo F, Gifford E, Shalaeva MY (2003) In silico ADME prediction: data, models, facts and myths. Mini-Rev Med Chem 3:861–875
23. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46:3–26
24. Yadav DK, Meena A, Srivastava A, Chanda D, Khan F, Chattopadhyay SK (2010) Development of QSAR model for

immunomodulatory activity of natural coumarinolignoids. Drug Des Dev Ther 4:173–186
25. Veber DF, Stephen R, Johnson, Cheng H-Y, Smith BR, Keith W, Ward, Kopple KD (2002) Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 45:2615–2623
26. Clark DE (1999) Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 1. Prediction of intestinal absorption. J Pharm Sci 88:807–814
27. Kutz SA, Ronald DH, Deborah JW (1980) Evaluation of chemicals for immunomodulatory effects using an in vitro antibody-producing assay. Environ Res 22:368–376
28. Muzio M, Muzio M, Daniela B, Polentarutti N, D'Amico G, Antonella S, Roberta M, Cornelis V, Penton-Rol G, Luigi PR, Paola A, Alberto M (2000) Differential expression and regulation of Toll-like receptors (TLR) in human leukocytes: selective expression of TLR3 in dendritic cells. J Immunol 164:5998–6004
29. Kariko K, Drew W, Rank A, Welsh (2004) Inhibition of Toll-like receptor and cytokine signaling—a unifying theme in ischemic tolerance. J Cereb Blood Flow Metab 24:1288–1304
30. Selzman CH (1998) Ovarian ablation alone promotes aortic intimal hyperplasia and accumulation of fibroblast growth factor. Ann Thorac Surg 98:2049–2054
31. Williams G, Giroir BP (1995) Regulation of cytokine gene expression: tumor necrosis factor, interleukin-1, and the emerging biology of cytokine receptors. New Horiz 3:276–287
32. Sasaki T, Soh H, Kimura T, Hasegawa T, Okada A, Fukuzawa M (2005) Recurrent acute pancreatitis caused by malrotation of the intestine and effective treatment with laparoscopic Ladd's procedure. Pediatric Surgery Intl 21:994–996
33. Philbinl VJ, Levy O (2007) Immunostimulatory activity of Toll-like receptor 8 agonists towards human leucocytes: basic mechanisms and translational opportunities. Biochem Soc Trans 35:1485–1497
34. Rachmilewitz D, Katakura K, Karmeli F, Hayashi T, Reinus C, Rudensky B, Akira S, Takeda K, Lee J, Takabayashi K, Raz E (2004) Toll-like receptor 9 signaling mediates the anti-inflammatory effects of probiotics in murine experimental colitis. Gastroenterology 126:520–528
35. Fenton MJ, Golenbock DT (1998) LPS-binding proteins and receptors. J Leukocyte Biol 64:25–32
36. Sau K, Mambula SS, Latz E, Henneke P, Golenbock DT, Levitz SM (2003) The antifungal drug amphotericin B promotes inflammatory cytokine release by a Toll-like receptor- and CD14-dependent mechanism. J Biol Chem 278:37561–37568