IJMS
Vol 50, No 3, March 2025

Original Article

# Targeting Trefoil Factor Family 3 in Obstructive Airway Diseases: A Computational Approach to Novel Therapeutics

Alireza Shahriary[1], PhD; Mohsen Sisakht[2], PhD; Masoud Arabfard[1], PhD; Esmaeil Behmard[3], PhD; Ali Najafi[4], PhD

[1]Chemical Injuries Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran;
[2]Department of Molecular Medicine, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran;
[3]School of Advanced Technologies in Medicine, Fasa University of Medical Sciences, Fasa, Iran;
[4]Molecular Biology Research Center, Biomedicine Technologies Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran

**Correspondence:**
Ali Najafi, PhD;
Molasadra St., Baqiyatallah University of Medical Sciences,
Postal code: 14376-65454, Tehran, Iran
**Tel:** +98 21 87554807
**Email:** najafi74@bmsu.ac.ir
Received: 26 February 2024
Revised: 06 May 2024
Accepted: 01 June 2024

## What's Known

- Trefoil Factor Family 3 (TFF3) plays a crucial role in airway remodeling, a key feature of obstructive airway diseases such as chronic obstructive pulmonary disease (COPD) and mustard lung.
- Inhibition of TFF3 expression and dimerization has been proposed as a potential therapeutic strategy.

## What's New

- Our study utilizes computational methods to identify genistein as a potent inhibitor of TFF3.
- We provide novel insights into the binding dynamics and efficacy of genistein against TFF3, highlighting its low toxicity and potential as a therapeutic agent.

## Abstract

**Background:** Airway remodeling, a hallmark of chronic obstructive pulmonary disease (COPD) and mustard lung disease, is influenced by the Trefoil Factor 3 (TFF3). This study sought to pinpoint a compound with minimal toxicity that can effectively suppress TFF3 expression and activity.

**Methods:** We employed an integrative approach, combining gene expression analysis, molecular docking, and molecular dynamics simulations to identify potential TFF3 inhibitors. Gene expression analysis utilized Z-scores from the Library of Integrated Network-Based Cellular Signatures (LINCS) database to identify compounds altering TFF3 expression. Drug-like properties were assessed through Lipinski's "Rule of Five." Molecular docking was conducted with AutoDock Vina (version 1.1.2), and molecular dynamics simulations were performed using Groningen Machine for Chemical Simulations (GROMACS) version 5.1. Toxicity evaluation leveraged a Graph Convolutional Network (GCN). Statistical significance was set at $P<0.05$.

**Results:** Eight of the compounds assessed significantly reduced TFF3 expression, with binding affinities (ΔG) ranging from -7 to -9.4 kcal/mol. Notably, genistein emerged as the frontrunner, showcasing potent TFF3 downregulation, minimal toxicity, and a robust inhibitory profile, as evidenced by molecular dynamics simulations. The significance of gene expression changes was indicated by Z-scores provided by the LINCS database rather than exact P values.

**Conclusion:** Genistein holds promise as a therapeutic agent for TFF3-mediated conditions, including mustard lung disease. Its potential to address the current therapeutic gaps is evident, but its clinical utility necessitates further *in vitro* and *in vivo* validation. A preprint of this article has already been published (https://assets.researchsquare.com/files/rs-3907985/v1/41b7e6e6-4d70-4573-81e6-4d5a913950bd.pdf?c=1707752778).

Please cite this article as: Shahriary AR, Sisakht M, Arabfard M, Behmard E, Najafi A. Targeting Trefoil Factor Family 3 in Obstructive Airway Diseases: A Computational Approach to Novel Therapeutics. Iran J Med Sci. 2025;50(3): 159-170. doi: 10.30476/ijms.2024.101737.3435.

**Keywords** ● Trefoil factor 3, airway remodeling ● Lung diseases, obstructive ● Genistein

## Introduction

Airway remodeling is a defining feature of respiratory diseases, notably chronic obstructive pulmonary disease (COPD)[1] and mustard lung (ML).[2] This remodeling, characterized by airway

wall thickening, mucous gland hypertrophy, and increased smooth muscle mass, leads to airflow obstruction and compromised respiratory function. The molecular underpinnings of these diseases are complex, with disruptions in specific pathways causing pathological changes in the airway.[3, 4] Recent studies highlight Trefoil Factor 3 (TFF3) as a central figure in airway remodeling and associated pathologies.[5, 6] Elevated TFF3 levels correlate with chronic inflammatory respiratory diseases, including COPD, ML, and various adenocarcinomas.[7-10] Beyond respiratory contexts, TFF3 is implicated in diverse conditions such as Type 2 Diabetes Mellitus (T2DM), Non-Alcoholic Fatty Liver Disease (NAFLD), neurodegeneration, gastric ulcers, colitis, and several malignancies.[11, 12] TFF3, part of the human TFF peptide family, is primarily found in mucosal environments.[13] Its dimer form, more potent than its monomer counterpart, is particularly significant for its anti-apoptotic properties.[14, 15] This dimerization enhances mucus thickness by interacting with soluble mucins such as Mucin 5AC (MUC5AC) and Mucin 6 (MUC6), emphasizing its role in mucus overproduction and airway obstruction.[13] All human TFFs, including Trefoil Factor 1 (TFF1), Trefoil Factor 2 (TFF2), and TFF3, are dual-capacity lectins recognizing the GlcNAc-α-1,4-Gal disaccharide, crucial for their mucus-thickening ability.[13-15] Given TFF3's broad roles, targeted therapeutic strategies are essential. Advanced techniques, including the Library of Integrated Network-Based Cellular Signatures (LINCS),[16] molecular docking, molecular dynamics simulations, and artificial intelligence (AI) algorithms, offer avenues for drug discovery. Our study aims to identify a low-toxicity compound that can downregulate TFF3 expression and inhibit its dimerization, providing a novel therapeutic approach for obstructive airway diseases.

## Materials and Methods

### *Ethical Approval*

This study was conducted in accordance with the institution's ethical standards and with the approval of the relevant ethics committee (Ethics Code: IR.BMSU.BLC.1400.005).

### *TFF3 Gene Expression Analysis*

This study, conducted in 2024, employed an integrative computational approach. We utilized the Library of LINCS dataset (GSE92742),[16] which captures changes in expression for approximately 12,000 genes across various human cell lines post-chemical exposure as our primary resource. Compounds that reduced *TFF3* expression by more than twofold were focused. All experiments utilized the A549 cell line, a standard *in vitro* model for airway remodeling derived from human lung carcinoma.

### *Preparation of 3D Structures*

Compounds' 3D structures were initiated by extracting PubChem IDs (CID) from the LINCS dataset API service.[17] These structures were then sourced from the PubChem database (https://pubchem.ncbi.nlm.nih.gov). Concurrently, the 3D structure of the TFF3 protein was retrieved from the Protein Data Bank (PDB ID: 1PE3) (https://www.rcsb.org).

### *Phytochemical Compound Analysis*

We sourced 2,845 phytochemical compounds from the PubChem database. Their druglike properties were assessed using the SwissADME service (http://www.swissadme.ch), which operates with the Simplified Molecular Input Line Entry System (SMILES) format.

### *Molecular Docking Analysis*

Molecular docking was employed to estimate binding affinity between ligands and the *TFF3* protein, specifically at the interface between the two monomers of its dimeric form (1PE3) structure. Preparation involved removing water molecules and the co-crystal ligand. Both *TFF3* and compound structures were converted to PDBQT format, with added Gasteiger partial charges for docking analyses. Docking simulations utilized AutoDock Vina on a high-performance computing system with a 192-core processor, enabling efficient and rapid processing of complex calculations.

### *Toxicity Assessment of Compounds*

The toxicity of the identified compounds was assessed Using Python (version 3.10), TensorFlow (version 2.21), and DeepChem (version 2.7.1). The Tox21 dataset guided the training, validation, and testing of the Graph Convolutional Network (GCN) architecture.

### *Dataset*

Tox21 encompasses 7,831 chemicals and 12 toxicological endpoints. These include five stress responses (SR-ARE, SR-ATAD5, SR-HSE, SR-MMP, SR-p53) and seven nuclear receptor signals (e.g., NR-AR, NR-ER). Chemicals, in SMILES format, have binary toxicity labels. To prevent overfitting, scaffold splitting clustered the compounds by molecular fingerprints. DeepChem's butina splitter method partitioned the dataset.

*Graph Convolutional Network*

While Convolutional Neural Networks (CNNs) excel with Euclidean data, they falter with non-Euclidean data such as chemical structures. Graph Convolutional Networks (GCNs), designed for non-Euclidean data, represent compounds as graphs: atoms as nodes and bonds as edges. Convolutional and pooling layers extract molecular patterns. DeepChem's GraphConv featurizer prepared chemical features. Training parameters included seven hidden layer units, a 0.4 dropout rate, a 0.0007 learning rate, and 100 epochs. Given the dataset imbalance, the Receiver Operating Characteristic-Area Under Curve (ROC-AUC) was the primary metric. Training used a Linux Operating System (OS) with Graphics Processing Unit (GPU) support and SMILES representations as input.

*Feature Importance in Tox21*

We measured the contribution of each Tox21 task to toxicity predictions using feature importance. Both random forest and permutation methods calculated scores. The AUC impact of each task was assessed by contrasting the importance of both methods.

*Comprehensive Bio-Evaluation through Integrated Analyses*

We combined data from gene expression, molecular docking, and toxicity assessments to pinpoint chemicals that i) reduce *TFF3* gene expression (Z-score<-2), ii) hinder TFF3 dimerization (ΔG<-7), and iii) exhibit low toxicity. While many chemicals affecting *TFF3* also influence other genes, certain ones show increased selectivity. Using the LINCS dataset's Signature Strength (SS), molecules with pronounced specificity to the *TFF3* gene were discerned. SS was gauged by the count of genes measurements with an absolute Z-score≥2.

*Molecular Dynamics Simulation*

To determine the optimal binding conformation of the compound with 1PE3, molecular dynamics simulations using GROMACS (version 5.1) on a Linux server were employed. Force field parameters were defined using the Charmm27 force field. SwissParam (https://www.swissparam.ch/) generated coordinate and topology files. The complex was solvated in a Transferable Intermolecular Potential 3-Point (TIP3P) water cubic box, ensuring a 1.0 nm distance from each edge, and neutralized with sodium chloride. Energy minimization and system equilibration were achieved using appropriate algorithms, Number of particles, Volume, and Temperature (NVT) and Number of particles, Pressure, and Temperature (NPT) ensembles, over 100 ps at 300 Kelvin and 1 atm. Simulations used the Particle Mesh Ewald (PME) method and LINear Constraint Solver algorithm constraints, running for 100 ns. Root Mean Square Deviation (RMSD) assessed protein atom stability. Ligand binding conformation was analyzed using Chimera, and binding energies were computed via the Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method.

*Computational Techniques Employed*

Our investigation integrated statistical and deep learning methodologies to explore TFF3 inhibition for treating obstructive airway diseases. Gene expression analysis utilized Z-scores from the LINCS database to identify compounds altering *TFF3* expression, providing a standardized comparison across experiments. Lipinski's "Rule of Five" assessed drug-like properties, guiding our selection of compounds with favorable pharmacokinetic profiles. Molecular docking, conducted with AutoDock Vina, employed empirical scoring functions to predict compound binding affinities to TFF3. In contrast, molecular dynamics simulations via GROMACS offered insights into the dynamic interactions within the protein-ligand complex based on physical principles. Toxicity evaluation leveraged a GCN, transcending traditional statistical models to predict compound toxicity through deep learning. This model's efficacy, including specificity, sensitivity, and precision, was detailed in our results, particularly in section 3.4, and visually summarized in figure 1.

## Results

*TFF3 Gene Expression Modulation by Chemicals*

From the LINCS database, 15,300 experimental conditions indicated *TFF3* downregulation. Of these, 6,001 matched compounds in the PubChem database fit our criteria. Focusing on the A549 cell line, 798 experiments were pertinent. These formed our analysis foundation. Table 1 lists the top 10 experiments.

The Z-score in table 1 quantifies gene expression deviation from a reference group mean:

$$Z = \frac{x - \mu}{\sigma}$$

Where x is the gene expression, μ is the reference mean, and σ is the standard deviation. The Z-score standardizes gene expression comparisons in LINCS, minimizing technical and batch effect variations.[18]

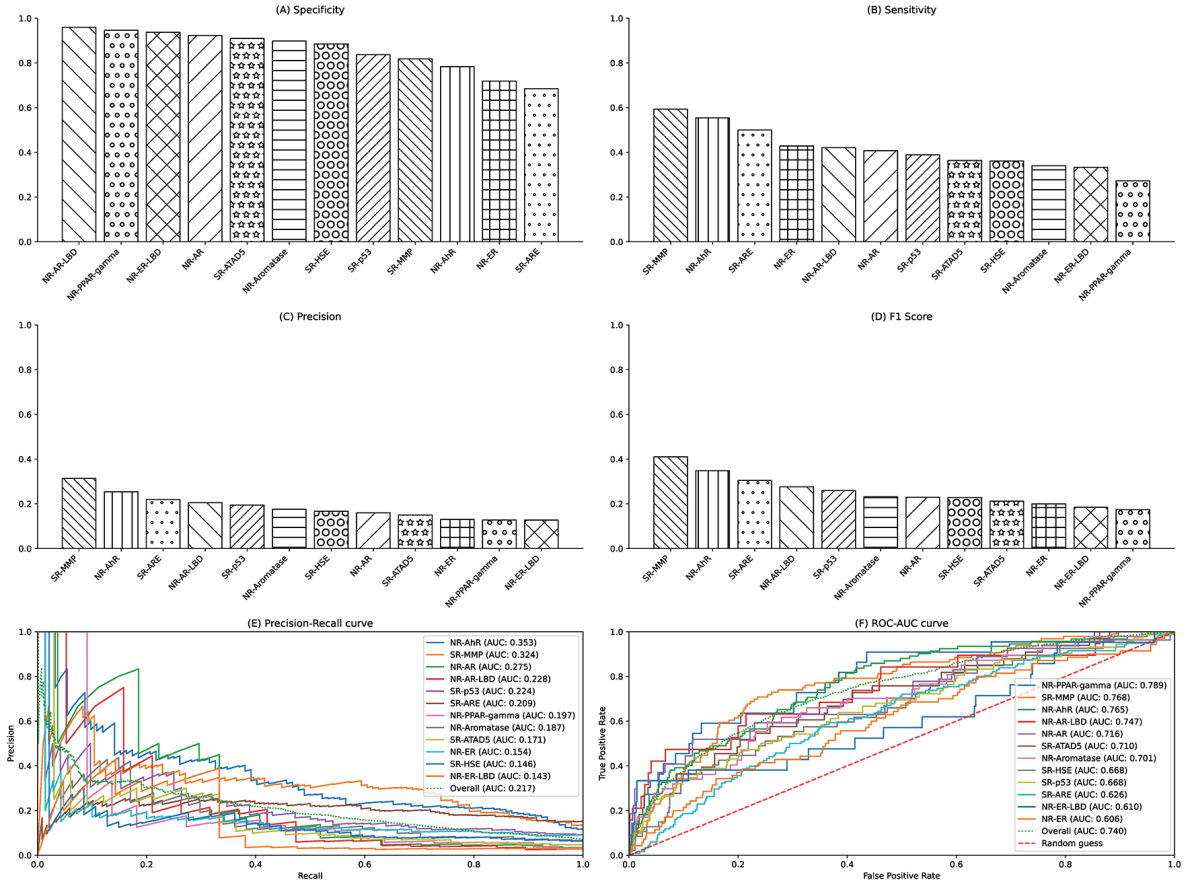


**Figure 1:** A comprehensive evaluation of model performance in toxicity prediction. Data were analyzed using key metrics and AUC analysis. (A) Specificity: True negative rate across different toxicity types. (B) Sensitivity: True positive rate detection of toxic cases. (C) Precision: Proportion of correct optimistic predictions. (D) F1 Score: Harmonic mean of precision and sensitivity. (E) PR-AUC: Precision-Recall Area Under Curve for 12 toxicological tasks, ranked by AUC values. (F) ROC-AUC: Receiver Operating Characteristic Area Under Curve for each task, illustrating the relationship between accurate positive and false favorable rates ranked by AUC values. The dotted lines indicate overall model performance across all functions.

**Table 1:** Leading chemicals down-regulating Trefoil Factor Family 3 in A549 cells

| No. | Signature ID | CID | Name | Dose | Time | Z-score |
|---|---|---|---|---|---|---|
| 1 | DOS041_ A549_24H:BRD-K36747900-001-01-8:5.01 | 54638598 | BRD-K36747900 | 5 | 24 | -3.9 |
| 2 | CPC016_A549_6H:BRD-K82731415-001-05-4:10 | 4592 | Olomoucine | 10 | 6 | -3.8 |
| 3 | PCLB002_A549_24H:BRD-A70731303:10 | 73707610 | Avrainvillamide-analog-5 | 10 | 24 | -3.8 |
| 4 | CPC010_A549_24H:BRD-K39462424-050-07-2:10 | 33036 | Dexchlorpheniramine | 10 | 24 | -2.9 |
| 5 | CPC010_A549_6H:BRD-A72180425-001-10-6:10 | 3689416 | K784-3188 | 10 | 6 | -2.9 |
| 6 | CPC015_A549_6H:BRD-A84174393-236-03-0:10 | 10291556 | Meloxicam | 10 | 6 | -2.8 |
| 7 | CPC016_A549_6H:BRD-K33396764-001-02-0:10 | 5280934 | Alpha-linolenic-acid | 10 | 6 | -2.8 |
| 8 | CPC015_A549_6H:BRD-A78391468-001-01-0:10 | 9847023 | Prednisolone-hemisuccinate | 10 | 6 | -2.7 |
| 9 | CPC010_A549_24H:BRD-A66435872-050-04-0:10 | 124846 | HTMT | 10 | 24 | -2.6 |
| 10 | DOS042_A549_24H:BRD-K49434056-001-01-0:4.94 | 54645999 | BRD-K49434056 | 5 | 24 | -2.6 |

The dose is presented in μM, and the time is indicated in hours (h).

### *Phytochemicals as Potential Drug Candidates*

Following Lipinski's rule via SWISSADME (http://www.swissadme.ch), 2,223 phytochemicals exhibited drug-like characteristics. Cross-referencing with PubChem IDs identified 101 as genuine phytochemicals. Notably, 13 from this subset were relevant to A549 cell line experiments.

### *Compound Interactions with TFF3 Dimer*

Docking studies targeting the TFF3 dimer interface assessed compound binding affinities. The top 10 compounds exhibited ΔG values between -8.6 and -9.4 (table 2). A negative ΔG suggests spontaneous, energetically favorable binding, with larger negative values indicating stronger affinities.

**Table 2:** Leading molecular docking scores for Trefoil Factor Family 3, expressed as ΔG values (Kcal/mol)

| No. | CID | Name | ΔG |
|---|---|---|---|
| 1 | 124846 | HTMT | -9.4 |
| 2 | 118221163 | BRD-K26510616 | -9 |
| 3 | 54654640 | SA-1472514 | -9 |
| 4 | 409805 | NSC-23766 | -8.8 |
| 5 | 44142121 | BRD-K11611839 | -8.8 |
| 6 | 73707542 | BG-1024 | -8.7 |
| 7 | 54654197 | BRD-K18511213 | -8.7 |
| 8 | 15301607 | VU-0415012 | -8.7 |
| 9 | 73707610 | Avrainvillamide-analog-5 | -8.6 |
| 10 | 7217941 | BRD-K55186349 | -8.6 |

*Model Performance*

Figure 1 encapsulates the model's performance on the Tox21 dataset, highlighting its specificity, sensitivity, precision, and F1 score. The dataset's imbalance, with a predominance of negative samples, is reflected in the lower sensitivity and accuracy for various toxicities. The model's precision-recall and ROC-AUC analyses are also presented. Figure 1E shows the precision-recall curve, and figure 1F displays the ROC-AUC curve for each toxicity alongside an overall test dataset ROC-AUC score of 74%.

*Feature Importance in Tox21 Dataset: Random Forest vs. Permutation*

Toxicological endpoint contributions were assessed using the importance of the random forest and permutation feature. Both methods highlighted Stress Response to Matrix Metalloproteinase (SR-MMP), Nuclear Receptor for Aryl Hydrocarbon Receptor (NR-AhR), and Stress Response to Antioxidant Response Element (SR-ARE) as key contributors. While there was agreement between the techniques, the permutation method, through value permutation, provided a more robust evaluation. Figure 2 visualizes the relative importance of both methods.

*Compound Toxicity Assessment*

The established model was employed to ascertain the toxicity profiles of compounds across 12 distinct toxicological endpoints. Table 3 showcases the top 10 compounds exhibiting the most favorable toxicity profiles.

*Integrative Analysis of Gene Expression, Molecular Docking, and Toxicity*

Our comprehensive assessment of compounds on *TFF3* gene expression,

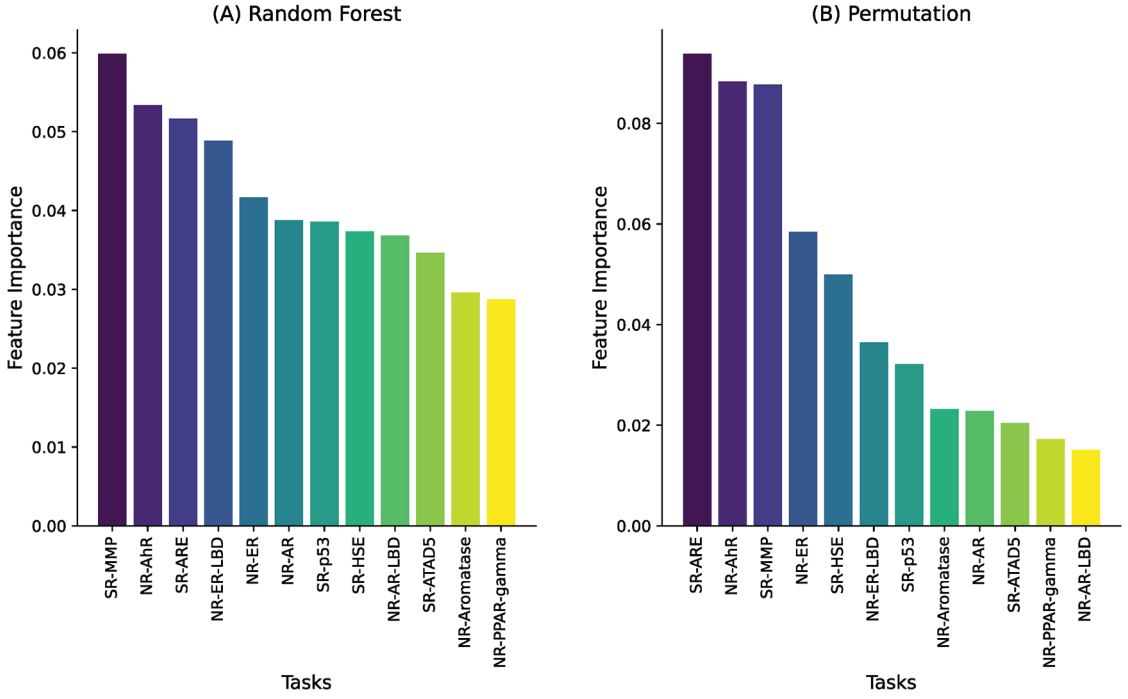


**Figure 2:** Feature importance assessment. This is a juxtaposition of feature importance derived from (A) Random Forest and (B) Permutation methods across multiple tasks. The most elevated bars denote the predominant features for each task.

**Table 3:** Top 10 compounds with favorable toxicity profiles

| No. | CID | Name | NR-AR | NR-AR-LBD | NR-AhR | NR-Aromatase | NR-ER | NR-ER-LBD | NR-PPAR-gamma | SR-ARE | SR-ATAD5 | SR-HSE | SR-MMP | SR-p53 | Mean |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 125519 | Aminogenistein | 0.05 | 0.15 | 0.00 | 0.15 | 0.07 | 0.01 | 0.40 | 0.02 | 0.01 | 0.33 | 0.01 | 0.13 | 0.11 |
| 2 | 230748 | INCA-6 | 0.44 | 0.07 | 0.14 | 0.05 | 0.06 | 0.24 | 0.10 | 0.01 | 0.26 | 0.06 | 0.02 | 0.04 | 0.12 |
| 3 | 68296 | 3H-1,2-Dithiole-3-thione | 0.59 | 0.03 | 0.00 | 0.30 | 0.49 | 0.05 | 0.01 | 0.02 | 0.01 | 0.00 | 0.00 | 0.04 | 0.13 |
| 4 | 2326992 | GNF-PF-254 | 0.36 | 0.07 | 0.02 | 0.05 | 0.15 | 0.11 | 0.24 | 0.08 | 0.37 | 0.05 | 0.11 | 0.05 | 0.14 |
| 5 | 5281707 | COUMESTROL | 0.04 | 0.02 | 0.01 | 0.42 | 0.01 | 0.01 | 0.72 | 0.00 | 0.00 | 0.58 | 0.00 | 0.02 | 0.15 |
| 6 | 5280443 | Apigenin | 0.10 | 0.44 | 0.04 | 0.50 | 0.10 | 0.04 | 0.48 | 0.00 | 0.03 | 0.10 | 0.00 | 0.01 | 0.16 |
| 7 | 5035 | Raloxifene | 0.78 | 0.08 | 0.02 | 0.03 | 0.00 | 0.01 | 0.21 | 0.02 | 0.03 | 0.73 | 0.01 | 0.01 | 0.16 |
| 8 | 222515 | Brazilin | 0.11 | 0.08 | 0.37 | 0.95 | 0.03 | 0.02 | 0.01 | 0.04 | 0.33 | 0.00 | 0.01 | 0.00 | 0.16 |
| 9 | 824226 | Ro 90-7501 | 0.53 | 0.13 | 0.00 | 0.32 | 0.06 | 0.48 | 0.11 | 0.00 | 0.00 | 0.44 | 0.01 | 0.02 | 0.18 |
| 10 | 5741425 | BRD-K30715099-001-01-2 | 0.58 | 0.47 | 0.12 | 0.08 | 0.19 | 0.14 | 0.31 | 0.03 | 0.11 | 0.15 | 0.01 | 0.07 | 0.19 |

The table enumerates compounds based on their mean toxicity scores derived from evaluations across 12 distinct toxicological endpoints. The mean represents the average score across all tasks for each chemical.

dimerization inhibition, and toxicity identified genistein as the prime candidate. As detailed in table 4, genistein demonstrated significant TFF3 downregulation (Z-score of -2.02), a favorable binding affinity (ΔG of -7.6) for TFF3 dimerization inhibition, and a low toxicity score (0.19). A graphical representation of the 2D structures of the final compounds from table 4 is depicted in figure 3A. Additionally, a 2D image of genistein and its associated toxicities across the 12 tasks is illustrated in figure 3B. Given its phytochemical origin, genistein's potential for reduced side effects underscores its promise for further exploration.

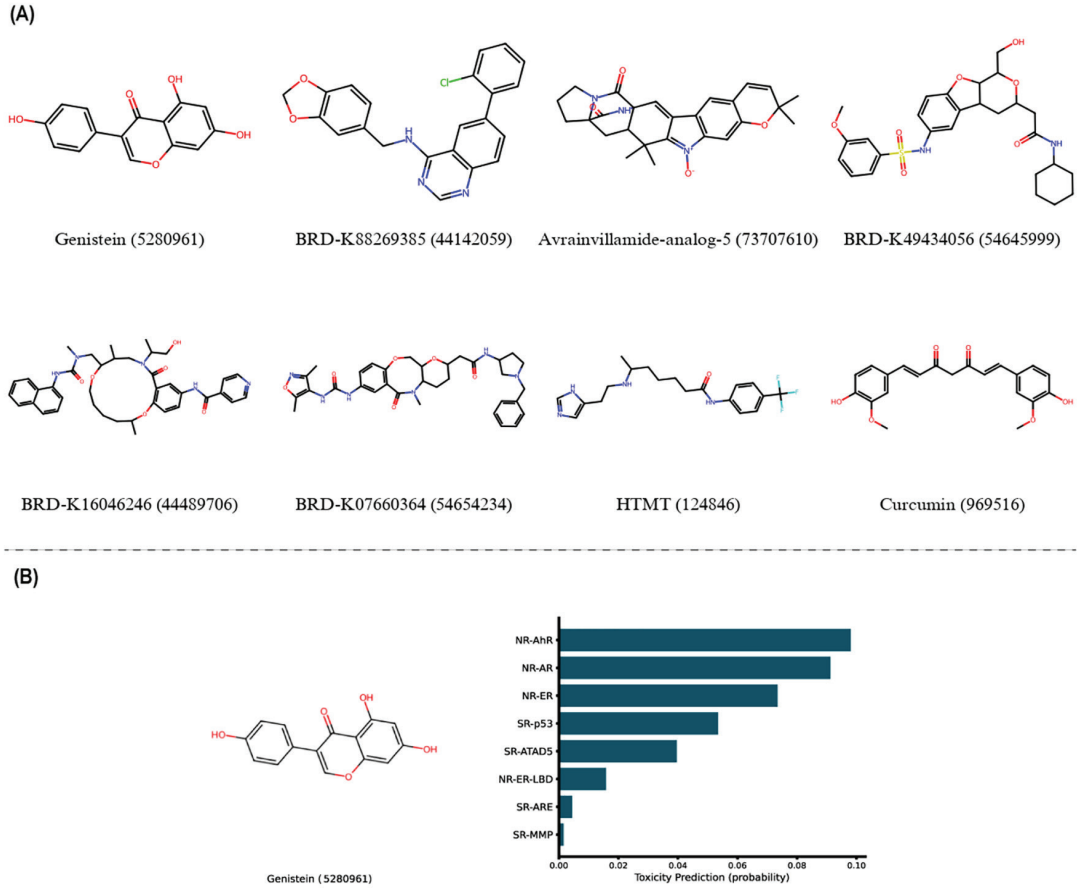


**Figure 3:** Structural representations and toxicity profile. (A) 2D structural representations of the final compounds listed in table 4, arranged in ascending order of their toxicity, with reactive atoms distinctly colored. (B) Bottom left: Graphical representation of genistein's molecular structure (CID: 5280961), with reactive atoms distinctly colored. (B) Bottom right: Bar chart illustrating the predicted toxicity levels of genistein across various biological tasks, providing insights into its safety profile.

**Table 4:** Assessment of compounds for Trefoil Factor Family 3 modulation and inhibition. Summary of compounds' effects on *TFF3* gene expression and inhibitory potential, tested on the A549 cell line over 6 or 24-hour durations at varied dosages

| No. | CID | Name | Dose | Time | Expression | ΔG | Toxicity | SS |
|---|---|---|---|---|---|---|---|---|
| 1 | 5280961 | Genistein | 0.04 | 24 | -2.02 | -7.6 | 0.20 | 410 |
| 2 | 44142059 | BRD-K88269385 | 10 | 6 | -2.17 | -7.2 | 0.58 | 538 |
| 3 | 73707610 | Avrainvillamide-analog-5 | 10 | 24 | -3.84 | -8.6 | 0.63 | 652 |
| 4 | 54645999 | BRD-K49434056 | 5 | 24 | -2.63 | -7.2 | 0.75 | NA |
| 5 | 44489706 | BRD-K16046246 | 10 | 6 | -2.14 | -8.1 | 0.78 | NA |
| 6 | 54654234 | BRD-K07660364 | 5 | 24 | -2.12 | -7 | 0.84 | NA |
| 7 | 969516 | Curcumin | 100 | 24 | -2.58 | -8.1 | 0.85 | NA |
| 8 | 124846 | HTMT | 10 | 24 | -2.69 | -9.4 | 0.97 | NA |

Toxicity, averaged from 12 scenarios, ranges from 0 (least toxic) to 1 (most toxic). Genistein, shown in bold fonts, is chosen for further study. "Signature Strength" (SS) indicates the specificity of a compound's gene expression impact. "NA" denotes broad gene effects. Gene expression is represented by Z-scores, with dosages in μM and time in hours.

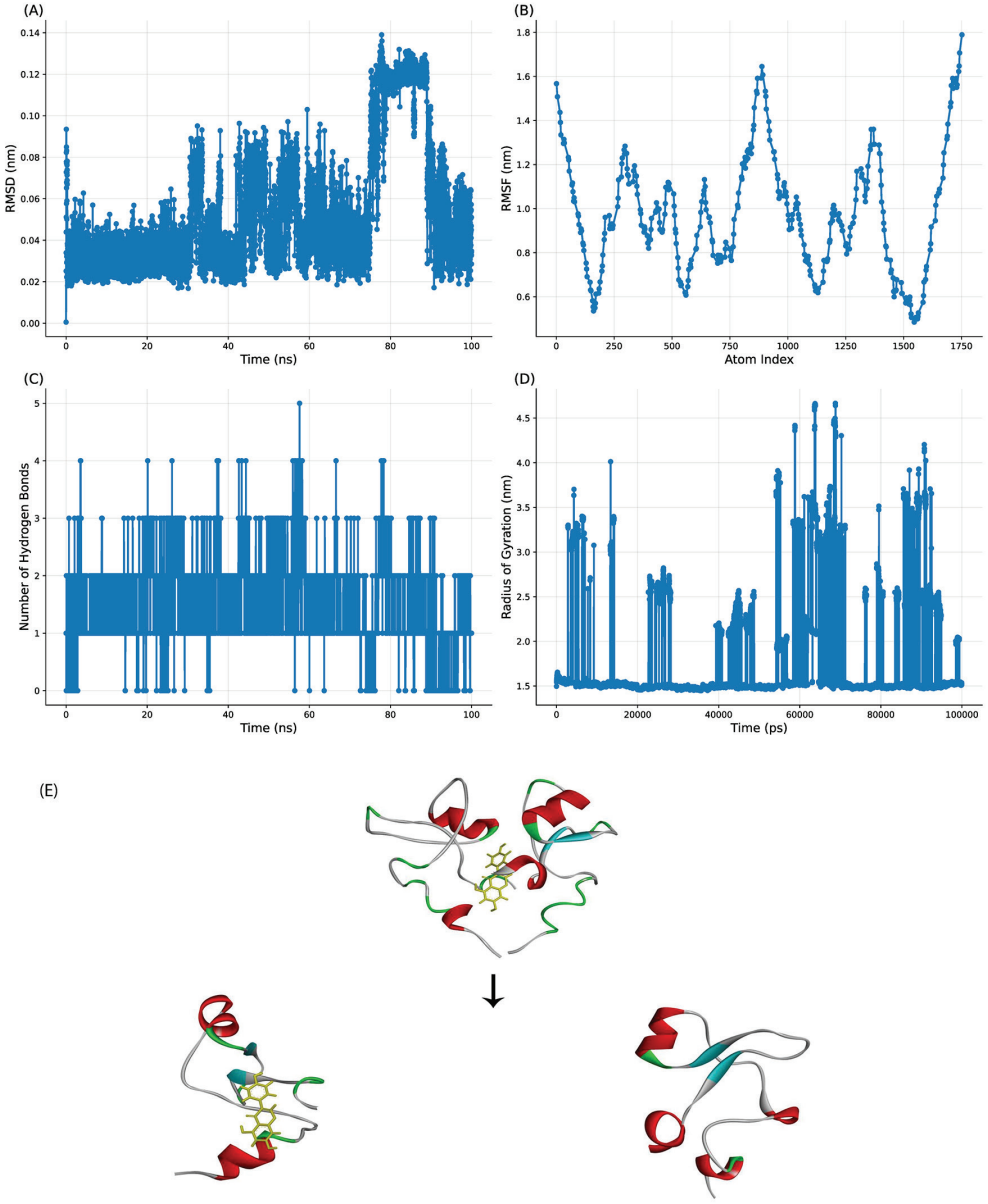


**Figure 4:** A comprehensive analysis of genistein-TFF3 complex dynamics and structural evolution. (A) Root Mean Square Deviation (RMSD) trajectory indicating the temporal stability of the genistein-TFF3 complex. (B) Root Mean Square Fluctuation (RMSF) of individual atoms, highlighting residue-specific flexibility and genistein's influence on TFF3 dimerization. (C) The temporal evolution of hydrogen bonds reflects the stability of genistein-TFF3 interactions throughout the simulation. (D) The radius of the Gyration trajectory suggests TFF3 monomer dissociation upon genistein binding. (E) 3D visualization at two key simulation frames: the initial frame shows the TFF3 dimer with genistein, and the concluding frame depicts the dissociation of TFF3 monomers, with genistein in yellow, emphasizing its role in dimeric separation.

*Molecular Dynamics Insights into Ligand-Receptor Interactions*

The binding energy components of the genistein-TFF3 complex, as derived from MM-PBSA calculations, are as follows: The overall binding energy is -19.62±3.32 kcal/mol, indicating a strong affinity between genistein and the TFF3 receptor (figure 4). This binding energy comprises electrostatic and van der Waals contributions, which are -18.05±6.14 kcal/mol and -34.26±2.55 kcal/mol, respectively. RMSF analysis (figure 4B) highlighted significant fluctuations indicative of TFF3 dimer dissociation upon genistein binding. This is further supported by the Gyration trajectory (figure 4D), where an increased gyration radius suggests TFF3 monomer separation. Figure 4E offers a 3D depiction of the TFF3 dimer at the simulation's start and end, with genistein distinctly marked in yellow.

## Discussion

Our study identified genistein as a potent inhibitor of TFF3, demonstrating significant down-regulation of TFF3 expression, effective inhibition of its function, and minimal toxicity. The pivotal role of TFF3 in the pathophysiology of airway remodeling, especially in obstructive airway diseases, has been previously highlighted.[6] This understanding set the stage for our comprehensive quest to pinpoint a compound that can effectively downregulate *TFF3* gene expression, inhibit its function, and simultaneously exhibit minimal toxicity.

Our initial approach was rooted in gene expression analysis. Utilizing the expansive LINCS database, we meticulously measured the impact of myriad compounds on *TFF3* gene expression. This preliminary analysis played a crucial role in shortlisting candidates for further evaluation. Building on this, we recognized the significance of TFF3's dimerization for its functional stability.[13, 19] The potential to disrupt TFF3's activity by targeting its dimerization interface became a focal point of our investigation. Our molecular docking studies, tailored specifically to TFF3's dimeric form, furnished critical insights into the binding affinities of the shortlisted compounds. This dual-faceted strategy led to identifying eight compounds that showcased pronounced downregulation of *TFF3* expression and also manifested favorable ΔG values, signaling strong binding affinities. To ensure the holistic evaluation of these compounds, it was imperative to assess their biological safety. We embarked on a rigorous toxicity evaluation, harnessing the cutting-edge GCN method within the DeepChem library. This assessment was pivotal in ensuring that the identified compounds, while effective, also adhered to safety benchmarks. Genistein emerged as a standout candidate among the evaluated compounds, ranking 13$^{th}$ in terms of low toxicity. Intriguingly, its derivative, amino-genistein, had the lowest toxicity. This naturally occurring phytochemical not only demonstrated prowess in downregulating *TFF3* expression and inhibiting its dimerization but also showcased minimal toxicity. The inherent nature of Genistein, being a phytochemical, augments its appeal, suggesting a potential for enhanced biocompatibility and a reduced spectrum of side effects compared to synthetic counterparts. Our molecular dynamics simulations further endorsed genistein's potential, revealing a stable genistein-TFF3 complex and suggesting genistein's capability to disrupt TFF3 dimerization.

Beyond our findings, genistein's broader pharmacological profile is noteworthy. Derived from soybeans, this isoflavone has been extensively researched for its myriad health benefits, ranging from its antioxidant and anti-inflammatory properties to its demonstrated antiproliferative effects on diverse cancer cell lines.[20-23] Additionally, genistein has been shown to have a potential role in preventing and treating osteoporosis by increasing bone mineral density.[24] Genistein has also shown potential benefits in treating cardiovascular disease, diabetes, and neurodegenerative disorders.[25-27] Furthermore, research suggests that genistein may have a positive impact in preventing obesity and enhancing insulin sensitivity.[28, 29] Additionally, genistein has been found to have potential therapeutic effects in the treatment of liver injury and inflammation.[30, 31] The exact mechanisms by which genistein exerts its biological effects are still under investigation, but it is thought to modulate the activity of various signaling pathways, including a tyrosine kinase, nuclear factor kappa B (NF-κB), nuclear factor erythroid 2-related factor 2 (Nrf2), and oxidative stress.[32-35] These findings suggest that genistein may also have potential therapeutic benefits in treating obstructive airway diseases such as ML, COPD, and asthma. In fact, studies have demonstrated that genistein can inhibit the activity of various inflammatory pathways associated with these diseases. For example, one study found that genistein was able to inhibit the NF-κB, tumor necrosis factor (TNF-α), and matrix metalloproteinase-9 (MMP9) associated pathways in lymphocytes from patients with COPD, leading to a reduction in inflammatory

markers.[36] Similarly, genistein was shown to reduce genotype-specific plasminogen activator inhibitor 1 (PAI-1) production in cultured human bronchial epithelial cells and mast cells. This was demonstrated by blocking increased promoter activities induced by transforming growth factor beta (TGFβ1) or immunoglobulin E (IgE) stimulation in cells transfected with the PAI-1 4G vector, suggesting genistein's potential as a treatment option for high PAI-1-related diseases such as asthma.[31, 37-39] A nationwide cohort study on over one million Korean infants shows that soy formula intake in newborns is significantly associated with a reduced risk of allergic asthma but not all-cause or non-allergic asthma.[38] Additionally, in a murine model, Bio-300 containing genistein significantly reduced inflammatory cell infiltration, goblet cell hyperplasia, collagen deposition, and airway hyperresponsiveness induced by ovalbumin.[38] Therefore, by inhibiting TFF3, the primary regulator in airway remodeling, genistein may have potential therapeutic benefits in treating obstructive airway diseases such as COPD, asthma, and ML. While our study identifies genistein as a promising inhibitor of TFF3 with potential therapeutic benefits for obstructive airway diseases, several limitations must be acknowledged. First, our findings are based on computational analyses and *in vitro* experiments using the A549 cell line, which may not fully replicate the *in vivo* environment. Second, the long-term effects and bioavailability of genistein in clinical settings remain to be investigated. Finally, further studies are needed to validate these results in animal models and clinical trials to confirm the efficacy and safety of genistein as a therapeutic agent for conditions such as COPD, ML, and asthma.

## Conclusion

In conclusion, our multidimensional investigation, encompassing gene expression analysis, molecular docking, toxicity assessment, and molecular dynamics simulations, illuminated the potential of genistein as a promising therapeutic candidate for obstructive airway diseases. By targeting the pivotal TFF3 molecule, genistein showcases the ability to downregulate its expression and inhibit its dimerization, an essential aspect of its functionality.

## Acknowledgement

We would like to express our gratitude to the Chemical Injuries Research Center, Molecular Biology Research Center, and Systems Biology and Poisonings Institute at Baqiyatallah University of Medical Sciences for their valuable assistance and contributions to this research.

## Authors' Contribution

A.S: Led project administration and workflow design; study conception, drafting, and reviewing the manuscript; M.S: Developed software and methodology; data acquisition and analysis, drafting, and reviewing; M.A: Study concept; drafting and reviewing the manuscript; E.B.: Study concept; data interpretation, manuscript drafting, and reviewing; A.N.: Study concept; drafting and reviewing. All authors read and approved the final manuscript and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved.

**Conflict of Interest:** None declared.

## References

1 Liu P, Wang Y, Chen C, Liu H, Ye J, Zhang X, et al. Research trends on airway remodeling: A bibliometrics analysis. Heliyon. 2024;10:e24824. doi: 10.1016/j.heliyon.2024.e24824. PubMed PMID: 38333835; PubMed Central PMCID: PMCPMC10850909.

2 Shahriary A, Seyedzadeh MH, Ahmadi A, Salimian J. The footprint of TGF-beta in airway remodeling of the mustard lung. Inhal Toxicol. 2015;27:745-53. doi: 10.3109/08958378.2015.1116645. PubMed PMID: 26606948.

3 Fang L, Sun Q, Roth M. Immunologic and Non-Immunologic Mechanisms Leading to Airway Remodeling in Asthma. Int J Mol Sci. 2020;21. doi: 10.3390/ijms21030757. PubMed PMID: 31979396; PubMed Central PMCID: PMCPMC7037330.

4 Wang Y, Xu J, Meng Y, Adcock IM, Yao X. Role of inflammatory cells in airway remodeling in COPD. Int J Chron Obstruct Pulmon Dis. 2018;13:3341-8. doi: 10.2147/COPD.S176122. PubMed PMID: 30349237; PubMed Central PMCID: PMCPMC6190811.

5 Marzouni ET, Dorcheh SP, Nejad-Moghaddam A, Ghanei M, Goodarzi H, Hosseini SE, et al. Adipose-derived mesenchymal stem cells ameliorate lung epithelial injury through mitigating of oxidative stress in mustard lung. Regen Med. 2020. doi: 10.2217/rme-2020-0051. PubMed PMID: 32935623.

6 Najafi A, Masoudi-Nejad A, Ghanei M, Nourani MR, Moeini A. Pathway reconstruction

of airway remodeling in chronic lung diseases: a systems biology approach. PLoS One. 2014;9:e100094. doi: 10.1371/journal.pone.0100094. PubMed PMID: 24978043; PubMed Central PMCID: PMCPMC4076832.
7 Weste J, Houben T, Harder S, Schluter H, Lucke E, Schreiber J, et al. Different Molecular Forms of TFF3 in the Human Respiratory Tract: Heterodimerization with IgG Fc Binding Protein (FCGBP) and Proteolytic Cleavage in Bronchial Secretions. Int J Mol Sci. 2022;23. doi: 10.3390/ijms232315359. PubMed PMID: 36499686; PubMed Central PMCID: PMCPMC9737082.
8 Viby NE, Pedersen L, Lund TK, Kissow H, Backer V, Nexo E, et al. Trefoil factor peptides in serum and sputum from subjects with asthma and COPD. Clin Respir J. 2015;9:322-9. doi: 10.1111/crj.12146. PubMed PMID: 24720774.
9 Mihalj M, Bujak M, Butkovic J, Zubcic Z, Tolusic Levak M, Ces J, et al. Differential Expression of TFF1 and TFF3 in Patients Suffering from Chronic Rhinosinusitis with Nasal Polyposis. Int J Mol Sci. 2019;20. doi: 10.3390/ijms20215461. PubMed PMID: 31683988; PubMed Central PMCID: PMCPMC6862153.
10 Doubkova M, Karpisek M, Mazoch J, Skrickova J, Doubek M. Prognostic significance of surfactant protein A, surfactant protein D, Clara cell protein 16, S100 protein, trefoil factor 3, and prostatic secretory protein 94 in idiopathic pulmonary fibrosis, sarcoidosis, and chronic pulmonary obstructive disease. Sarcoidosis Vasc Diffuse Lung Dis. 2016;33:224-34. PubMed PMID: 27758987.
11 Yang Y, Lin Z, Lin Q, Bei W, Guo J. Pathological and therapeutic roles of bioactive peptide trefoil factor 3 in diverse diseases: recent progress and perspective. Cell Death Dis. 2022;13:62. doi: 10.1038/s41419-022-04504-6. PubMed PMID: 35039476; PubMed Central PMCID: PMCPMC8763889.
12 Zhu Y, Zhao S, Deng Y, Gordillo R, Ghaben AL, Shao M, et al. Hepatic GALE Regulates Whole-Body Glucose Homeostasis by Modulating Tff3 Expression. Diabetes. 2017;66:2789-99. doi: 10.2337/db17-0323. PubMed PMID: 28877911; PubMed Central PMCID: PMCPMC5652600.
13 Jarva MA, Lingford JP, John A, Soler NM, Scott NE, Goddard-Borger ED. Trefoil factors share a lectin activity that defines their role in mucus. Nat Commun. 2020;11:2265. doi: 10.1038/s41467-020-16223-7. PubMed PMID: 32404934; PubMed Central PMCID: PMCPMC7221086.
14 Braga Emidio N, Brierley SM, Schroeder CI, Muttenthaler M. Structure, Function, and Therapeutic Potential of the Trefoil Factor Family in the Gastrointestinal Tract. ACS Pharmacol Transl Sci. 2020;3:583-97. doi: 10.1021/acsptsci.0c00023. PubMed PMID: 32832864; PubMed Central PMCID: PMCPMC7432662.
15 Hoffmann W. Trefoil factor family (TFF) peptides. Encyclopedia. 2021;1:974-87. doi: 10.3390/encyclopedia1030074.
16 Keenan AB, Jenkins SL, Jagodnik KM, Koplev S, He E, Torre D, et al. The Library of Integrated Network-Based Cellular Signatures NIH Program: System-Level Cataloging of Human Cells Response to Perturbations. Cell Syst. 2018;6:13-24. doi: 10.1016/j.cels.2017.11.001. PubMed PMID: 29199020; PubMed Central PMCID: PMCPMC5799026.
17 Pilarczyk M, Fazel-Najafabadi M, Kouril M, Shamsaei B, Vasiliauskas J, Niu W, et al. Connecting omics signatures and revealing biological mechanisms with iLINCS. Nat Commun. 2022;13:4678. doi: 10.1038/s41467-022-32205-3. PubMed PMID: 35945222; PubMed Central PMCID: PMCPMC9362980.
18 Colan SD. The why and how of Z scores. J Am Soc Echocardiogr. 2013;26:38-40. doi: 10.1016/j.echo.2012.11.005. PubMed PMID: 23261367.
19 Bastholm SK, Samson MH, Becher N, Hansen LK, Stubbe PR, Chronakis IS, et al. Trefoil factor peptide 3 is positively correlated with the viscoelastic properties of the cervical mucus plug. Acta Obstet Gynecol Scand. 2017;96:47-52. doi: 10.1111/aogs.13038. PubMed PMID: 27731893.
20 Sohel M, Biswas P, Al Amin M, Hossain MA, Sultana H, Dey D, et al. Genistein, a potential phytochemical against breast cancer treatment-insight into the molecular mechanisms. Processes. 2022;10:415.
21 Yu X, Yan J, Li Y, Cheng J, Zheng L, Fu T, et al. Inhibition of castration-resistant prostate cancer growth by genistein through suppression of AKR1C3. Food Nutr Res. 2023;67. doi: 10.29219/fnr.v67.9024. PubMed PMID: 36794010; PubMed Central PMCID: PMCPMC9899042.
22 Alorda-Clara M, Torrens-Mas M, Morla-Barcelo PM, Roca P, Sastre-Serra J, Pons DG, et al. High Concentrations of Genistein Decrease Cell Viability Depending on Oxidative Stress and Inflammation in Colon Cancer Cell Lines. Int J Mol Sci. 2022;23. doi: 10.3390/ijms23147526. PubMed PMID: 35886874; PubMed Central PMCID: PMCPMC9323408.

23 Zhao L, Wang Y, Liu J, Wang K, Guo X, Ji B, et al. Protective Effects of Genistein and Puerarin against Chronic Alcohol-Induced Liver Injury in Mice via Antioxidant, Anti-inflammatory, and Anti-apoptotic Mechanisms. J Agric Food Chem. 2016;64:7291-7. doi: 10.1021/acs.jafc.6b02907. PubMed PMID: 27609057.
24 Kushwaha AC, Mohanbhai SJ, Sardoiwala MN, Jaganathan M, Karmakar S, Roy Choudhury S. Nanoemulsified Genistein and Vitamin D Mediated Epigenetic Regulation to Inhibit Osteoporosis. ACS Biomater Sci Eng. 2022;8:3810-8. doi: 10.1021/acsbiomaterials.2c00165. PubMed PMID: 36005299.
25 Li R, Robinson M, Ding X, Geetha T, Al-Nakkash L, Broderick TL, et al. Genistein: A focus on several neurodegenerative diseases. J Food Biochem. 2022;46:e14155. doi: 10.1111/jfbc.14155. PubMed PMID: 35460092.
26 Li R, Ding X, Geetha T, Fadamiro M, St Aubin CR, Shim M, et al. Effects of Genistein and Exercise Training on Brain Damage Induced by a High-Fat High-Sucrose Diet in Female C57BL/6 Mice. Oxid Med Cell Longev. 2022;2022:1560435. doi: 10.1155/2022/1560435. PubMed PMID: 35620577; PubMed Central PMCID: PMCPMC9129997.
27 Rasheed S, Rehman K, Shahid M, Suhail S, Akash MSH. Therapeutic potentials of genistein: New insights and perspectives. J Food Biochem. 2022;46:e14228. doi: 10.1111/jfbc.14228. PubMed PMID: 35579327.
28 Gan M, Chen X, Chen Z, Chen L, Zhang S, Zhao Y, et al. Genistein Alleviates High-Fat Diet-Induced Obesity by Inhibiting the Process of Gluconeogenesis in Mice. Nutrients. 2022;14. doi: 10.3390/nu14081551. PubMed PMID: 35458112; PubMed Central PMCID: PMCPMC9032493.
29 Nasimi Doost Azgomi R, Moini Jazani A, Karimi A, Pourreza S. Potential roles of genistein in polycystic ovary syndrome: A comprehensive systematic review. Eur J Pharmacol. 2022;933:175275. doi: 10.1016/j.ejphar.2022.175275. PubMed PMID: 36108737.
30 Ding Q, Pi A, Hao L, Xu T, Zhu Q, Shu L, et al. Genistein Protects against Acetaldehyde-Induced Oxidative Stress and Hepatocyte Injury in Chronic Alcohol-Fed Mice. J Agric Food Chem. 2023;71:1930-43. doi: 10.1021/acs.jafc.2c05747. PubMed PMID: 36653166.
31 Goh YX, Jalil J, Lam KW, Husain K, Premakumar CM. Genistein: A Review on its Anti-Inflammatory Properties. Front Pharmacol. 2022;13:820969. doi: 10.3389/fphar.2022.820969. PubMed PMID: 35140617; PubMed Central PMCID: PMCPMC8818956.
32 Li Y, Zhang JJ, Chen RJ, Chen L, Chen S, Yang XF, et al. Genistein mitigates oxidative stress and inflammation by regulating Nrf2/HO-1 and NF-kappaB signaling pathways in hypoxic-ischemic brain damage in neonatal mice. Ann Transl Med. 2022;10:32. doi: 10.21037/atm-21-4958. PubMed PMID: 35282070; PubMed Central PMCID: PMCPMC8848430.
33 Jafari A, Esmaeilzadeh Z, Khezri MR, Ghasemnejad-Berenji H, Pashapour S, Sadeghpour S, et al. An overview of possible pivotal mechanisms of Genistein as a potential phytochemical against SARS-CoV-2 infection: A hypothesis. J Food Biochem. 2022;46:e14345. doi: 10.1111/jfbc.14345. PubMed PMID: 35866873; PubMed Central PMCID: PMCPMC9350103.
34 Li Y, Ou S, Liu Q, Gan L, Zhang L, Wang Y, et al. Genistein improves mitochondrial function and inflammatory in rats with diabetic nephropathy via inhibiting MAPK/NF-kappaB pathway. Acta Cir Bras. 2022;37:e370601. doi: 10.1590/acb370601. PubMed PMID: 35976278; PubMed Central PMCID: PMCPMC9377651.
35 Shirvanian K, Vali R, Farkhondeh T, Abderam A, Aschner M, Samarghandian S. Genistein Effects on Various Human Disorders Mediated via Nrf2 Signaling. Curr Mol Med. 2024;24:40-50. doi: 10.2174/1566524023666221128162753. PubMed PMID: 36443970.
36 Liu XJ, Bao HR, Zeng XL, Wei JM. Effects of resveratrol and genistein on nuclear factorkappaB, tumor necrosis factoralpha and matrix metalloproteinase9 in patients with chronic obstructive pulmonary disease. Mol Med Rep. 2016;13:4266-72. doi: 10.3892/mmr.2016.5057. PubMed PMID: 27035424; PubMed Central PMCID: PMCPMC4838123.
37 Han JY, Jo A, Fukahori S, Tabor M, Kumar R, Yon DK, et al. Genotype-specific PAI-1 Regulation in Human Epithelial Cells and Mast Cells by Genistein. Journal of Allergy and Clinical Immunology. 2024;153:AB141. doi: 10.1016/j.jaci.2023.11.466.
38 Cho S, Oh J, Han JY, Kim M, Cho J, Fukahori S, et al. Impact of soy isoflavones (genistein) on allergic asthma in epidemiologic and mouse models: national birth cohort in South Korea and murine model. Journal of Allergy and Clinical Immunology. 2024;153:AB20. doi: 10.1016/j.jaci.2023.11.085.

39 Nazari-Khanamiri F, Ghasemnejad-Berenji M. Cellular and molecular mechanisms of genistein in prevention and treatment of diseases: An overview. J Food Biochem. 2021;45:e13972. doi: 10.1111/jfbc.13972. PubMed PMID: 34664285.