


 Cite this: *RSC Adv.*, 2025, **15**, 41648

# Structure-based design of orthosteric and allosteric CCR2 inhibitors for potential IPF therapy

 Gaofei Yan,<sup>†</sup> Jianliang Huang,<sup>†</sup> Ziyuan Huang,<sup>†</sup> Shufang Luo,<sup>f</sup> Haiying Guo,<sup>f</sup> Shun Tao,<sup>b</sup> Enping Li,<sup>ab</sup> Ying Liu,<sup>ab</sup> Jianguan Huang,<sup>c</sup> Yanlin Liu,<sup>e</sup> Zimeng Liu,<sup>h</sup> Shuang Zhao,<sup>\*a</sup> Jing Chen<sup>\*g</sup> and Mingsheng Lei<sup>\*ab</sup>

Idiopathic pulmonary fibrosis (IPF) is a fatal respiratory disease with an extremely poor prognosis. Our previous studies revealed that CCR2 (C–C chemokine receptor type 2) expression is significantly upregulated in bleomycin-induced pulmonary fibrosis in mice. Moreover, bioinformatics analysis indicated that higher CCR2 expression is associated with poorer prognosis in patients. Therefore, we proposed a novel dual-target intervention strategy focused on the orthosteric and allosteric binding sites of the CCR2 receptor. By integrating structure-based pharmacophore modeling, 3D-QSAR, common feature pharmacophore modeling, and large-scale virtual screening (covering 152 406 molecules), we successfully identified two candidate small molecules, compound **17** and compound **67**, that exhibit high site selectivity. Molecular dynamics (MD) simulations, principal component analysis (PCA), and potential energy surface analyses *via* umbrella sampling confirmed that both compounds attain stable binding conformations at their respective target sites. MM/PBSA calculations revealed that compound **17** binds at the orthosteric site with a free energy of  $-30.91$  kcal mol<sup>-1</sup>, while compound **67** binds at the allosteric site with a free energy of  $-26.11$  kcal mol<sup>-1</sup>. Surface plasmon resonance (SPR) confirmed compound **17**'s direct binding to murine CCR2 ( $K_D = 3.46$  μM), while co-administration with compound **67** synergistically enhanced binding affinity. Simultaneously, we analyzed the CCK8 results and found that both compounds **17**, **67** and positive control nintedanib, exhibited a concentration-dependent increase in their inhibitory effects on pulmonary fibrosis. Furthermore, in a TGF-β-induced pulmonary fibrosis cell model, both compounds significantly reduced hydroxyproline and COL1A1 levels and upregulated ELN expression, with compound **17** exhibiting comparable antifibrotic efficacy to the positive control nintedanib. Collectively, our integrative computational-experimental approach reveals a therapeutic framework for precision-targeting CCR2-driven pathologies.

 Received 14th July 2025  
 Accepted 13th October 2025

DOI: 10.1039/d5ra05026j

[rsc.li/rsc-advances](http://rsc.li/rsc-advances)
<sup>a</sup>Zhangjiajie College, Zhangjiajie, Hunan, China. E-mail: mingshenglei@163.com

<sup>b</sup>Department of Pulmonary and Critical Care Medicine, Zhangjiajie Hospital Affiliated to Hunan Normal University, Zhangjiajie, Hunan, China

<sup>c</sup>Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, PR China

<sup>d</sup>The Key Laboratory of Carcinogenesis of the Chinese Ministry of Health, The Key Laboratory of Carcinogenesis and Cancer Invasion of the Chinese Ministry of Education, Cancer Research Institute, School of Basic Medicine, Central South University, Changsha, China

<sup>e</sup>Computational Medicine and Epidemiology Laboratory (CMEL), Guangdong Medical University, Zhanjiang, China

<sup>f</sup>Department of Respiratory and Critical Care Medicine, The Fifth Affiliated Hospital, Southern Medical University, China

<sup>g</sup>Department of Cardiovascular Disease, The First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong, China

<sup>h</sup>Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou Campus, Guangzhou, China

<sup>†</sup> These three authors made equal contributions to this work.

## 1 Introduction

IPF is a common interstitial lung disease of unknown etiology characterized by extensive fibrosis in both lungs, leading to restrictive ventilatory defects. A five-year follow-up study of a multi-institutional patient registry revealed a five-year survival rate of only 53.7%. Patients with IPF often succumb within three months of an acute exacerbation,<sup>1</sup> leading a serious threat to human health. Current treatment methods primarily focus on palliative care, utilizing drugs like nintedanib and pirfenidone, and high doses of acetylcysteine. Nintedanib and pirfenidone are the most commonly approved drugs for treating IPF. They have been shown to slow the decline in lung function, highlighting their clinical value.<sup>2</sup> However, both are expensive, placing a significant economic burden on patients, and their clinical efficacy is modest, further compromised by the occurrence of adverse events. In one study, some patients experienced multiple adverse effects, predominantly gastrointestinal symptoms (*e.g.*, diarrhea, dyspepsia, and vomiting), as well as



photosensitivity and skin rashes, particularly with pirfenidone. In cases where these side effects were exceedingly severe and posed a greater risk than the disease itself, discontinuation of therapy became necessary,<sup>3</sup> indicating a need for new drug development.

The underlying mechanisms of IPF remain unclear, Mohanan *et al.* demonstrated that organelle stress and oxidative stress play pivotal roles in pulmonary fibrosis development,<sup>4</sup> chemokines are key mediators of the inflammatory and immune responses in IPF. Mehrnaz *et al.*<sup>5</sup> demonstrated that CCR2 deficiency attenuates pulmonary fibrosis, thereby confirming the involvement of CCR2 in the pathogenesis of pulmonary fibrosis. Moreover, analysis of The Cancer Genome Atlas (TCGA) data revealed that higher expression of CCR2 is associated with poorer prognosis in patients.<sup>6</sup> Receiver operating characteristic (ROC) curve analysis further suggested that CCR2 could serve as a potential biomarker for diagnosing pulmonary fibrosis. Additionally, limited existing studies have demonstrated that CCR2 can influence the progression of pulmonary fibrosis by recruiting inflammatory cells.<sup>7–9</sup> However, to our knowledge, despite CCR2 playing a significant role in the pathogenesis of IPF and serving as a potential biomarker and therapeutic target, there are currently very few drugs targeting CCR2 for the treatment of pulmonary fibrosis. Therefore, designing anti-fibrotic drugs targeting CCR2 holds substantial scientific significance and clinical promise, depicted as Abstract Fig. A.

Computer-Aided Drug Design (CADD) represents a transformative approach in drug discovery and development, utilizing computational tools to design and optimize novel therapeutic compounds. Key aspects of CADD include Pharmacophore Modeling, Molecular Docking, Quantitative Structure–Activity Relationship (QSAR), and MD Simulations. CADD offers numerous advantages over traditional drug discovery methods, including speed, efficiency, precision, and innovation. SPR is a highly sensitive optical technique used to study molecular interactions, particularly suitable for real-time monitoring of the binding kinetics between biomolecules and their ligands. Therefore, in this study, we employed CADD and MD simulations to develop drugs targeting CCR2, followed by experimental validation to develop CCR2 dual-pocket inhibitors with potential therapeutic efficacy for IPF.

To address the urgent need for more effective and better-tolerated treatments for IPF, especially in China where the disease burden is high and the only approved drugs—nintedanib and pirfenidone—are not only associated with adverse effects but also prohibitively expensive, this study aims to develop a safer, affordable, and effective alternative to reduce the medical and economic burden on patients and the healthcare system, this study focuses on the chemokine receptor CCR2, which has been implicated in the pathogenesis and progression of pulmonary fibrosis. Despite its biological relevance, CCR2 remains an underexplored target in IPF drug development. The key research question of this study is whether novel small-molecule inhibitors targeting CCR2 can be rationally designed and validated using computational and experimental approaches. Accordingly, the primary objective is to

identify and characterize potential CCR2 inhibitors with dual-pocket binding properties by integrating CADD, MD simulations, and experiments. Through this approach, we aim to provide new candidate compounds with therapeutic promise for the treatment of IPF, thereby contributing to the development of more effective anti-fibrotic strategies.

The primary objective of this study is to identify and characterize novel CCR2-targeted small molecules using a structure-based drug discovery strategy, supported by molecular simulations and *in vitro* binding validation.

## 2 Materials and methods

### 2.1 Experimental validation and in-depth analysis

**2.1.1 Data collection.** In this study, RNA-seq data and related clinical information for idiopathic pulmonary fibrosis patients were obtained from the GEO database (GSE70866). This dataset includes bronchoalveolar lavage (BAL) samples from 112 idiopathic pulmonary fibrosis (IPF) patients and 20 healthy controls, generated based on the GPL14550 platform (Table S1). Data normalization was performed using the Robust Multi-array Average (RMA) algorithm from the limma R package (version 4.4.3). The effectiveness of models was validated using ROC curves. Prognostic analysis of survival curves was performed using the Kaplan–Meier test and Log-rank test to evaluate differences between groups. Statistical significance was set at  $p < 0.05$ .

**2.1.2 Establishment of idiopathic pulmonary fibrosis model.** This study was approved by the Animal Ethics Committee of Hunan Normal University. The idiopathic pulmonary fibrosis model was established using C57BL/6J mice (6–8 weeks, 18–22 g).<sup>10</sup> Mice were randomly divided into a control group ( $n = 6$ ) and an experimental group ( $n = 6$ ). Mice in the experimental group were anesthetized with 1% sodium pentobarbital (50 mg kg<sup>-1</sup>), and a single dose of 5 mg kg<sup>-1</sup> bleomycin (BLM) solution (Selleck, USA) was administered intratracheally. The BLM solution was prepared in normal saline (1 mg mL<sup>-1</sup>), and healthy mice served as controls. All procedures were conducted following the guidelines for the care and use of laboratory animals published by the US National Institutes of Health and in accordance with the protocols approved by the Animal Ethics Committee of Hunan Normal University.

After 4 weeks of feeding, the mice were sacrificed, and lung tissues were collected for hematoxylin and eosin (H&E) staining (Servicebio, China) and Masson's trichrome staining (Servicebio, China) to evaluate the modeling effects. H&E staining was employed for histopathological examination,<sup>11</sup> and the lung sections were scored based on alveolar congestion, hemorrhage, and infiltration (4: Extremely serious, 3: Serious, 2: Moderate, 1: Mild, 0: Normal). Masson's trichrome staining was used to assess collagen fiber deposition. Images of the tissue sections were captured at 200× magnification, and the same shade of blue was selected using Image-Pro Plus 6.0 software as the standard for identifying positive collagen areas. The collagen pixel area and corresponding tissue pixel area were measured in



each image, and the collagen area percentage was calculated (collagen pixel area/tissue pixel area  $\times$  100).

### 2.1.3 CCR2 detection

**2.1.3.1 RT-qPCR detection of CCR2 mRNA.** Total RNA was extracted from the tissues using Trizol reagent (Thermo Fisher, USA) following the manufacturer's instructions, and reverse transcription was performed using an mRNA reverse transcription kit (CWBio, China). RT-qPCR (Quantitative Reverse Transcription-polymerase Chain Reaction) amplification and analysis were conducted using the SYBR Green method. GAPDH mRNA was used as an endogenous control. The primer sequences (purchased from Beijing Tsingke Biotech Co., Ltd, China) were as follows:

Mouse-CCR2 forward: TGTGATTGACAAGCACTTAGACC, reverse: TGGAGAGATACCTTCGGAAGCTT

Mouse-GAPDH forward: GCGACTTCAACAGCAACTCCC, reverse: CACCCTGTTGCTGTAGCCGTA

Human-COL1A1 forward: GAGGGCCAAGACGAAGACATC, reverse: CAGATCACGTCATCGACAAC

Human-ELN forward: GCAGGAGTTAAGCCCAAGG, reverse: TGTAGGGCAGTCCATAGCCA.

**2.1.3.2 Western Blot detection of CCR2 protein.** Tissue samples were lysed using RIPA lysis buffer (Coolaber, China), and protein concentrations were measured using a BCA (Bicinchoninic Acid) protein assay kit (Biosharp, China). SDS-PAGE was performed for protein separation (ACE, China), followed by transfer onto a nitrocellulose membrane (Millipore, USA). The membrane was blocked with 5% milk at room temperature for 2 hours and incubated with the CCR2 antibody (Proteintech, China) overnight at 4 °C. After washing three times with TBST (tris-buffered saline with Tween 20), the membrane was incubated with a secondary antibody at room temperature for 1 hour on a shaker, using  $\beta$ -actin as the endogenous control. Finally, the signal was developed using a chemiluminescent imaging system, and quantitative analysis was performed using Image J software.

**2.1.3.3 Data analysis.** Independent samples *t*-tests and one-way analysis of variance were used for group comparisons, with a significance level set at  $P < 0.05$ . All data were analyzed using GraphPad Prism 9 software.

## 2.2 Preliminary exploration in drug discovery

**2.2.1 Protein structure preparation.** The 3D structure of CCR2 (PDB ID: 5T1A) was obtained from the Protein Data Bank (<https://www.rcsb.org/>). According to the study by Y. Zheng *et al.*,<sup>12</sup> CCR2 features two distinct binding pockets: an orthosteric site, located on the extracellular side of the membrane, and an allosteric site, located intracellularly. The orthosteric site is the primary binding pocket where endogenous ligands typically bind, directly influencing receptor activation and signaling. In contrast, the allosteric site serves as a secondary binding pocket that modulates receptor activity indirectly by inducing conformational changes upon ligand binding. Moreover, Y. Zheng *et al.*<sup>12</sup> mentioned that the simultaneous binding of two inhibitors to both the orthosteric and allosteric sites of CCR2 exhibits

stronger inhibitory potency compared to the binding of a single inhibitor to the orthosteric site alone. These two sites are crucial in receptor function and serve as reference points for constructing pharmacophores and docking molecules. The orthosteric and allosteric mechanisms are particularly significant as they allow for the identification of small molecules that can either directly inhibit receptor activation or modulate receptor function through indirect pathways, thereby offering multiple avenues for therapeutic intervention.

**2.2.2 Protein construction.** The CCR2 structure underwent amino acid repair, hydrogen atom addition, and energy minimization. Protein preparation included amino acid residue repair, hydrogen addition, and water removal using PyMOL software. The prepared protein structure was then subjected to energy minimization in the absence of the membrane, employing the AMBER ff14SB force field,<sup>13,14</sup> using a steepest descent algorithm until the maximum force was below 1000 kJ mol<sup>-1</sup> nm<sup>-1</sup>. This minimized structure was subsequently used for receptor pharmacophore modeling. The resulting structure was saved in pdb format for further use. Since CCR2 is a seven-transmembrane GPCR (G protein-coupled receptor), constructing a biomolecular membrane model is essential. The protein was embedded in the membrane system along with other relevant molecules. This process involved importing the optimized protein structure, adjusting the phospholipid membrane orientation, and assembling the membrane with DOPC as the membrane component, ensuring that CCR2 correctly spans the two layers of DOPC. Na<sup>+</sup> and Cl<sup>-</sup> ions were added to maintain system charge balance. Finally, the structure was exported in a format compatible with GROMACS, with appropriate force fields applied, including AMBER19SB for proteins, Lipid 17 for phospholipids,<sup>15</sup> and GAFF2 for ligands.<sup>16,17</sup> The more specific protocols were presented in 2.4.2 section.

**2.2.3 Structure-based pharmacophore (SBP) modeling.** We downloaded the SMILES strings and corresponding IC<sub>50</sub> values of small molecules targeting CCR2 from the BindingDB database (<https://www.bindingdb.org/bind/index.jsp>). The molecules were classified based on their activity levels, ensuring the IC<sub>50</sub> values spanning 4 orders of magnitude.<sup>18</sup>

Optimized protein structures and co-crystallized ligands' coordination were used to define binding pockets for SBP modeling. Small molecules were classified into active and inactive based on their IC<sub>50</sub> values and used as a test set for pharmacophore validation. Compounds with IC<sub>50</sub> values  $\leq$  10 nM were classified as active, while those with IC<sub>50</sub>  $\geq$  1000 nM were considered inactive. Compounds with intermediate potency (IC<sub>50</sub> between 10 and 1000 nM) were excluded from the validation set to ensure a clear distinction between actives and inactives (Tables S5A and S5B). Excluded volume pharmacophores were also constructed for the binding region of small molecules, addressing issues such as molecular size to ensure that the small molecule is correctly positioned within the designated binding pocket with Discovery Studio 2019 (DS). The generated pharmacophores were evaluated based on scores, sensitivity, and specificity. Additionally, the SBP consists of certain pharmacophore features, such as hydrogen bond



donors, hydrogen bond acceptors, hydrophobic features, and cationic features, *etc.* The best SBP pharmacophore was selected for virtual screening. Separate SBP models were created for the internal and external binding pockets of CCR2, named SBPi and SBPo, respectively.

## 2.3 The further optimization of drug design

### 2.3.1 Ligand-based pharmacophore (LBP) modeling

**2.3.1.1 3D-QSAR modeling.** The collected small molecule data were utilized for feature extraction and evaluation (Table S5C). 3D-QSAR pharmacophores were generated and assessed based on Maximum Fit (Maximum Fit Value), Total Cost (Total Cost Function Value), RMS (Root Mean Square Deviation), Correlation (Correlation Coefficient). The best 3D-QSAR model was selected after comprehensive evaluation with DS. Using the training set of small molecules (Table S5C), we performed 3D-QSAR modeling, selecting pharmacophore features such as HB ACCEPTOR, HB DONOR, HYDROPHOBIC, RING AROMATIC, and HBD.

**2.3.1.2 Common feature pharmacophore (CFP) modeling.** CCR2 inhibitors in clinical phases I, II, and III were collected and aligned to identify common structural features (Table S5D). These features were used to generate CFP pharmacophores. Active and inactive molecules were imported for validation. The best CFP model was chosen based on scores, sensitivity, specificity, and quality with DS. Based on drugs that have entered clinical trials, we selected HB ACCEPTOR, HB ACCEPTOR lipid, HB DONOR, HYDROPHOBIC, RING AROMATIC, HBA PROJECTION, and HBD1 as pharmacophore features to establish the CFP model.

**2.3.1.3 Pharmacophore-based virtual screening (PBVS).** We collected small molecules from the Chemdiv database (<https://www.chemdiv.com/cn/>), including Anti-Infective Library, Anti-Inflammatory Library, Antiviral Concentric Library, Antiviral Library, Chemokine Receptor Targeted Library, GPCR Targeted Library, Immunological Library, and hGPCR Complete, totaling 152 406 molecules. The generated SBPo, SBPi, 3D-QSAR, and CFP pharmacophores were used for screening with DS. Molecules meeting the pharmacophore criteria were retained for further ADMET analysis.

**2.3.2 ADMET screening.** Small molecules were evaluated using the ADME and Toxicity modules with DS, including Absorption Level (human intestinal absorption level), Blood–Brain Barrier Level (BBB Level, indicating the molecule's ability to cross the blood–brain barrier), Solubility (aqueous solubility level), CYP2D6 inhibition (cytochrome P450 2D6 inhibition potential), Hepatotoxicity (liver toxicity potential), Plasma Protein Binding (PPB, proportion of the drug bound to plasma proteins), AlogP98 (a measure of lipophilicity based on the Ghose–Crippen method), and Polar Surface Area 2D (PSA\_2D, 2-dimensional polar surface area reflecting molecular polarity) were assessed. Toxicity analysis included Mouse Female NTP probability (National Toxicology Program carcinogenicity probability for female mice), Mouse Male NTP probability (for male mice), Rat Female NTP probability (for female rats), Rat Male NTP probability (for male rats), Weight of Evidence

probability (WOE probability, indicating overall toxicological risk), Ames probability (probability of mutagenicity based on Ames test), and Aerobic Biodegradability probability (environmental biodegradability potential under aerobic conditions). Molecules passing ADMET validation underwent molecular docking.

**2.3.3 Batch molecular docking.** We imported the small molecules for the internal and external pockets using the Hermite® Platform, utilizing its GPU-accelerated molecular docking and binding free energy calculation capabilities. We performed batch molecular docking and MM/PBSA binding energy calculations for both pockets (Fig. 2A).

Protein and ligands were selected for docking in virtual screening, targeting both external and internal binding pockets. Docked molecules were exported for MM/PBSA scoring using parameters like Protein Force Field (amber99sb), Ligand Force Field (gaff2). The position of docked molecules were used for molecular dynamics simulations.

Docking calculations were performed using grid boxes defined for each small molecule. For compound 17, the docking grid box was centered at  $X = 5.1935 \text{ \AA}$ ,  $Y = 27.6990 \text{ \AA}$ , and  $Z = 187.8815 \text{ \AA}$ . For compound 67, the grid box was centered at  $X = 6.82 \text{ \AA}$ ,  $Y = 26.76 \text{ \AA}$ , and  $Z = 155.90 \text{ \AA}$ , with box dimensions of  $16.721 \times 17.450 \times 17.253 \text{ \AA}^3$  and  $18.91 \times 18.23 \times 24.12 \text{ \AA}^3$ , respectively.

Subsequently, the optimal docking position was validated using Vina software, and the 3D docking conformation was visualized with PyMOL (Table S12). Furthermore, the finally selected small molecules were compared with the corresponding co-crystallized ligands within their binding pockets to ensure binding plausibility.

## 2.4 Comprehensive validation and multilevel analysis

### 2.4.1 Small molecule and protein structure preparation.

We collected a total of 70 inhibitors related to CCR2 from BindingDB. The specific classification methods will be detailed in the subsequent pharmacophore modeling section. The structure of CCR2 was obtained from RCSB PDB. According to the article by Zheng Y, CCR2 has two binding pockets: orthosteric (external) and allosteric (internal). This study indicated that simultaneous inhibition of both sites results in better stability compared to single-site inhibition. The authors also provided molecular coordinates for the two pocket antagonists (Fig. 2A).

Because Zheng Y. *et al.*<sup>12</sup> grafted T4 lysozyme (T4L) onto the original CCR2 protein structure to improve solubility, we removed this grafted structure to restore the physiological state of CCR2.

**2.4.2 Molecular dynamics simulations.** To explore the drug action mechanism in more detail and for further screening, we conducted MD simulations on the CCR2 protein. Considering that CCR2 is a membrane protein, we performed membrane addition using the Charmm-gui website (<https://www.charmm-gui.org/>)<sup>19,20</sup> (Fig. 3A). The CCR2 molecule was positioned in the center of the complex, embedded in a DOPC bilayer, and placed in a water box. This structure was prepared for subsequent MD simulations as a membrane protein.



Using the Charmm-gui website, we generated GROMACS-compatible files and imported them into GROMACS for energy minimization, multi-step equilibration, and a production MD simulation lasting 100 ns. By analyzing the subsequent results based on the analysis of PCA results and molecular docking techniques, we preliminarily identified the optimal binding mode. Following this, two consecutive 500 ns molecular dynamics simulations were performed. We selected the best-performing results for presentation in the main text.

These were specific details below. We performed residue repair and molecular dynamics simulation to extract the stable structure of the protein, which was used as the original structure for subsequent operations.

The specific steps are as follows: PDB Reader & Manipulator: (1) the optimized protein pdb file was imported for pre-processing and then exported. (2) Membrane builder: the optimized protein was imported, and the protein and small molecules were selected. The orientation of the phospholipid membrane relative to the protein was adjusted to form an “embedded” relationship, ensuring the correct position of the protein within the membrane. POPC was selected as the membrane component with a 1 : 1 ratio of the upper and lower layers, and the appropriate membrane size was set. (3) Ion addition:  $K^+$  and  $Cl^-$  ions were added to balance the charge of the system. (4) Membrane assembly: the membrane was assembled, and the output structure was checked for accuracy. AMBER force fields were added, with AMBER19SB used for proteins, Lipid 17 for phospholipids, and GAFF2 for ligands. Other settings were kept as default, and the structure was exported in a format compatible with GROMACS for further processing.

We collected the sdf files of molecules 17, 41, 62, 64, 67 and positive control nintedanib. These selected molecules were hydrogenated using UCSF Chimera (<https://www.cgl.ucsf.edu/chimera/>) and saved in mol2 format. Optimization was performed at the B97-3c level using ORCA,<sup>21</sup> followed by single-point energy and molecular surface electrostatic potential calculations at the B3LYP-D3(BJ)/def2-TZVP level. RESP2 charges and topology files were generated using Multiwfn<sup>22</sup> and sobtop (<http://sobereva.com/soft/Sobtop>). Processed them according to the methodology, and then combined the small molecules with the protein. Integrated the membrane, protein and ligands to complex system. Simulations were run using Gromacs2022.1 (ref. 23) with specific settings for energy minimization, system equilibration, and long-range interaction handling. The TIP3P water model<sup>24</sup> was used, maintaining a minimum distance of 1 Å between solute atoms and the periodic box edge. Energy minimization was performed using the steepest descent method, with a maximum force of 1000.0 kJ mol<sup>-1</sup> nm<sup>-1</sup>. The system was equilibrated through three steps of constrained dynamics for 125 ps and 500 ps, with time steps of 1 fs and 2 fs, respectively. Temperature and pressure were controlled using the velocity-rescale thermostat and Berendsen barostat, set at 298.5 K and 1.01325 bar. Long-range interactions were handled using the particle mesh Ewald (PME) method, with a van der Waals cutoff of 10 Å. Due to significant interactions between the ligand and protein, they

were constrained to a single temperature control group and a single group for translational and rotational motion elimination. Molecular dynamics simulations were conducted with a 2 fs time step for 100 ns. Based on the subsequent analysis, the optimal binding mode was preliminarily identified. Following this approach, two consecutive 500 ns molecular dynamics simulations were conducted.

**2.4.3 Stability and conformation analysis.** Protein stability and conformation were analyzed using RMSD (Root Mean Square Deviation), RMSF (Root Mean Square Fluctuation), radius of gyration (Rg), and solvent-accessible surface area (SASA). Post-simulation, periodic correction and trajectory stability analysis were performed. RMSD is a crucial metric for assessing system stability and is fundamental for further analysis of small molecule binding. Hence, the averages and standard deviations assessed system stability. RMSF measures the fluctuation of each atom over the 500 ns simulation and is valuable for identifying active residues, binding pockets, and detailed interaction relationships. Rg indicates the compactness of a molecule. Smaller Rg values suggest a more compact molecule. SASA Reflects the surface area of a molecule that is accessible to solvent molecules. Larger SASA values are typical of unfolded or partially folded proteins. Rg and SASA evaluated enzyme conformation changes. Principal Component Analysis (PCA) identified the lowest potential energy and optimal system conformation. RMSD and Rg values were fitted, and free energy landscape maps were plotted to extract the most stable conformation. Hydrogen bond (H-bond) analysis was conducted using gromacs through the gmx hbond command. The hbnum.vgx file plots the number of hydrogen bonds over the simulation time. This data reflects the dynamic nature of hydrogen bond formation and breakage. A higher count may indicate stronger ligand–protein interactions, while fluctuations suggest transient bonding. Additionally, the analysis considers hydrogen bonds formed within 0.35 nm between donor and acceptor atoms, reflecting the standard geometric criteria for H-bond detection. This threshold ensures that both classical hydrogen bonds and closely interacting atom pairs are identified. The analysis was performed to quantify hydrogen bonds between the protein and ligand during MD simulation. In addition, the combined application of umbrella sampling molecular dynamics and binding free energy calculations using molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) was employed to identify thermodynamically favorable binding poses of the ligands docked to CCR2. This multi-methodological approach enabled systematic evaluation of both binding pathway energetics and site-specific interaction thermodynamics, revealing preferential binding conformations within CCR2's dual-pocket architecture.

**2.4.4 Umbrella Sampling Analysis.** To generate equilibrated starting structures for the pulling simulations, we selected the most stable configuration and processed MD simulations accordingly. The final structures from each trajectory served as the initial configurations for the pulling simulations. For each ligand, the system was placed in a sufficiently large rectangular box to satisfy the minimum image convention while allowing ample space for pulling simulations along the z-



axis. As in previous simulations, TIP3P water was used as the solvent, and 100 mM NaCl was included in the simulation cell. Pulling simulations were conducted using GROMACS, taking advantage of enhanced pull code features implemented during the project.

Equilibration for the pulling simulations was performed over 100 ps under a constant-pressure (NPT) equilibration ensemble, where weak coupling<sup>25</sup> was used to maintain isotropic pressure at 1.0 bar. After equilibration, all protein and ligands were restrained. Position restraints are commonly used in protein–ligand simulations to mimic the stability of larger structures. In this setup, the ligand was pulled away from the core structure along the z-axis over 500 ps, using a spring constant of 1500 kJ mol<sup>-1</sup> nm<sup>-2</sup> and a pull rate of 0.005 nm ps<sup>-1</sup> (0.05 Å ps<sup>-1</sup>). This approach, adapted from Justin Lemkul's study,<sup>26</sup> and was slightly changed to apply a faster pulling rate to expedite data collection while preserving result reliability.

Snapshots from these trajectories were used to generate starting configurations for umbrella sampling.<sup>27–29</sup> To optimize sampling resolution, an asymmetric distribution of sampling windows was applied: a spacing of 0.1 nm was used for COM (Center of Mass) separations up to 2 nm, while a 0.2 nm spacing was applied beyond 2 nm. This approach provided finer detail at smaller COM distances. The weighted histogram analysis method (WHAM)<sup>30</sup> was employed to analyze the results.

**2.4.5 Binding free energy calculation using gmx\_MM/PBSA.** Following periodic boundary condition removal and trajectory stabilization assessment through RMSD analysis, the 10 ns equilibrium phase trajectories were selected for binding free energy calculations using the MM/PBSA method. A strong correlation was observed between the MM/PBSA-calculated binding free energies and the experimentally determined binding affinities.<sup>31</sup> Within the MM/PBSA framework, the binding free energy ( $\Delta G_{\text{bind}}$ ) is calculated as the difference between the free energy of the complex and the sum of the free energies of the unbound receptor and ligand.<sup>32,33</sup> This is approximated by:

$$\Delta G_{\text{bind}} \approx \Delta E_{\text{gas}} + \Delta G_{\text{solv}} - T\Delta S,$$

where  $\Delta E_{\text{gas}}$  (gas-phase interaction energy),  $\Delta G_{\text{solv}}$  (solvation free-energy change), and  $-T\Delta S$  (entropic contribution) collectively define the thermodynamic driving forces of binding.

In gmx\_MMPBSA,  $\Delta E_{\text{gas}}$  is derived from molecular mechanics terms, including bonded interactions ( $\Delta\text{BOND}$ ,  $\Delta\text{ANGLE}$ ,  $\Delta\text{DIHED}$ ) and non-bonded interactions ( $\Delta\text{VDWAALS}$ ,  $\Delta\text{EEL}$ ,  $\Delta 1-4$  VDW,  $\Delta 1-4$  EEL). These terms are summed and reported as  $\Delta G_{\text{GAS}}$ . The solvation free energy ( $\Delta G_{\text{solv}}$ ) comprises polar ( $\Delta E_{\text{PB}}$  or  $\Delta E_{\text{GB}}$ , depending on Poisson–Boltzmann or Generalized Born solvation models) and nonpolar components ( $\Delta\text{ENPOLAR}$  from solvent-accessible surface area [SASA] calculations, and optionally  $\Delta\text{EDISPER}$  for dispersion corrections).

The final output term  $\Delta\text{TOTAL}$  represents the sum of  $\Delta G_{\text{GAS}}$  and  $\Delta G_{\text{SOLV}}$ , providing an estimate of  $\Delta G_{\text{bind}}$  without entropic contributions. To obtain the complete binding free energy, entropy corrections must be subtracted from  $\Delta\text{TOTAL}$ .

Following MD analysis, one compound was selected from the orthosteric site and one from the allosteric site as final candidates for chemical synthesis and subsequent experimental validation.

## 2.5 Chemical synthesis

The synthetic procedures for compounds **17** and **67** were outlined in the SI (Section 1). Their chromatographic purity was confirmed by quantitative proton Nuclear Magnetic Resonance (<sup>1</sup>H NMR) and Liquid Chromatography–Mass Spectrometry (LC–MS), with compound **17** exhibiting a purity of >99.8% and compound **67** demonstrating a purity of >99.4%.

## 2.6 Surface plasmon resonance (SPR)

**2.6.1 Experimental principle.** SPR<sup>34,35</sup> was performed to evaluate real-time, label-free molecular interactions. One binding partner was immobilized on the biosensor chip, and the analyte was injected over the surface. Binding and dissociation events were recorded as changes in resonance angle, allowing kinetic parameters to be determined. The interaction assay between Mouse CCR2 protein and the substance to be tested was modeled using the surface plasmon resonance technique, and the substance provided by us was assayed (Tables S2A–C).

**2.6.2 Experimental material.** Details of the experimental materials used for SPR validation are provided in Table S2.

**2.6.3 Experimental methods.** Protein immobilization was performed using the amino coupling method; protein coupling buffer solution: 1.0 × PBS-P+ (pH 7.4); interaction buffer solution: 1.0 × PBS-P+ (pH 7.4), 5% (v/v) DMSO.

**2.6.3.1 Protein coupling.** (1) Placed the running buffer (200 mL 1 × PBS Buffer), water bottle, and waste bottle in the left and right trays, respectively, and inserted the appropriate feed tubes.

(2) Held the CM5 chip in hand with the lettered side facing up. Followed the direction of the arrow on the chip, gently pushed the chip into the slot, and finally closed the hatch of the chip compartment.

(3) Activated chip channel 4 with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC, GE Healthcare) and *N*-hydroxysuccinimide (NHS, GE Healthcare) at a flow rate of 10 μL min<sup>-1</sup>.

(4) Diluted the ligand protein to 50 μg mL<sup>-1</sup> with sodium acetate and immobilized the protein in the 4-channel of the chip at a flow rate of 10 μL min<sup>-1</sup> to generate the coupling map.

(5) Closed the channels with ethanolamine at a flow rate of 10 μL min<sup>-1</sup>.

(6) Repeated steps (3)–(5) for channel 3 as a reference, with the difference that protein-free acetate buffer was used in step (4).

**2.6.3.2 Protein-to-be-tested interaction tests.** Solvent correction

The 5% DMSO concentration calibration curve was configured by mixing the 4.5% and 5.8% master batches (Table S3).

Determination of the Substance to Be Measured



(1) Each substance to be assayed was diluted to a range of concentrations in a 96-well plate and coupled to the target proteins by microarray from low to high concentrations. The flow rate was  $30 \mu\text{L min}^{-1}$  for a duration of 150 s.

(2) After each concentration point had flowed, the chip was regenerated with 10 mM glycine hydrochloride (pH 2.0) solution for 5 min. This process was repeated until the corresponding concentrations of all the substances to be measured had been run.

(3) Binding and dissociation constants were obtained by globally fitting the data to a 1 : 1 Langmuir binding model using the Biacore Insight assessment software (Cytiva, Marlborough, MA, USA).<sup>36</sup>

## 2.7 Cell modeling and viability assay

Initially, 2000 cells were seeded into each well of a 96-well plate. Human diploid fibroblasts (MRC-5) (iCell Bioscience, Inc., China) were induced with TGF- $\beta$  ( $10 \text{ ng mL}^{-1}$ ) for 48 hours to establish a pulmonary fibrosis cell model.<sup>37</sup> Subsequently, compounds **17** and **67** were administered to the model cells. These drugs were dissolved in DMSO, diluted in DMEM medium, and applied at gradient concentrations of 0, 1, 1.25, 2.5, 5, 10, 25, 50, and  $100 \mu\text{mol}$ . In addition, we used nintedanib as a positive control. Following a 48-hour drug treatment,  $10 \mu\text{L}$  (10% volume) of CCK8 reagent was added to each well. The 96-well plate was incubated in the dark within an incubator for 3 hours. Absorbance at 450 nm was measured using a microplate reader to assess cell viability.

# 3 Results

## 3.1 Experimental validation and in-depth analysis

**3.1.1 Significant upregulation of CCR2 in idiopathic pulmonary fibrosis.** In this study, we utilized bioinformatics analysis to examine the expression levels of CCR2 in patients with IPF. In Fig. 1A, the violin plot reveals that CCR2 gene expression is significantly higher in the IPF group than in the normal control group ( $p < 0.05$ ), suggesting a strong association between CCR2 upregulation and IPF onset. This finding indicates that CCR2 may actively contribute to the progression of fibrosis. Survival analysis (Fig. 1B) indicates that patients with elevated CCR2 expression have significantly lower survival rates compared to those with lower expression ( $p = 0.0049$ ). Additionally, Fig. 1C shows the ROC curve, with an area under the curve (AUC) of 0.693, suggesting that CCR2 is an independent diagnostic marker for pulmonary fibrosis. These results underscore the potential of CCR2 as a negative prognostic factor for IPF patients, highlighting its value as a clinical biomarker.

After identifying CCR2 upregulation in patient-derived transcriptomic data (Fig. 1A–C), we further validated these findings in a bleomycin-induced mouse model of pulmonary fibrosis.

Four weeks after BLM administration, H&E staining of lung tissue revealed severe alveolar congestion and exudation, with localized alveolar wall proliferation of connective tissue, accompanied by alveolar collapse and loss of alveolar structure. Inflammatory cell infiltration was observed within the alveolar spaces and around some bronchioles, with a total inflammation

score of 10. In contrast, the Control group displayed clear alveolar structures with intact alveolar walls, without significant inflammatory cell infiltration or fibrosis, yielding a total inflammation score of 2 (Fig. 1D and E). Masson's trichrome staining showed prominent collagen fiber proliferation in the alveolar walls of the BLM group, with some areas forming fibrous masses, indicating significant alveolar wall fibrosis. The calculated average collagen area percentage in the BLM group was 21.13%, derived from the ratio of collagen pixel area to tissue pixel area. In the Control group, no significant blue-stained collagen fibers were observed, with minimal collagen proliferation or fibrosis, resulting in an average collagen area percentage of 6.13% (Fig. 1F). Collectively, these findings confirm the successful induction of idiopathic pulmonary fibrosis in mice *via* BLM intratracheal instillation. In the bleomycin (BLM)-induced pulmonary fibrosis mouse model, hydroxyproline levels were significantly elevated, indicating enhanced collagen deposition in the lung tissue (Fig. 1G).

To elucidate the expression of CCR2 in idiopathic pulmonary fibrosis, we examined the mRNA and protein levels of CCR2 in lung tissues from both the Control group and BLM-induced idiopathic pulmonary fibrosis mice using RT-qPCR and western blotting. The results demonstrated that the expression levels of CCR2 mRNA and protein were significantly elevated in the BLM group compared to the Control group (Fig. 1H–J).

## 3.2 Structure-based pharmacophore (SBP) screening

**3.2.1 Membrane protein construction.** During membrane system construction, the CHARMM-GUI web server was used to generate a membrane-protein system. The appropriate orientation of the CCR2 protein within the lipid bilayer was determined, the length, width were both  $60 \text{ \AA}$  and the protein was inserted accordingly. To neutralize the system,  $84 \text{ K}^+$  ions and  $98 \text{ Cl}^-$  ions were added. Once the structural setup was confirmed to be correct, the system was exported in GROMACS format. After importing into GROMACS, the structures of the protein, membrane, and small molecules were further integrated and processed.

**3.2.2 Structure-based pharmacophore (SBP) modeling.** For the external pocket, we constructed pharmacophore models and found that SBPo-1 to SBPo-5 shared the same six pharmacophore features, characterized by AADHHP, and each scored 12.047 (Table S4A). We validated 10 SBPo pharmacophores using active and inactive small molecules from the test set (Table S5), assessing their sensitivity and specificity, and calculated Youden's  $J$  statistic (Table S4B). Specifically,  $J = \text{Sensitivity} + \text{Specificity} - 1$ .

Sensitivity (true positive rate): the proportion of actual positives correctly identified by the test.

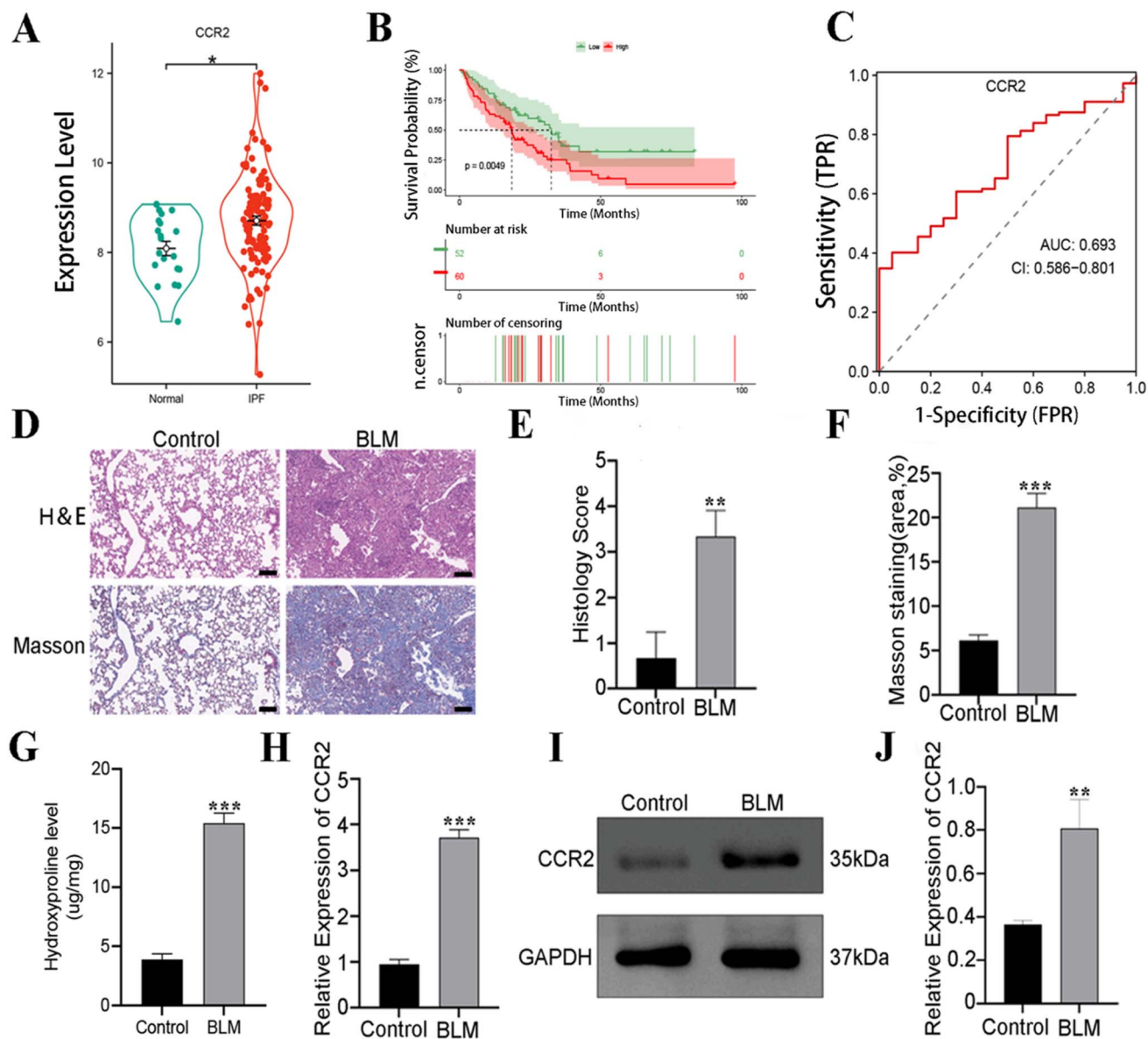
Specificity (true negative rate): the proportion of actual negatives correctly identified by the test.

The value of Youden's  $J$  statistic ranges from  $-1$  to  $1$ :

$1$ : Indicates a perfect test (100% sensitivity and 100% specificity).

$0$ : Indicates a test with no diagnostic ability (equivalent to random chance).





**Fig. 1** CCR2 Expression and Pathological Validation in IPF Progression. (A–C) *In silico* analysis based on patient data from dataset GSE70866. (A) The violin plot depicts CCR2 gene expression levels in the normal control and IPF groups (GSE70866) ( $*P < 0.1$ ). (B) The survival analysis plot outlines survival probabilities over time for groups with high and low CCR2 expression. (C) The ROC curve demonstrates the diagnostic performance of CCR2 in identifying pulmonary fibrosis. (D–J) *In vivo* validation using a bleomycin-induced pulmonary fibrosis mouse model. (D) Representative images of H&E and Masson's trichrome staining of lung tissues 4 weeks after intratracheal instillation of BLM in mice. (E) Pathological scoring results of H&E-stained lung tissues at week 4 ( $**P < 0.01$ ). (F) Analysis of collagen area in lung tissues at week 4 ( $***P < 0.001$ ). (G) Hydroxyproline content in lung tissue of bleomycin (BLM)-induced pulmonary fibrosis mice and control mice. The hydroxyproline level was significantly increased in BLM-treated mice compared to controls ( $***P < 0.001$ ). (H) RT-qPCR analysis of CCR2 expression levels in BLM-induced idiopathic pulmonary fibrosis ( $***P < 0.001$ ). (I and J) Western blot analysis of CCR2 expression levels in lung tissues of each group ( $**P < 0.01$ ).

Negative values: indicate that the test performs worse than random chance.

The results showed that SBPo scored 12.047, with a sensitivity of 0.750, a specificity of 0.619, and a  $J$  value of 0.369. Therefore, SBPo-1 was selected as the best SBPo for further analysis (Fig. 2B). The fitting of the external reference molecule with SBPo-1 is shown in Fig. 2C.

Similarly, for the internal pocket, we constructed pharmacophore models and found that SBPi-1 had six pharmacophore

features, characterized by AAHHN, and scored 10.582 (Table S6A). Calculating Youden's  $J$  statistic, we found that SBPi-1 and SBPi-2 had positive  $J$  statistic values of 0.006 and 0.176, respectively (Table S6B). Considering both the score and feature values, SBPi-1 was selected as the best SBPi pharmacophore (Fig. 2D). The fitting of the internal reference molecule with SBPi-1 is shown in Fig. 2E.



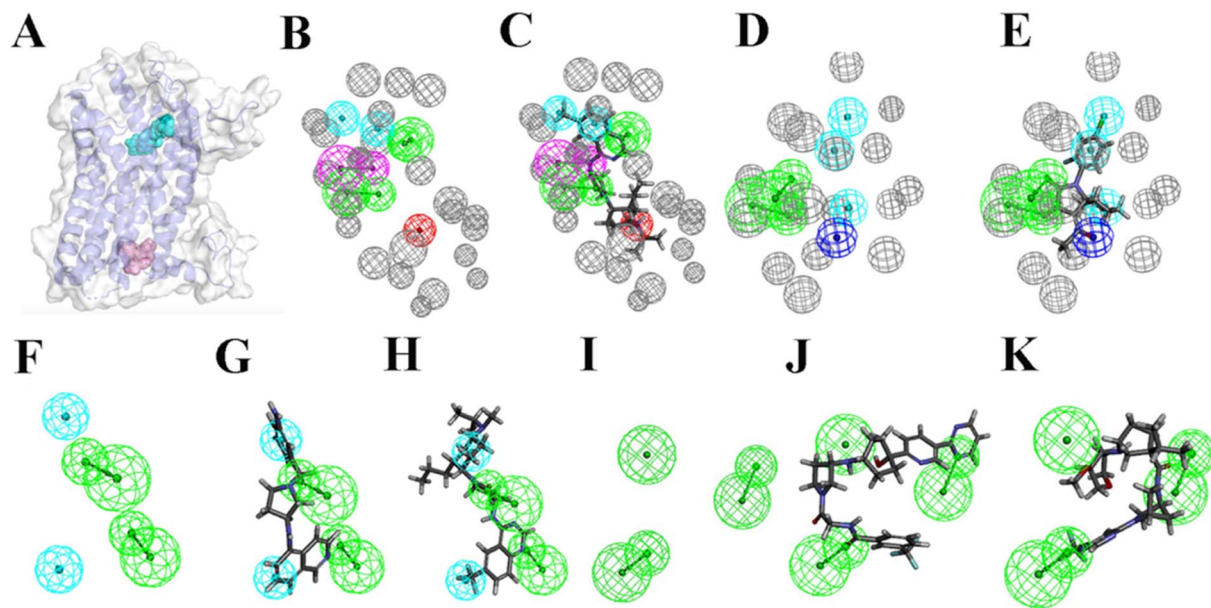


Fig. 2 Structure-based pharmacophore modeling of CCR2 dual-pocket inhibitors. (A) The optimized 3D structure of CCR2, with binding sites visualized using Hermite software. The cyan pocket represents the orthosteric (external) site, while the violet pocket represents the allosteric (internal) site. (B) SBPo-1 pharmacophore model. The purple spheres represent hydrogen bond donors, green spheres represent hydrogen bond acceptors, light blue spheres represent hydrophobic features, red spheres represent cationic features, and grey spheres represent exclusion volumes. (C) The fitting of the reference small molecule with SBPo-1. (D) SBPi-1 pharmacophore model. The green spheres represent hydrogen bond acceptors, light blue spheres represent hydrophobic features, dark blue spheres represent anionic features, and grey spheres represent exclusion volumes. (E) The fitting of the reference small molecule with SBPi-1. (F) 3D-QSAR pharmacophore model. (G) Molecule with the lowest activity value, BDBM50088301. (H) Molecule with the highest activity value, BDBM50197912. (I) CFP-1 pharmacophore model. (J) Fitting of the molecule PF-04136309 with the highest Fit value to CFP1. (K) Fitting of the molecule PF-04634817 with the lowest Fit value to CFP1.

### 3.3 The further development of drug design

**3.3.1 3D-QSAR modeling.** We exported the parameters for 10 3D-QSAR models and found that QSAR1 had the highest correlation between actual and predicted IC<sub>50</sub> (0.913), the highest fit value (16.040), a total cost of 240.708, and an RMS of 3.737 (Table S7).

Based on the predicted and actual activity values of the training set molecules, we derived a regression curve, with QSAR1 showing an  $R^2$  value of 0.83 (Fig. S1), indicating excellent correlation. Analyzing the 3D-QSAR pharmacophore model (Fig. 2F), the molecule with the lowest activity value was BDBM50088301, with an IC<sub>50</sub> of 0.03 nM and a fit value of 15.492 (Fig. 2G), while the molecule with the highest activity value was BDBM50197912, with an IC<sub>50</sub> of 30 000 nM and a fit value of 8.908 (Fig. 2H).

**3.3.2 Common feature modeling.** Analyzing the pharmacophore parameters, we found that CFP1 and CFP2 had the same parameter information: both featured HHA with a rank of 63.731 (Table S8). We then predicted the activity of known active small molecules (Fig. S4A) and calculated the absolute difference between the summed active and inactive values for each pharmacophore, resulting in a discriminative value, DIF (Fig. S4B) with DS. This analysis showed that CFP1 had a higher discriminative value than CFP2.

We analyzed the CFP pharmacophore model (Fig. 2I). The molecule with the best fit, PF-04136309, had a fit value of 3.000

(Fig. 2J), while the molecule with the worst fit, PF-04634817, had a fit value of 1.431 (Fig. 2K).

**3.3.3 Multi-pharmacophore virtual screening.** After successfully completing the modeling of SBP, 3D-QSAR, and CFP pharmacophores, we utilized the ChemDiv database to perform virtual screening using these three pharmacophores. The process was as follows: (1) SBPo screening: initially, we used the SBPo pharmacophore for screening, which resulted in 1722 candidate molecules remaining. (2) SBPi screening: these candidates were further screened using the SBPi pharmacophore, narrowing down the list to 69 small molecules. (3) 3D-QSAR screening: the same 69 small molecules were then screened using the 3D-QSAR pharmacophore, with all 69 molecules being retained. (4) CFP screening: finally, the CFP pharmacophore was used for screening, and the same 69 small molecules were retained as the final result of the virtual screening process.

This multi-pharmacophore virtual screening approach ensured a comprehensive evaluation of potential candidates, leveraging the strengths of each pharmacophore model.

**3.3.4 ADMET screening.** We collected the SMILES strings of the 69 small molecules and analyzed their ADME properties, including Absorption Level, BBB, Solubility, CYP2D6, Hepatotoxicity, PPB, AlogP98, and PSA\_2D (Tables S9A and S10A). Additionally, we assessed their toxicity properties, such as Mouse Female NTP probability, Mouse Male NTP probability, Rat Female NTP probability, Rat Male NTP probability, WOE probability, Ames probability, and Aerobic Biodegradability probability with DS (Tables S9B and S10B).



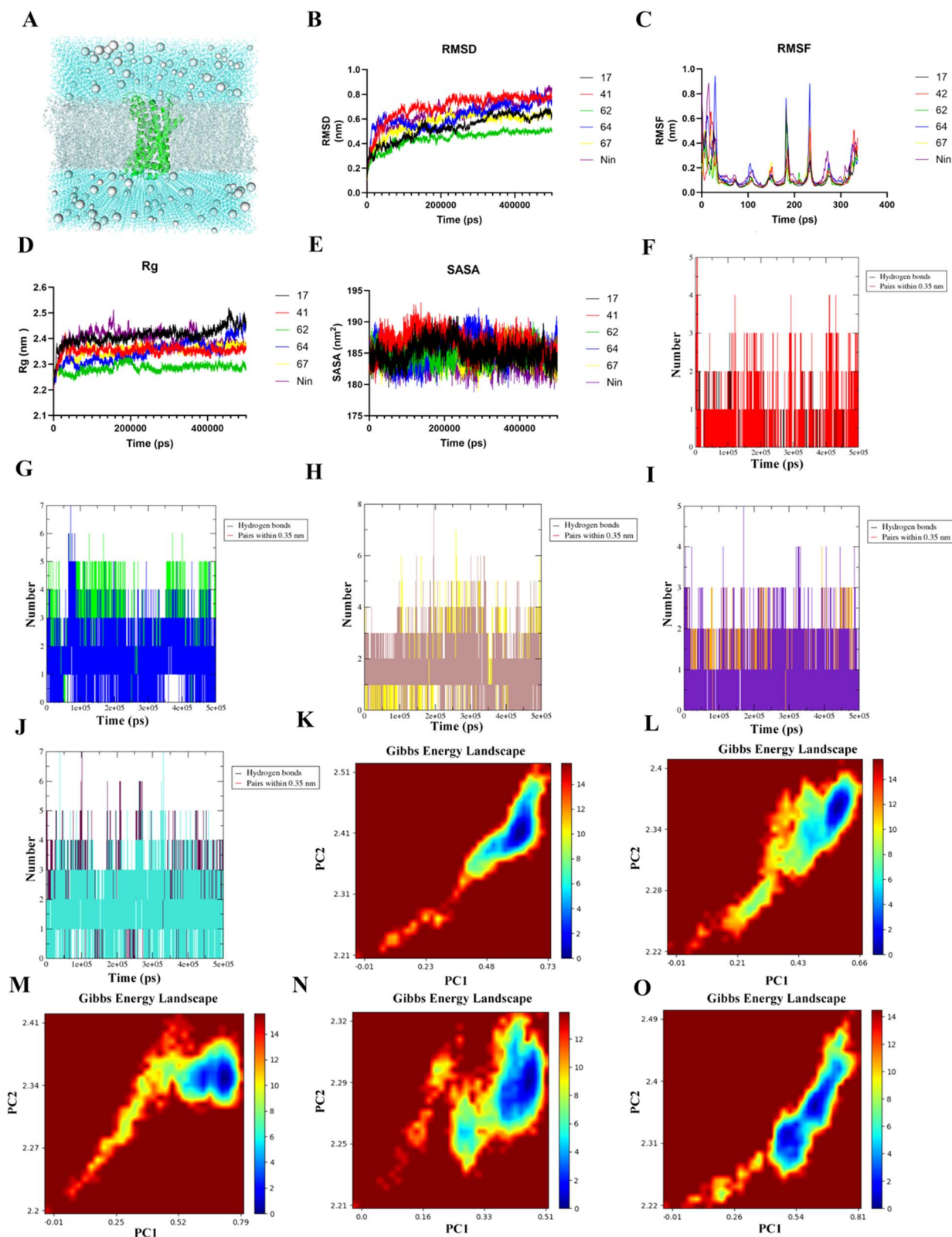


Fig. 3 Multidimensional profiling of receptor–ligand complexes via molecular dynamics simulations. (A) Full-atom model of CCR2 receptor embedded in a phospholipid bilayer (green: protein; gold: lipids; blue: water; grey: ions). (B) RMSD trajectories for 500 ns simulation, compounds 17 (black); 41 (red); 62 (green); 64 (blue); 67 (yellow) and nintedanib (purple). (C). Per-residue RMSF heatmap (the color representation was consistent with RMSD). (D) Rg reflecting structural compactness (the color representation was consistent with RMSD). (E) SASA quantifying hydrophobic core exposure (the color representation was consistent with RMSD). (F–J) Hydrogen-bond network fluctuations (both bond count fluctuations and short-range contacts ( $\leq 3.5$  Å) and represented compounds 17, 41, 62, 64 and 67). (K–O) Principal component analysis with the regions highlighting dominant conformational clusters and represented compounds 17, 41, 62, 64 and 67.



Molecules exhibiting toxic, carcinogenic, or otherwise undesirable properties were excluded. We retained the following molecules for further screening and analysis: External pocket: ligands 17, 40, 68; internal pocket: ligands 41, 62, 64, 67. These retained molecules will be used in the subsequent steps of screening and analysis.

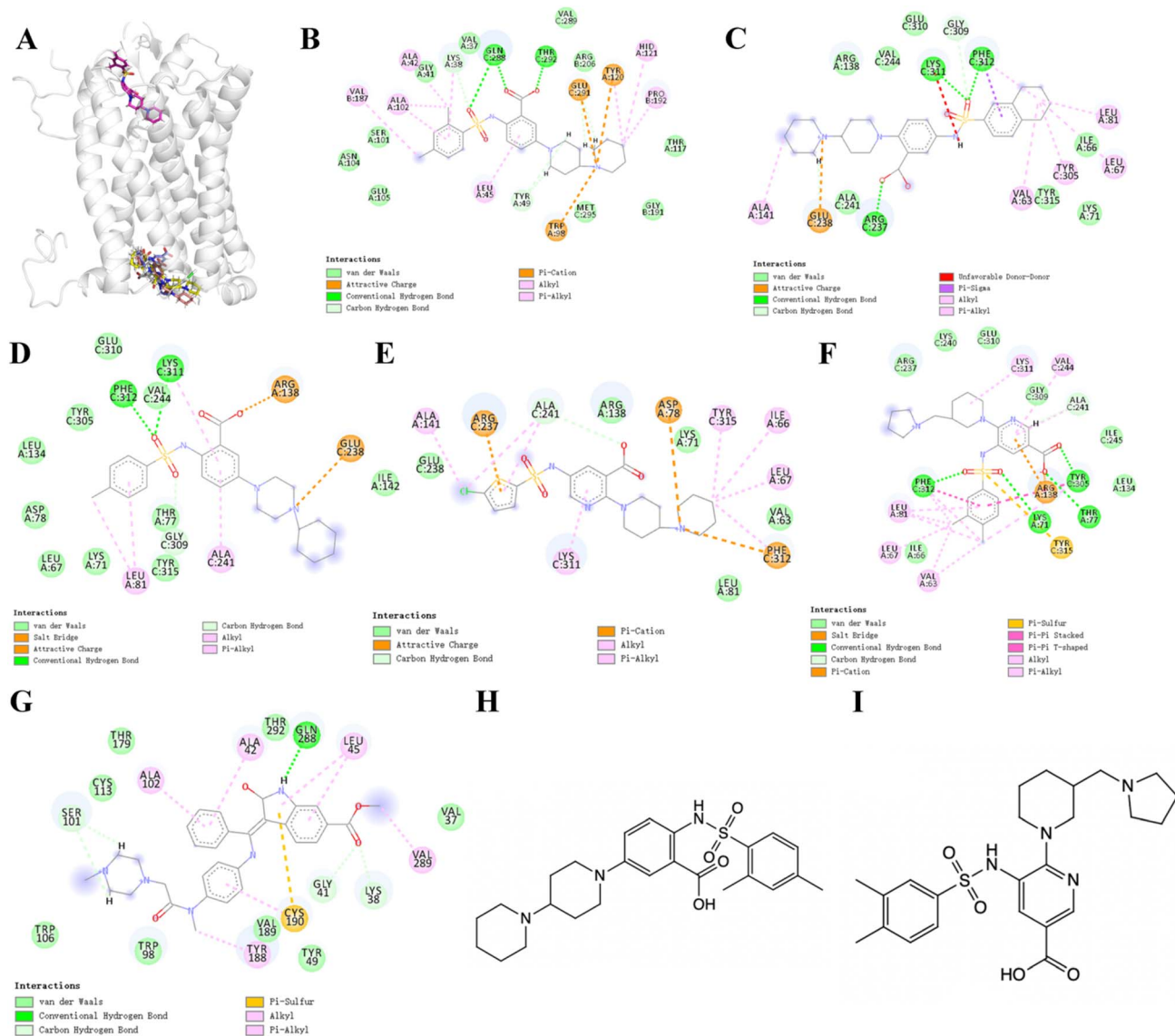
**3.3.5 Batch molecular docking.** The results indicated that molecules 17, 40, and 68 were the candidates for the external pocket, while molecules 41, 62, 64, and 67 were the candidates for the internal pocket (Tables S11A and S11B). However, considering that the binding free energies ( $dG$ ) of molecules 40 and 68 were too low, we selected molecule 17 as the binding molecule for the external pocket. The candidates for the internal pocket remained unchanged.

### 3.4 The further optimization of drug design

#### 3.4.1 Analysis results

**3.4.1.1 Root mean square deviations (RMSD).** We analyzed the RMSD over 500 ns for the system (Fig. 3B). The results showed that all five small molecules exhibited good stability, with average RMSD values ranging from 4 to 4.5 nm. The mean RMSD values for molecules 17, 41, 62, 64, 67 and nintedanib were 0.639, 0.699, 0.443, 0.622, 0.565 and 0.648, respectively. Based on RMSD analysis, molecules 62 and 67 demonstrated better stability.

**3.4.1.2 Root mean square fluctuations (RMSF).** RMSF analysis revealed six regions with significant amino acid fluctuations. The mean RMSF values for molecules 17, 41, 62, 64, 67 and nintedanib were 0.125, 0.128, 0.106, 0.162, 0.118 and 0.174. These regions may indicate structural conformational changes



**Fig. 4** Overview of Molecular Docking and Umbrella Sampling Analysis. (A) Model diagram illustrating the optimal binding conformations of five small molecules. (B–G) Docking poses of compounds 17, 41, 62, 64, 67 and nintedanib in their optimal binding orientations. (H–I) Molecular structures of the finally selected compounds 17 and 67. Ligand 17: SMILES: Cc1cc(S(Nc(ccc(N(CC2)CCN2C2CCCC2)c2)c2C(O)=O)(=O)=O)c(C)cc1 ligand 67: SMILES: Cc(cc1)C(C)cc1S(Nc1c(N2CC(CN3CCCC3)CCC2)ncc(C(O)=O)c1)(=O)=O.



and suggest potential active pockets. Compounds **64** and **67** exhibited extensive interactions with amino acid residues within the binding pocket (Fig. 3C).

Further analysis identified and recorded amino acids with fluctuation values greater than 0.5 nm:

- Molecule **17**: residues 182–185
- Molecule **41**: residues 19–24
- Molecule **62**: residues 182–183
- Molecule **64**: residues 11–16, 28–31, 232–234
- Molecule **67**: residues 9, 19, 233

These results indicate that the approximate regions around these residues are critical for structural changes upon interaction with these molecules.

**3.4.1.3 Gyration ( $R_g$ ) and solvent-accessible surface area (SASA) analysis.** The average  $R_g$  values for molecules **17**, **41**, **62**, **64**, **67** and nintedanib were 2.411, 2.352, 2.185, 2.348, 2.352 and 2.400 nm, respectively. Compound **62** and **64** exhibited smaller  $R_g$ , indicating a relatively stable binding conformation (Fig. 3D).

All six molecules showed stable SASA fluctuations around 185 nm<sup>2</sup>. The average SASA values for molecules **17**, **41**, **62**, **64**,

**67** and nintedanib were 185.135, 186.624, 184.854, 185.437, 184.744 and 183.480 nm<sup>2</sup>, respectively. Compounds **62** and **67** exhibited lower SASA values, indicating greater system stability (Fig. 3E).

**3.4.1.4 Hydrogen bond analysis.** The average hydrogen bond number for molecules **17**, **41**, **62**, **64**, and **67** were 0.388, 2.728, 1.828, 1.130, 2.153. The average numbers of hydrogen bonds formed within 0.35 nm for molecules **17**, **41**, **62**, **64**, and **67** were 0.373, 1.531, 1.642, 0.643, 1.687. These results indicate that molecules **41** and **67** may have stronger connection with amino acids in the pocket (Fig. 3F–J).

**3.4.2 Principal component analysis (PCA).** We also performed PCA, a dimensionality reduction technique used to find the lowest energy states of a system. The lowest points on the free energy landscape represent the most likely stable conformations. We first removed the periodic boundary conditions, extracted the RMSD and  $R_g$  values at the same time scale, and derived the Gibbs free energy (Fig. 3K–O).

**3.4.3 Optimal molecular docking.** Based on the PCA analysis, the protein–ligand complex corresponding to the minimum energy conformation was extracted (Fig. 4A) and

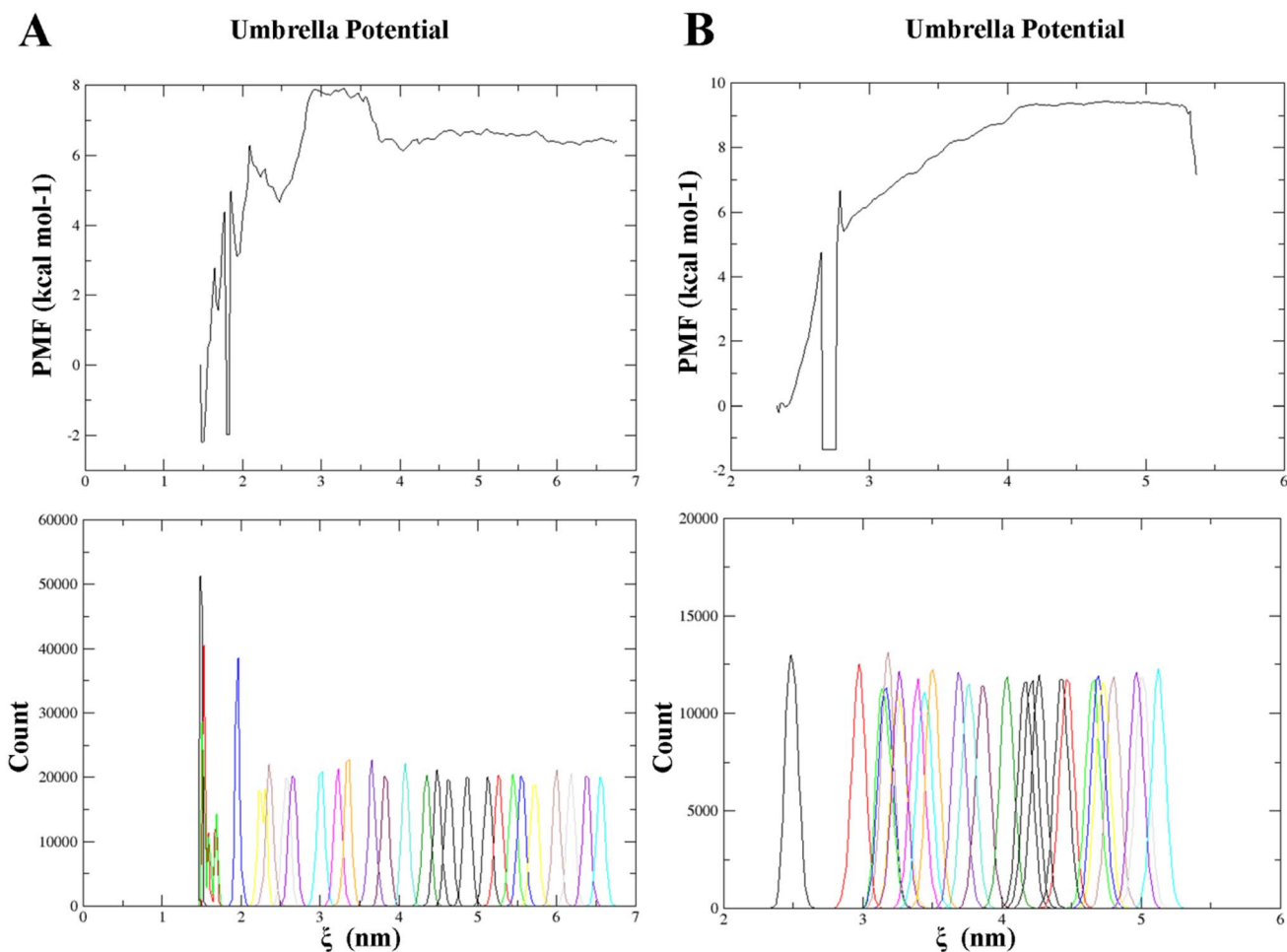


Fig. 5 Umbrella sampling analysis. (A) Umbrella sampling results for compound **17**: the top panel shows the PMF analysis, and the bottom panel displays the distribution of sampling windows. (B) Umbrella sampling results for compound **67**: the top panel presents the PMF analysis, while the bottom panel depicts the distribution of sampling windows.



analyzed for its docking interactions (Fig. 4B–G). Analysis showed that molecules **41**, **62**, **64**, and **67** interacted more than twice with the amino acids VAL63, LEU67, LEU81, ALA141, ARG237, GLU238, TYR305, LYS311, PHE312, ARG138, ALA241, VAL244 and TYR315. Molecule **17** interacted with amino acids LYS38, ALA42, LEU45, TYR49, TRP98, ALA102, TYR120, HID121, VAL187, PRO192, GLN288, GLU291 and THR292.

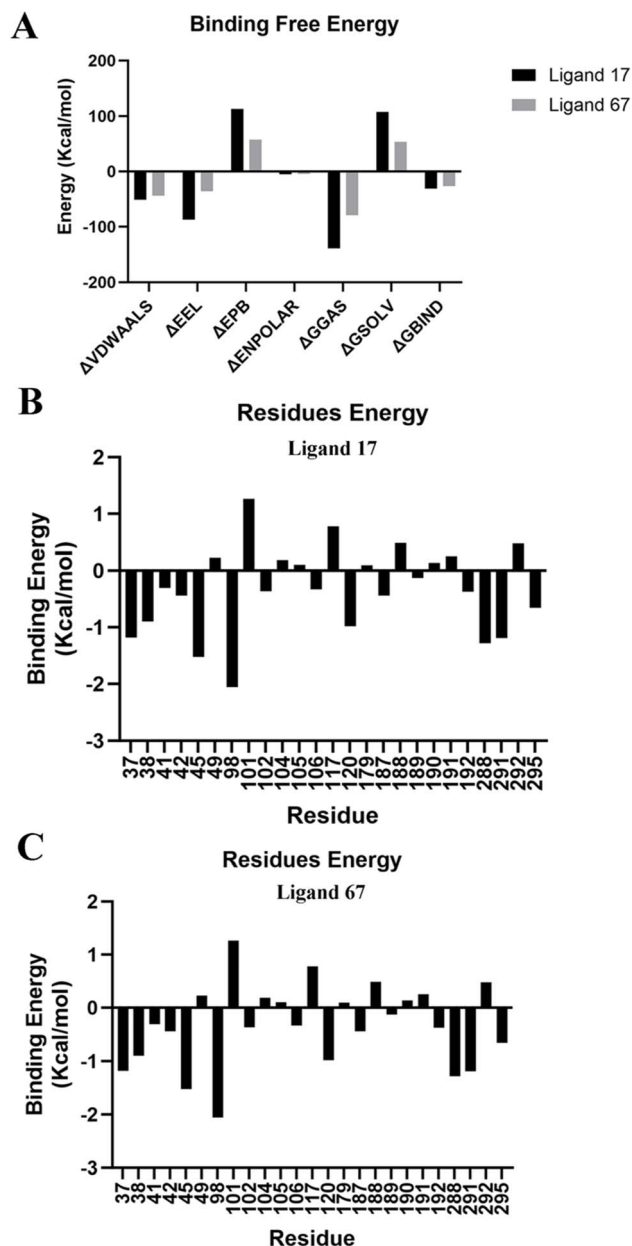


Fig. 6 Binding free energy with gmx\_MMPBSA. (A) Binding free energy analysis for compounds **17** and **67**, including the energy terms  $\Delta$ VDDWAALS,  $\Delta$ EEL,  $\Delta$ EPB,  $\Delta$ BOND,  $\Delta$ ANGLE,  $\Delta$ DIHED,  $\Delta$ 1–4 VDW,  $\Delta$ 1–4 EEL,  $\Delta$ EDISPER,  $\Delta$ ENPOLAR,  $\Delta$ G<sub>GAS</sub>,  $\Delta$ G<sub>SOLV</sub>, and  $\Delta$ G<sub>BIND</sub>. Energy contributions with a value of zero (*i.e.*,  $\Delta$ BOND,  $\Delta$ ANGLE,  $\Delta$ DIHED,  $\Delta$ 1–4 VDW,  $\Delta$ 1–4 EEL, and  $\Delta$ EDISPER) were omitted. (B) Amino acid binding energy contributions within approximately 4 Å of the pocket surrounding compound **17**; negative values indicate a spontaneous binding tendency. (C) Amino acid binding energy contributions within a 4 Å range of the pocket containing compound **67**, with negative values signifying a spontaneous binding trend.

The molecular docking scores (in kcal mol<sup>-1</sup>) for compounds **17**, **41**, **62**, **64**, **67**, and nintedanib are as follows, respectively: -8.480, -8.500, -7.682, -7.570, -8.540 and -8.728. Compared to the initial binding poses, the binding energies were significantly improved. The docking result for validation was presented as Table S12. Compared with the co-crystallized ligands in the same binding pocket, both compound **17** and compound **67** exhibited lower binding energies, indicating that the selected small molecules form more stable interactions than those observed in the crystal structures.

**3.4.4 Final results.** Based on the previous analysis, compound **17** was selected from the orthosteric pocket as it uniquely met the selection criteria in that site. For compounds in allosteric pocket, compounds **62**, **64**, and **67** demonstrated competitive performance. Although compound **64** showed promising interactions, it exhibited limitations in hydrogen bond formation. Compounds **62** and **67** displayed highly similar behavior; however, considering the number of hydrogen bonds, RMSF analysis, and subsequent molecular docking results, compound **67** showed more extensive interactions with amino acid residues within the pocket. This interaction richness is advantageous for future target mechanism studies. Therefore, compound **67** was selected as the final candidate from the allosteric pocket.

Based on these analyses, we selected molecule **17** for the orthosteric pocket and molecule **67** for the allosteric pocket as dual-pocket binders for the CCR2 protein (Fig. 4H and I).

**3.4.5 Umbrella Sampling Analysis.** The umbrella sampling simulations were conducted over 27 sampling windows (Fig. 5A and B), encompassing two ligand systems (**17** and **67**), with each window simulated for 10 ns, yielding a total cumulative simulation time of 270 ns. The final center-of-mass (COM) distances between protein and ligand were determined as 6.562 nm for system **17** and 6.649 nm for system **67**, respectively. The WHAM confirmed that the sampling process maintained uniform statistical distribution across reaction coordinates.

The potential of mean force (PMF) profiles revealed distinct energy landscapes along the reaction coordinate ( $\xi$ ), which quantifies the spatial displacement from initial binding to complete dissociation. Both systems exhibited similar dissociation pathways, though with notable energy differences. For ligand **17**, the PMF minimum (-2.2 kcal mol<sup>-1</sup>) occurred at  $\xi \approx$

Table 1 Binding free energy calculation (kcal mol<sup>-1</sup>)<sup>a</sup>

| Energy                     | Compound 17 | Compound 67 |
|----------------------------|-------------|-------------|
| $\Delta$ VDDWAALS          | -51.41      | -44.17      |
| $\Delta$ EEL               | -87.04      | -35.48      |
| $\Delta$ EPB               | 112.75      | 57.6        |
| $\Delta$ ENPOLAR           | -5.21       | -4.05       |
| $\Delta$ G <sub>GAS</sub>  | -138.45     | -79.65      |
| $\Delta$ G <sub>SOLV</sub> | 107.54      | 53.54       |
| $\Delta$ G <sub>BIND</sub> | -30.91      | -26.11      |

<sup>a</sup> The energy components  $\Delta$ BOND,  $\Delta$ ANGLE,  $\Delta$ DIHED,  $\Delta$ 1–4 VDW,  $\Delta$ 1–4 EEL, and  $\Delta$ EDISPER exhibited negligible contributions (averaging zero) and were therefore omitted from the final analysis.



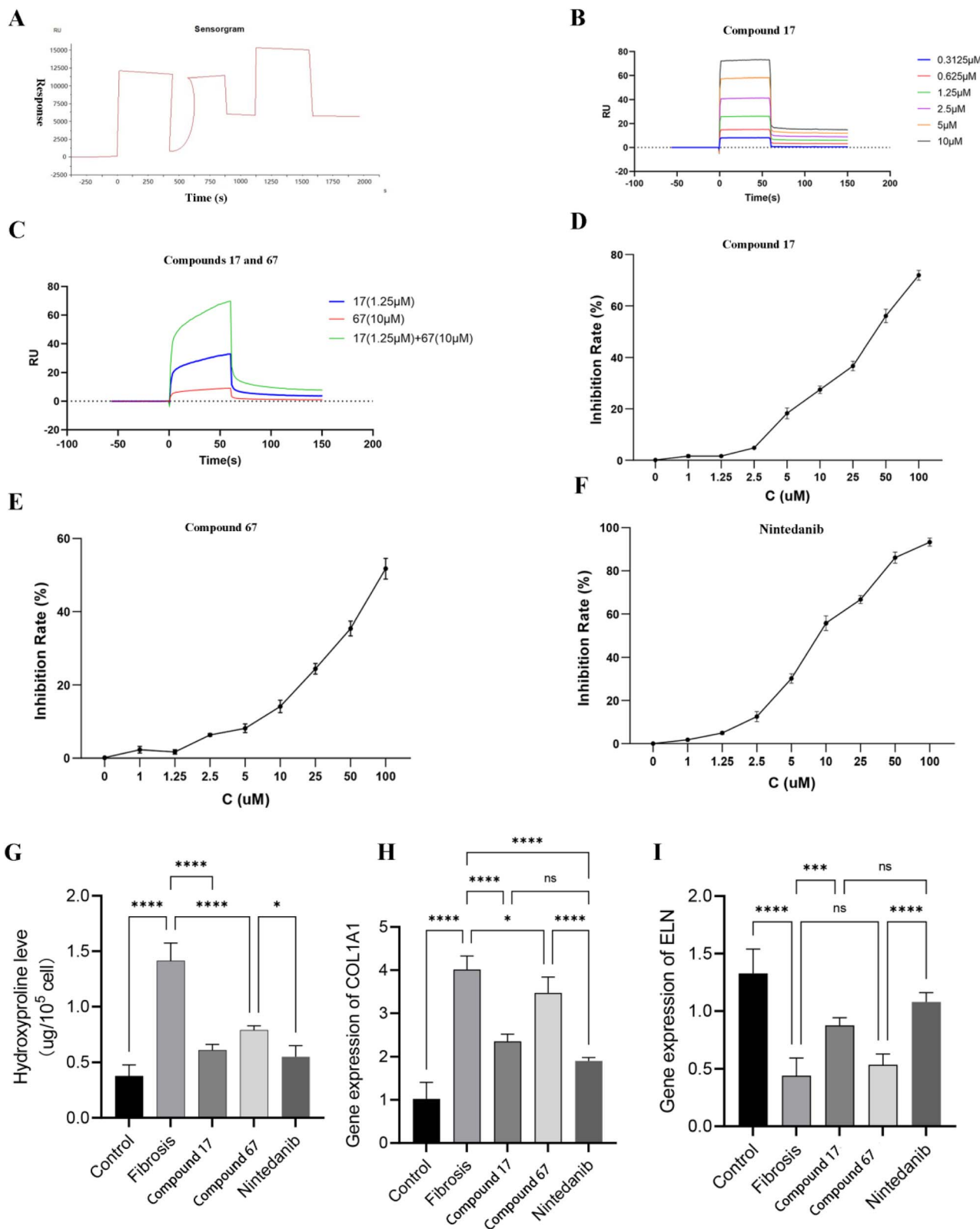


Fig. 7 Evaluation of compound 17 and compound 67 in molecular interaction, cell viability, and anti-fibrotic activity *in vitro*. (A) Mouse CCR2 protein coupling map. (B) Mouse CCR2 protein interaction assay with molecule 17. (C) Interaction assay of molecule 17, molecule 67 alone and co-injection with Mouse CCR2 protein. (D) The viability of compound 17-treated cells was measured using the CCK-8 assay. (E) The viability of compound 67-treated cells was measured using the CCK-8 assay. (F) The viability of compound nintedanib-treated cells was measured using the CCK-8 assay. (G) Hydroxyproline content in MRC-5 cells after TGF- $\beta$  stimulation and treatment with compounds 17, 67 and nintedanib. (\*\*\*\* $P < 0.0001$ ). (H) Relative mRNA expression levels of COL1A1 in each group. (\*\*\*\* $P < 0.0001$ ). (I) Relative mRNA expression levels of ELN in each group. (\*\* $P < 0.01$ ).



Table 2 The results of the assay with mouse CCR2 protein

| Receptor           | Ligand      | $K_D$ (M)             | $K_a$ (1 Ms <sup>-1</sup> ) | $K_d$ (1 s <sup>-1</sup> ) |
|--------------------|-------------|-----------------------|-----------------------------|----------------------------|
| Mouse CCR2 protein | Molecule 17 | $3.46 \times 10^{-6}$ | $1.05 \times 10^5$          | $3.63 \times 10^{-1}$      |

1.5 nm, indicating the most stable binding conformation. In contrast, ligand 67 demonstrated its global energy minimum ( $-1.35$  kcal mol<sup>-1</sup>) at  $\xi \approx 2.7$  nm.

Critical dissociation thresholds were identified at  $\xi \approx 4.5$  nm for ligand 17 and  $\xi \approx 4.2$  nm for ligand 67, corresponding to their respective equilibrium points for complete pocket egress. The PMF profiles revealed that ligand 17 maintains stronger binding affinity compared to 67, as evidenced by its deeper energy well ( $\Delta G = -2.2$  vs.  $-1.35$  kcal mol<sup>-1</sup>). These energy landscapes provide quantitative guidance for optimizing ligand binding geometries through targeted modifications of molecular interactions along the dissociation pathway.

#### 3.4.6 Binding free energy calculation using gmx\_MMPBSA.

The computed binding free energies revealed significant thermodynamic stability for both ligands:  $-30.91$  kcal mol<sup>-1</sup> for ligand 17 and  $-26.11$  kcal mol<sup>-1</sup> for ligand 67. These results confirm strong binding interactions between the ligands and CCR2's dual binding pockets, with ligand 17 demonstrating superior orthosteric binding affinity compared to the allosteric binding mode of ligand 67 (Fig. 6A and Table 1).

Per-residue energy decomposition was performed for amino acids within 4 Å of the ligand binding cores (Fig. 6B and C). Critical binding contributors ( $\Delta G < -1$  kcal mol<sup>-1</sup>) were identified as:

Ligand 17 (Table S13A): Trp98 (Chain A), Leu45 (Chain A), Gln288 (Chain B), Glu291 (Chain B), Val37 (Chain A)

Ligand 67 (Table S13B): Lys71 (Chain A), Phe306 (Chain A), Leu67 (Chain A), Tyr309 (Chain A), Ile66 (Chain A), Leu81 (Chain A)

Comparative analysis with PCA-derived binding poses revealed conserved interaction patterns:

Orthosteric site (ligand 17): recurrent residues Leu45, Trp98, Gln288, and Glu291

Allosteric site (ligand 67): conserved residues Leu67 and Leu81

Notably, Trp98 (Chain A) in the orthosteric pocket and Lys71 (Chain A) in the allosteric pocket exhibited the strongest binding contributions ( $-3.2$  kcal mol<sup>-1</sup> and  $-2.8$  kcal mol<sup>-1</sup>, respectively), suggesting their pivotal roles in ligand–receptor stabilization. These key residues likely mediate pharmacophore interactions essential for ligand efficacy, as evidenced by their consistent appearance across complementary analytical approaches.

### 3.5 SPR results

**3.5.1 Mouse CCR2 protein coupling.** Mouse CCR2 protein coupling amounted to a total of 5890 RU, of which the coupled sensing map is shown in Fig. 7A.

**3.5.2 Affinity determination.** Determination of the substance to be measured was carried out according to the

dissolved concentration, and the results of the determination of the kinetic fit of the substance to be measured are shown below.

(1) Mouse CCR2 protein and drug 17 in Fig. 7B.

(2) Compounds 17 and 67 were injected separately and co-injected in Fig. 7C.

**3.5.3 Wrap-up.** The purpose of this experiment was to determine the affinity of the to-be-tested substance to interact with Mouse CCR2 protein using Biacore, which was performed by using CM5 chip coupled protein. The results of the assay are shown in the table below (Table 2).

In result,<sup>2</sup> the affinity response of 17 + 67 co-injection was superior to that of 17 or 67 alone.

### 3.6 Viability assay results

The CCK8 assay results demonstrated that both compounds 17, 67 and nintedanib effectively suppressed the proliferation of pulmonary fibrosis cells induced by TGF- $\beta$ . Additionally, with increasing drug concentrations, the 48-hour cell inhibition rate showed a corresponding increase, as depicted in Fig. 7D–F. Furthermore, it is evident that compound 17 exhibits a stronger cytotoxic effect on the pulmonary fibrosis cell model compared to compound 67; whereas, nintedanib demonstrates an even stronger cytotoxic effect than compound 17.

### 3.7 Compounds 17 and 67 effectively inhibit TGF- $\beta$ -induced pulmonary fibrosis cell model

Finally, we evaluated the inhibitory effects of compounds 17 and 67 on pulmonary fibrosis at the cellular level, using nintedanib as a positive control. First, we established a pulmonary fibrosis cell model by stimulating MRC-5 cells with TGF- $\beta$ . Subsequently, we treated the cell model with compounds 17 (50  $\mu$ M), 67 (100  $\mu$ M), and nintedanib (10  $\mu$ M) for 48 hours. The levels of hydroxyproline, collagen COL1A1, and elastin ELN were measured in each group. We found that TGF- $\beta$  stimulation upregulated the expression of COL1A1 and hydroxyproline in MRC-5 cells, while elastin ELN expression was downregulated, indicating successful induction of the pulmonary fibrosis cell model. Following treatment with compounds 17, 67 and nintedanib, the expression of COL1A1 and hydroxyproline significantly decreased (Fig. 7G–I). Moreover, elastin ELN expression was upregulated upon treatment with compounds 17 and nintedanib. Compound 17 exhibited antifibrotic efficacy similar to that of nintedanib, while compound 67 demonstrated lower efficacy in inhibiting pulmonary fibrosis compared to nintedanib. However, the effective concentration of nintedanib for antifibrotic activity was significantly lower than that of compounds 17 and 67, indicating a stronger cytotoxic effect *in vitro*, which warrants further investigation (Fig. 7D–I). Nonetheless, due to its lower effective dose, nintedanib may potentially exert lower systemic toxicity *in vivo*. Hence, Further



optimization of compounds **17** and **67** and comprehensive *in vivo* evaluations are needed to advance their therapeutic potential. In summary, the cellular model suggests that compounds **17** and **67** have inhibitory effects on pulmonary fibrosis.

## 4 Discussion

This study presents a comprehensive structure-based approach for developing both orthosteric and allosteric CCR2 inhibitors as potential treatments for IPF. By integrating computational drug discovery with multi-layered pharmacological screening and molecular dynamics simulations, we systematically explored both binding sites and their inhibitory mechanisms.

### 4.1 CADD strategy and virtual screening results

Our computational approach successfully identified promising CCR2 modulators through a systematic drug discovery pipeline. The structure-based pharmacophore (SBP) modeling yielded highly specific models for both binding sites. For the external (orthosteric) pocket, SBPo-1 demonstrated optimal performance with a score of 12.047, sensitivity of 0.750, specificity of 0.619, and Youden's *J* statistic of 0.369, characterized by AADHHP pharmacophore features. The internal (allosteric) pocket pharmacophore SBPi-1 scored 10.582 with AAHHHN features and a *J* statistic of 0.176, indicating good discriminative ability.

The 3D-QSAR modeling further validated our approach, with QSAR1 showing excellent predictive capability (correlation coefficient of 0.913, fit value of 16.040, total cost of 240.708, and RMS of 3.737). The regression analysis yielded an  $R^2$  value of 0.83, demonstrating strong correlation between predicted and actual IC50 values across a 4-order magnitude range (from 0.03 nM for BDBM50088301 to 30 000 nM for BDBM50197912).

The multi-pharmacophore virtual screening of 152 406 molecules from the ChemDiv database was highly efficient: SBPo screening retained 1722 candidates, subsequent SBPi screening narrowed this to 69 molecules, and both 3D-QSAR and CFP screening maintained all 69 candidates. Following ADMET filtering, we identified 7 promising compounds (external pocket: ligands **17**, **40**, **68**; internal pocket: ligands **41**, **62**, **64**, **67**) for further analysis.

### 4.2 Molecular dynamics simulation insights

The 500 ns molecular dynamics simulations provided comprehensive insights into ligand–receptor interactions and stability. RMSD analysis revealed excellent system stability for all tested compounds, with mean values ranging from 0.443 nm (compound **62**) to 0.699 nm (compound **41**), indicating stable binding conformations. Notably, compounds **62** and **67** demonstrated superior stability compared to other candidates.

RMSF analysis identified six critical regions of amino acid fluctuations, with mean values ranging from 0.106 nm (compound **62**) to 0.174 nm (nintedanib). The analysis revealed specific residue regions crucial for binding: compound **17** affected residues 182–185, compound **41** influenced residues

19–24, compound **62** impacted residues 182–183, compound **64** affected multiple regions (11–16, 28–31, 232–234), and compound **67** influenced residues **9**, **19**, and **233**.

Radius of gyration analysis showed that compound **62** exhibited the most compact binding ( $R_g = 2.185$  nm), followed by compounds **64** and **41** (both 2.348 nm), while compound **17** showed slightly higher compactness (2.411 nm) compared to nintedanib (2.400 nm). SASA analysis demonstrated stable fluctuations around 185 nm<sup>2</sup> for all compounds, with compounds **62** (184.854 nm<sup>2</sup>) and **67** (184.744 nm<sup>2</sup>) showing the lowest values, indicating greater system stability.

Hydrogen bond analysis revealed significant differences in binding interactions: compounds **41** (2.728 bonds) and **67** (2.153 bonds) formed the most extensive hydrogen bond networks, while compound **17** showed moderate interaction (0.388 bonds). The short-range contact analysis ( $\leq 3.5$  Å) confirmed these patterns, with compound **67** (1.687 contacts) and compound **62** (1.642 contacts) demonstrating strong local interactions.

### 4.3 Principal component analysis and binding mode optimization

PCA analysis successfully identified the lowest energy conformations for each compound, providing crucial insights into optimal binding geometries. The free energy landscape analysis revealed distinct binding preferences and conformational stability patterns, guiding our final compound selection.

The optimized molecular docking based on PCA-derived conformations showed significantly improved binding scores compared to initial poses. Final docking scores (in kcal mol<sup>-1</sup>) were: compound **17** (−8.480), compound **41** (−8.500), compound **62** (−7.682), compound **64** (−7.570), compound **67** (−8.540), and nintedanib (−8.728). Interaction analysis revealed that compounds targeting the allosteric site (**41**, **62**, **64**, **67**) frequently interacted with key residues including VAL63, LEU67, LEU81, ALA141, ARG237, GLU238, TYR305, LYS311, PHE312, ARG138, ALA241, VAL244, and TYR315, while compound **17** at the orthosteric site primarily engaged LYS38, ALA42, LEU45, TYR49, TRP98, ALA102, TYR120, HID121, VAL187, PRO192, GLN288, GLU291, and THR292.

### 4.4 Thermodynamic validation and binding energetics

Umbrella sampling analysis provided detailed thermodynamic profiles for the dissociation pathways. The potential of mean force (PMF) calculations revealed that compound **17** achieved its most stable binding conformation at  $\xi \approx 1.5$  nm with a PMF minimum of  $-2.2$  kcal mol<sup>-1</sup>, while compound **67** reached its global energy minimum ( $-1.35$  kcal mol<sup>-1</sup>) at  $\xi \approx 2.7$  nm. Critical dissociation thresholds were identified at  $\xi \approx 4.5$  nm for compound **17** and  $\xi \approx 4.2$  nm for compound **67**, indicating compound **17**'s stronger binding affinity.

MM/PBSA calculations revealed exceptionally strong binding free energies:  $-30.91$  kcal mol<sup>-1</sup> for compound **17** and  $-26.11$  kcal mol<sup>-1</sup> for compound **67**, confirming thermodynamically favorable interactions. Per-residue energy decomposition identified critical binding contributors ( $\Delta G < -1$  kcal mol<sup>-1</sup>): for compound **17**, Trp98 ( $-3.2$  kcal mol<sup>-1</sup>),



Leu45, Gln288, Glu291, and Val37; for compound **67**, Lys71 ( $-2.8 \text{ kcal mol}^{-1}$ ), Phe306, Leu67, Tyr309, Ile66, and Leu81. The consistency between PCA-derived binding poses and MM/PBSA analysis, particularly the conservation of key residues (Leu45, Trp98, Gln288, Glu291 for compound **17**; Leu67 and Leu81 for compound **67**), validates our computational approach.

#### 4.5 Experimental validation and therapeutic efficacy

SPR analysis confirmed compound **17**'s direct binding to murine CCR2 with a dissociation constant ( $K_D$ ) of  $3.46 \mu\text{M}$  ( $K_a = 1.05 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ ,  $K_d = 3.63 \times 10^{-1} \text{ s}^{-1}$ ), demonstrating strong binding affinity consistent with computational predictions. The observed synergistic effect when compounds **17** and **67** were co-administered supports our dual-pocket inhibition strategy, with enhanced suppression compared to individual treatments.

In the BLM-induced IPF mouse model, CCR2 expression was significantly upregulated ( $p < 0.001$  for both mRNA and protein levels), with hydroxyproline content substantially increased ( $p < 0.001$ ), confirming the pathophysiological relevance of CCR2 in pulmonary fibrosis progression. Survival analysis revealed that patients with elevated CCR2 expression had significantly lower survival rates ( $p = 0.0049$ ), with ROC analysis yielding an AUC of 0.693, supporting CCR2's role as a diagnostic biomarker.

Cellular efficacy studies in TGF- $\beta$ -induced MRC-5 fibrosis models demonstrated concentration-dependent antifibrotic effects for both compounds. Treatment with compound **17** ( $50 \mu\text{M}$ ) and compound **67** ( $100 \mu\text{M}$ ) significantly reduced hydroxyproline content ( $p < 0.0001$ ) and COL1A1 expression ( $p < 0.0001$ ) while upregulating ELN expression ( $p < 0.001$  for compound **17**). Notably, compound **17** exhibited antifibrotic efficacy comparable to nintedanib ( $10 \mu\text{M}$ ), while requiring higher concentrations, suggesting potential for optimization to achieve therapeutic efficacy at lower doses.

#### 4.6 Clinical implications and future perspectives

The structural simplicity and lower development costs of compounds **17** and **67** position them as promising candidates for affordable IPF treatment, particularly valuable for resource-limited settings where current therapies like nintedanib and pirfenidone remain prohibitively expensive. The dual-pocket targeting strategy offers a novel therapeutic paradigm that could potentially overcome resistance mechanisms associated with single-site inhibition.

However, several limitations warrant consideration. The higher effective concentrations required compared to nintedanib indicate the need for structural optimization to enhance potency while maintaining selectivity. The absence of comprehensive *in vivo* pharmacological validation leaves questions regarding bioavailability, tissue distribution, and long-term safety unresolved.

## 5 Conclusions and future directions

Our integrated computational-experimental approach successfully identified compounds **17** and **67** as promising dual-pocket

CCR2 modulators with demonstrated antifibrotic efficacy. The comprehensive CADD strategy, validated through rigorous MD simulations and experimental studies, provides a robust framework for CCR2-targeted drug development. While current efficacy profiles require optimization, the cost-effectiveness potential and novel dual-site mechanism of action make these compounds attractive candidates for further development.

Future research priorities include: (1) structural optimization to enhance binding affinity and reduce effective concentrations, (2) comprehensive *in vivo* pharmacokinetic and toxicity studies, (3) development of nanomaterial-based delivery systems to improve bioavailability, and (4) integration of multi-omics approaches for personalized therapeutic strategies. The ultimate goal remains the development of cost-effective, accessible CCR2-targeted therapies for IPF and related fibrotic diseases.

## Author contributions

Gaofei Yan, Jianliang Huang, Ziyuan Huang and Mingsheng Lei developed the research idea. Gaofei Yan is responsible for the computational and analytical methods. Jianliang Huang and Jing Chen are responsible for the experimental or analytical methods. Shufang Luo, Haiying Guo, Shun Tao, Enping Li, Zimeng Liu, Jiangyuan Huang, Ying Liu and Yanlin Liu are responsible for managing and curating the data. Shuang Zhao, Jing Chen and Mingsheng Lei supervised the research project and provided materials, funding, or other resources.

## Conflicts of interest

The authors hereby declare that there are no conflicts of interest.

## Abbreviations

|       |   |
|-------|---|
| CCR   | Chemokine receptor  |
| SBPi  | Structure-based pharmacophore-inner pocket                |
| SBPo  | Structure-based pharmacophore-outer pocket                |
| CFP   | Common feature pharmacophore                              |
| PBVS  | Pharmacophore-based virtual screening                     |
| ADMET | Absorption, distribution, metabolism, excretion, toxicity |
| HB    | Hydrogen bond   |

## Data availability

The code is open source at <https://github.com/gaofeiyan/Molecular-Dynamics-Materials-for-Membrane-Protein.git> and [https://github.com/gaofeiyan/MD\\_Umbrella-Sampling.git](https://github.com/gaofeiyan/MD_Umbrella-Sampling.git). In molecular dynamics, we have extracted the structures corresponding to important frames. The codes and part of the trajectory files from the dynamics process have also can be found in the links.

Consent for publication: all authors have read and approved the submission of the manuscript.



This manuscript and supplementary information (SI) includes the details about chemical synthesis, supporting figures and tables that was collected or analyzed throughout the course of the present research. See DOI: <https://doi.org/10.1039/d5ra05026j>.

## Acknowledgements

This study received financial support from the Hunan Provincial Respiratory Disease Rehabilitation and Nursing Engineering Research Center Innovation Capacity Building Project (No. 202012), the Zhangjiajie Science and Technology Development Key Special Project (No. 202304), the National Key Clinical Specialty Major Scientific Research Project (No. 20230382).

## References

- 1 K. Tsubouchi, N. Hamada, S. Tokunaga, K. Ichiki, S. Takata, H. Ishii, *et al.*, Survival and acute exacerbation for patients with idiopathic pulmonary fibrosis (IPF) or non-IPF idiopathic interstitial pneumonias: 5-year follow-up analysis of a prospective multi-institutional patient registry, *BMJ Open Respir. Res.*, 2023, **10**(1), e001864.
- 2 X. Wu, W. Li, Z. Luo and Y. Chen, A comprehensive comparison of the safety and efficacy of drugs in the treatment of idiopathic pulmonary fibrosis: a network meta-analysis based on randomized controlled trials, *BMC Pulm. Med.*, 2024, **24**(1), 58.
- 3 M. Chianese, G. Screm, F. Salton, *et al.*, Pirfenidone and nintedanib in pulmonary fibrosis: lights and shadows, *Pharmaceuticals*, 2024, **17**(6), 709.
- 4 A. Mohanan, K. R. Washimkar and M. N. Mugale, Unraveling the interplay between vital organelle stress and oxidative stress in idiopathic pulmonary fibrosis, *Biochim. Biophys. Acta, Mol. Cell Res.*, 2024, **1871**(3), 119676.
- 5 M. Gharaee-Kermani, R. E. McCullumsmith, I. F. Charo, S. L. Kunkel and S. H. Phan, CC-chemokine receptor 2 required for bleomycin-induced pulmonary fibrosis, *Cytokine*, 2003, **24**(6), 266–276.
- 6 Y. Wan, X. Wang, T. Liu, *et al.*, Prognostic value of CCR2 as an immune indicator in lung adenocarcinoma: A study based on tumor-infiltrating immune cell analysis, *Cancer Med.*, 2021, **10**(12), 4150–4163.
- 7 C. He and A. B. Carter, C (C) Learning the role of chemokines in pulmonary fibrosis, *Am. J. Respir. Cell Mol. Biol.*, 2020, **62**(5), 546–547.
- 8 M. P. Keane, The role of chemokines and cytokines in lung fibrosis, *Eur. Respir. Rev.*, 2008, **17**(109), 151–156.
- 9 R. M. Strieter, B. N. Gomperts and M. P. Keane, The role of CXC chemokines in pulmonary fibrosis, *J. Clin. Invest.*, 2007, **117**(3), 549–556.
- 10 J. Gao, S. Peng, X. Shan, G. Deng, L. Shen, J. Sun, *et al.*, Inhibition of AIM2 inflammasome-mediated pyroptosis by Andrographolide contributes to amelioration of radiation-induced lung inflammation and fibrosis, *Cell Death Dis.*, 2019, **10**(12), 957.
- 11 Q. Lv, J. Wang, C. Xu, X. Huang, Z. Ruan and Y. Dai, Pirfenidone alleviates pulmonary fibrosis in vitro and in vivo through regulating Wnt/GSK-3 $\beta$ / $\beta$ -catenin and TGF- $\beta$ 1/Smad2/3 signaling pathways, *Mol. Med.*, 2020, **26**(1), 49.
- 12 Y. Zheng, L. Qin, N. V. Zacarias, H. de Vries, G. W. Han, M. Gustavsson, *et al.*, Structure of CC chemokine receptor 2 with orthosteric and allosteric antagonists, *Nature*, 2016, **540**(7633), 458–461.
- 13 S. J. Park, N. Kern, T. Brown, *et al.*, CHARMM-GUI PDB manipulator: various PDB structural modifications for biomolecular modeling and simulation, *J. Mol. Biol.*, 2023, **435**(14), 167995.
- 14 C. Tian, K. Kasavajhala, K. A. A. Belfon, *et al.*, ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution, *J. Chem. Theory Comput.*, 2019, **16**(1), 528–552.
- 15 T. P. Brown, D. E. Santa, B. A. Berger, L. Kong, N. J. Wittenberg and W. Im, CHARMM GUI Membrane Builder for oxidized phospholipid membrane modeling and simulation, *Curr. Opin. Struct. Biol.*, 2024, **86**, 102813.
- 16 S. Kim, J. Lee and S. Jo, CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules, *J. Comput. Chem.*, 2017, **38**(21), 1879–1886.
- 17 J. Wang, R. M. Wolf, J. W. Caldwell, *et al.*, Development and testing of a general amber force field, *J. Comput. Chem.*, 2004, **25**(9), 1157–1174.
- 18 H. Lit and R. Hoffmann, Pharmacophore perception, development, and use in drug design, *Pharmacophore Models*, 2000, **2**, 173.
- 19 B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr, *et al.*, CHARMM: the biomolecular simulation program, *J. Comput. Chem.*, 2009, **30**(10), 1545–1614.
- 20 J. Lee, X. Cheng, S. Jo, *et al.*, CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field, *Biophys. J.*, 2016, **110**(3), 641a.
- 21 F. Neese, F. Wennmohs, U. Becker, *et al.*, The ORCA quantum chemistry program package, *J. Chem. Phys.*, 2020, **152**(22), 224108.
- 22 T. Lu and F. Chen, Multiwfn: a multifunctional wavefunction analyzer, *J. Comput. Chem.*, 2012, **33**(5), 580–592.
- 23 P. Bauer, B. Hess and E. Lindahl, *GROMACS 2022.1 Manual*, 2022, DOI: [10.5281/zenodo.6637571](https://doi.org/10.5281/zenodo.6637571).
- 24 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, *et al.*, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 1983, **79**(2), 926–935.
- 25 H. Yu, S. Yang, Z. Chen, Z. Xu, X. Quan and J. Zhou, Orientation and Conformation of Hydrophobin at the Oil–Water Interface: Insights from Molecular Dynamics Simulations, *Langmuir*, 2022, **38**(19), 6191–6200, DOI: [10.1021/acs.langmuir.2c00614](https://doi.org/10.1021/acs.langmuir.2c00614).
- 26 J. A. Lemkul and D. R. Bevan, *J. Phys. Chem. B*, 2010, **114**(4), 1652–1660.
- 27 Y.-L. Han, H.-H. Yin, C. Xiao, M. T. Bernards, Yi He and Yi-X. Guan, Understanding the Molecular Mechanisms of Polyphenol Inhibition of Amyloid  $\beta$  Aggregation, *ACS*



- Chem. Neurosci.*, 2023, **14**(22), 4051–4061, DOI: [10.1021/acchemneuro.3c00586](https://doi.org/10.1021/acchemneuro.3c00586).
- 28 Y. Gu, L. Yan, B. Ma, K. Ren, C. Cao and N. Gu, Probing Conformational Transition of TRPV5 Induced by Mechanical Force Using Coarse-Grained Molecular Dynamics, *J. Chem. Inf. Model.*, 2023, **63**(21), 6768–6777, DOI: [10.1021/acs.jcim.3c00614](https://doi.org/10.1021/acs.jcim.3c00614).
- 29 J. T. Wiemann, D. Nguyen, S. Bhattacharyya, Y. Li and Y. Yu, Antibacterial Activity of Amphiphilic Janus Nanoparticles Enhanced by Polycationic Ligands, *ACS Appl. Nano Mater.*, 2023, **6**(21), 20398–20409, DOI: [10.1021/acsanm.3c04486](https://doi.org/10.1021/acsanm.3c04486).
- 30 D. Tim, K. Ahmad, J. R. Burns, Q. H. Nguyen, Z. S. Siwy, M. Tornow, P. V. Coveney, R. Tampé and H. Stefan, Principles of Small-Molecule Transport through Synthetic Nanopores, *ACS Nano*, 2021, **15**(10), 16194–16206, DOI: [10.1021/acsnano.1c05139](https://doi.org/10.1021/acsnano.1c05139).
- 31 P. A. Kollman, I. Massova, C. Reyes, *et al.*, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, *Acc. Chem. Res.*, 2000, **33**(12), 889–897.
- 32 M. S. Valdés-Tresanco, M. E. Valdés-Tresanco, P. A. Valiente and E. Moreno, gmx\_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS, *J. Chem. Theory Comput.*, 2021, **17**(10), 6281–6291, DOI: [10.1021/acs.jctc.1c00645](https://doi.org/10.1021/acs.jctc.1c00645).
- 33 B. R. Miller, T. Dwight McGee, J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, MMPBSA.py: An Efficient Program for End-State Free Energy Calculations, *J. Chem. Theory Comput.*, 2012, **8**(9), 3314–3321, DOI: [10.1021/ct300418h](https://doi.org/10.1021/ct300418h).
- 34 D. J. Oshannessy, M. Brighamburke, K. K. Sonesson, *et al.*, Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: use of nonlinear least squares analysis methods, *Anal. Biochem.*, 1993, **212**(2), 457–468.
- 35 Y. Liu, J. Jiang, S. Du, *et al.*, Artemisinins ameliorate polycystic ovarian syndrome by mediating LONP1-CYP11A1 interaction, *Science*, 2024, **384**(6701), eadk5382.
- 36 Cytiva, (n.d.), *Biacore T200 Evaluation Software Handbook*, (Edition AA, Doc. 28-9768-78) [PDF]. Retrieved August 14, 2025, from, <https://rrc.uic.edu/wp-content/uploads/sites/112/2020/02/BiacoreT200SoftwareHandbook.pdf>.
- 37 D. Li, X. Zhang, Z. Song, *et al.*, Advances in common in vitro cellular models of pulmonary fibrosis, *Immunol. Cell Biol.*, 2024, **102**(7), 557–569.

