ACCSCIENCE PUBLISHING

Tumor Discovery

ORIGINAL RESEARCH ARTICLE

Bioinformatics approach to identifying potential cancer-associated mutations in *CCL2*

Shah Kamal[1†], Najeeb Ullah[1†], Amanullah Amanullah[1], Mariam Ahmed Mujtaba[2], Kashif Ali Khan[3], Cheng Deng[1*], Shanshan Lai[1*], and Mohammad Amjad Kamal[4,5,6,7,8*]

[1]Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing, China

[2]Department of Biotechnology, Women University Mardan, East Canal Road Mardan, 23200, Pakistan

[3]Department of Pharmacy, Shaheed Benazir Bhutto University Sheringal, Dir Upper, Pakistan

[4]Institutes for Systems Genetics, Frontiers Science Center for Disease-related Molecular Network, West China School of Nursing, West China Hospital, Sichuan University, Chengdu, China

[5]King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

[6]Department of Pharmacy, Faculty of Allied Health Sciences, Daffodil International University, Dhaka, Bangladesh

Abstract

Chemokine C-C motif ligand 2 (CCL2), also known as monocyte chemoattractant protein 1 or small inducible cytokine A2, is a cytokine from the CC chemokine family. It plays a crucial role in recruiting monocytes, memory T-cells, and dendritic cells to inflammatory sites resulting from tissue injury or infection. C-C motif chemokine receptor 2 (CCR2) is a chemokine receptor that may influence lymphocyte function. The interaction between CCL2 and CCR2 is essential for inflammatory responses and cancer regulation, as it attracts monocytes and macrophages to tumor sites, facilitating tumor growth and metastasis. Given the importance of CCL2 in regulating cell trafficking and cancer progression, we employed a bioinformatics approach to examine the effects of oncogenic missense mutations in *CCL2* on CCR2 activation. We used precise computational methods to identify the molecular characteristics responsible for the altered activity and interactions, thereby enhancing our understanding of the molecular mechanisms underlying disease progression. We generated a three-dimensional model of the CCL2 protein with the identified mutations using the I-TASSER algorithm. The effects of these mutations on the protein's stability and functional properties were evaluated using various prediction tools, and molecular dynamics simulations were conducted using WebGro software. Our analysis of 83 *CCL2* missense mutations identified 10 disease-causing mutations, including C59G, which was directly linked to cancer. The C59G mutation increases the binding interaction between CCL2 and CCR2. The C59G position was determined to be highly conserved, and substitutions of cysteine (C) 59 with glycine (G) altered CCL2 activity. Our results suggest that this mutation alters the CCL2–CCR2 binding activity, lowering the risk of developing cancer and defending against pathogen invasion during the neutrophil-mediated innate immune response.

Keywords: CCL2; Molecular dynamic simulation; Colorectal cancer; Point mutation

†These authors contributed equally to this work.

***Corresponding authors:**
Cheng Deng
(dengcheng2014@126.com)
Mohammad Amjad Kamal
(ma.kamal@wchscu.cn)
Shanshan Lai
(lss7259@163.com)

Citation: Kamal S, Ullah N, Amanullah A, *et al*. Bioinformatics approach to identifying potential cancer-associated mutations in *CCL2*. *Tumor Discov*. 2024;3(4):3891.
doi: 10.36922/td.3891

Received: June 7, 2024

Accepted: September 5, 2024

Published Online: October 21, 2024

Copyright: © 2024 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[7]Centre for Global Health Research, Saveetha Medical College and Hospital, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India
[8]Enzymoics, 7 Peterlee Place, Hebersham, Novel Global Community Educational Foundation, NSW, Australia
(This article belongs to the *Special Issue: Colorectal Cancer: Best Tools for Diagnosis to Management Strategies*)

1. Introduction

Chemokine C-C motif ligand 2 (CCL2), also known as monocyte chemoattractant protein 1 (MCP-1) or small inducible cytokine A2, is an important cytokine in the CC chemokine family.[1] CCL2 adopts a β-trefoil fold, which is a common protein fold that provides plasticity for recognizing various targets.[2] Monomeric CCL2 binds to its receptor, C-C motif chemokine receptor 2 (CCR2), through a single, topologically novel binding site.[3] CCL2 penetrates the extracellular portion of the CCR2 transmembrane domain and interacts with the receptor through its N-terminal glutamine. There are numerous hydrophobic and polar interactions between CCL2 and the Gα-protein of the CCR2 receptor complex, which contribute to the receptor's constitutive activity.[4]

CCL2 plays a crucial role in regulating cellular mechanics by attracting monocytes, memory T-cells, and dendritic cells to inflammation sites resulting from tissue injury or infection.[5] This small cytokine is primarily secreted by monocytes, macrophages, and dendritic cells and is involved in various physiological processes, including immune cell recruitment, antitumor activity, and granuloma formation.[6] CCL2 interacts with cell surface receptors, such as CCR2 and CCR4, and exhibits chemotactic activity for monocytes and basophils.[7] Its effects extend to diseases characterized by monocytic infiltrates, such as psoriasis, rheumatoid arthritis, atherosclerosis, and certain cancers. In cancer, CCL2 has been demonstrated to promote cell proliferation, migration, and metastasis, underscoring its potential as a therapeutic target for cancer treatment.[8,9]

Chemokines can induce the release of stored promalignant chemokines from breast tumor cells, highlighting the substantial tumor-promoting role of chemokine coexpression in breast cancer progression.[10-12] Epithelial ovarian cancer cells produce MCP-1, which is likely the major source of this chemokine. This production may contribute to the accumulation of tumor-associated macrophages, which can subsequently influence tumor behavior.[13,14] Cancer cells and macrophages interact to promote tumor angiogenesis, a crucial process in cancer progression. Specifically, MCP-1 has been linked to the progression of human esophageal carcinoma by facilitating angiogenesis, enhancing macrophage recruitment, and promoting the development of new blood vessels.[15,16] Moreover, MCP-1 produced by human gastric carcinoma cells plays a crucial role in angiogenesis. It attracts and activates macrophages, which then aid in forming new blood vessels and promote tumor growth.[17] MCP-1 has a limited role in eliciting an immune response in renal cell carcinomas, as assessed by the parameters analyzed in this study.[18] The simultaneous induction of tumor necrosis factor-alpha (TNF-α), urokinase-type plasminogen activator, interleukin (IL)-8, and MCP-1 can enhance the interaction of tumor and inflammatory cells, thereby increasing cytotoxicity and promoting tumor suppression.[19] MCP-1 can potentiate the effects of bacterial endotoxins by activating macrophages to become tumoricidal, potentially suppressing metastasis.[20] The chemokine MCP-1 plays an important role in promoting tumor progression and angiogenesis across various cancer types, including breast, ovarian, esophageal, and gastric carcinomas. MCP-1 production by tumor cells contributes to the recruitment and activation of tumor-associated macrophages, supporting cancer growth and metastasis.

Computational analysis has become increasingly vital for identifying harmful mutations and predicting the effects of missense mutations on protein structures and functions. This type of analysis is now a key tool in designing new drugs and therapies for genetic diseases.[21] In this study, we evaluated the potential significance of mutations in *CCL2* that can impact its protein functions and ligand interactions. Initially, we used various computational algorithms, including meta-single nucleotide polymorphism (SNP),[22] PolyPhen-2,[23] and Pmut,[24] to categorize harmful mutations. We further employed the Functional Analysis through Hidden Markov Models (FATHMM) server to identify specific genetic variants associated with cancer development or initiation among disease-causing mutations. This approach allowed us to distinguish between cancer-promoting mutations and those related to other disease phenotypes, providing valuable insights into the specific role of certain variants in cancer pathogenesis.[25] In addition to the sequence-based analysis, we employed structure-based computational tools, such as mCSM,[26] SDM,[27] and DUET,[28] to predict the impact of the identified mutations on the structure and stability of the corresponding proteins.

One missense mutation in *CCL2* was selected based on its cancer association for further investigation. The

I-TASSER server was used to generate a structural model incorporating the specified missense mutation in *CCL2*.[29] The model underwent molecular dynamics (MD) simulations and three-dimensional (3D) modeling through WebGro to evaluate how the mutation affected the stability and functionality of the CCL2 protein.[30] WebGro software was then used to perform MD simulations on the mutant CCL2 model, enabling the assessment of its stability and potential functional consequences compared to the native protein. A thorough analysis of the impact of the missense mutation on the protein structure and function was facilitated by this integrated computational approach that combines structure prediction with MD[31] (Figure 1).

Molecular docking studies offer valuable insights into the binding dynamics and interactions between CCL2 and CCR2, highlighting the potential for therapeutic interventions in diseases mediated by this signaling pathway. Molecular docking analyses have identified the critical residues in CCL2, such as T16, N17, and R18, which are essential for forming stable complexes with CCR2. These interactions involve both hydrophobic interactions and hydrogen bonds, which collectively contribute to the binding affinity of CCL2 (ligand) with CCR2 (receptor).[3] For instance, the binding of the compound baicalin (BA) to CCL2 is facilitated by substantial hydrophobic interactions with residues such as I20 and V51 and hydrogen bonds with residues such as T16 and S17. The binding energy calculated for CCL2 interactions with CCR2, that is, approximately −6.3 kcal/mol for the interaction between BA and human CCL2, indicates the formation of a favorable and stable ligand–receptor complex.[3] Understanding these molecular interactions through docking studies provides valuable insights into drug design. By targeting the CCL2–CCR2 interaction, researchers can develop antagonists that disrupt this pathway. These antagonists hold promise as therapeutic agents for the treatment of inflammatory and autoimmune diseases as well as certain cancers where CCL2/CCR2 signaling plays an important role.

CCL2 mutations can substantially alter the protein's structure and function and potentially affect its binding affinity to the CCR2 receptor. This may impact the recruitment and activation of tumor-associated macrophages, thereby influencing angiogenesis and tumor progression. This mutation could also modulate the immune response to CCL2, potentially influencing antitumor immunity and macrophage tumoricidal activity. Overall, *CCL2* mutations could have profound implications for the tumorigenic potential of this chemokine in cancer development. However, further research is required to elucidate the specific genetic and epigenetic factors that influence the expression and activity of *CCL2* mutations.

2. Material and methods

2.1. Collection of datasets

In this study, we collected data on *CCL2* missense mutations from the Ensembl database (release 109) (https://www.

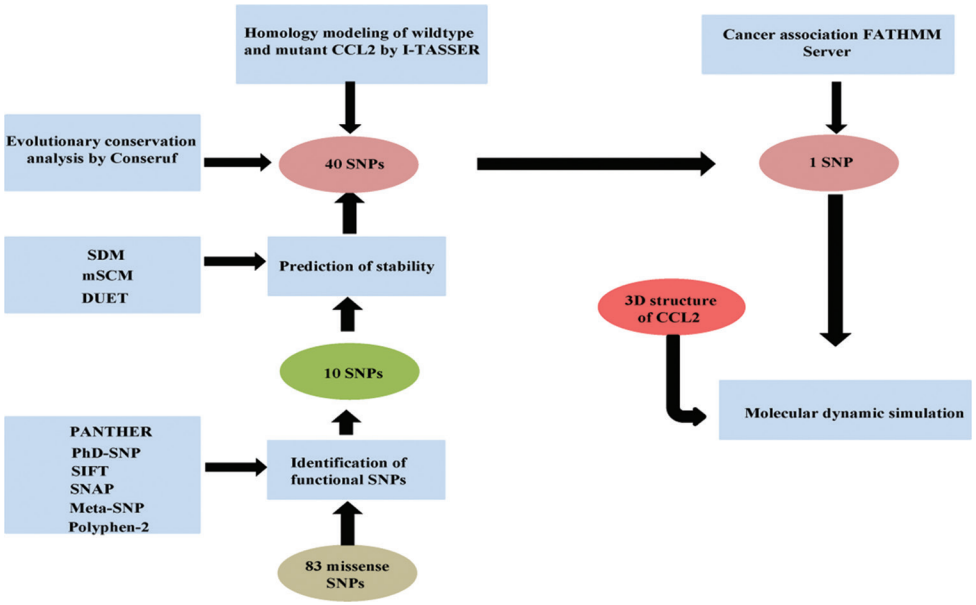


Figure 1. Flowchart illustrating the stepwise analysis of missense SNPs in human CCL2, employing computational tools to investigate 83 missense SNPs. Initially, six tools identified 10 SNPs with functional effects, including 40 highly destabilizing SNPs. One mutation, C59G, was predicted as the most high-risk missense SNP and subjected to molecular dynamics simulations.
Abbreviations: SNP: Single-nucleotide polymorphism; CCL2: Chemokine C-C motif ligand 2.

ensembl.org), which provides comprehensive genomic information (Ensembl ID: ENSG00000108691).[32] This dataset was chosen due to its extensive collection of genomic variations. In addition, data were obtained from the UniProtKB database (https://www.uniprot.org), which includes sequence details in FASTA format (UniProt ID: P13500), and the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov).[33,34] Structural information regarding CCL2 was retrieved from the Protein Data Bank (PDB) (CCL2 PDB ID: 2LIE, and CCR2 PDB ID: 2LWL) through the website http://www.rcsb.org/.[35]

2.2. Computational prediction of mutation pathogenicity

To assess the effects of mutations on CCL2 proteins, we used computational methods essential for distinguishing between single-nucleotide variations that are detrimental, benign, or cancerous. Accurate evaluation of the functional impacts of these genetic alterations is essential for identifying disease-causing mutations. Meta-SNP (http://snps.biofold.org/meta-snp/) is a web-based tool that uses multiple predictive algorithms to assess the potential pathogenicity of protein mutations.[22] We employed meta-SNP to assess mutation pathogenicity by integrating the predictions from algorithms such as SNAP,[36] PhD-SNP, Panther,[37] and SIFT.[38]

Meta-SNP is a reliable tool for predicting how mutations affect protein functions. It combines results from various algorithms using a random forest technique, which involves 100 decision trees within the WEKA library. This method ensures the precise assessment of a mutation's impact on its protein function.[39] The meta-SNP algorithm was trained on the SV-2009 dataset through a 20-fold cross-validation procedure and provides dependable predictions of disease-causing variants.[22] It outputs a probability score between 0 and 1, with scores above 0.5 indicating a higher likelihood of the mutation being pathogenic.[40]

Next, we used the Polyphen-2 web server (http://genetics.bwh.harvard.edu/pph2/), which is an automated tool designed to predict the potential effects of amino acid substitutions on the structure and function of human proteins.[41] This site features an input form that enables users to query either a single amino acid substitution or a coding, non-synonymous SNP that is annotated in the dbSNP database.[23]

2.3. Computational assessment of protein stability

Mutations can affect a protein's structure and stability, resulting in changes in free energy (ΔΔG in Kcal/mol). The difference between the mutated (ΔGm) and original (ΔGw) proteins' free energy indicates a change in stability, where negative ΔΔG indicates destabilization, whereas positive ΔΔG indicates stabilization.[27,42] Various computational tools, including mCSM, SDM, and DUET, were employed to assess the impact of mutations on CCL2 protein stability and function, providing valuable insights into how specific mutations alter protein behavior.[27]

The SDM online tool, available at http://marid.bioc.cam.ac.uk/sdm2, accurately predicts the impact of mutations on protein stability through *in silico* analysis. It examines variations in amino acid substitutions tolerated by homologous proteins with known 3D structures in specific environments, providing valuable insights into how mutations can affect protein behavior and stability.[27,43,44]

The mCSM computational technique uses machine learning to predict the effects of missense mutations on CCL2 protein stability and function. It employs graph-based signatures to depict how mutations impact the interaction network among residues. By integrating evolutionary information, complex network metrics, and energetic factors, mCSM creates an accurate predictive model. It is available at http://structure.bioc.cam.ac.uk/mcsm.[21,26] The mCSM method evaluates how mutations affect protein structure and interactions with other proteins and nucleic acids. It provides insights into how mutations contribute to disease development, improving our understanding of the molecular mechanisms of various diseases. This computational approach clarifies the effects of mutations on protein stability, function, and interactions, enhancing our understanding of disease pathogenesis.[45,46]

DUET (http://structure.bioc.cam.ac.uk/duet) is a computational tool that predicts changes in protein stability due to individual mutations. It combines two methods to calculate the change in folding free energy (ΔΔG in Kcal/mol), with negative ΔΔG values indicating destabilizing mutations and positive values indicating stabilizing mutations.[28,47] DUET uses support vector machines trained with sequential minimal optimization to integrate the predictions from mCSM and SDM. This machine learning technique leverages the strengths of both methods to provide a comprehensive assessment of how missense mutations affect protein stability. By combining these predictive capabilities, DUET enhances the accuracy and reliability of mutation impact predictions on protein stability.[48]

2.4. Evaluating cancer-related mutations using the FATHMM server

The FATHMM server, available at http://fathmm.biocompute.org.uk, was used to assess the cancer relevance of deleterious mutations in *CCL2*. This web server predicts the functional consequences of both coding and noncoding

variants, including their potential association with cancer, where scores below the default threshold of −0.75 indicate a potential cancer link. FATHMM helps distinguish between cancer-related mutations and non-pathogenic germline polymorphisms.[25,49] Using FATHMM, we aimed to deepen our understanding of the genetic basis of cancer by evaluating the impact of harmful mutations on disease progression, identifying clinically significant mutations, and advancing our knowledge of cancer development.[25]

2.5. Sequence conservation analysis of CCL2

The ConSurf web server (http://consurf.tau.ac.il) analyzes the evolutionary patterns of amino acids in proteins to identify regions critical for structure and function. Using a query sequence or structure, the server automatically gathers homologous sequences, performs multiple sequence alignment, and constructs a phylogenetic tree that illustrates their evolutionary relationships. This information is then utilized within a probabilistic framework to estimate the evolutionary rates for each position in the sequence.[50]

2.6. Structural modeling of CCL2 point mutation sites

To create a mutant model of CCL2 featuring the targeted mutation C59C (native) and C59G (mutant), the crystal structure of the native peptides was initially downloaded from the PDB. Subsequently, the I-TASSER server (https://zhanglab.ccmb.med.umich.edu/I-TASSER/), which uses a multiple threading approach, was employed to construct the intended mutant model.[29]

2.7. Quality evaluation of the protein model

The CCL2 protein is a complex molecule composed of amino acids, with its function intricately linked to its specific 3D structure. Understanding CCL2's structure can provide valuable insights into its function, interactions with other molecules, and potential roles in biological processes, such as the immune system and inflammation.[51] The I-TASSER server was used to generate protein models of mutant and native CCL2.[29] The quality assessment process included an analysis of the stereochemical characteristics of the models, such as the bond lengths, bond angles, and torsion angles. This evaluation aimed to confirm that the protein models exhibit a conformation that is both geometrically and chemically reasonable. Deviations from expected values in the stereochemical features of the model could suggest potential errors in the model.[52] The protein models were validated by comparing them with experimental data to ensure that they aligned with the known structural properties of CCL2 protein, which verified the accuracy of the predicted models by examining how accurately they matched with the actual protein structure, thereby ensuring the reliability of the models.[53]

The reliability of the mutant and native CCL2 protein models was ensured by validating their stereochemical quality using the Rampage server (https://servicesn.mbi.ucla.edu/PROCHECK). This involved analyzing the residue-by-residue geometry and overall structure geometry to assess the accuracy of models.[54]

2.8. MD simulation

MD simulation is a computational technique used to simulate the dynamic movement of atoms and molecules throughout a period. By simulating movements and interactions of biological molecules, such as proteins, MD provides detailed insights into their molecular structures, dynamics, and interactions, which are often difficult to obtain through experimental methods alone.[55,56] The accuracy of MD simulation results relies heavily on the quality of the protein model.[57] Key factors for a high-quality protein model include an accurate representation of the native structure, selection of a suitable force field, inclusion of surrounding solvent and relevant environmental factors, and facilitation of post-simulation analysis.[58,59] Addressing these aspects improves the credibility and biological significance of the findings derived from MD simulations.

The native and mutant models of CCL2 proteins with the highest I-TASSER score, indicating their optimal conformation, were selected for further analysis.[29] To examine the effects of the identified mutations on the CCL2 protein structure at the molecular level, we performed an MD simulation using WebGro software.[60] This revealed how native C59C and mutant C59G influenced the structure of CCL2 during a 100-ns simulation, and the MD simulation system was neutralized by introducing sodium or chlorine ions to balance the total charges. Before MD simulation, the steepest descent algorithm was used for system minimization through 5000 iterations[61] (The simulations were performed at a constant temperature of 300 K and a pressure of 1.0 bar, with a 0.15 M NaCl concentration. Approximately 1000 frames were generated per 100-ns simulation.[62]

As a crucial initial step, we performed protein docking to predict the binding location of the molecule within the protein's active site. This process provided essential insights into the subsequent MD simulations, allowing for the detailed examination of the protein's structure and function.[63,64] MD simulations produced multiple trajectories, including root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), and hydrogen bonds.[65,66]

The RMSD profiles of protein–ligand complexes were analyzed using the WebGRO server. The best-docked

complexes were prepared for MD simulations using the GROMOS 43a1 force field.[65] RMSD measures the average distance between the atoms of two superimposed molecular structures. It is commonly used in bioinformatics to assess the similarity of 3D structures, particularly in the study of protein conformations.[61] RMSD is calculated using the formula:

$$\text{RMSD} = \sqrt{1/N \sum_{i=1}^{n} (\delta i)^2} \quad \text{(I)}$$

where δi represents the distance between the corresponding atoms in the two structures, and *N* is the number of atoms considered. Typically, RMSD is calculated for the backbone heavy atoms or alpha carbon (Cα) atoms.

RMSF is a measure of the average deviation of atomic positions from their mean positions over time in MD simulations. It quantifies the fluctuations of individual atoms or residues, providing insights into the flexibility and mobility of different regions within a protein structure. RMSF is calculated as the square root of the variance of the fluctuations around the average position[62] using the formula:

$$\text{RMSF} = \sqrt{(\gamma i - \gamma i)^2} \quad \text{(II)}$$

where γi represents the coordinates of atom *i,* and ⟨γi⟩ is its ensemble average position. High RMSF values indicate regions with considerable mobility, whereas low values suggest rigidity.

To gain insights into the hydrogen bonding characteristics of the proteins, we analyzed the hydrogen bond geometries. As the positions of hydrogen atoms are uncertain in these structures, they were excluded from the analysis. In total, 4195 hydrogen bonds were identified and categorized into four types based on the donor and acceptor atom origins: main-chain to main-chain (MC-MC), main-chain to side-chain, side-chain to main-chain, and side-chain to side-chain. The MC-MC hydrogen bonds were further classified according to the secondary structures, including alpha helices, beta strands, and non-secondary structures.

Furthermore, we investigated the correlations between hydrogen bond geometry and environmental factors, such as atomic depth, secondary structure, and partial charge. Atomic depth was defined as the distance to the nearest bulk water molecule. This analysis provides a comprehensive understanding of the characteristics and determinants of hydrogen bonding in high-resolution protein structures.

2.9. Examination of alterations in CCL2 protein secondary structure

The Database of Secondary Structure in Proteins tool analyzes hydrogen bonding and other secondary structures of CCL2 proteins – indicators to classify protein residues – as well as changes in secondary structures between native and mutant CCL2 proteins.[66]

2.10. Docking of CCL2-native and -mutant proteins with receptor CCR2

The objective of molecular docking for the ligand–receptor (CCL2–CCR2) binding process is to construct an accurate binding free energy landscape and estimate the binding affinity and binding kinetics. We introduced a novel hybrid docking (HDOCK) protocol that integrates template-based and template-free approaches to enhance protein–protein docking accuracy. Unlike traditional methods that focus on docking monomer proteins, HDOCK utilizes homologous complex structures of the proteins involved. Initially, homologous complexes were identified using BLAST searches of the PDB, with a sequence identity cutoff set at 25%. Among multiple hits, the optimal template was selected based on the highest sequence coverage, sequence similarity, and resolution. This selected homologous complex structure was then used as a docking template, incorporating potential binding interface information. By leveraging the structural information from the homologous complexes, HDOCK substantially improves the prediction accuracy of protein–protein interactions.[67] The individual protein structures are modeled using the homologous complexes as templates, either by superimposing them onto available homologous complexes or through homology modeling.[68] In our study, the individual component structures of the modeled complex were separated and subjected to a global sampling of potential binding modes using our hierarchical protein–protein docking algorithm. This approach allows for a comprehensive exploration of possible interaction interfaces between the proteins.[69,70] Symmetric docking was achieved by constraining the translational degrees of freedom to ensure symmetry along a specific axis. In cases where no complex structure templates were available, the individual protein structures for docking were modeled using their homologous monomer structures with Modeller.[71] The binding modes obtained from the docking simulations were evaluated and ranked according to their binding energy scores. To group similar conformations, an RMSD cutoff of 5.0 Å was applied for clustering. From the resulting clusters, the top 100 binding modes were examined and the best 10 models were selected for submission to the Critical Assessment of Predicted Interactions initiative.

3. Results

The results of this study demonstrated the considerable impact of mutations in the *CCL2* gene on the molecular stability, functionality, and structure of the resulting

protein. The altered protein might not perform its normal functions effectively, potentially contributing to disease development or progression. Advanced computational methods were employed to assess protein stability, identify cancer-causing mutations, and predict pathogenicity.[72,73]

3.1. Collection of missense mutations

To explore the potential effects of *CCL2* mutation on protein function, we collected a comprehensive dataset comprising 83 mutations from a publicly accessible database. These mutations were located across various coding regions of the gene, enabling a thorough assessment of their potential effects on protein functionality.

3.2. Missense mutation analysis

The deleterious effects of mutations can be anticipated by evaluating the importance of the affected peptides. To analyze the 83 missense mutations in *CCL2*, computational tools such as Polyphen-2 and Meta-SNP (comprising Panther, PhD-SNP, SIFT, and SNAP) were employed (Table 1). The results revealed that 10 mutations were likely disease-causing and distributed across various protein domains.

3.3. CCL2 mutations affect protein stability

Three algorithms, mCSM, SDM, and DUET, were used to predict the effects of 83 mutations on CCL2 protein stability. The results are shown in Table 2. mCSM predicted that 40 mutations in CCL2 were likely highly destabilizing.

3.4. Identification of cancer-causing mutations in CCL2

Using the FATHMM server, 83 CCL2 mutations were analyzed to assess their potential involvement in cancer development or progression. The analysis results suggest that the C59G mutation alters the binding dynamics between CCL2 and CCR2, potentially reducing the risk of cancer and enhancing the immune response to pathogen invasion.

3.5. Conservation analysis of CCL2 mutations

The C59G mutation, identified as a cancer-causing mutation, was selected for further analysis. We utilized the ConSurf server to investigate the conservation and spatial distribution of the C59G mutation, which offered valuable insights into its potential functional implications.[50] ConSurf analysis revealed that the C59G mutation occurred at a highly conserved position, obtaining the maximum conservation score of 9 on the ConSurf scale (Figure 2A and B). These findings prompted further computational investigations into the potential effects of this mutation on the stability and function of the CCL2 protein.

3.6. Structure validation of native and mutant CCL2 proteins

The reference model for the native protein was the 3D structure of CCL2 (PDB ID 2LIE). The I-TASSER server was used to identify the template proteins with similar folds in the PDB database and generate 10,000 conformations (decoys) based on a mutant sequence.[74] I-TASSER used the SPICKER tool to group the generated conformations through pairwise structural alignments, identifying models with structural similarities. This process typically results in five clusters, each containing multiple models with identical structures. I-TASSER then selects a representative model from each cluster to generate five final models that best represent the diverse conformations predicted.[74] The model displaying the highest confidence score (C-score) was selected for subsequent analysis.[75] Ramachandran plots were used to assess the model's quality by examining the distribution of the protein backbone torsion angles (ϕ and ψ) and identifying any structural issues.[76] The Ramachandran plot analysis revealed that 90% of native residues and 89% of mutant residues were positioned in the favored region of the native CCL2 structure, indicating a high degree of structural similarity (Figure 3).

3.7. MD simulation of CCL2

To investigate the effects of the C59G mutation on the CCL2 protein, a 100-ns MD simulation was performed. The simulation tracked various structural parameters, including secondary structure elements, to monitor the protein's behavior and identify any changes in its conformational dynamics.

The RMSD was computed for the backbone residues of both the native and mutant CCL2 proteins to assess their deviation from their initial structures throughout the simulation and quantify the extent of structural variation. The RMSD values for the mutant CCL2 protein (C59G) were substantially higher than those of the native protein (Figure 3A), indicating a higher degree of structural instability in the mutant structure. This suggests that the C59G mutation introduced substantial conformational changes, potentially affecting the protein's stability and function. MD simulations revealed that the native CCL2 protein remained structurally stable throughout the simulation and that the mutant proteins, particularly C59G, exhibited a longer stabilization period. The RMSD values for the mutant proteins steadily increased between 15 and 30 ns, indicating a higher degree of structural deviation from the initial conformation. In contrast, the RMSD values for the native C59C–CCL2 protein remained relatively stable, which suggests that the C59G mutation

Table 1. Harmful mutation predictions using in silico online tools

Serial number	Mutants	Prediction of deleterious mutation													
		Panther		PhD-SNP		SIFT		SNAP		Meta-SNP		PolyPhen-2			
		Analysis	Score	Analysis	Score	Analysis	Score	Analysis	Score	Analysis	Score	Analysis	Score	Sensitivity	Specificity
1	M1L	Disease	0.982	Normal	0.153	Normal	0.15	Normal	0.25	Normal	0.248	Benign	0.4	0.84	0.79
2	K2E	Normal	0.466	Normal	0.12	Normal	0.41	Normal	0.34	Normal	0.128	Possibly damaging	0.637	0.8	0.84
3	K2T	Disease	0.567	Normal	0.075	Normal	0.42	Normal	0.275	Normal	0.085	Possibly damaging	0.86	0.72	0.89
4	K2R	Normal	0.255	Normal	0.05	Normal	0.44	Normal	0.095	Normal	0.082	Benign	0.209	0.88	0.74
5	V3I	Normal	0.196	Normal	0.033	Normal	0.27	Normal	0.13	Normal	0.071	Benign	0.181	0.89	0.73
6	V3G	Disease	0.696	Normal	0.242	Disease	0.01	Normal	0.46	Normal	0.422	Possibly damaging	0.959	0.63	0.92
7	S4F	Disease	0.871	Normal	0.208	Normal	0.11	Normal	0.485	Normal	0.383	Possibly damaging	0.99	0.51	0.95
8	A5V	Normal	0.266	Normal	0.037	Normal	0.42	Normal	0.175	Normal	0.075	Benign	0.139	0.9	0.71
9	A5T	Normal	0.269	Normal	0.047	Normal	0.48	Normal	0.145	Normal	0.074	Benign	0.098	0.91	0.68
10	I13T	Normal	0.169	Normal	0.05	Normal	0.44	Normal	0.16	Normal	0.076	Benign	0.002	0.99	0.18
11	I13V	Normal	0.127	Normal	0.082	Normal	0.38	Normal	0.145	Normal	0.089	Benign	0.004	0.98	0.35
12	A14E	Normal	0.482	Disease	0.571	Normal	0.08	Disease	0.585	Normal	0.498	Possibly damaging	0.839	0.73	0.88
13	A15P	Disease	0.727	Disease	0.749	Normal	0.08	Disease	0.62	Disease	0.653	Possibly damaging	0.921	0.68	0.91
14	T16S	Normal	0.122	Normal	0.103	Normal	0.55	Normal	0.13	Normal	0.096	Benign	0.086	0.91	0.67
15	I18N	Normal	0.119	Normal	0.346	Normal	0.31	Normal	0.37	Normal	0.293	Benign	0.001	0.99	0.09
16	I18T	Normal	0.285	Normal	0.088	Normal	0.41	Normal	0.19	Normal	0.056	Benign	0.001	0.99	0.09
17	P19L	Normal	0.148	Normal	0.134	Normal	0.76	Normal	0.245	Normal	0.167	Benign	0.003	0.98	0.26
18	P19Q	Normal	0.262	Normal	0.123	Normal	0.35	Normal	0.32	Normal	0.136	Possibly damaging	0.449	0.83	0.8
19	Q20R	Normal	0.351	Normal	0.498	Normal	0.47	Disease	0.55	Disease	0.609	Possibly damaging	0.691	0.78	0.85
20	A23T	Normal	0.374	Normal	0.369	Normal	0.55	Normal	0.37	Normal	0.302	Benign	0.348	0.85	0.78
21	A23S	Normal	0.384	Normal	0.213	Normal	0.88	Normal	0.155	Normal	0.249	Possibly damaging	0.603	0.81	0.83
22	A27G	Normal	0.183	Normal	0.081	Normal	1	Normal	0.37	Normal	0.092	Benign	0.028	0.94	0.59
23	A27V	Disease	0.619	Normal	0.324	Normal	0.17	Disease	0.585	Normal	0.484	Possibly damaging	0.876	0.71	0.89
24	N29K	Normal	0.286	Normal	0.425	Normal	0.69	Disease	0.58	Normal	0.304	Benign	0.098	0.91	0.68
25	P31A	Normal	0.274	Normal	0.355	Normal	0.11	Disease	0.66	Normal	0.314	Benign	0.119	0.9	0.7
26	Y36C	Disease	0.903	Disease	0.717	Disease	0.01	Disease	0.745	Disease	0.795	Possibly damaging	0.976	0.59	0.93
27	N37D	Normal	0.19	Normal	0.2	Normal	0.26	Disease	0.56	Normal	0.25	Benign	0.21	0.88	0.74
28	T39I	Normal	0.222	Normal	0.18	Normal	0.6	Disease	0.515	Normal	0.145	Benign	0.021	0.95	0.56
29	T39A	Normal	0.178	Normal	0.229	Normal	0.59	Disease	0.62	Normal	0.244	Benign	0.006	0.97	0.45
30	N40D	Normal	0.226	Normal	0.384	Normal	0.45	Normal	0.415	Normal	0.33	Benign	0.022	0.95	0.57

(Cont'd...)

Table 1. *(Continued)*

Serial number	Mutants	Prediction of deleterious mutation													
		Panther		PhD-SNP		SIFT		SNAP		Meta-SNP		PolyPhen-2			
		Analysis	Score	Analysis	Score	Analysis	Score	Analysis	Score	Analysis	Score	Analysis	Score	Sensitivity	Specificity
31	N40K	Normal	0.128	Normal	0.245	Normal	0.65	Normal	0.415	Normal	0.248	Benign	0.023	0.95	0.57
32	R41K	Normal	0.182	Normal	0.072	Normal	0.77	Normal	0.405	Normal	0.084	Benign	0.002	0.99	0.18
33	K42I	Disease	0.703	Disease	0.634	Normal	0.12	Disease	0.665	Disease	0.669	Possibly damaging	0.936	0.66	0.91
34	S44L	Normal	0.468	Normal	0.482	Disease	0.02	Disease	0.53	Normal	0.49	Benign	0.02	0.95	0.56
35	V45M	Normal	0.11	Normal	0.054	Normal	0.25	Normal	0.41	Normal	0.069	Benign	0.026	0.94	0.58
36	V45A	Normal	0.161	Normal	0.116	Normal	0.53	Normal	0.475	Normal	0.116	Benign	0.103	0.91	0.69
37	R47K	Normal	0.173	Normal	0.223	Normal	0.36	Normal	0.43	Normal	0.241	Benign	0.1	0.91	0.68
38	L48F	Disease	0.653	Normal	0.416	Disease	0.02	Disease	0.68	Disease	0.582	Possibly damaging	0.997	0.27	0.98
39	A49E	Normal	0.207	Normal	0.08	Normal	0.96	Normal	0.41	Normal	0.076	Benign	0.007	0.97	0.46
40	A49V	Normal	0.241	Normal	0.084	Normal	0.4	Normal	0.41	Normal	0.08	Benign	0.015	0.95	0.54
41	A49T	Normal	0.152	Normal	0.074	Normal	0.53	Normal	0.345	Normal	0.085	Benign	0.015	0.95	0.54
42	A49P	Normal	0.291	Normal	0.404	Normal	0.23	Disease	0.635	Normal	0.328	Possibly damaging	0.587	0.81	0.83
43	S50N	Normal	0.434	Normal	0.311	Normal	0.16	Disease	0.58	Normal	0.325	Possibly damaging	0.651	0.79	0.84
44	Y51H	Disease	0.845	Disease	0.647	Disease	0.01	Disease	0.86	Disease	0.733	Possibly damaging	0.998	0.18	0.98
45	R53G	Disease	0.743	Disease	0.539	Normal	0.21	Disease	0.685	Disease	0.661	Possibly damaging	0.906	0.69	0.9
46	I54V	Normal	0.253	Normal	0.085	Normal	0.28	Disease	0.51	Normal	0.089	Benign	0.075	0.92	0.66
47	T55S	Disease	0.662	Normal	0.455	Normal	0.53	Disease	0.585	Disease	0.6	Possibly damaging	0.718	0.78	0.85
48	S56N	Normal	0.234	Normal	0.216	Normal	0.26	Disease	0.62	Normal	0.236	Benign	0.107	0.91	0.69
49	S56R	Disease	0.627	Normal	0.485	Normal	0.31	Disease	0.685	Disease	0.612	Possibly damaging	0.965	0.62	0.93
50	S57C	Disease	0.812	Disease	0.518	Disease	0.01	Disease	0.64	Disease	0.637	Possibly damaging	0.976	0.59	0.93
51	C59G	Disease	0.992	Disease	0.857	Disease	0	Disease	0.83	Disease	0.916	Possibly damaging	1	0	1
52	K61I	Disease	0.525	Normal	0.223	Normal	0.53	Normal	0.445	Normal	0.296	Possibly damaging	0.487	0.83	0.81
53	K67R	Normal	0.244	Normal	0.043	Normal	0.41	Normal	0.42	Normal	0.086	Benign	0.029	0.94	0.59
54	K67D	Normal	0.498	Normal	0.211	Normal	0.2	Disease	0.605	Normal	0.26	Possibly damaging	0.712	0.78	0.85
55	K67T	Normal	0.464	Normal	0.105	Normal	0.61	Normal	0.435	Normal	0.111	Benign	0.178	0.89	0.73
56	T68I	Disease	0.98	Disease	0.626	Disease	0.02	Disease	0.685	Disease	0.56	Possibly damaging	0.967	0.61	0.93
57	I69V	Normal	0.196	Normal	0.124	Normal	0.14	Normal	0.34	Normal	0.12	Benign	0.006	0.97	0.45
58	A71T	Normal	0.25	Normal	0.128	Disease	0.04	Disease	0.63	Normal	0.175	Benign	0.41	0.84	0.79
59	A71V	Normal	0.369	Normal	0.159	Disease	0.01	Disease	0.645	Normal	0.21	Possibly damaging	0.653	0.79	0.84
60	E73G	Disease	0.581	Normal	0.454	Normal	0.1	Disease	0.64	Disease	0.509	Benign	0.208	0.88	0.74

(Cont'd...)

Table 1. ***(Continued)***

Serial number	Mutants	Prediction of deleterious mutation													
		Panther		PhD-SNP		SIFT		SNAP		Meta-SNP		PolyPhen-2			
		Analysis	Score	Analysis	Score	Analysis	Score	Analysis	Score	Analysis	Score	Analysis	Score	Sensitivity	Specificity
61	I74N	Disease	0.885	Disease	0.771	Disease	0.01	Disease	0.725	Disease	0.772	Possibly damaging	0.964	0.62	0.92
62	D77E	Normal	0.269	Normal	0.409	Disease	0.04	Disease	0.685	Normal	0.325	Benign	0.311	0.86	0.77
63	D77G	Disease	0.661	Disease	0.67	Disease	0.01	Disease	0.71	Disease	0.691	Possibly damaging	0.933	0.67	0.91
64	P78S	Disease	0.548	Disease	0.517	Disease	0.01	Disease	0.76	Disease	0.728	Possibly damaging	0.941	0.66	0.91
65	P78H	Disease	0.886	Disease	0.557	Disease	0	Disease	0.805	Disease	0.732	Possibly damaging	0.998	0.18	0.98
66	K79E	Normal	0.289	Normal	0.219	Normal	0.4	Normal	0.47	Normal	0.245	Benign	0.049	0.93	0.63
67	K79R	Normal	0.347	Normal	0.154	Normal	0.19	Normal	0.43	Normal	0.162	Benign	0.285	0.86	0.76
68	K81N	Disease	0.554	Normal	0.234	Normal	0.21	Normal	0.465	Normal	0.342	Benign	0.364	0.85	0.78
69	W82L	Disease	0.986	Disease	0.622	Disease	0	Disease	0.795	Disease	0.75	Possibly damaging	1	0	1
70	V83A	Disease	0.977	Normal	0.406	Disease	0.01	Disease	0.695	Disease	0.62	Possibly damaging	0.992	0.49	0.95
71	V83L	Disease	0.965	Normal	0.466	Disease	0.01	Disease	0.675	Disease	0.63	Possibly damaging	0.941	0.66	0.91
72	Q84R	Normal	0.304	Normal	0.289	Normal	0.13	Normal	0.48	Normal	0.309	Benign	0.097	0.91	0.68
73	D85Y	Disease	0.714	Normal	0.334	Disease	0.01	Disease	0.56	Disease	0.519	Possibly damaging	0.98	0.57	0.94
74	M87V	Normal	0.139	Normal	0.061	Normal	0.4	Disease	0.525	Normal	0.1	Benign	0.02	0.95	0.56
75	M87K	Disease	0.596	Normal	0.252	Disease	0.02	Disease	0.745	Normal	0.372	Benign	0.128	0.9	0.7
76	D88N	Normal	0.132	Normal	0.024	Normal	0.32	Normal	0.23	Normal	0.063	Benign	0.004	0.98	0.35
77	D88V	Normal	0.191	Normal	0.061	Disease	0.03	Normal	0.465	Normal	0.112	Benign	0.103	0.91	0.69
78	H89P	Normal	0.443	Normal	0.322	Normal	0.21	Disease	0.65	Normal	0.343	Possibly damaging	0.722	0.78	0.85
79	K92R	Normal	0.16	Normal	0.042	Normal	0.38	Normal	0.16	Normal	0.066	Benign	0.008	0.96	0.48
80	T96A	Normal	0.232	Normal	0.027	Normal	0.42	Disease	0.56	Normal	0.084	Benign	0.071	0.92	0.66
81	P97L	Normal	0.168	Normal	0.035	Normal	0.41	Normal	0.48	Normal	0.055	Benign	0.012	0.96	0.52
82	P97T	Normal	0.245	Normal	0.034	Normal	0.68	Normal	0.45	Normal	0.05	Benign	0.033	0.94	0.6
83	T99S	Normal	0.091	Normal	0.006	Normal	0.42	Disease	0.51	Normal	0.081	Benign	0	1	0

Table 2. Prediction of protein stability using in silico tools

Serial number	Mutants	Protein stability prediction					
		SDM		mCSM		DUET	
		ΔΔGkcal/mol	Prediction	ΔΔGkcal/mol	Prediction	ΔΔGkcal/mol	Prediction
1	M1L	0	Destabilizing	−0.528	Destabilizing	−0.079	Destabilizing
2	K2E	−0.08	Destabilizing	0.15	Stabilizing	0.411	Stabilizing
3	K2T	−0.39	Destabilizing	−0.419	Destabilizing	−0.251	Destabilizing
4	K2R	−0.24	Destabilizing	−0.297	Destabilizing	−0.066	Destabilizing
5	V3I	0.11	Stabilizing	−0.864	Destabilizing	−0.547	Destabilizing
6	V3G	−0.38	Destabilizing	−1.907	Destabilizing	−1.798	Destabilizing
7	S4F	0.58	Stabilizing	−1.339	Destabilizing	−1.047	Destabilizing
8	A5V	−0.18	Destabilizing	−0.527	Destabilizing	−0.257	Destabilizing
9	A5T	−0.37	Destabilizing	−1.443	Destabilizing	−1.2	Destabilizing
10	I13T	−0.96	Destabilizing	−1.451	Destabilizing	−1.33	Destabilizing
11	I13V	−0.1	Destabilizing	−1.056	Destabilizing	−0.804	Destabilizing
12	A14E	−0.13	Destabilizing	−2.421	Destabilizing	−2.148	Destabilizing
13	A15P	−1.96	Destabilizing	−0.606	Destabilizing	−0.767	Destabilizing
14	T16S	−0.78	Destabilizing	−1.357	Destabilizing	−1.302	Destabilizing
15	I18N	−1.39	Destabilizing	−2.458	Destabilizing	−2.435	Destabilizing
16	I18T	−2.02	Destabilizing	−2.715	Destabilizing	−2.85	Destabilizing
17	P19L	1.29	Stabilizing	−0.789	Destabilizing	−0.197	Destabilizing
18	P19Q	−0.23	Destabilizing	−0.973	Destabilizing	−0.68	Destabilizing
19	Q20R	0.56	Stabilizing	0.009	Stabilizing	0.357	Stabilizing
20	A23T	−0.81	Destabilizing	−1.028	Destabilizing	−0.853	Destabilizing
21	A23S	−1.39	Destabilizing	−0.951	Destabilizing	−0.886	Destabilizing
22	A27G	−2.55	Destabilizing	−1.264	Destabilizing	−1.617	Destabilizing
23	A27V	−1.46	Destabilizing	−0.774	Destabilizing	−0.831	Destabilizing
24	N29K	0.44	Stabilizing	0.144	Stabilizing	0.606	Stabilizing
25	P31A	−0.2	Destabilizing	−0.91	Destabilizing	−0.671	Destabilizing
26	Y36C	0.09	Stabilizing	0.42	Stabilizing	0.633	Stabilizing
27	N37D	0.15	Stabilizing	0.276	Stabilizing	0.634	Stabilizing
28	T39I	1.33	Stabilizing	−0.053	Destabilizing	0.493	Stabilizing
29	T39A	1.34	Stabilizing	−0.876	Destabilizing	−0.322	Destabilizing
30	N40D	−0.38	Destabilizing	0.119	Stabilizing	0.347	Stabilizing
31	N40K	−0.2	Destabilizing	0.131	Stabilizing	0.451	Stabilizing
32	R41K	−0.19	Destabilizing	−0.038	Destabilizing	0.192	Stabilizing
33	K42I	0.47	Stabilizing	0.158	Stabilizing	0.452	Stabilizing
34	S44L	1.24	Stabilizing	0.062	Stabilizing	0.585	Stabilizing
35	V45M	−0.26	Destabilizing	−0.396	Destabilizing	−0.295	Destabilizing
36	V45A	0.19	Destabilizing	−0.971	Stabilizing	−0.641	Destabilizing
37	R47K	−0.1	Destabilizing	−0.316	Destabilizing	−0.058	Destabilizing
38	L48F	−0.66	Destabilizing	−1.439	Destabilizing	−1.493	Destabilizing
39	A49E	0.72	Stabilizing	−0.528	Destabilizing	0.007	Stabilizing
40	A49V	1.41	Stabilizing	−0.416	Destabilizing	0.201	Stabilizing
41	A49T	0.29	Stabilizing	−0.942	Destabilizing	−0.546	Destabilizing

(Cont'd...)

Table 2. ***(Continued)***

Serial number	Mutants	Protein stability prediction					
		SDM		mCSM		DUET	
		ΔΔGkcal/mol	Prediction	ΔΔGkcal/mol	Prediction	ΔΔGkcal/mol	Prediction
42	A49P	−1.14	Destabilizing	−0.416	Destabilizing	−0.366	Destabilizing
43	S50N	0.07	Stabilizing	−1.123	Destabilizing	−0.858	Destabilizing
44	Y51H	−1.23	Destabilizing	−2.402	Destabilizing	−2.457	Destabilizing
45	R53G	0.37	Stabilizing	−1.393	Destabilizing	−1.084	Destabilizing
46	I54V	−2.38	Destabilizing	−1.695	Destabilizing	−2.05	Destabilizing
47	T55S	−0.71	Destabilizing	−0.468	Destabilizing	−0.319	Destabilizing
48	S56N	0.28	Stabilizing	−0.564	Destabilizing	−0.226	Destabilizing
49	S56R	0.9	Stabilizing	−0.645	Destabilizing	−0.226	Destabilizing
50	S57C	0.62	Stabilizing	−0.491	Destabilizing	−0.162	Destabilizing
51	C59G	−0.6	Destabilizing	−1.767	Destabilizing	−1.667	Destabilizing
52	K61I	1.4	Stabilizing	0.437	Stabilizing	0.915	Stabilizing
53	K67R	0	Destabilizing	−0.699	Destabilizing	−0.435	Destabilizing
54	K67D	−0.83	Destabilizing	−1.817	Destabilizing	−1.832	Destabilizing
55	K67T	−0.14	Destabilizing	−1.275	Destabilizing	−1.086	Destabilizing
56	T68I	0.13	Stabilizing	−0.148	Destabilizing	0.111	Stabilizing
57	I69V	−0.03	Destabilizing	−0.347	Destabilizing	−0.033	Destabilizing
58	A71T	−1.12	Destabilizing	−0.861	Destabilizing	−0.738	Destabilizing
59	A71V	0.38	Stabilizing	−0.472	Destabilizing	−0.071	Destabilizing
60	E73G	0.56	Stabilizing	−1.152	Destabilizing	−0.846	Destabilizing
61	I74N	−1.35	Destabilizing	−1.567	Destabilizing	−1.522	Destabilizing
62	D77E	−0.5	Destabilizing	−0.784	Destabilizing	−0.609	Destabilizing
63	D77G	0.22	Stabilizing	−1.093	Destabilizing	−0.874	Destabilizing
64	P78S	−0.77	Destabilizing	−1.976	Destabilizing	−1.844	Destabilizing
65	P78H	1.14	Stabilizing	−1.898	Destabilizing	−1.331	Destabilizing
66	K79E	0.25	Stabilizing	−0.115	Destabilizing	0.229	Stabilizing
67	K79R	−0.11	Destabilizing	−0.144	Destabilizing	0.11	Stabilizing
68	K81N	−0.88	Destabilizing	−0.001	Destabilizing	0.044	Stabilizing
69	W82L	0.63	Stabilizing	−2.029	Destabilizing	−1.49	Destabilizing
70	V83A	−0.69	Destabilizing	−1.948	Destabilizing	−1.924	Destabilizing
71	V83L	0.14	Stabilizing	−0.428	Destabilizing	−0.079	Destabilizing
72	Q84R	0.13	Stabilizing	−0.043	Destabilizing	0.245	Stabilizing
73	D85Y	−0.37	Destabilizing	−0.479	Destabilizing	−0.514	Destabilizing
74	M87V	0.14	Stabilizing	−1.727	Destabilizing	−1.309	Destabilizing
75	M87K	−0.19	Destabilizing	−1.477	Destabilizing	−1.02	Destabilizing
76	D88N	−0.34	Destabilizing	−0.639	Destabilizing	−0.469	Destabilizing
77	D88V	−0.11	Destabilizing	0.312	Stabilizing	0.518	Stabilizing
78	H89P	−2.03	Destabilizing	0.226	Stabilizing	−0.113	Destabilizing
79	K92R	−0.37	Destabilizing	−0.162	Destabilizing	0.048	Stabilizing
80	T96A	1.15	Stabilizing	−0.191	Destabilizing	0.318	Stabilizing
81	P97L	0.84	Stabilizing	−0.258	Destabilizing	0.246	Stabilizing
82	P97T	0.28	Stabilizing	−0.404	Destabilizing	−0.002	Destabilizing
83	T99S	0	Destabilizing	−0.626	Destabilizing	−0.32	Destabilizing

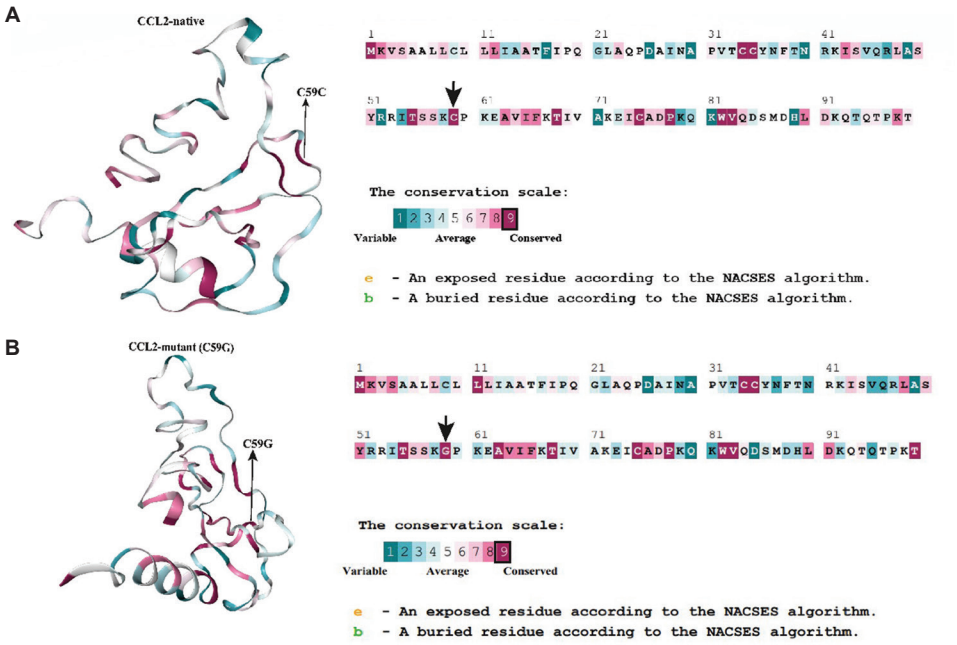


Figure 2. The ConSurf server was used to analyze the residue conservation of the C59G mutation in CCL2. The results are displayed in a colored format to visualize the evolutionary conservation of these residues within the protein structure. (A) represents the C59 residue, while (B) indicates the position of the C59G mutation. This mutation lies within the most conserved region of the CCL2 protein structure. The conservation of this region suggests its critical role in the protein's function and stability.
Abbreviation: CCL2: Chemokine C-C motif ligand 2.

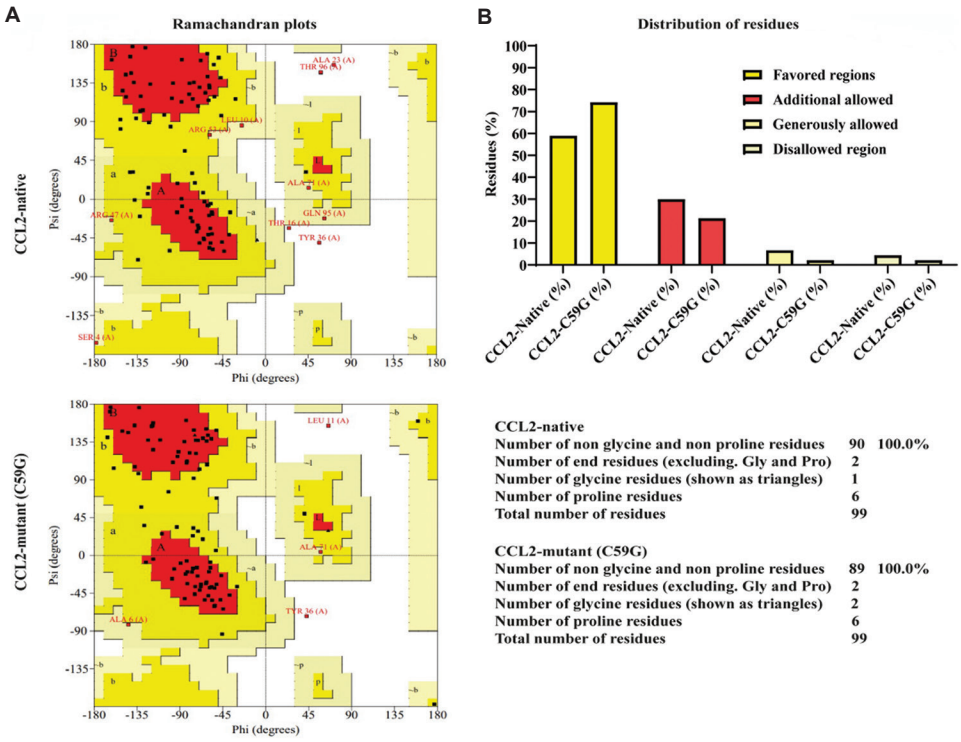


Figure 3. (A and B) Ramachandran plot analysis was used to evaluate the structural quality of the native and mutant CCL2 proteins by examining the distribution of residues within the plot. The results are presented in a colored format to facilitate easier visualization and interpretation of the evolutionary conservation of these residues within the protein structure.
Abbreviation: CCL2: Chemokine C-C motif ligand 2.

introduced considerable conformational changes that required more time to reach equilibrium (Figure 4A).

To assess the impact of mutations on the stability and dynamics of the CCL2 protein, the RMSF of the Cα atoms was calculated. This analysis aimed to understand how mutations affect the dynamics of the residues and compare these effects with those of the native protein. In general, the RMSF values of the mutant structures demonstrated higher flexibility than those of the native structures, particularly in regions spanning residues 5 – 80. Specifically, these regions exhibited significant fluctuations in RMSF values, ranging from 0.1 to 0.7 Å, which indicated increased dynamic behavior in these areas (Figure 4B).

After analyzing the intramolecular hydrogen bonds, a noticeable decrease in the number of hydrogen bonds formed during the simulation in the mutant protein was observed compared to the native structure. Assessment of the intramolecular hydrogen bonds indicated that the native protein maintained a relatively stable number of hydrogen bonds, ranging from 40 to 80, whereas the mutant protein exhibited more variability, with the number of hydrogen bonds fluctuating between 40 and 69. This implies that the mutant protein underwent increased flexibility and structural alterations compared to the native CCL2 protein (Figure 5A and B).

Essential dynamics (ED) analysis was conducted to identify correlated movements within both the native and mutant peptides of CCL2. The cumulative sum of the eigenvalues revealed that significant fluctuations were mainly observed in the first two eigenvectors for both types of proteins. This concentration of fluctuations in the first two eigenvectors indicates that these principal components (PCs) capture the primary motions of the system. Subsequently, the trajectories of both native and mutant simulations were plotted onto phase space along PC1 and PC2 at 300 K to visualize these motions (Figure 5C and D).

The Rg analysis revealed that the mutant protein was less compact compared to the native protein. The Rg value for the mutant protein ranged from 1.3 nm to 1.45 nm, which reflected increased flexibility and structural changes. Conversely, the native protein demonstrated a higher compactness with Rg values between 1.4 nm and 1.3 nm. This suggests that the mutations notably affected the protein's structural compactness and flexibility (Figure 6A and B).

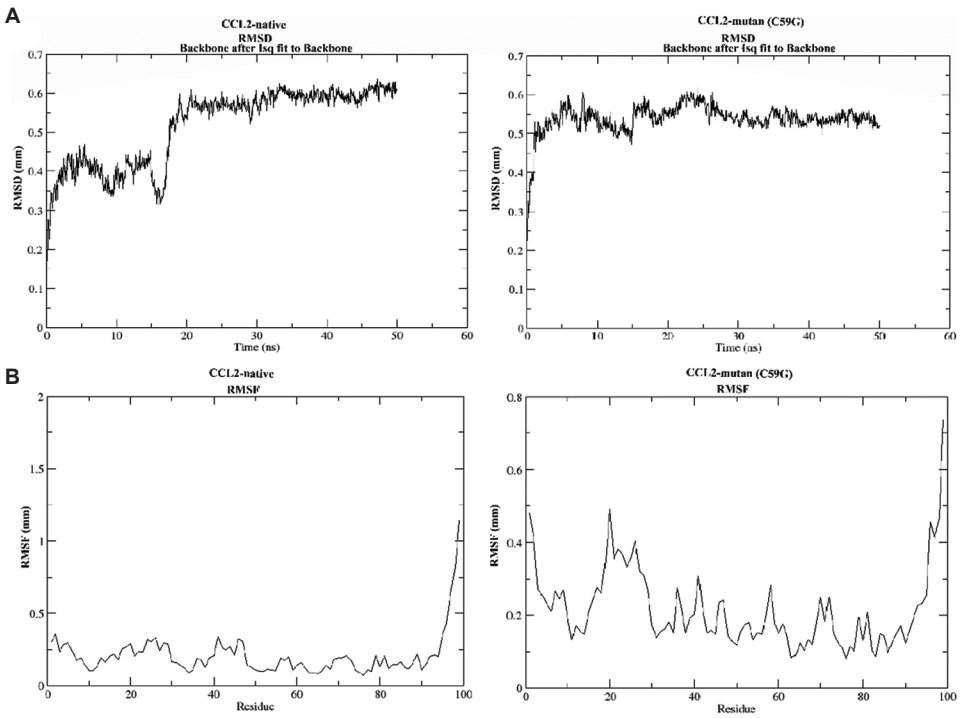


Figure 4. The RMSD and RMSF were used as measures to assess the structural alterations and dynamic fluctuations in the proteins. (A) Molecular dynamics simulations were conducted on both native and C59G mutant CCL2 proteins for 100 ns to explore the impacts of mutations on protein stability. The RMSD of the backbone was calculated over time to quantify the structural changes. (B) The RMSF of backbone alpha carbon (Cα) atoms in the native CCL2 protein structure was examined throughout the simulation duration. The results are depicted on a plot, where the Y-axis represents RMSF values and the X-axis corresponds to each atom in the protein structure.
Abbreviations: RMSD: root mean square deviation; RMSF: root mean square fluctuation; CCL2: Chemokine C-C motif ligand 2.

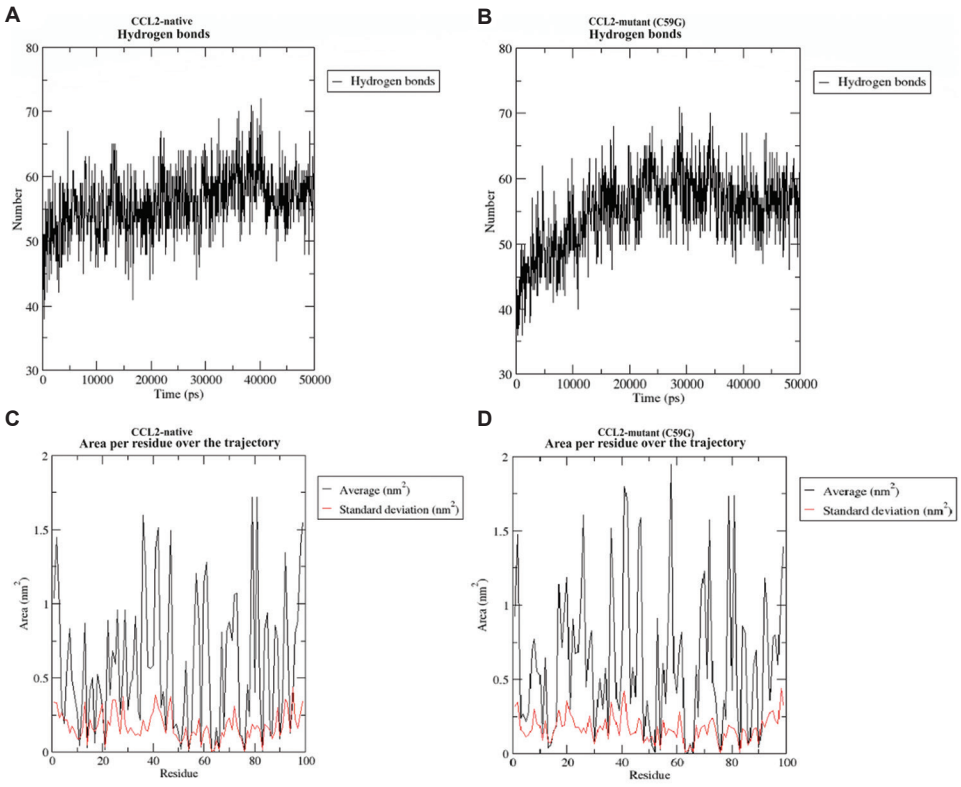


Figure 5. Hydrogen bonding and two-dimensional projection of native and mutant CCL2 proteins. (A) represents the average number of intermolecular hydrogen bonds in native CCL2, while (B) shows the average number of intermolecular hydrogen bonds in the CCL2-C59G mutant structures as a function of time. (C and D) depict the projection of the motion of the protein in phase space for CCL2-native and CCL2-C59G, respectively.
Abbreviation: CCL2: Chemokine C-C motif ligand 2.

Throughout the simulation, the SASA of both the native and mutant CCL2 proteins was monitored. The native protein exhibited a consistent SASA value, ranging from 55 to 65 nm² throughout the simulation period. In contrast, the mutant protein displayed a higher and more variable SASA, ranging from 60 to 75 nm², with significant fluctuations observed over time. This variation in SASA suggests that the mutant protein underwent substantial changes in its solvent accessibility, potentially affecting its stability and interactions with other molecules (Figure 6C and D).

During the simulation, the native protein maintained a stable SASA value, indicating its stability. Conversely, the mutant structures exhibited higher SASA values with notable fluctuations, suggesting increased solvent accessibility and alterations in its interactions with other molecules.

3.8. Hydrogen bonding analysis of native and mutant CCL2 proteins

MD simulations were used to monitor the interactions between CCR2 and CCL2. The majority of the important protein–peptide interactions were hydrogen bonds (Figure 7). A timeline of these interactions was generated to highlight the key hydrogen bonds formed during the simulations. The number of hydrogen bonds formed between CCR2 and CCL2 is an essential parameter to assess the stability of the complex. Our results revealed that the native CCL2–C59C complex exhibits a lower average number of hydrogen bonds (total hydrogen bonds: 3350) (Figure 7A). In contrast, the C59G mutation in CCL2 leads to an increase in hydrogen bond formation (total hydrogen bonds: 4195) (Figure 7B). This directly correlates with the binding affinity and stability of the receptor–ligand interaction, which suggests that the C59G mutation enhances receptor–ligand interactions by promoting stronger and more stable hydrogen bonding.

3.9. Effects of CCL2 mutants on the protein secondary and tertiary structures

The I-TASSER-modeled structure of CCL2 (native and mutant) was evaluated using UCSF Chimera calculations for homology modeling. This evaluation involved comparative modeling by satisfaction of spatial restraints, which requires a known related structure (i.e., a template) and a target-template sequence alignment. The

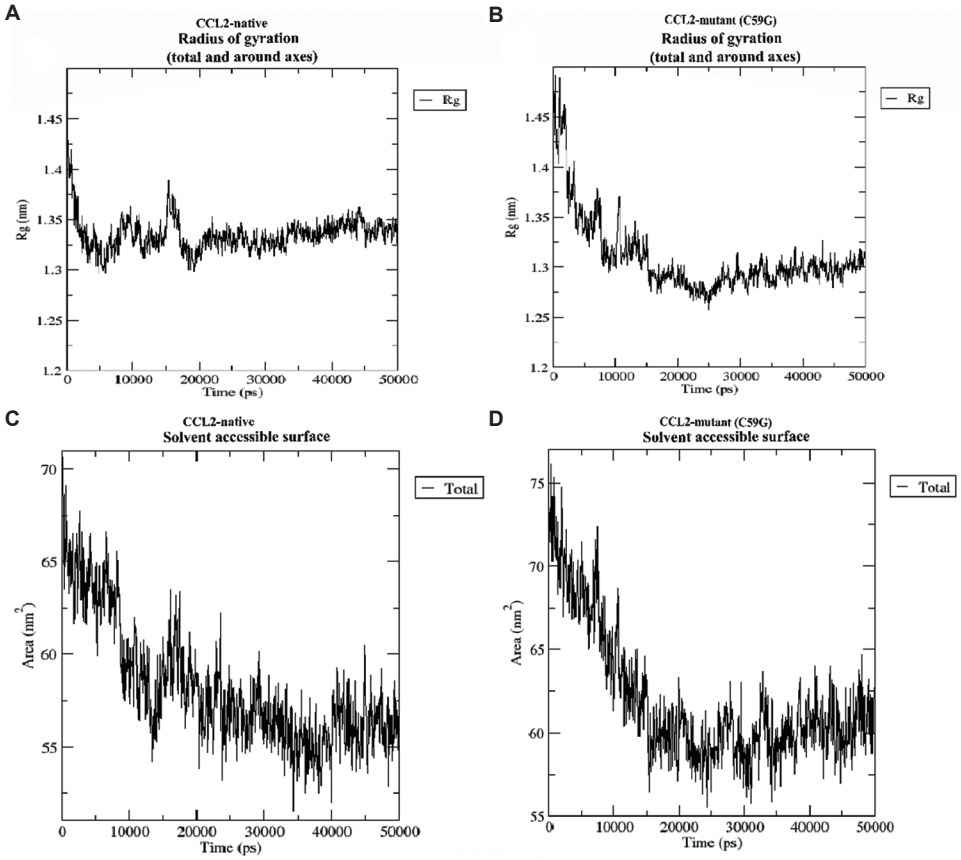


Figure 6. Rg and SASA analysis. (A and B) The Rg of Cα atoms in the native and mutant structures was monitored throughout the simulation. The native protein consistently maintained a stable Rg value, whereas the mutant structures displayed notable fluctuations, suggesting alterations in compactness. (C and D) SASA of both the native and mutant structures over time
Abbreviations: Rg: radius of gyration; SASA: solvent-accessible surface area; Cα: alpha carbon.

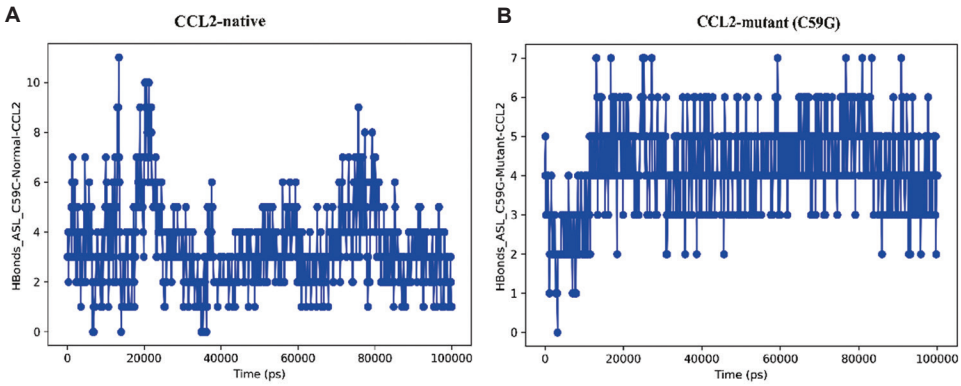


Figure 7. Protein contact hydrogen bond. (A) Representation of the hydrogen bonds of the native CCL2–C59C. (B) Representation of the hydrogen bonds of the mutant CCL2–C59G. Simulation time is presented in picoseconds.
Abbreviation: CCL2: Chemokine C-C motif ligand 2.

Chimera interface to Modeller allows for the generation of theoretical models of a protein using at least one known related structure and a target-template sequence alignment.[77] Homology modeling was performed on the native and mutant CCL2 proteins to investigate the spatial arrangement of the conserved amino acids, particularly cysteine (C59) and glycine (G59). This approach allowed for the assessment of how the C59G mutation affects

the local structural environment and the overall protein conformation. I-TASSER was used to generate a 3D model of the CCL2 protein based on the cysteine (C59) and glycine (G59) structures. The model was selected due to its high C-score, indicating a high confidence level in the predicted structure. This computational method enabled the prediction of the protein tertiary structure. According to I-TASSER modeling results, the predicted structure of the native CCL2 protein has a C-score of −1.20, a TM-score of 0.56 ± 0.15, and an RMSD of 6.4 ± 3.9 Å. Similarly, the mutant CCL2 protein was predicted to have a comparable structure, with a C-score of −1.10, a TM-score of 0.54 ± 0.14, and an RMSD of 6.2 ± 3.7 Å. These values suggest that both models have a moderate confidence level and a correct global topology, despite some local structural variations.

The mutation of a cysteine (C) to glycine (G) in the CCL2 protein substantially affected both its secondary and tertiary structures. This is mainly due to the differing side-chain properties of these amino acids, which can alter local interactions and disrupt hydrogen bonding networks. In addition, the mutation's impact on electrostatic interactions and hydrophobic forces can alter the protein's overall folding and stability. These structural alterations can significantly influence the protein's function, solubility, and susceptibility to aggregation.[78,79]

In the present study, we observed changes in the bond distances between the amino acids in the CCL2 protein. The native CCL2 structure with K58 and C59 had bond distances of 9.6 and 6.5 Å, respectively. In contrast, the mutant G59 structure with K58 and P60 had bond distances of 11.7 and 7.1 Å, respectively. The observed changes in bond distances suggest notable structural differences in the mutant protein compared to the native structure. The mutant protein displays structural variations compared to the native protein, as indicated by distance measurements and C-scores (Figure 8A and B). Cysteine at position 59 is a positively charged amino acid that participates in electrostatic interactions. Substituting cysteine with glycine eliminates this positive charge, disrupting electrostatic interactions. Furthermore, cysteine can form disulfide bonds, which can alter the protein's overall structure. These changes can significantly affect the protein's function, stability, and interactions with other molecules.[80,81] These findings highlight the structural variations in CCL2 among proteins and critical mutations, providing crucial insights into their functional implications.

3.10. Molecular docking of native and mutant CCL2 with CCR2

The molecular docking results of native and mutant CCL2 with the CCR2 receptor were obtained using the HDOCK online tool, which revealed significant insights into their binding interactions. Native CCL2 displayed binding affinity to CCR2, with the top 10 minimum docking scores being −176.10, −160.30, −159.82, −156.20, −155.16, −153.10, −152.99, −151.07, −150.53, and −150.21 kcal/mol (Figure 9A). The lowest score of −176.10 kcal/mol indicates the most favorable binding interaction. Similarly, mutant CCL2 also exhibited robust binding affinity to CCR2, with top docking scores of −175.86, −172.30, −168.52, −161.03, −160.64, −159.83, −159.64, −158.47, −155.39, and −153.79 kcal/mol (Figure 9B). Although the mutant CCL2 had slightly higher minimum docking scores than the native CCL2, the scores indicate strong binding interactions. The comparison of the native and mutant CCL2 forms revealed that the mutant CCL2 had a marginally better minimum docking score (−175.86 kcal/mol) than the native form (−176.10 kcal/mol) (Figure 9), suggesting a slightly higher binding affinity for CCR2. However, the overall similarity in the docking scores indicates that the mutation alters the binding affinity of CCL2 to CCR2. These findings suggest that mutant CCL2 has a strong binding affinity for the CCR2 receptor and that the mutation impacts CCL2's ability to interact with its receptor. Further experimental validation and structural analysis are required to fully elucidate the functional implications of this mutation on CCL2–CCR2 signaling.

4. Discussion

CCL2 interacts with CCR2 through its N-terminal β strand, which is a critical component of the chemokine receptor binding process.[82] The CCL2–CCR2 signaling pathway plays a crucial role in the pathophysiology of various disorders, including chronic obstructive pulmonary disease, asthma, cystic fibrosis, and cancer. In the context of cancer, the interaction between CCL2, which is released by cancer cells, and CCR2 in the tumor microenvironment is essential for the progression and metastasis of cancer.[9,83] Previous research has shown that CCL2 promotes tumor growth and metastasis by recruiting macrophages and immune cells to the tumor microenvironment.[84] The CCL2–CCR2 axis is a critical factor in the progression of various cancers, including prostate, pancreatic, breast, lung, kidney, bladder, and colorectal cancers. In prostate cancer, CCL2 promotes tumor growth and migration through the PI3K/Akt signaling pathway and contributes to resistance against castration and chemotherapy, making it a potential therapeutic target.[85] In pancreatic ductal adenocarcinoma, high levels of CCL2 correlate with a poor prognosis and facilitate angiogenesis, and CCR2 inhibitors have shown promise in enhancing treatment outcomes.[86] In breast cancer, CCL2 enhances cell migration and is implicated in triple-negative breast cancer, suggesting that targeting

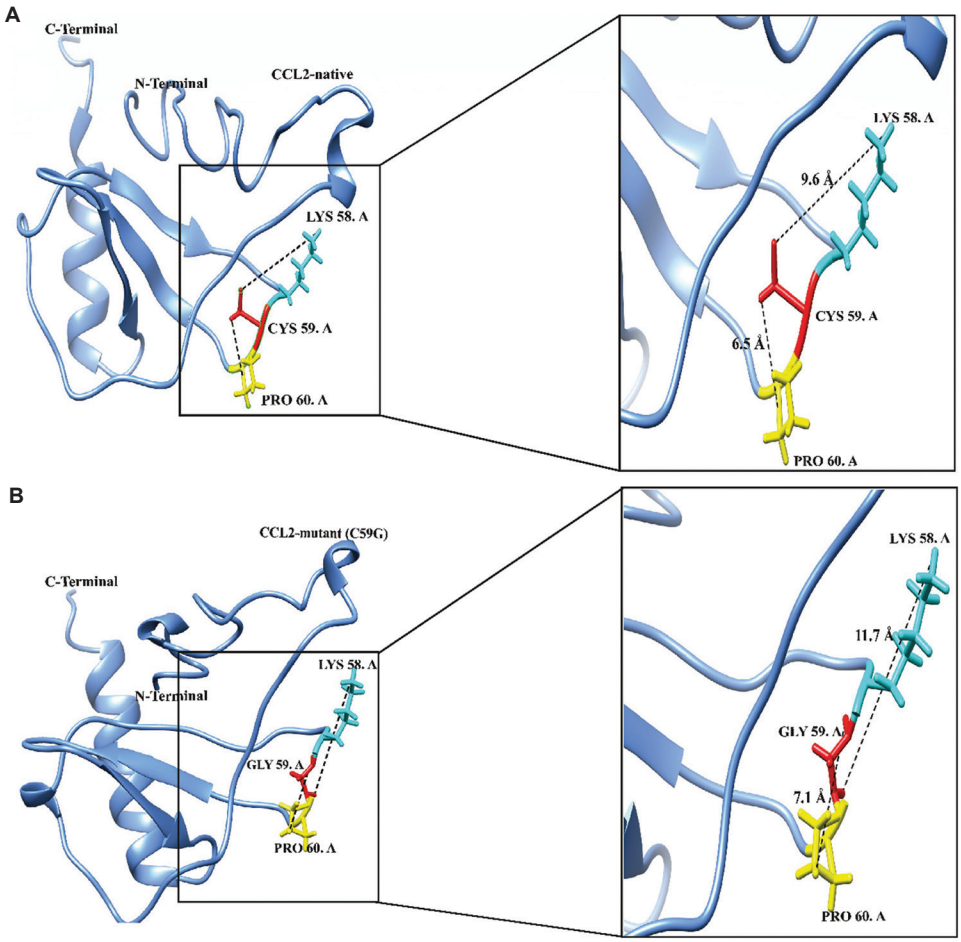
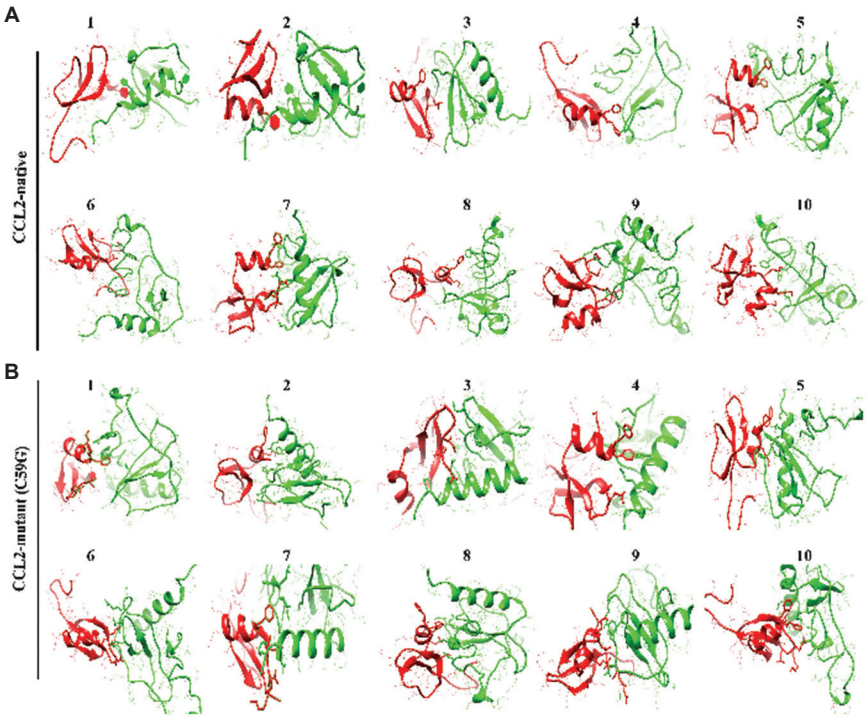


Figure 8. Mutation of cysteine (C59) to glycine (G59) in the CCL2 protein led to substantial structural changes. (A) Native CCL2 structure, where the bond distances between C59 and the neighboring amino acids K58 and P60 were 9.6 and 6.5 Å, respectively. (B) CCL2 mutant G59 structure, where these bond distances increased to 11.7 Å between G59 and K58 and 7.1 Å between G59 and P60. These changes in bond distances indicate that the mutation disrupted the local interactions and altered the overall folding of the protein.
Abbreviation: CCL2: Chemokine C-C motif ligand 2.

this axis helps overcome therapy resistance.[84] In lung cancer, CCL2 recruits myeloid-derived suppressor cells (MDSCs) and tumor-associated macrophages, reducing the effectiveness of immune checkpoint inhibitors, whereas its inhibition increases chemotherapy sensitivity.[87] In renal cell carcinoma, the elevated expression of CCL2 is linked to poor survival, promoting angiogenesis, and the accumulation of MDSCs.[88] CCL2 also enhances bladder cancer cell migration and infiltration, indicating that its blockade may improve responses to both chemotherapy and immunotherapy.[89] In colorectal cancer, CCL2 is involved in metastasis and angiogenesis, highlighting its potential as a prognostic marker and therapeutic target.[9] Overall, targeting the CCL2–CCR2 axis presents a promising strategy for cancer treatment across multiple tumor types. Our findings on the C59G mutation offer insights into its potential role in modulating CCL2 activity and influencing tumor progression. Further investigation in cohorts of patients with cancer is needed to validate this mutation's clinical relevance and potential as a therapeutic target.

CCL2 plays a critical role in macrophage recruitment and activation. Mutations that affect CCL2's ability to bind to its receptor CCR2 could alter macrophage behavior, impacting inflammatory responses and tumor-associated macrophage dynamics.[90] For example, mutations that impair CCL2 function might reduce macrophage infiltration into tumors, potentially affecting tumor growth and metastasis.[5] Monocytes migrate into tissues by activating integrins on their surface through the CCL2–CCR2 axis. This activation facilitates the processes of rolling and adhesion, which allow monocytes to pass through the vascular endothelium and enter the tissues.[91] The CCL2–CCR2 axis activates monocytes, macrophages,



Rank	Docking Score	Confidence Score	Ligand rmsd (Å)
1	-176.10	0.6276	94.85
2	-160.30	0.5513	96.67
3	-159.82	0.5489	85.09
4	-156.70	0.5310	93.02
5	-155.16	0.5238	103.75
6	-153.10	0.5155	74.74
7	-152.99	0.5149	87.08
8	-151.07	0.5053	101.22
9	-150.53	0.5026	117.04
10	-150.21	0.5010	83.08

Rank	Docking Score	Confidence Score	Ligand rmsd (Å)
1	-175.86	0.6265	97.01
2	-172.30	0.6097	83.29
3	-168.52	0.5916	83.48
4	-161.03	0.5549	89.16
5	-160.64	0.5530	86.36
6	-159.83	0.5490	115.62
7	-159.64	0.5481	88.42
8	-158.47	0.5422	106.61
9	-155.39	0.5269	115.72
10	-153.79	0.5189	115.28

Figure 9. Molecular docking results of native and mutant CCL2 with the CCR2 receptor using the HDOCK online tool. (A) The native form of CCL2 displayed a strong binding affinity to CCR2, with the top 10 minimum docking scores ranging from −176.10 to −150.21 kcal/mol. The lowest score of −176.10 kcal/mol indicates the most favorable binding interaction. (B) The mutant form of CCL2 also demonstrated robust binding affinity to CCR2, with top docking scores between −175.86 and −153.79 kcal/mol.
Abbreviations: CCL2: Chemokine C-C motif ligand 2; CCR2: C-C motif chemokine receptor 2.

memory T lymphocytes, and natural killer cells, resulting in the release of pro-inflammatory cytokines such as IL-1, IL-6, and TNF. In addition, macrophages activated by CCL2 secrete tissue repair factors, including vascular endothelial growth factor, platelet-derived growth factor, and transforming growth factor.[92] Functional assays are required to evaluate how these mutations influence activities such as macrophage migration and cytokine production. Understanding these effects could reveal new therapeutic opportunities for modulating macrophage responses in inflammatory diseases and cancer.

Using computational tools, our analysis of 83 mutations extracted from a public database predicted that ten of these mutations were likely to be deleterious and potentially impact the function of the CCL2 protein (Table 1). FATHMM analysis only identified one (C59G) mutation as a cancer-causing or cancer-promoting mutation, indicating its potential role in tumorigenesis (Table 3). The C59G mutation was selected for further analysis, and a conservation study revealed that the C and G positions were highly conserved, with a conservation score of 9 (Figure 2A and B). A previous study discovered that the double mutants R199H/D265H and R203H/D265H as well as the triple mutant R199H/R203H/D265H in CXCR1 were unable to trigger an IL-8-dependent calcium response.[93] Identifying the specific locus of a mutated amino acid can provide valuable insights into how the mutation alters the protein's overall structure and function.[94] Many studies have extensively explored how the location of mutations within a protein structure impacts its function. By examining the relationship between the mutation location and its effects on protein structure, stability, and function, these studies aim to elucidate disease development and progression.[95,96] Using I-TASSER, we modeled the structures of native and mutant CCL2 proteins (Figures 3-5) to perform a MD simulation analysis. This approach allowed us to investigate the structural changes caused by the mutation and its potential impact on the protein's function and overall structure. The RMSD data from the simulations revealed that the mutant CCL2 protein exhibited a distinct and significant pattern of deviation throughout the simulation period, differing from the behavior observed in the native protein (Figure 4A). The simulation results indicated that the mutation substantially destabilizes the CCL2 protein, resulting in structural changes distinct from those observed in the native protein, thereby impacting its stability and overall structure. The RMSF data supported the hypothesis that the mutation destabilizes the protein structure. RMSF

Table 3. Cancer association prediction of CCL2 mutations using the FATHMM server

S. No.	Cancer association			S. No.	Cancer association		
	Mutation	Prediction	Score		Mutation	Prediction	Score
1	M1L	Passenger/Other	3.83	43	S50N	Passenger/Other	3.33
2	K2E	Passenger/Other	3.71	44	Y51H	Passenger/Other	3.02
3	K2T	Passenger/Other	3.68	45	R53G	Passenger/Other	3.51
4	K2R	Passenger/Other	3.8	46	I54V	Passenger/Other	3.44
5	V3I	Passenger/Other	4.03	47	T55S	Passenger/Other	3.48
6	V3G	Passenger/Other	3.85	48	S56N	Passenger/Other	3.41
7	S4F	Passenger/Other	3.99	49	S56R	Passenger/Other	3.4
8	A5V	Passenger/Other	4.17	50	S57C	Passenger/Other	3.36
9	A5T	Passenger/Other	4.08	51	C59G	Cancer	−2.04
10	I13T	Passenger/Other	4.13	52	K61I	Passenger/Other	3.54
11	I13V	Passenger/Other	4.22	53	K67R	Passenger/Other	3.53
12	A14E	Passenger/Other	3.9	54	K67D	Passenger/Other	3.49
13	A15P	Passenger/Other	3.85	55	K67T	Passenger/Other	3.8
14	T16S	Passenger/Other	4.08	56	T68I	Passenger/Other	3.14
15	I18N	Passenger/Other	4.03	57	I69V	Passenger/Other	3.61
16	I18T	Passenger/Other	4.1	58	A71T	Passenger/Other	3.78
17	P19L	Passenger/Other	4.13	59	A71V	Passenger/Other	3.74
18	P19Q	Passenger/Other	4.13	60	E73G	Passenger/Other	3.54
19	Q20R	Passenger/Other	4.07	61	I74N	Passenger/Other	3.35
20	A23T	Passenger/Other	3.97	62	D77E	Passenger/Other	3.43
21	A23S	Passenger/Other	4.06	63	D77G	Passenger/Other	3.43
22	A27G	Passenger/Other	3.72	64	P78S	Passenger/Other	2.23
23	A27V	Passenger/Other	3.71	65	P78H	Passenger/Other	2.19
24	N29K	Passenger/Other	3.74	66	K79E	Passenger/Other	3.58
25	P31A	Passenger/Other	3.63	67	K79R	Passenger/Other	3.42
26	Y36C	Passenger/Other	3.53	68	K81N	Passenger/Other	3.44
27	N37D	Passenger/Other	3.52	69	W82L	Passenger/Other	2.93
28	T39I	Passenger/Other	3.73	70	V83A	Passenger/Other	2.91
29	T39A	Passenger/Other	3.71	71	V83L	Passenger/Other	2.96
30	N40D	Passenger/Other	3.79	72	Q84R	Passenger/Other	3.6
31	N40K	Passenger/Other	3.83	73	D85Y	Passenger/Other	3.45
32	R41K	Passenger/Other	3.6	74	M87V	Passenger/Other	3.49
33	K42I	Passenger/Other	3.34	75	M87K	Passenger/Other	3.3
34	S44L	Passenger/Other	3.53	76	D88N	Passenger/Other	3.64
35	V45M	Passenger/Other	3.49	77	D88V	Passenger/Other	3.56
36	V45A	Passenger/Other	3.54	78	H89P	Passenger/Other	3.53
37	R47K	Passenger/Other	3.57	79	K92R	Passenger/Other	3.66
38	L48F	Passenger/Other	3.37	80	T96A	Passenger/Other	3.72
39	A49E	Passenger/Other	3.72	81	P97L	Passenger/Other	3.87
40	A49V	Passenger/Other	3.63	82	P97T	Passenger/Other	3.83
41	A49T	Passenger/Other	3.57	83	T99S	Passenger/Other	4.1
42	A49P	Passenger/Other	3.52				

analysis revealed distinct fluctuation patterns in the mutant protein compared to the native structure, indicating that the mutation significantly impacted protein stability and dynamics (Figure 4B). The analysis showed a substantial increase in residue-level variation in the mutant protein, suggesting that the mutation caused notable changes in the protein's conformation and increased its flexibility.

Analysis of the intramolecular hydrogen bonds in the CCL2 protein revealed a notable decrease in the number of hydrogen bonds formed in the mutant protein compared to the native structure during the simulation. This reduction suggests an increase in the flexibility of the mutant protein and significant structural differences between the two forms (Figure 5A and B). To elucidate the structural dynamics of the native and mutant CCL2 proteins, an ED analysis was performed. This approach enabled the identification of correlated movements within both protein structures, providing insights into the molecular mechanisms underlying their functional differences. Examination of the sum of the eigenvalues revealed that the primary fluctuations were confined to the first two eigenvectors for both the native and mutant proteins. This finding suggests that these PCs capture the main motions within the system (Figure 5C and D).

Analysis of the Rg plot indicated that the amino acid substitution at position G59 had a destabilizing effect on the protein's structure, which led to reduced protein compactness (Figure 6A and B). This finding was supported by SASA analysis, which revealed higher SASA values for the mutant protein, indicating a substantial structural transition (Figure 6C and D). The observed conformational changes strongly suggest that the amino acid substitution in the mutant protein significantly altered its overall structure. Furthermore, the mutation resulted in a less stable, more flexible, and less compact structure compared to the native protein. Our results collectively indicate that the mutation had a profound impact on protein structure and stability.

Previous studies have shown the synergistic value of combining *in silico* approaches and wet laboratory experiments to gain a comprehensive understanding of biological processes.[97,98] The combination of computational methods and experimental techniques has emerged as a potent strategy for predicting and confirming hypotheses across various biological systems.[98] Integrating computational mutation predictions with MD simulation analysis has been pivotal in pinpointing disease-causing mutations.[99] This approach has allowed for the identification of the most harmful mutations from a large pool of mutations, offering valuable insights into the molecular mechanisms underlying various diseases.[100] Computational methods are essential in providing the foundation for genetic research and enabling the understanding of the molecular basis of diseases.[101] Moreover, computational methods play a crucial role in identifying disease-causing mutations and exploring potential therapeutic strategies.[102,103] These methods also aid in the clinical assessment of genetic variants, enabling more informed and effective treatment decisions.

5. Conclusion

This investigation into the oncogenic missense mutations of *CCL2*, particularly the C59G mutation, revealed that this specific mutation may impair CCL2 activity, potentially lower the risk of cancer, and enhance the immune response. Through bioinformatics and computational analyses, we found that the C59G mutation significantly affects the stability and function of CCL2, potentially hindering its interaction with the CCR2 receptor. This mutation, one of the 83 *CCL2* missense mutations, is associated with disease causation and has implications for cancer development. MD simulations and structural analyses confirmed that these mutations induce substantial conformational changes, decrease stability, and alter the ED. Despite the increased binding affinities of mutant CCL2 to CCR2, the C59G mutation markedly affects the protein's stability and function. This study emphasizes the importance of understanding the molecular characteristics of CCL2 mutations, particularly C59G, in influencing disease processes, such as that of colorectal cancer, and highlights the potential implications for immune responses and cancer regulation.

Acknowledgments

None.

Funding

This work was supported by National Natural Science Foundation of China (32270438, 32170498, 31970388), the National Key Research and Development Program of China (2021YFF0702000, 2018YFD0900602), 1.3+.5 project for disciplines of excellence, West China Hospital, Sichuan University (ZYJC21050), the Science and Technology Department of Sichuan Province (2022YFH0116), the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), the National Clinical Research Center for Geriatrics, West China Hospital, and Sichuan University (Z2023JC003).

Conflict of interest

Mohammad Amjad Kamal is the Guest Editor of this special issue, but was not in any way involved in the editorial and peer-review process conducted for this paper,

directly or indirectly. Separately, other authors declared that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Author contributions

Conceptualization: Shah Kamal, Cheng Deng, Shanshan Lai
Formal analysis: Shah Kamal, Amanullah Amanullah, Najeeb Ullah, Mariam Ahmed Mujtaba, Kashif Ali Khan
Investigation: Shah Kamal, Mohammad Amjad Kamal, Shanshan Lai, Cheng Dend
Methodology: Najeeb Ullah, Shah Kamal, Mariam Ahmed Mujtaba, Kashif Ali Khan
Writing–original draft: Shah Kamal, Cheng Deng, Mohammad Amjad Kamal, Shanshan Lai
Writing–review & editing: Shanshan Lai, Mohammad Amjad Kamal, Najeeb Ullah, Cheng Deng

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

All datasets generated for this study are available from the corresponding authors on reasonable request.

References

1. Zhu S, Liu M, Bennett S, Wang Z, Pfleger KDG, Xu J. The molecular structure and role of CCL2 (MCP-1) and C-C chemokine receptor CCR2 in skeletal biology and diseases. *J Cell Physiol*. 2021;236(10):7211-7222.

   doi: 10.1002/jcp.30375

2. Schubert M, Bleuler-Martinez S, Butschi A, *et al*. Plasticity of the β-trefoil protein fold in the recognition and control of invertebrate predators and parasites by a fungal defence system. *PLoS Pathog*. 2012;8(5):e1002706.

   doi: 10.1371/journal.ppat.1002706

3. Joshi N, Kumar D, Poluri KM. Elucidating the molecular interactions of chemokine CCL2 orthologs with flavonoid baicalin. *ACS Omega*. 2020;5(35):22637-22651.

   doi: 10.1021/acsomega.0c03428

4. Shao Z, Tan Y, Shen Q, *et al*. Molecular insights into ligand recognition and activation of chemokine receptors CCR2 and CCR3. *Cell Discov*. 2022;8(1):44.

   doi: 10.1038/s41421-022-00403-4

5. Li H, Wu M, Zhao X. Role of chemokine systems in cancer and inflammatory diseases. *MedComm* (*2020*). 2022;3(2):e147.

   doi: 10.1002/mco2.147

6. Turner MD, Nedjai B, Hurst T, Pennington DJ. Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease. *Biochim Biophys Acta*. 2014;1843(11):2563-2582.

   doi: 10.1016/j.bbamcr.2014.05.014

7. White GE, Iqbal AJ, Greaves DR. CC chemokine receptors and chronic inflammation-therapeutic opportunities and pharmacological challenges. *Pharmacol Rev*. 2013;65(1):47-89.

   doi: 10.1124/pr.111.005074

8. Dommel S, Blüher M. Does C-C motif chemokine ligand 2 (CCL2) link obesity to a pro-inflammatory state? *Int J Mol Sci*. 2021;22(3):1500.

   doi: 10.3390/ijms22031500

9. Xu M, Wang Y, Xia R, Wei Y, Wei X. Role of the CCL2-CCR2 signalling axis in cancer: Mechanisms and therapeutic targeting. *Cell Prolif*. 2021;54(10):e13115.

   doi: 10.1111/cpr.13115

10. Ben-Baruch A. Host microenvironment in breast cancer development: Inflammatory cells, cytokines and chemokines in breast cancer progression: Reciprocal tumor-microenvironment interactions. *Breast Cancer Res*. 2002;5(1):31.

    doi: 10.1186/bcr554

11. Yoshimura T, Li C, Wang Y, Matsukawa A. The chemokine monocyte chemoattractant protein-1/CCL2 is a promoter of breast cancer metastasis. *Cell Mol Immunol*. 2023;20(7):714-38.

    doi: 10.1038/s41423-023-01013-0

12. Soria G, Yaal-Hahoshen N, Azenshtein E, *et al*. Concomitant expression of the chemokines RANTES and MCP-1 in human breast cancer: A basis for tumor-promoting interactions. *Cytokine*. 2008;44(1):191-200.

    doi: 10.1016/j.cyto.2008.08.002

13. Furukawa S, Soeda S, Kiko Y, *et al*. MCP-1 promotes invasion and adhesion of human ovarian cancer cells. *Anticancer Res*. 2013;33(11):4785-4790.

14. Hefler L, Tempfer C, Heinze G, *et al*. Monocyte chemoattractant protein-1 serum levels in ovarian cancer patients. *Br J Cancer*. 1999;81(5):855-829.

    doi: 10.1038/sj.bjc.6690776

15. Bhat AA, Nisar S, Maacha S, *et al*. Cytokine-chemokine network driven metastasis in esophageal cancer; promising avenue for targeted therapy. *Mol Cancer*. 2021;20:2.

    doi: 10.1186/s12943-020-01294-3

16. Singh S, Anshita D, Ravichandiran V. MCP-1: Function, regulation, and involvement in disease. *Int*

Immunopharmacol. 2021;101(Pt B):107598.

doi: 10.1016/j.intimp.2021.107598

17. Kitadai Y. Cancer-stromal cell interaction and tumor angiogenesis in gastric cancer. *Cancer Microenviron*. 2010;3:109-116.

doi: 10.1007/s12307-009-0032-9

18. Tesch GH, Schwarting A, Kinoshita K, Lan HY, Rollins BJ, Kelley VR. Monocyte chemoattractant protein-1 promotes macrophage-mediated tubular injury, but not glomerular injury, in nephrotoxic serum nephritis. *J Clin Invest*. 1999;103(1):73-80.

doi: 10.1172/jci4876

19. Niiya M, Niiya K, Kiguchi T, *et al*. Induction of TNF-alpha, uPA, IL-8 and MCP-1 by doxorubicin in human lung carcinoma cells. *Cancer Chemother Pharmacol*. 2003;52(5):391-398.

doi: 10.1007/s00280-003-0665-1

20. Huang S, Singh RK, Xie K, *et al*. Expression of the JE/MCP-1 gene suppresses metastatic potential in murine colon carcinoma cells. *Cancer Immunol Immunother*. 1994;39(4):231-238.

doi: 10.1007/bf01525986

21. Pandurangan AP, Blundell TL. Prediction of impacts of mutations on protein structure and interactions: SDM, a statistical approach, and mCSM, using machine learning. *Protein Sci*. 2020;29(1):247-257.

doi: 10.1002/pro.3774

22. Capriotti E, Altman RB, Bromberg Y. Collective judgment predicts disease-associated single nucleotide variants. *BMC Genomics*. 2013;14 Suppl 3:S2.

doi: 10.1186/1471-2164-14-s3-s2

23. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013;Chapter 7:Unit7.20.

doi: 10.1002/0471142905.hg0720s76

24. López-Ferrando V, Gazzo A, de la Cruz X, Orozco M, Gelpí JL. PMut: A web-based tool for the annotation of pathological variants on proteins, 2017 update. *Nucleic Acids Res*. 2017;45(W1):W222-W228.

doi: 10.1093/nar/gkx313

25. Rogers MF, Shihab HA, Mort M, Cooper DN, Gaunt TR, Campbell C. FATHMM-XF: Accurate prediction of pathogenic point mutations via extended features. *Bioinformatics*. 2018;34(3):511-513.

doi: 10.1093/bioinformatics/btx536

26. Pires DE, Ascher DB, Blundell TL. mCSM: Predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics*. 2014;30(3):335-342.

doi: 10.1093/bioinformatics/btt691

27. Worth CL, Preissner R, Blundell TL. SDM--a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res*. 2011;39(Web Server issue):W215-W222.

doi: 10.1093/nar/gkr363

28. Pires DE, Ascher DB, Blundell TL. DUET: A server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res*. 2014;42(Web Server issue):W314-W319.

doi: 10.1093/nar/gku411

29. Yang J, Zhang Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res*. 2015;43(W1):W174-W181.

doi: 10.1093/nar/gkv342

30. Gorai S, Junghare V, Kundu K, *et al*. Synthesis of dihydrobenzofuro[3,2-b]chromenes as potential 3CLpro inhibitors of SARS-CoV-2: A molecular docking and molecular dynamics study. *ChemMedChem*. 2022;17(8):e202100782.

doi: 10.1002/cmdc.202100782

31. Farajzadeh-Dehkordi M, Mafakher L, Harifi A, Samiee-Rad F, Rahmani B. Computational analysis of the functional and structural impact of the most deleterious missense mutations in the human Protein C. *PLoS One*. 2023;18(11):e0294417.

doi: 10.1371/journal.pone.0294417

32. Martin FJ, Amode MR, Aneja A, *et al*. Ensembl 2023. *Nucleic Acids Res*. 2023;51(D1):D933-D941.

doi: 10.1093/nar/gkac958

33. Sherry ST, Ward MH, Kholodov M, *et al*. dbSNP: The NCBI database of genetic variation. *Nucleic Acids Res*. 2001;29(1):308-311.

doi: 10.1093/nar/29.1.308

34. UniProt Consortium. UniProt: A hub for protein information. *Nucleic Acids Res*. 2015;43(Database issue):D204-D212.

doi: 10.1093/nar/gku989

35. Goodsell DS, Zardecki C, Di Costanzo L, *et al*. RCSB Protein Data Bank: Enabling biomedical research and drug discovery. *Protein Sci*. 2020;29(1):52-65.

doi: 10.1002/pro.3730

36. Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants. *BMC Genomics*. 2015;16 Suppl 8(Suppl 8):S1.

doi: 10.1186/1471-2164-16-s8-s1

37. Mi H, Lazareva-Ulitsky B, Loo R, *et al*. The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res*. 2005;33(Database issue):D284-D288.

doi: 10.1093/nar/gki078

38. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003;31(13):3812-3814.

doi: 10.1093/nar/gkg509

39. Bendl J, Stourac J, Salanda O, *et al*. PredictSNP: Robust and accurate consensus classifier for prediction of disease-related mutations. *PLoS Comput Biol*. 2014;10(1):e1003440.

doi: 10.1371/journal.pcbi.1003440

40. Pejaver V, Mooney SD, Radivojac P. Missense variant pathogenicity predictors generalize well across a range of function-specific prediction challenges. *Hum Mutat*. 2017;38(9):1092-108.

doi: 10.1002/humu.23258

41. Adzhubei IA, Schmidt S, Peshkin L, *et al*. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.

doi: 10.1038/nmeth0410-248

42. Vila JA. Proteins' evolution upon point mutations. *ACS Omega*. 2022;7(16):14371-14376.

doi: 10.1021/acsomega.2c01407

43. Pandurangan AP, Ochoa-Montaño B, Ascher DB, Blundell TL. SDM: A server for predicting effects of mutations on protein stability. *Nucleic Acids Res*. 2017;45(W1):W229-W235.

doi: 10.1093/nar/gkx439

44. Choudhury A, Mohammad T, Anjum F, *et al*. Comparative analysis of web-based programs for single amino acid substitutions in proteins. *PLoS One*. 2022;17(5):e0267084.

doi: 10.1371/journal.pone.0267084

45. Rodrigues CHM, Myung Y, Pires DEV, Ascher DB. mCSM-PPI2: Predicting the effects of mutations on protein-protein interactions. *Nucleic Acids Res*. 2019;47(W1):W338-W344.

doi: 10.1093/nar/gkz383

46. Rodrigues CHM, Pires DEV, Ascher DB. mmCSM-PPI: Predicting the effects of multiple point mutations on protein-protein interactions. *Nucleic Acids Res*. 2021;49(W1):W417-W424.

doi: 10.1093/nar/gkab273

47. Chen CW, Lin J, Chu YW. iStable: Off-the-shelf predictor integration for predicting protein stability changes. *BMC Bioinformatics*. 2013;14 Suppl 2(Suppl 2):S5.

doi: 10.1186/1471-2105-14-s2-s5

48. Panchal NK, Bhale A, Verma VK, Beevi SS. Computational and molecular dynamics simulation approach to analyze the impact of XPD gene mutation on protein stability and function. 2020.

doi: 10.1101/2020.07.18.209841

49. Shihab HA, Gough J, Cooper DN, Day IN, Gaunt TR. Predicting the functional consequences of cancer-associated amino acid substitutions. *Bioinformatics*. 2013;29(12):1504-1510.

doi: 10.1093/bioinformatics/btt182

50. Ashkenazy H, Abadi S, Martz E, *et al*. ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res*. 2016;44(W1):W344-W350.

doi: 10.1093/nar/gkw408

51. Ishimoto N, Park JH, Kawakami K, *et al*. Structural basis of CXC chemokine receptor 1 ligand binding and activation. *Nat Commun*. 2023;14(1):4107.

doi: 10.1038/s41467-023-39799-2

52. Lovell SC, Davis IW, Arendall WB 3rd, *et al*. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*. 2003;50(3):437-50.

doi: 10.1002/prot.10286

53. Konopka BM, Nebel JC, Kotulska M. Quality assessment of protein model-structures based on structural and functional similarities. *BMC Bioinformatics*. 2012;13(1):242.

doi: 10.1186/1471-2105-13-242

54. Ravikumar A, Ramakrishnan C, Srinivasan N. Stereochemical assessment of (ϕ,ψ) outliers in protein structures using bond geometry-specific Ramachandran steric-maps. *Structure*. 2019;27:1875-1884.

doi: 10.1016/j.str.2019.09.009

55. Durrant JD, McCammon JA. Molecular dynamics simulations and drug discovery. *BMC Biol*. 2011;9(1):71.

doi: 10.1186/1741-7007-9-71

56. Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. *Neuron*. 2018;99(6):1129-1143.

doi: 10.1016/j.neuron.2018.08.011

57. Muhammed MT, Aki-Yalcin E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem Biol Drug Des*. 2019;93(1):12-20.

doi: 10.1111/cbdd.13388

58. Binbay FA, Rathod DC, George AAP, Imhof D. Quality assessment of selected protein structures derived from homology modeling and AlphaFold. *Pharmaceuticals* (*Basel*). 2023;16(12):1662.

doi: 10.3390/ph16121662

59. Brooks BR, Brooks CL 3rd, Mackerell AD Jr., *et al*. CHARMM: The biomolecular simulation program. *J Comput Chem*. 2009;30(10):1545-1614.

doi: 10.1002/jcc.21287

60. Vieira IHP, Botelho EB, de Souza Gomes TJ, Kist R, Caceres RA, Zanchi FB. Visual dynamics: A WEB application for molecular dynamics simulation using GROMACS. *BMC Bioinformatics*. 2023;24(1):107.

doi: 10.1186/s12859-023-05234-y

61. Bagewadi ZK, Khan TM, Gangadharappa B, Kamalapurkar A, Shamsudeen SM, Yaraguppi DM. Molecular dynamics and simulation analysis against superoxide dismutase (SOD) target of *Micrococcus luteus* with secondary metabolites from *Bacillus licheniformis* recognized by genome mining approach. *Saudi J Biol Sci*. 2023;30:103753.

doi: 10.1016/j.sjbs.2023.103753

62. Shahraki O, Zargari F, Edraki N, Khoshneviszadeh M, Firuzi O. Molecular dynamics simulation and molecular docking studies of 1,4- dihydropyridines as P-glycoprotein's allosteric inhibitors. *J Biomol Struct Dyn*. 2016;36:112-125.

doi: 10.1080/07391102.2016.1268976

63. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. *Molecules*. 2015;20(7):13384-13421.

doi: 10.3390/molecules200713384

64. Hildebrand PW, Rose AS, Tiemann JKS. Bringing molecular dynamics simulation data into view. *Trends Biochem Sci*. 2019;44(11):902-913.

doi: 10.1016/j.tibs.2019.06.004

65. Tumskiy RS, Tumskaia AV. Multistep rational molecular design and combined docking for discovery of novel classes of inhibitors of SARS-CoV-2 main protease 3CLpro. *Chem Phys Lett*. 2021;780:138894.

doi: 10.1016/j.cplett.2021.138894

66. Carter P, Andersen CA, Rost B. DSSPcont: Continuous secondary structure assignments for proteins. *Nucleic Acids Res*. 2003;31(13):3293-3295.

doi: 10.1093/nar/gkg626

67. Yan Y, Wen Z, Wang X, Huang SY. Addressing recent docking challenges: A hybrid strategy to integrate template-based and free protein-protein docking. *Proteins*. 2016;85:497-512.

doi: 10.1002/prot.25234

68. Case DA, Cheatham TE 3rd, Darden T, *et al*. The Amber biomolecular simulation programs. *J Comput Chem*. 2005;26(16):1668-1688.

doi: 10.1002/jcc.20290

69. Huang SY, Zou X. MDockPP: A hierarchical approach for protein-protein docking and its application to CAPRI rounds 15-19. *Proteins*. 2010;78(15):3096-3103.

doi: 10.1002/prot.22797

70. Huang SY, Yan C, Grinter SZ, Chang S, Jiang L, Zou X. Inclusion of the orientational entropic effect and low-resolution experimental information for protein-protein docking in Critical Assessment of PRedicted Interactions (CAPRI). *Proteins*. 2013;81(12):2183-2191.

doi: 10.1002/prot.24435

71. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct*. 2000;29:291-325.

doi: 10.1146/annurev.biophys.29.1.291

72. Shah K, Amanullah A, Qingqing W, *et al*. Bioinformatics analysis of missense mutations in CXCR1 implicates altered protein stability and function. *Tumor Discov*. 2024;3:2512.

doi: 10.36922/td.2512

73. Sobitan A, Edwards W, Jalal MS, *et al*. Prediction of the effects of missense mutations on human myeloperoxidase protein stability using *in silico* saturation mutagenesis. *Genes* (*Basel*). 2022;13(8):1412.

doi: 10.3390/genes13081412

74. Yang J, Zhang Y. Protein structure and function prediction using I-TASSER. *Curr Protoc Bioinformatics*. 2015;52:5.8.1-5.8.15.

doi: 10.1002/0471250953.bi0508s52

75. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. 2008;9:40.

doi: 10.1186/1471-2105-9-40

76. Sobolev OV, Afonine PV, Moriarty NW, *et al*. A global Ramachandran score identifies protein structures with unlikely stereochemistry. *Structure*. 2020;28(11):1249-1258.e2.

doi: 10.1016/j.str.2020.08.005

77. Yang Z, Lasker K, Schneidman-Duhovny D, *et al*. UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *J Struct Biol*. 2012;179(3):269-278.

doi: 10.1016/j.jsb.2011.09.006

78. Wang L, Ding MY, Wang J, Gao JG, Liu RM, Li HT. Effects of site-directed mutagenesis of cysteine on the structure of sip proteins. *Front Microbiol*. 2022;13:805325.

doi: 10.3389/fmicb.2022.805325

79. Konrat R, Kräutler B, Weiskirchen R, Bister K. Structure of cysteine-and glycine-rich protein CRP2. Backbone dynamics reveal motional freedom and independent spatial orientation of the lim domains. *J Biol Chem.* 1998;273(36):23233-23240.

doi: 10.1074/jbc.273.36.23233

80. Qiu H, Honey DM, Kingsbury JS, *et al*. Impact of cysteine variants on the structure, activity, and stability of recombinant human α-galactosidase A. *Protein Sci*. 2015;24(9):1401-1411.

doi: 10.1002/pro.2719

81. Gerlza T, Winkler S, Atlic A, *et al*. Designing a mutant CCL2-

HSA chimera with high glycosaminoglycan-binding affinity and selectivity. *Protein Eng Des Sel*. 2015;28(8):231-240.

doi: 10.1093/protein/gzv025

82. Hao Q, Vadgama JV, Wang P. CCL2/CCR2 signaling in cancer pathogenesis. *Cell Commun Signal*. 2020;18(1):82.

doi: 10.1186/s12964-020-00589-8

83. Lim SY, Yuzhalin AE, Gordon-Weeks AN, Muschel RJ. Targeting the CCL2-CCR2 signaling axis in cancer metastasis. *Oncotarget*. 2016;7(19):28697-28710.

doi: 10.18632/oncotarget.7376

84. Kadomoto S, Izumi K, Mizokami A. Roles of CCL2-CCR2 axis in the tumor microenvironment. *Int J Mol Sci*. 2021;22(16):8530.

doi: 10.3390/ijms22168530

85. Iwamoto H, Izumi K, Mizokami A. Is the C-C motif ligand 2–C-C chemokine receptor 2 axis a promising target for cancer therapy and diagnosis? *Int J Mol Sci*. 2020;21(23):9328.

doi: 10.3390/ijms21239328

86. Anderson EM, Thomassian S, Gong J, Hendifar A, Osipov A. Advances in pancreatic ductal adenocarcinoma treatment. *Cancers* (*Basel*). 2021;13(21):5510.

doi: 10.3390/cancers13215510

87. Li X, Zhong J, Deng X, *et al*. Targeting myeloid-derived suppressor cells to enhance the antitumor efficacy of immune checkpoint blockade therapy. *Front Immunol*. 2021;12:754196.

doi: 10.3389/fimmu.2021.754196

88. Chen W, Pan X, Cui X. RCC immune microenvironment subsequent to targeted therapy: A friend or a foe? *Front Oncol*. 2020;10:573690.

doi: 10.3389/fonc.2020.573690

89. Jin J, Lin J, Xu A, *et al*. CCL2: An important mediator between tumor cells and host cells in tumor microenvironment. *Front Oncol*. 2021;11:722916.

doi: 10.3389/fonc.2021.722916

90. Toya M, Zhang N, Tsubosaka M, *et al*. CCL2 promotes osteogenesis by facilitating macrophage migration during acute inflammation. *Front Cell Dev Biol*. 2023;11:1213641.

doi: 10.3389/fcell.2023.1213641

91. Butcher MJ, Galkina EV. wRAPping up early monocyte and neutrophil recruitment in atherogenesis via Annexin A1/FPR2 signaling. *Circ Res*. 2015;116(5):774-777.

doi: 10.1161/circresaha.115.305920

92. Deshmane SL, Kremlev S, Amini S, Sawaya BE. Monocyte chemoattractant protein-1 (MCP-1): An overview. *J Interferon Cytokine Res*. 2009;29(6):313-326.

doi: 10.1089/jir.2008.0027

93. Suetomi K, Rojo D, Navarro J. Identification of a signal transduction switch in the chemokine receptor CXCR1. *J Biol Chem*. 2002;277(35):31563-31566.

doi: 10.1074/jbc.M204713200

94. Studer RA, Dessailly BH, Orengo CA. Residue mutations and their impact on protein structure and function: Detecting beneficial and pathogenic changes. *Biochem J*. 2013;449:581-594.

doi: 10.1042/BJ20121221

95. Gerasimavicius L, Livesey BJ, Marsh JA. Loss-of-function, gain-of-function and dominant-negative mutations have profoundly different effects on protein structure. *Nat Commun*. 2022;13(1):3895.

doi: 10.1038/s41467-022-31686-6

96. Arshad M, Bhatti A, John P. Identification and in silico analysis of functional SNPs of human TAGAP protein: A comprehensive study. *PLoS One*. 2018;13(1):e0188143.

doi: 10.1371/journal.pone.0188143

97. Schwaigerlehner L, Pechlaner M, Mayrhofer P, Oostenbrink C, Kunert R. Lessons learned from merging wet lab experiments with molecular simulation to improve mAb humanization. *Protein Eng Des Sel*. 2018;31(7-8):257-265.

doi: 10.1093/protein/gzy009

98. Cárdenas R, Martínez-Seoane J, Amero C. Combining experimental data and computational methods for the non-computer specialist. *Molecules*. 2020;25(20):4783.

doi: 10.3390/molecules25204783

99. Kamaraj B, Purohit R. Computational screening of disease-associated mutations in OCA2 gene. *Cell Biochem Biophys*. 2013;68:97-109.

doi: 10.1007/s12013-013-9697-2

100. Stefl S, Nishi H, Petukh M, Panchenko A, Alexov E. Molecular mechanisms of disease-causing missense mutations. *J Mol Biol*. 2013;425:3919-3936.

doi: 10.1016/j.jmb.2013.07.014

101. Peterson TA, Doughty E, Kann MG. Towards precision medicine: Advances in computational approaches for the analysis of human variants. *J Mol Biol*. 2013;425(21):4047-4063.

doi: 10.1016/j.jmb.2013.08.008

102. Mugumbate G, Nyathi B, Zindoga A, Munyuki G. Application of computational methods in understanding mutations in *Mycobacterium tuberculosis* drug resistance. *Front Mol Biosci*. 2021;8:643849.

doi: 10.3389/fmolb.2021.643849

103. Luo P, Chen B, Liao B, Wu FX. Predicting disease-associated genes: Computational methods, databases, and evaluations. *WIREs Data Min Knowl Discov*. 2020;11:e1383.

doi: 10.1002/widm.1383