Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

In Silico Identification of Natural Compound-Derived IL-18 Receptor Alpha Antagonists as Candidate Therapeutics for Graft-versus-Host Disease

Lahoussaine Aserrar

laserrar@um6ss.ma

Mohammed VI University of Sciences and Health (UM6SS), Laboratory of Onco-Pathology, Biology and Cancer Environment, Faculty of Medicine https://orcid.org/0009-0001-7355-6830

Hicham Charoute

Research Unit of Epidemiology, Biostatistics and Bioinformatics, Institut Pasteur du Maroc https://orcid.org/0000-0002-9338-6744

Salaheddine Redouane

Genomic Sequencing Laboratory, Institut Pasteur du Maroc https://orcid.org/0000-0002-5945-0386

Mounia Bendari

Mohammed VI University of Sciences and Health (UM6SS), Laboratory of Onco-Pathology, Biology and Cancer Environment, Faculty of Medicine https://orcid.org/0000-0001-8414-2855

Fadila Guessous

Mohammed VI University of Sciences and Health (UM6SS), Laboratory of Onco-Pathology, Biology and Cancer Environment, Faculty of Medicine

Research Article

Keywords: Graft-versus-host disease; Interleukin-18 receptor; Molecular docking; Molecular dynamics simulation; Natural products; Drug discovery; ADMET

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9440789/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

In Silico Identification of Natural Compound-Derived IL-18 Receptor Alpha Antagonists as Candidate Therapeutics for Graft-versus-Host Disease

Lahoussaine Aserrar[1], Hicham Charoute[4,5], Salaheddine Redouane[5], Mounia Bendari[3], Fadila Guessous[1,2]

Lahoussaine Aserrar: https://orcid.org/0009-0001-7355-6830

Hicham Charoute: https://orcid.org/0000-0002-9338-6744

Salaheddine Redouane: https://orcid.org/0000-0002-5945-0386

Mounia Bendari: https://orcid.org/0000-0001-8414-2855

[1] Laboratory of Onco-Pathology, Biology and Cancer Environment, Faculty of Medicine, Mohammed VI University of Sciences and Health (UM6SS), Casablanca, Morocco

[2] Department of Microbiology, Immunology and Cancer Biology, School of Medicine, University of Virginia, Charlottesville, VA 22908, USA.

[3] Department of Hematology, Faculty of Medicine, Mohammed VI University of Sciences and Health (UM6SS), Casablanca, Morocco

[4] Research Unit of Epidemiology, Biostatistics and Bioinformatics, Institut Pasteur du Maroc, Casablanca, Morocco

[5] Genomic Sequencing Laboratory, Institut Pasteur du Maroc, Casablanca, Morocco

Corresponding author: Lahoussaine Aserrar, Laboratory of Onco-Pathology, Biology and Cancer Environment, Faculty of Medicine, Mohammed VI University of Sciences and Health (UM6SS), Casablanca, Morocco. Email: laserrar@um6ss.ma

Abstract

Graft-versus-host disease (GvHD) is a life-threatening complication of allogeneic hematopoietic stem cell transplantation, driven by donor-derived immune cells that recognise and attack host tissues. The interleukin-18 receptor alpha chain (IL-18Rα), a member of the IL-1 receptor superfamily, has a pivotal proinflammatory role in GvHD pathogenesis by mediating IL-18 signalling and promoting IFN-γ production.

We applied an integrated computational approach combining molecular docking with AutoDock Vina and molecular dynamics simulations with GROMACS to identify candidate natural inhibitors of IL-18Rα. A library of phytochemicals from Moroccan medicinal species with documented anti-inflammatory properties, including *Nigella sativa*, *Curcuma longa* and *Rosmarinus officinalis*, was screened against the receptor ectodomain, and the three top-ranked complexes were simulated for 340 ns, followed by in silico ADMET profiling.

CNP0188167.1 showed the strongest predicted affinity (-11.5 kcal/mol), ahead of CNP0075206.3 (-10.0) and CNP0196376.4 (-9.8). The three complexes displayed distinct dynamic behaviours. CNP0196376.4 maintained a continuous hydrogen bond network with the receptor, present in 91.8 % of frames, whereas CNP0075206.3 and CNP0188167.1 reached only 29.1 % and 32.2 %, the latter following a biphasic profile with complete loss of hydrogen bonding between 68 and 136 ns and subsequent recovery above 98 %. ADMET profiling identified unfavourable predictions for all three compounds, including hepatotoxicity and assay interference alerts.

These compounds therefore represent structural starting points rather than ready candidates, and targeted chemical modification would be required to address the predicted liabilities. All findings are computational and require experimental validation.

Keywords:

Graft-versus-host disease; Interleukin-18 receptor; Molecular docking; Molecular dynamics simulation; Natural products; Drug discovery; ADMET

Introduction

IL-18R is a receptor for the interleukin-18 (IL-18), a member of the IL-1 family of cytokines, and plays a proinflammatory role with the ability to induce IFNγ production (Dinarello 1999, 2001). It was initially termed as interferon-gamma-inducing factor (Nakamura et al. 1989). This interleukin binds to the alpha chain of its receptor (IL-18Rα), and then the IL-18/IL-18Rα complex binds to IL-18Rβ to form the IL-18/IL-18Ra/IL-18Rβ ternary complex that juxtaposes the intracellular Toll-Interleukin-1 receptor (TIR) domains of IL-18Rα and IL-18Rβ, enabling the recruitment and the activation of MyD88/NF-κB pathway (**Figure 1**) (Tsutsumi et al. 2014; Rex et al. 2020).

IL-18R plays an important role in Graft-versus-host disease (GvHD), a condition that occurs after allogeneic hematopoietic stem cell transplantation (allo-HSCT), when donor immune cells recognize and attack recipient organs, depending on the type of GvHD (Teshima et al. 2016; Zhang et al. 2016; Kanate et al. 2020; Reddy et al. 2020).

Acute GvHD typically occurs within the first 100 days following transplantation and is primarily driven by early alloimmune responses, commonly affecting the skin, liver, and gastrointestinal tract. In contrast, chronic GvHD usually occurs after day 100, and it involves more complex immune dysregulation. It often resembles autoimmune disorders and affects multiple organs, such as the eyes, mouth, lungs, and joints (Pusic and Pavletic 2019; Malard et al. 2023).

Studies show multiple links between the IL-18/IL-18R pathway and GvHD. Increased expression of interleukin-18 receptor on T lymphocytes has been observed in patients with acute graft-versus-host disease after allogeneic bone marrow transplantation (Fujimori et al. 2002). Additionally, upregulation of the interleukin-18 receptor alpha chain was observed in both CD4+ and CD8+ T cells in acute graft-versus-host disease mice (Itoi et al. 2004). Caspase-1-mediated IL-18 release in T cells triggers inflammasome activation, which in turn exacerbates inflammation and increases GvHD severity (Talley et al. 2023). Furthermore, the upregulation of the IL-18R on T cells, caused by IL-33/ST2, results in increased IFN-γ production. This further exacerbates GvHD (Reichenbach et al. 2015). Importantly, Neutralizing IL-18Rα with an anti-IL-18Rα monoclonal antibody provides evidence of protective effects against acute graft-versus-host disease in a mouse model (Li et al. 2015).

Current treatments for graft-versus-host disease (GvHD) are mainly based on non-specific immunosuppressors, such as corticosteroids. These treatments, often prescribed empirically, are very toxic and increase the risk of relapse. Targeted options remain limited, although JAK inhibitors such as ruxolitinib have demonstrated efficacy

in steroid-refractory disease and are now approved for this indication. Precision medicine approaches will require validated biomarkers to guide treatment selection. Although the National Institutes of Health consensus criteria have improved clinical trial design, more effective and better tolerated therapies remain necessary (Pusic and Pavletic 2019).

In this study, we combined molecular docking and molecular dynamics simulations to identify candidate inhibitors of IL-18Rα. Docking was performed on phytochemicals from Moroccan medicinal plants known for their anti-inflammatory activities, including *Nigella sativa*, *Curcuma longa* and *Rosmarinus officinalis* (Table 1 lists all species included). Molecular docking was used to predict the binding mode and binding affinity scores of the candidate ligands (Trott and Olson 2010). In contrast, molecular dynamics simulations were performed to evaluate the stability of the protein-ligand complexes in a simulated physiological environment (Chakravorty et al. 2024).

This workflow was designed to prioritise natural compounds for subsequent experimental evaluation **(Figure 2).**

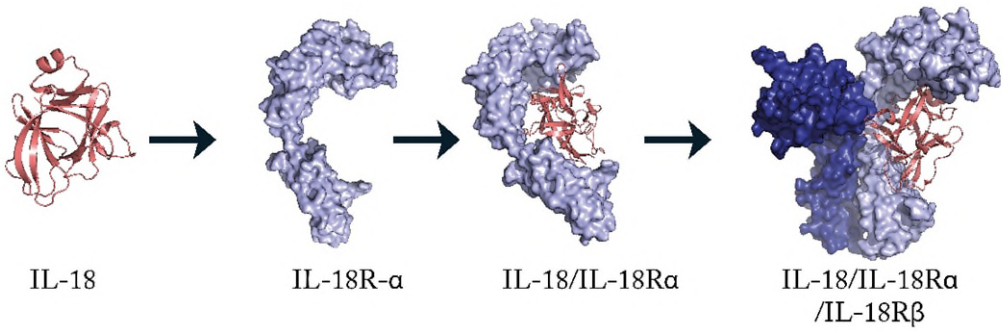


Figure 1. Stepwise assembly of the IL-18 receptor signaling complex. From left to right: free IL-18 (red ribbon), showing its β-trefoil fold; the unbound IL-18Rα ectodomain (light blue surface); the binary IL-18/IL-18Rα complex, in which IL-18 engages Site I of the receptor; and the ternary IL-18/IL-18Rα/IL-18Rβ complex, with IL-18Rβ shown in dark blue. Structures rendered from PDB entries 3WO3 and 3WO4 with PyMOL version 3.1.7.2.

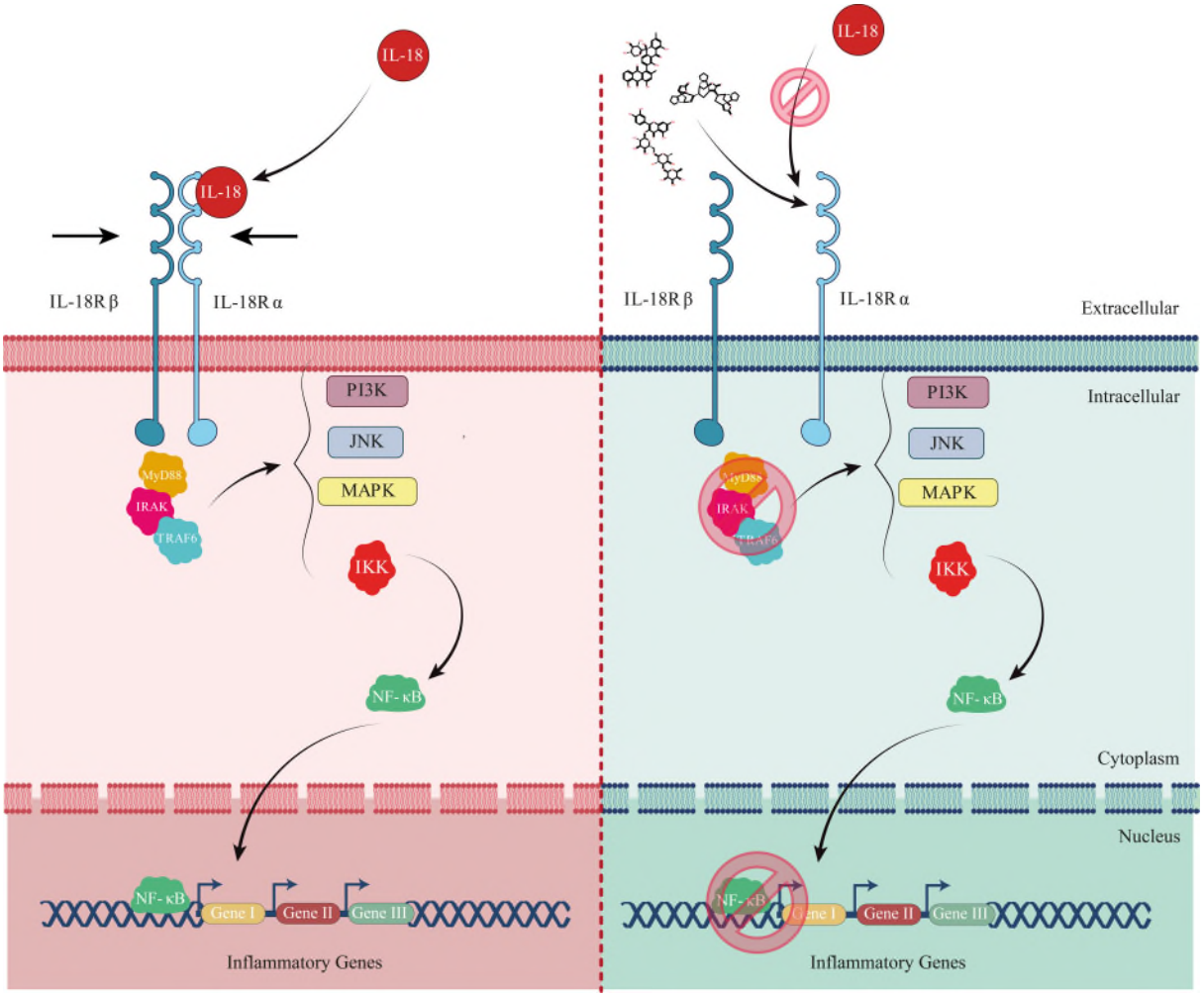


Figure 2. Proposed mechanism of IL-18Rα inhibition by plant-derived phytochemicals. ***Left panel (pink)***: under physiological conditions, IL-18 binds IL-18Rα and recruits IL-18Rβ, assembling the MyD88/IRAK/TRAF6 complex at the intracellular TIR domains. This activates IKK and the PI3K, JNK and MAPK branches, leading to NF-κB translocation to the nucleus and transcription of inflammatory genes. ***Right panel (green)***: candidate phytochemicals occupy Site I of IL-18Rα, preventing IL-18 binding. Ternary complex formation and MyD88 recruitment are blocked, NF-κB is not activated, and inflammatory gene transcription is suppressed. Red crossed circles mark the inhibited steps.

Materials and Methods

Protein structure preparation

The IL-18Rα ectodomain (chain B) was extracted from the crystal structure of the IL-18/IL-18Rα complex (PDB ID: 3WO3, resolution 3.10 Å). Residues Ile49 to Thr51 and the C-terminal segment Lys316 to Arg329 were unresolved in the crystal structure and were not modelled. Side-chain geometry was checked with Swiss-PDB Viewer (Kaplan 2001). Energy minimization was performed by steepest descent using GROMACS, and polar hydrogen atoms and Kollman charges were added with AutoDock Tools (Morris et al. 2009; Goodsell et al. 2021).

Ligands preparation

Natural compounds reported in 50 Moroccan medicinal plant species **(Table 1)** were retrieved from the COCONUT database in SDF format for every plant, COCONUT is an open-access database cataloging natural products derived from plants, microorganisms, and marine organisms. It offers structural and biological information for discovering new bioactive molecules. Containing more than 695,000 molecules, COCONUT's free accessibility and diversity make it ideal for virtual screening (Sorokina et al. 2021; Chandrasekhar et al. 2025).

A total of 3,390 compound entries were retrieved across the 50 species, corresponding to 2,631 unique structures after removal of compounds present in more than one species. Ligands were then optimized with the MMFF94 force field and converted to PDBQT format using OpenBabel prior to docking.

Identification of the IL-18Rα interaction interface

The IL-18 binding interface was characterized using three independent approaches. Recurrent contact residues were extracted from PDBsum profiles of the free and IL-18-bound receptor structures (3WO3 and 3WO4). Topological pocket prediction was performed with PrankWeb (https://prankweb.cz/). Crystallographic Site I residues were taken from Tsutsumi et al. (2014). The residues identified by each approach are listed in Table 2, together with those recovered by at least two of the three. This reference interface was defined prior to docking and was used to evaluate the resulting poses.

Molecular docking

Molecular docking was performed with AutoDock Vina v1.1.2 (Trott and Olson 2010). The search space was not restricted to a predefined pocket: the grid box was centred at **x = -41.953, y = 80.723, z = -18.154**, with

dimensions of **64 × 102 × 80 Å**, encompassing the entire IL-18Rα ectodomain. This unbiased configuration was chosen so that binding sites would be identified by the calculation rather than imposed a priori. Exhaustiveness was left at the default value of 8. Docking poses were subsequently evaluated against the consensus interface defined in **Table 2**.

Interaction analysis

Receptor surfaces and ligand poses of the three selected candidates were rendered with PyMOL version 3.1.7.2 (Schrödinger LLC). Two-dimensional protein-ligand interaction diagrams were generated with Discovery Studio V21.0 (Dassault Systèmes BIOVIA, San Diego, CA, USA) and exported in SVG format.

ADMET Properties prediction

Drug-likeness characteristics and ADMET properties (absorption, distribution, metabolism, excretion and toxicity) were predicted with ADMETlab 3.0 (https://admetlab3.scbdd.com/, accessed 22 October 2026). Compound selection had initially been performed using ADMETlab 2.0; re-evaluation with version 3.0 gave identical outcomes for the four rule-based filters used in the selection score, and all values reported here are those of version 3.0.

Ligand Screening of Top 10 Candidates and Final Top 3 Selection

The top 10 ligands with the highest binding affinity were selected from all docked compounds. The selection of the three candidates for MD simulation was carried out in two stages. First, analogous molecules and those with fragmented structures were eliminated to ensure structural diversity (marked "ND"). Next, a weighted selection score was calculated as:

$$\textbf{Score} = 0.2 \times S_dock + 0.5 \times S_MW + 0.3 \times S_ADMET$$

where S_{dock}, S_{MW} and S_{ADMET} are ordinal scores derived from the docking affinity, the molecular weight and the drug-likeness filters respectively. S_{dock} was assigned as 1 for -9 kcal/mol, 2 for -10 kcal/mol and 3 for ≤ -11 kcal/mol. S_{MW} was assigned as 0 for > 800 Da, 1 for 701-800 Da, 2 for 621-700 Da and 3 for ≤ 620 Da. S_{ADMET} was assigned as 1 if the compound passed the drug-likeness filters and 0 otherwise **(Excel Supplementary Data)**.

Since MD simulation of all ten candidates was not feasible within the available computational resources, a weighted selection score was used to reduce the set to three structurally distinct compounds.

The weights were set a priori to favour compounds within a synthesizable molecular weight range while retaining strong predicted affinity, and the scheme was used as a pragmatic triage filter rather than as a validated predictive model. Multi-parameter approaches of this type are widely used to balance competing properties during early compound selection (Bickerton et al. 2012; D. Segall 2012).

Ligand preparation for MD simulation

Ligand structures in MOL2 format were prepared using OpenBabel to ensure proper hydrogenation and correct ionization state. The SwissParam (https://old.swissparam.ch/) server was used to generate the topology file (.itp) and coordinates file (.pdb), which is compatible with CHARMM. Data from the (.gro) file of the ligands were added to one of the IL-18Rα, and then the (.itp) file was incorporated into the topol.top file, all these parameters were performed to ensure the compatibility of the complex with GROMACS algorithms before minimisation and equilibration (Supplementary data).

Molecular dynamics simulations

Molecular dynamics simulations were performed with GROMACS 2021.3 compiled with CUDA 11.3.1, using the CHARMM27 force field for the receptor and SwissParam derived topologies for the ligands. Each complex was placed in a cubic box with a minimum distance of 1.0 nm between the solute and the box edge, solvated with SPC/E water, and neutralised by adding Na^+ and Cl^- ions.

The solvated systems were energy minimised by steepest descent for up to 50,000 steps, until the maximum force fell below 1000 kJ mol^{-1} nm^{-1}. Equilibration was carried out in two stages of 100 ps each, under position restraints on the protein: first in the NVT ensemble at 300 K using the velocity rescaling thermostat with a coupling constant of 0.1 ps and separate coupling groups for the solute and for water and ions, then in the NPT ensemble at 1 bar using the Parrinello-Rahman barostat with a coupling constant of 2.0 ps and an isothermal compressibility of 4.5×10^{-5} bar^{-1}.

Production runs were carried out without restraints at 300 K and 1 bar with a 2 fs integration time step. Long range electrostatics were treated by particle mesh Ewald with a fourth order interpolation and a Fourier grid spacing of 0.16 nm, short range electrostatic and van der Waals interactions were truncated at 1.0 nm using the

Verlet cutoff scheme, a long range dispersion correction was applied to energy and pressure, and all bonds involving hydrogen were constrained with the LINCS algorithm. Coordinates were written every 10 ps.

The production phase was initially scheduled for 400 ns per complex. Repeated interruptions of the computing cluster stopped the three simulations at different times. To allow a direct comparison between the three systems, all trajectories were truncated to 340 ns, the duration reached by the shortest simulation, and all analyses reported in this study therefore cover an identical 340 ns window for the three complexes.

Trajectory analysis

Periodic boundary artefacts were removed in two steps with gmx trjconv: the raw trajectory was first corrected with the -pbc nojump option, then the complex was made whole and centred in a compact unit cell representation. All analyses were performed on this corrected trajectory, over the full 340 ns window, using the analysis tools of GROMACS 2021.3.

Backbone root mean square deviation was computed with gmx rms after least squares fitting of the backbone atoms to the energy minimised starting structure. Ligand root mean square deviation was computed separately, with each ligand superimposed onto its own reference structure, and therefore reports internal conformational flexibility rather than displacement within the binding site. Root mean square fluctuation was obtained with gmx rmsf on all protein atoms, averaged per residue with the -res option. The radius of gyration of the protein was computed with gmx gyrate.

Intermolecular hydrogen bonds between the receptor and the ligand were counted with gmx hbond, using the default geometric criteria of a donor to acceptor distance below 0.35 nm and a hydrogen donor acceptor angle below 30 degrees. The solvent accessible surface area of the complex was computed with gmx sasa on a custom index group comprising the protein and the ligand.

Essential dynamics was performed with gmx covar on all protein atoms, after least squares fitting to the reference structure, and the resulting covariance matrix was diagonalised. The trajectory was projected onto the first two eigenvectors with gmx anaeig. Free energy landscapes were constructed with gmx sham from the joint probability distribution of these two principal components at 298.15 K, the Gibbs free energy being expressed relative to the most populated conformational state.

Additional plots were produced with Python scripts, provided as supplementary data: distributions of protein and ligand RMSD, and projections of each trajectory onto the first two principal components coloured by simulation time, used to follow the temporal evolution of the conformational clusters explored by each complex.

Table 1. Moroccan medicinal plant species included in the screening library, with the study reporting their anti-inflammatory properties. *Flueggea virosa*, listed last, was not part of the screened library: it is included as the documented botanical source of CNP0188167.1, identified from the literature after the screening.

Plantes	***Study***	***Plantes***	***Study***
Ammi visnaga	Bouyahya et al. (2022)	*Mentha pulegium*	Mohammadi et al. (2024)
Ammodaucus leucotrichus	Bouyahya et al. (2022)	*Nasturtium officinale*	Klimek-Szczykutowicz et al. (2018)
Anacyclus pyrethrum	Elazzouzi et al. (2022)	*Nigella sativa*	Bourgou et al. (2010)
Argania spinosa	Bouyahya et al. (2022)	*Olea europaea*	El & Karakaya (2009)
	Makbal et al. (2021)	*Opuntia ficus-indica*	Lu (2023)
Artemisia herba-alba	Khlifi et al. (2013)	*Origanum vulgare*	Avola et al. (2020)
Asphodelus microcarpus	Bouyahya et al. (2022)	*Papaver rhoeas*	Bouyahya et al. (2022)
Berberis vulgaris	Bouyahya et al. (2022)	*Pelargonium graveolens*	Bouyahya et al. (2022)
Calendula arvensis	Bouyahya et al. (2022)	*Pistacia lentiscus*	Sehaki et al. (2023)
Capparis spinosa	El Azhary et al. (2017)	*Retama monosperma*	El Yadini et al. (2023)
Centaurium erythraea	El Menyiy et al. (2021)	*Retama raetam*	Bouyahya et al. (2022)
Ceratonia siliqua	Bouyahya et al. (2022)	*Rosmarinus officinalis*	Sousa Borges et al. (2018)
	Atta et al. (2023)	*Sideritis arborescens*	Bouyahya et al. (2022)
Chamaerops humilis	Bouyahya et al. (2022)	*Syzygium aromaticum*	Bouyahya et al. (2022)
Cistus monspeliensis	Bouyahya et al. (2022)	*Tetraclinis articulata*	Bouyahya et al. (2022)
Cistus salviifolius	Bouyahya et al. (2022)	*Thymus Algeriensis*	Sobeh et al. (2020)
Citrus aurantium L	Bouyahya et al. (2022)	*Thymus hirtus*	Lahlou et al. (2022)
Citrus limon	Bouyahya et al. (2022)	*Thymus vulgaris*	Bouyahya et al. (2022)
Cupressus sempervirens	Bouyahya et al. (2022)		Abdelhamed et al. (2022)
Curcuma longa	Aggarwal & Harikumar (2009)	*Viscum album*	Bouyahya et al. (2022)
Equisetum arvense	Bouyahya et al. (2022)	*Withania frutescens*	Bouissane et al. (2024)
Eucalyptus globulus	Bouyahya et al. (2022)	*Zingiber officinale*	Ayustaningwarno al. (2024)
Foeniculum vulgare	Bouyahya et al. (2022)	*Ziziphus lotus*	Bouyahya et al. (2022)
Genista tridentata	Laranjeira et al. (2023)	*Zygophyllum gaetulum*	Bouyahya et al. (2022)
Haloxylon scoparium	Bouyahya et al. (2022)	*Marrubium vulgare*	Bouyahya et al. (2022)
Jasonia glutinosa	Valero et al. (2025)	*Melissa officinalis*	Bouyahya et al. (2022)

Lavandula stoechas	Algieri et al. (2016)	***Flueggea virosa****	**Diarra et al. (2024)**

Table 2. Residues of the IL-18Rα Site I interface identified by three independent approaches. The consensus column lists the residues recovered by at least two of the three approaches. The asterisk marks the residue identified by all three.

Recurrent residues according to PDBsum profiles of 3WO3 and 3WO4 for IL-18Rα	**According to the topological pocket predictions using (PrankWeb)**	**According to Tsutsumi et al. 2014 (Site I)**	**Consensus**
Cys22	Thr23	Ser127	**Arg25 ***
Thr23	Arg25	Gln124	**Cys22**
Ser24	Ile28	Arg25	**Thr23**
Arg25	Thr29	His27	**Ser24**
Pro26	Val30	Cys22	**His27**
His27	Glu34	Cys41	**Ile28**
Ile28	Pro35	Cys43	**Thr29**
Thr29	Phe36	Cys81	**Cys43**
Cys43	Tyr37	Ser24	**Gln124**
Arg123	Glu141	Lys128	**Ser127**
Gln124	Asn142	Ser42	
Val125	Tyr144		
Thr126	Tyr145		
Ser127	Ile198		

* Identified by all three approaches.

Results

Docking Results

A total of 3,390 compounds were retrieved from the COCONUT database for the 50 selected species. After removal of structures shared between species, 2,631 unique compounds were retained and converted to PDBQT (Protein Data Bank, Partial Charge and Atom Type) format for docking with AutoDock Vina. Of these, 2,611 were successfully docked against the IL-18Rα ectodomain; the remaining 20 failed during format conversion or pose generation.

Docking scores ranged from -11.80 to -2.50 kcal/mol, with a mean of -6.40 and a standard deviation of 1.40. Two compounds scored ≤ -11.0 kcal/mol, 6 between -11.0 and -10.0, 37 between -10.0 and -9.0, 292 between -9.0 and -8.0, and 2,274 above -8.0 kcal/mol **(Figure 3.a)**.

Although the search space was not restricted to a predefined pocket, the highest-scoring poses of the ten top-ranked compounds converged on a single cavity. The centroids of their best poses lay within 7.5 Å of one another, with a maximum deviation of 4.4 Å from their mean position (x = -43.47, y = 76.11, z = -18.25). This cavity corresponds to the interface defined in **Table 2**, which had been characterized independently of the docking calculation.

The unbiased nature of the search was confirmed by the distribution of lower-ranked poses. Across the 90 poses generated for the ten top-ranked compounds, 15 (17 %) were located more than 10 Å from the mean centroid, with maximum displacements of up to 35 Å, indicating that alternative surface regions were effectively sampled. In contrast, the best-scoring pose of every compound fell within the same cavity, with all ten centroids lying within 4.4 Å of their mean position.

The ten highest-affinity ligands were evaluated for drug-likeness and ADMET properties with ADMETlab 3.0, and their interaction profiles were examined with Discovery Studio. CNP0238863.1 and CNP0238863.4 were discarded because of their fragmented structures. CNP0115399.1, CNP0232905.1 and CNP0232905.2 were discarded because of their close structural and chemical similarity to CNP0188167.1, which was retained as the representative of this series on the basis of its higher binding affinity. Among the remaining candidates, CNP0253739.1 and CNP0076333.1 obtained the lowest selection scores because of their molecular weights of 818.43 and 1106.11 Da, which fell outside the range considered suitable for further development; CNP0253739.1 was not retained despite having the most favourable docking score of the entire screen (-11.80

kcal/mol). CNP0188167.1, CNP0075206.3 and CNP0196376.4 were therefore selected for molecular dynamics simulation **(Table 3)**.

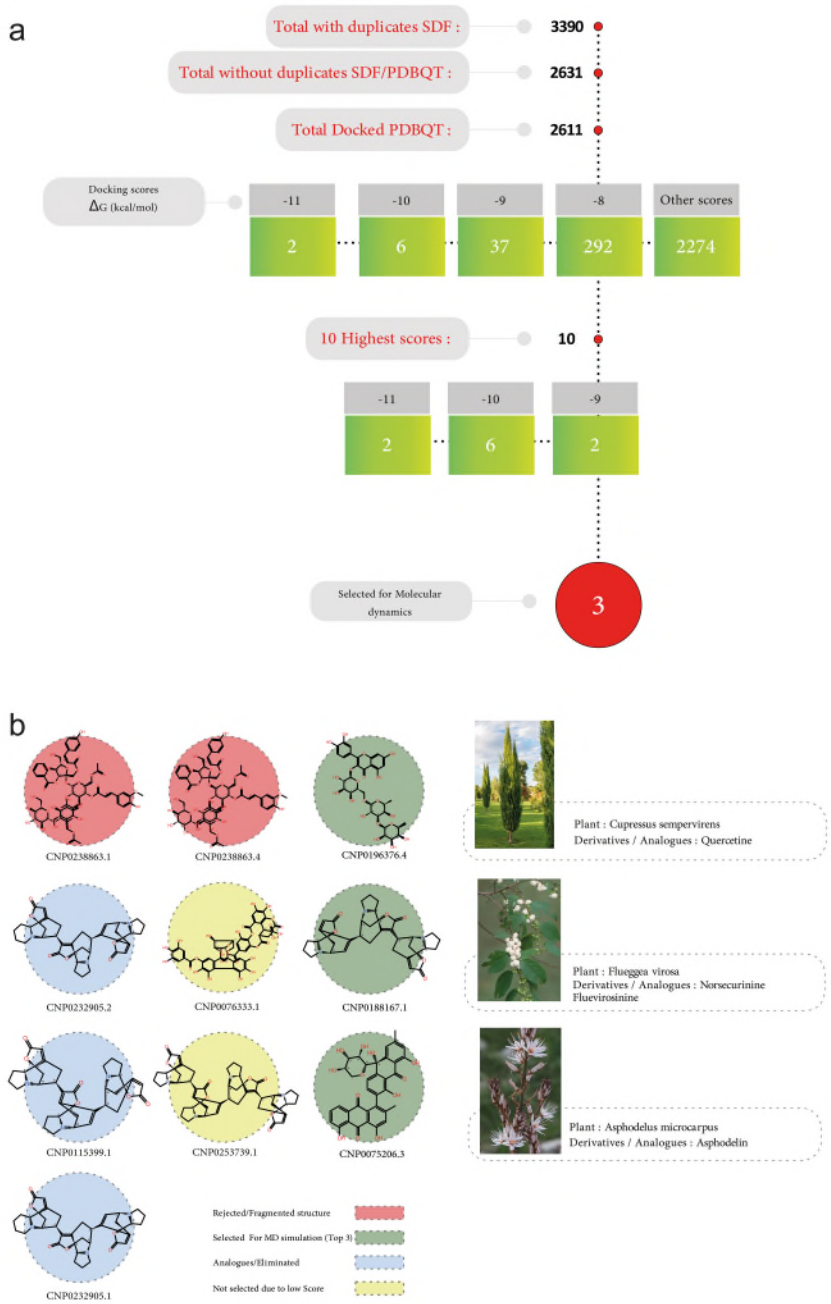


Figure 3. a) Virtual screening workflow and compound selection for molecular dynamics simulations.

b) Chemical structures and plant sources of selected natural compounds and their structural analogs. Among the 10 molecules with the highest binding affinities, three were selected for molecular dynamics simulation after elimination of fragmented structures (red), redundant analogues (blue), and compounds with low Selection scores (yellow). The three selected molecules (green) belong to distinct chemical classes: **CNP0188167.1**, a securinane-type alkaloid related to norsecurinine and the fluevirosinines, reported in *Flueggea virosa*; **CNP0075206.3**, a C-glycosylated anthraquinone–anthrone dimer reported in the genus *Asphodelus*; and **CNP0196376.4**, a triglycosylated quercetin derivative reported in *Cupressus sempervirens*.

Table 3. Docking scores, molecular weight, drug-likeness filter outcomes and selection scores for the ten highest-ranked compounds. ND indicates compounds eliminated before scoring, because of a fragmented structure or close similarity to a retained analogue.

Ligand ID	Docking Score	MW	Lipinski	Pfizer	GSK	Golden Triangle	Selection Score (Top 3)
CNP0238863.1	-10.10	660.23	Rejected	Accepted	Rejected	Rejected	ND
CNP0238863.4	-10.10	660.23	Rejected	Accepted	Rejected	Rejected	ND
CNP0115399.1	-10.00	609.28	Accepted	Accepted	Rejected	Rejected	ND
CNP0232905.2	-10.90	623.3	Accepted	Accepted	Rejected	Rejected	ND
CNP0232905.1	-10.60	623.3	Accepted	Accepted	Rejected	Rejected	ND
CNP0188167.1	-11.50	609.28	Accepted	Accepted	Rejected	Rejected	2.7
CNP0075206.3	-10.00	640.16	Rejected	Accepted	Rejected	Rejected	1.7
CNP0196376.4	-9.80	756.21	Rejected	Accepted	Rejected	Rejected	1.0
CNP0253739.1	-11.80	818.43	Rejected	Accepted	Rejected	Rejected	0.9
CNP0076333.1	-9.80	1106.11	Rejected	Accepted	Rejected	Rejected	0.5

ND: Not Determined

Table 4. Drug-likeness parameters of the three selected candidates, predicted with ADMETlab 3.0. QED ranges from 0 (all properties unfavourable) to 1 (all favourable); values above 0.67 are generally considered drug-like (Bickerton et al. 2012). Synthetic accessibility ranges from 1 (easy) to 10 (difficult).

Parameter	CNP0188167.1	CNP0075206.3	CNP0196376.4
LogP	1.59	3.40	0.84
LogD	1.79	2.92	1.31
LogS	-3.87	-6.34	-2.33
TPSA (Å²)	88.62	222.28	328.35
H-bond acceptors	9	12	20
H-bond donors	0	8	12
Rotatable bonds	2	2	8
Rings	12	7	6
QED	0.347	0.138	0.103
Synthetic accessibility	6.47	4.68	5.41
Fsp3	0.694	0.229	0.545

Table 5. ADMET profile of the three selected candidates, predicted with ADMETlab 3.0 (accessed 22 October 2026). Unless a unit is given, values are probabilities ranging from 0 to 1. Following the ADMETlab classification, 0-0.3 indicates a favourable prediction, 0.3-0.7 intermediate and 0.7-1.0 unfavourable. For human intestinal absorption and bioavailability, the probability refers to the unfavourable outcome, that is poor absorption or low bioavailability.

Category	Parameter	CNP0188167.1	CNP0075206.3	CNP0196376.4
Absorption	Caco-2 permeability (log cm/s)	-5.25	-5.73	-6.80
	MDCK permeability (cm/s)	1.37e-05	1.12e-05	9.96e-06
	Human intestinal absorption	0.000	0.043	0.859
	Bioavailability (F 20%)	0.000	0.008	0.986
	Bioavailability (F 30%)	0.005	0.982	1.000
	P-gp inhibitor	0.005	0.000	0.000
	P-gp substrate	0.147	0.297	0.987
Distribution	Plasma protein binding (%)	56.43	95.07	83.51
	Volume of distribution (L/kg)	1.199	1.952	0.888
	Fraction unbound (%)	37.51	6.77	15.57
	Blood-brain barrier penetration	0.031	0.000	0.000
Metabolism	CYP1A2 inhibitor	0.000	1.000	0.000
	CYP2C9 inhibitor	0.000	0.189	0.000
	CYP2C19 substrate	1.000	0.000	0.000
	CYP2D6 inhibitor	0.998	0.133	0.000
	CYP3A4 substrate	0.993	0.002	0.000
	Human liver microsomal stability	0.993	0.912	0.986

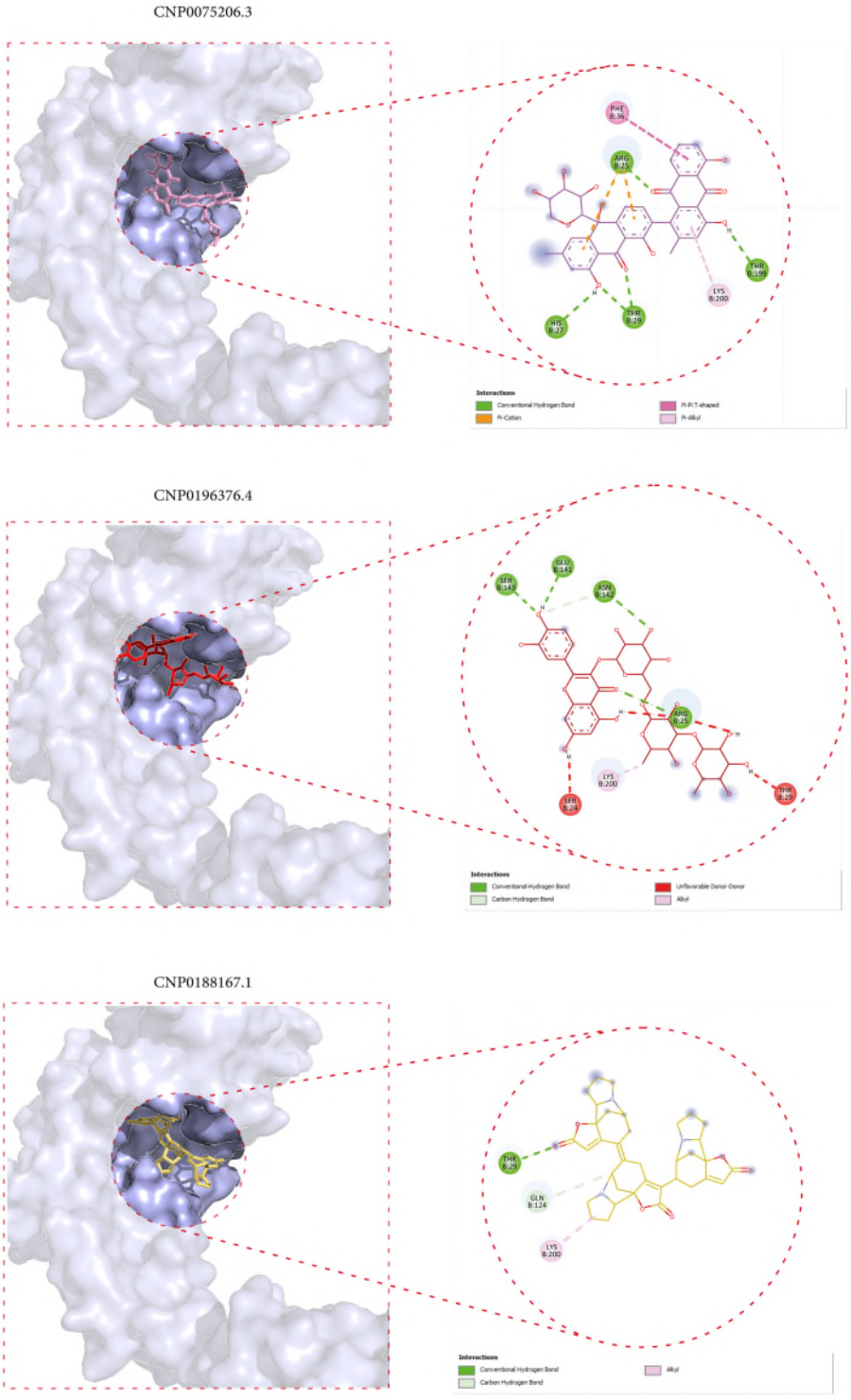
Category	Parameter	CNP0188167.1	CNP0075206.3	CNP0196376.4
Excretion	Plasma clearance (mL/min/kg)	3.64	1.43	0.95
	Half-life t1/2	1.791	3.477	5.485
Toxicity	hERG blockade	0.599	0.010	0.005
	Human hepatotoxicity	0.906	0.866	0.325
	Drug-induced liver injury	0.281	0.996	0.925
	AMES mutagenicity	0.855	0.989	0.936
	Genotoxicity	1.000	0.997	0.834
	Carcinogenicity	0.820	0.647	0.020
	Rat oral acute toxicity	0.957	0.182	0.036
	Maximum recommended daily dose	0.991	0.556	0.137
	Drug-induced nephrotoxicity	0.973	0.988	0.242
	Drug-induced neurotoxicity	0.967	0.046	0.001
	Ototoxicity	0.675	0.913	0.969
	Haematotoxicity	0.677	0.907	0.153
	Skin sensitisation	0.962	0.998	1.000
	Respiratory toxicity	0.898	0.840	0.110

Table 6. Structural alerts of the three selected candidates, predicted with ADMETlab 3.0. Values are the number of matching substructures identified by each rule set, except for colloidal aggregation, which is a probability ranging from 0 to 1. PAINS denotes pan-assay interference compounds.

Alert	CNP0188167.1	CNP0075206.3	CNP0196376.4
PAINS	0	1	1
Aquatic toxicity rule	3	4	2
Genotoxic carcinogenicity / mutagenicity rule	1	1	0
Non-genotoxic carcinogenicity rule	1	1	0
Skin sensitisation rule	3	3	8
FAF-Drugs4 rule	1	2	3
Colloidal aggregators (probability)	0.558	0.594	0.634

Discovery Studio Analysis:

Among the three selected ligands, **CNP0188167.1** showed the strongest predicted affinity but the sparsest interaction network, forming a conventional hydrogen bond with Thr29, a carbon-hydrogen bond with Gln124 and an alkyl contact with Lys200. **CNP0075206.3** displayed a more extensive profile, with conventional hydrogen bonds to Arg25, His27, Thr29 and Thr199, a π-cation interaction with Arg25, a π-alkyl contact with Lys200 and a π-π T-shaped contact with Phe36. **CNP0196376.4** formed hydrogen bonds with Arg25, Glu141, Asn142 and Ser143, together with carbon-hydrogen bonds, an alkyl contact with Lys200, and two unfavourable donor-donor contacts with Ser24 and Thr29 **(Figure 4)**.



335

336 **Figure 4.** Discovery Studio analysis of the three best IL-18Rα ligands

337

Drug-likeness and ADMET analysis

None of the three candidates satisfied all drug-likeness filters (**Tables 3 and 4**). **CNP0188167.1** was accepted by the Lipinski and Pfizer rules but rejected by the GSK and Golden Triangle rules, while **CNP0075206.3** and **CNP0196376.4** were accepted only by the Pfizer rule. QED values were low for all three (0.347, 0.138 and 0.103 respectively), well below the 0.67 threshold generally associated with drug-like compounds. Synthetic accessibility was rated difficult for **CNP0188167.1** (6.47) and easier for the other two (4.68 and 5.41).

Predicted absorption differed markedly between the compounds. **CNP0188167.1** showed a favourable profile, with low probabilities of poor intestinal absorption (0.000) and of being a P-glycoprotein substrate (0.147), although its Caco-2 permeability (-5.25 log cm/s) fell slightly below the -5.15 threshold. **CNP0196376.4** showed the opposite pattern, with high probabilities of poor absorption (0.859), low bioavailability (0.986 for F20%) and P-glycoprotein efflux (0.987), consistent with its triglycosylated structure, a topological polar surface area of 328.35 Å² and 12 hydrogen bond donors. **CNP0075206.3** was intermediate, with a very low predicted aqueous solubility (LogS -6.34) and high plasma protein binding (95.07 %).

Predicted metabolic liabilities were compound-specific. **CNP0188167.1** was predicted to be a CYP2C19 and CYP3A4 substrate (1.000 and 0.993) and a CYP2D6 inhibitor (0.998), whereas **CNP0075206.3** was predicted to inhibit CYP1A2 (1.000).

The predicted toxicity profiles were unfavourable for all three compounds, with a clear gradient **(Table 5)**. **CNP0075206.3** showed the highest number of unfavourable endpoints, including genotoxicity (0.997), drug-induced liver injury (0.996), AMES mutagenicity (0.989), nephrotoxicity (0.988) and skin sensitisation (0.998). **CNP0188167.1** showed genotoxicity (1.000), nephrotoxicity (0.973), neurotoxicity (0.967), acute oral toxicity (0.957), skin sensitisation (0.962), hepatotoxicity (0.906) and carcinogenicity (0.820). **CNP0196376.4** showed the most favourable profile of the three, with four unfavourable endpoints (skin sensitisation 1.000, AMES 0.936, drug-induced liver injury 0.925 and genotoxicity 0.834) and favourable predictions for carcinogenicity (0.020), hERG blockade (0.005), acute oral toxicity (0.036) and neurotoxicity (0.001).

Structural alert analysis **(Table 6)** identified pan-assay interference substructures in **CNP0075206.3** and CNP0196376.4, and flagged genotoxic carcinogenicity and skin sensitisation substructures in all three compounds. Predicted colloidal aggregation propensity was above 0.55 for all three.

Molecular Dynamics Results

Protein root mean square deviation

The backbone RMSD of the CNP0075206.3 and CNP0188167.1 complexes converged to similar values, with medians of 1.86 nm (IQR 1.82 to 1.90) and 1.83 nm (IQR 1.79 to 1.97) and a post-equilibration range of 1.71 to 2.19 nm, indicating comparable receptor behaviour with these two ligands. The CNP0196376.4 complex reached higher values, with a mean of 2.16 nm over the whole trajectory, a maximum of 2.48 nm and a mean of 2.36 nm after 100 ns, corresponding to approximately 17 % greater backbone deviation than the other two complexes. The main rearrangement occurred between 50 and 100 ns, after which the trajectory stabilised and remained stable from approximately 210 ns until the end of the simulation (Figure 5.a/5.c).

These absolute values are high for a single folded domain and reflect the modular architecture of the IL-18Rα ectodomain rather than any loss of secondary structure. The receptor consists of three immunoglobulin-like domains connected by flexible linkers, and in the absence of IL-18 and of the membrane anchor, nothing maintains the extended arrangement observed in the crystal structure. Two independent observations support this interpretation. First, the radius of gyration decreased from 3.2 to 3.4 nm at the beginning of the simulations to a plateau of 2.25 to 2.35 nm, and the solvent accessible surface area decreased in parallel, both indicating progressive compaction rather than unfolding (Figure 5.f, Figure 5.i). Second, the per-residue fluctuations are concentrated in the interdomain linkers and in the terminal segments, around residues 55 to 60, 220 and 290 to 315, while the domain cores remain below 0.5 nm (Figure 5.e). The global backbone RMSD therefore measures the reorientation of the three domains relative to one another, not a destabilisation of the receptor fold.

Ligand root mean square deviation

Ligand RMSD was computed after superimposing each ligand onto its own reference structure. These values therefore report the internal conformational flexibility of each molecule and not its displacement within the binding site. The three ligands show clearly separated distributions with non-overlapping interquartile ranges: 0.044 nm (IQR 0.036 to 0.056) for **CNP0188167.1**, 0.122 nm (IQR 0.089 to 0.146) for **CNP0075206.3** and 0.355 nm (IQR 0.326 to 0.397) for **CNP0196376.4**. **CNP0188167.1** behaves as a conformationally rigid molecule throughout the simulation, **CNP0075206.3** displays moderate internal motion with occasional peaks, and **CNP0196376.4**, the largest and most heavily glycosylated of the three, samples a broader range of internal con-

formations through rotation of its sugar moieties (**Figure 5.b/5.d**). Retention of each ligand within the binding site is addressed below, through the occupancy of protein-ligand hydrogen bonds.

Root Mean Square Fluctuations

The three complexes showed a common major peak around residue 60, indicating a flexible region of the receptor. **CNP0196376.4** induced the largest fluctuations, in several regions including residues 200 to 220 and 270 to 315. **CNP0188167.1** generated the lowest fluctuations, with a profile that was homogeneous along the sequence, and **CNP0075206.3** was intermediate **(Figure 5.e)**.

Four regions stood out in all three profiles. The N-terminal segment, residues 1 to 25, showed high initial flexibility. A second highly flexible region, probably a functional loop, was located between residues 40 and 60. A third region, close to the bound ligand, was centred on residues 200 to 220. Finally, residues 270 to 315 fluctuated markedly, particularly in the presence of **CNP0196376.4** and **CNP0188167.1** (**Figure 5.e**).

Radius of Gyration

CNP0075206.3 and **CNP0188167.1** reached a stable plateau of approximately 2.25 to 2.35 nm within the first 50 ns. **CNP0196376.4** showed a longer adaptation phase, with an initially high radius of gyration of 3.0 to 3.4 nm maintained between 0 and 100 ns, followed by a progressive decrease towards approximately 2.30 nm after 200 ns (Figure 5.f). As discussed above, this decrease reflects the closure of the three immunoglobulin-like domains of the isolated ectodomain rather than a loss of secondary structure.

Potential energy

The total potential energy of each system remained stable throughout the production run, fluctuating by approximately 1.7×10^3 kJ/mol around a constant mean, which confirms that the three simulations were properly equilibrated. Absolute potential energies were not compared between systems. These values are dominated by the solvent, and the offset of about 1.6×10^4 kJ/mol between the **CNP0196376.4** system and the other two corresponds to a difference of roughly 380 water molecules in the simulation box. They therefore carry no information on the quality of ligand binding.

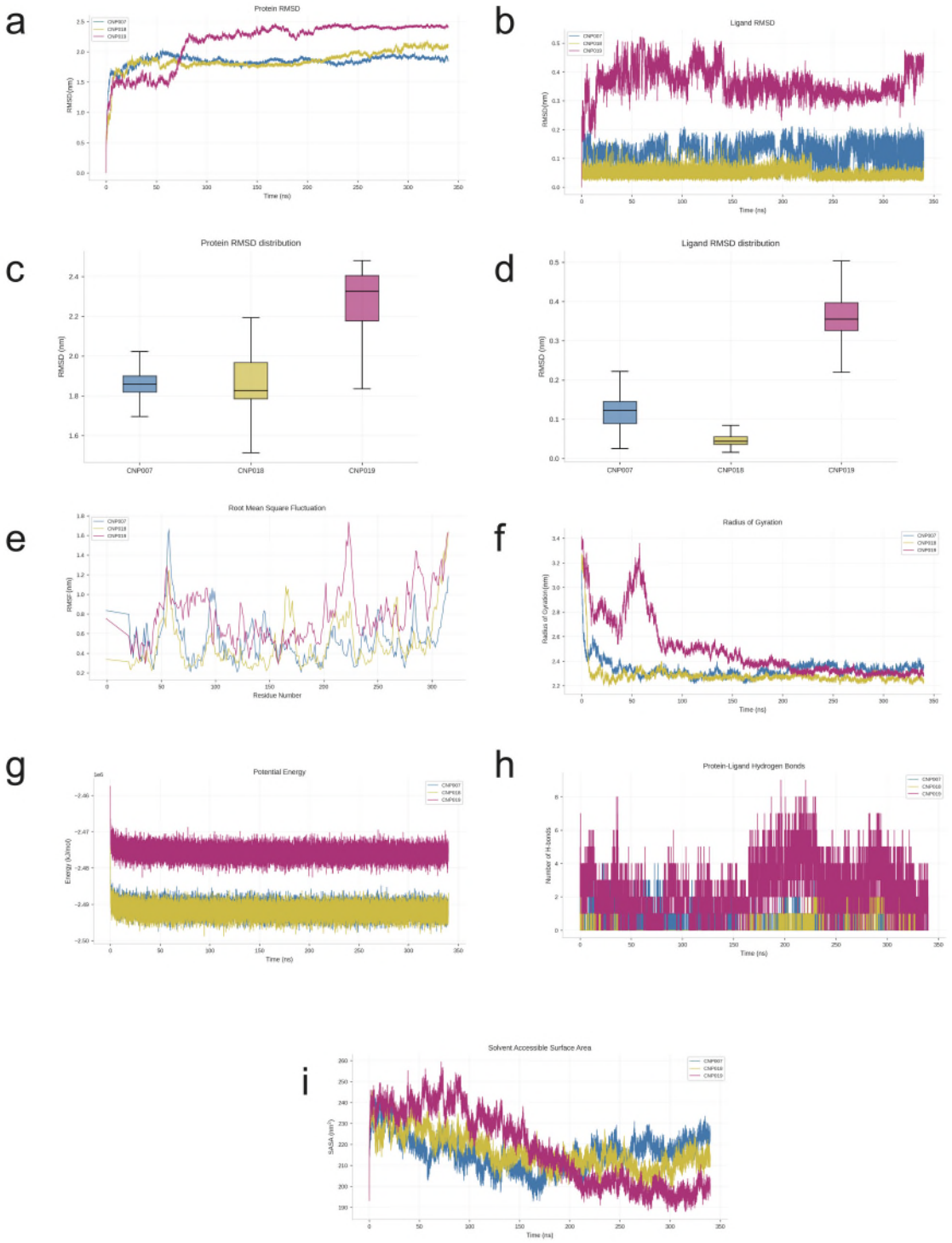
Protein-ligand hydrogen bonds

Hydrogen bond occupancy, defined as the proportion of frames in which at least one hydrogen bond is present between the receptor and the ligand, differs markedly between the three complexes. Over the full 340 ns trajectory, **CNP0196376.4** maintains an average of 2.49 hydrogen bonds with an occupancy of 91.8 %, whereas **CNP0075206.3** and **CNP0188167.1** average 0.37 and 0.46 hydrogen bonds with occupancies of 29.1 % and 32.2 % respectively, with transient maxima of up to nine simultaneous bonds for CNP0196376.4 **(Figure 5.h)**.

The temporal profiles are more informative than these averages. **CNP0196376.4** maintains occupancy between 72 and 99 % in every 34 ns window, indicating a continuously engaged hydrogen bond network. **CNP0188167.1** follows a biphasic profile: occupancy falls to zero between 68 and 136 ns, then rises progressively and reaches 98 % from 238 ns onward. A ligand released into bulk solvent would not reestablish a persistent hydrogen bond network, so this profile indicates that the ligand remained within the binding site and adopted a more favourable orientation, consistent with the reduction of its internal RMSD over the same interval. **CNP0075206.3** shows a monotonic decrease, from 68.9 % during the first 34 ns to 1.5 % over the last 34 ns. Because this compound is a polyaromatic dimer capable of binding through π stacking and hydrophobic contacts, the loss of hydrogen bonding does not establish that the ligand left the binding site, and this point is addressed in the limitations.

Solvent accessible surface areas

Solvent accessible surface area reflects the exposure of the protein-ligand complex to the aqueous solvent and provides a measure of system compactness. **CNP0196376.4** displayed the highest initial SASA, averaging 237.8 nm² over the first 50 ns and 241.5 nm² between 50 and 100 ns, followed by a progressive decrease to 222.5 nm² between 100 and 200 ns and 199.4 nm² over the final phase. **CNP0188167.1** and **CNP0075206.3** started from lower values (228.8 and 227.8 nm² respectively over the first 50 ns) and stabilized more rapidly, with overall means of 216.4 and 216.1 nm². The three complexes converged toward comparable mean values (216.4, 216.1 and 218.0 nm²), but **CNP0196376.4** required a markedly longer adaptation phase **(Figure 5.i)**.



443

444 **Figure 5.** Molecular dynamics trajectory analyses of the three IL-18Rα-ligand complexes over 340 ns. (a)
445 Backbone RMSD of the protein over time. (b) Internal ligand RMSD, computed after superimposing each ligand
446 onto its own reference structure. (c) Distribution of the backbone RMSD. (d) Distribution of the internal ligand
447 RMSD. (e) Per-residue root mean square fluctuation. (f) Radius of gyration of the protein. (g) Total potential
448 energy of the system. (h) Number of hydrogen bonds between the receptor and the ligand. (i) Solvent accessible

surface area of the complex. All analyses were performed on the trajectory corrected for periodic boundary conditions and cover an identical 340 ns window for the three systems. **CNP0075206.3** in blue, **CNP0188167.1** in yellow, **CNP0196376.4** in magenta.

Principal Component Analysis

Principal component analysis describes the distribution of conformational states explored by each protein-ligand complex during the simulation, the first two components capturing the dominant collective motions. The three complexes differed markedly in the extent of this exploration. **CNP0196376.4** occupied by far the largest region of the PC1-PC2 space, spanning PC1 from -40 to +121 and PC2 from -45 to +50, and formed several distinct clusters indicative of multiple conformational states. **CNP0188167.1** was intermediate, spanning PC1 from -67 to +42 and PC2 from -23 to +41, and followed a continuous path rather than occupying several separate regions. **CNP0075206.3** explored the narrowest region of the three, spanning PC1 from -36 to +59 and PC2 from -29 to +32 (**Figure 6.a**).

Separated PCA clusters with temporal evolution in the space

For **CNP0188167.1** (**Figure 6.b**), the trajectory began with a compact cluster between 0 and 50 ns, located at PC1 from -60 to -40 and PC2 from +10 to +40, followed by a continuous transition between 50 and 200 ns across PC1 from -40 to +20 and PC2 from +40 to -20. This transition followed a smooth, quasi-circular path without abrupt jumps. From 200 to 340 ns the system stabilized at PC1 from +20 to +40 and PC2 from -20 to +30, forming the tightest and densest distribution of the three complexes.

For CNP0075206.3 (**Figure 6.c**), an initial dense cluster was observed between 0 and 50 ns at PC1 from +20 to +60 and PC2 from +10 to -30, with only minor excursions consistent with thermal fluctuations. Between 50 and 150 ns the system migrated towards PC1 from -10 to +20 and PC2 from +10 to +30, and from 150 to 340 ns it settled at PC1 from -40 to -10 and PC2 from -20 to +10, with a visibly gradual transition and a broader conformational spread than **CNP0188167.1**.

For **CNP0196376.4** (**Figure 6.d**), the first 150 ns showed the widest dispersion of the three, spanning PC1 from +60 to +120 and PC2 from -45 to +20 without convergence, with several secondary clusters forming after 130 ns. The system then migrated towards PC1 from +40 to +80 and PC2 from -40 to +20, retaining a broad distribution with several visible trajectory branches. A major transition occurred between 150 and 250 ns towards PC1

from -10 to +40 and PC2 from +20 to +50, followed between 250 and 340 ns by a rapid shift to PC1 from -40 to -20 and PC2 from 0 to -25, where dense and stable clusters formed at the end of the simulation.

Free Energy landscape

The free energy landscape of **CNP0075206.3 (Figure 6.e)** was the most rugged of the three, with a global minimum at PC1 ≈ -27 and PC2 ≈ -13 and low-energy regions, defined here as $\Delta G \leq 2.5$ kJ/mol, distributed across a wide range of PC1 from -30 to 53. This fragmentation into several shallow minima separated by low barriers is consistent with more than one accessible binding geometry rather than a single dominant pose (Qiang et al. 2025).

CNP0188167.1 (Figure 6.f) showed an intermediate profile, with a global minimum at PC1 ≈ -7 and PC2 ≈ -15 and low-energy bins confined to PC1 from -7 to 35.

CNP0196376.4 (Figure 6.g) showed the most sharply defined landscape, with a single deep basin at PC1 ≈ -35 and PC2 ≈ -17 and a low-energy region restricted to PC1 from -35 to -29 and PC2 from -17 to -8. This basin coincides with the dense cluster occupied by the trajectory over the final 90 ns, indicating that the complex converged towards a single well-defined conformational state after its initial adaptation phase.

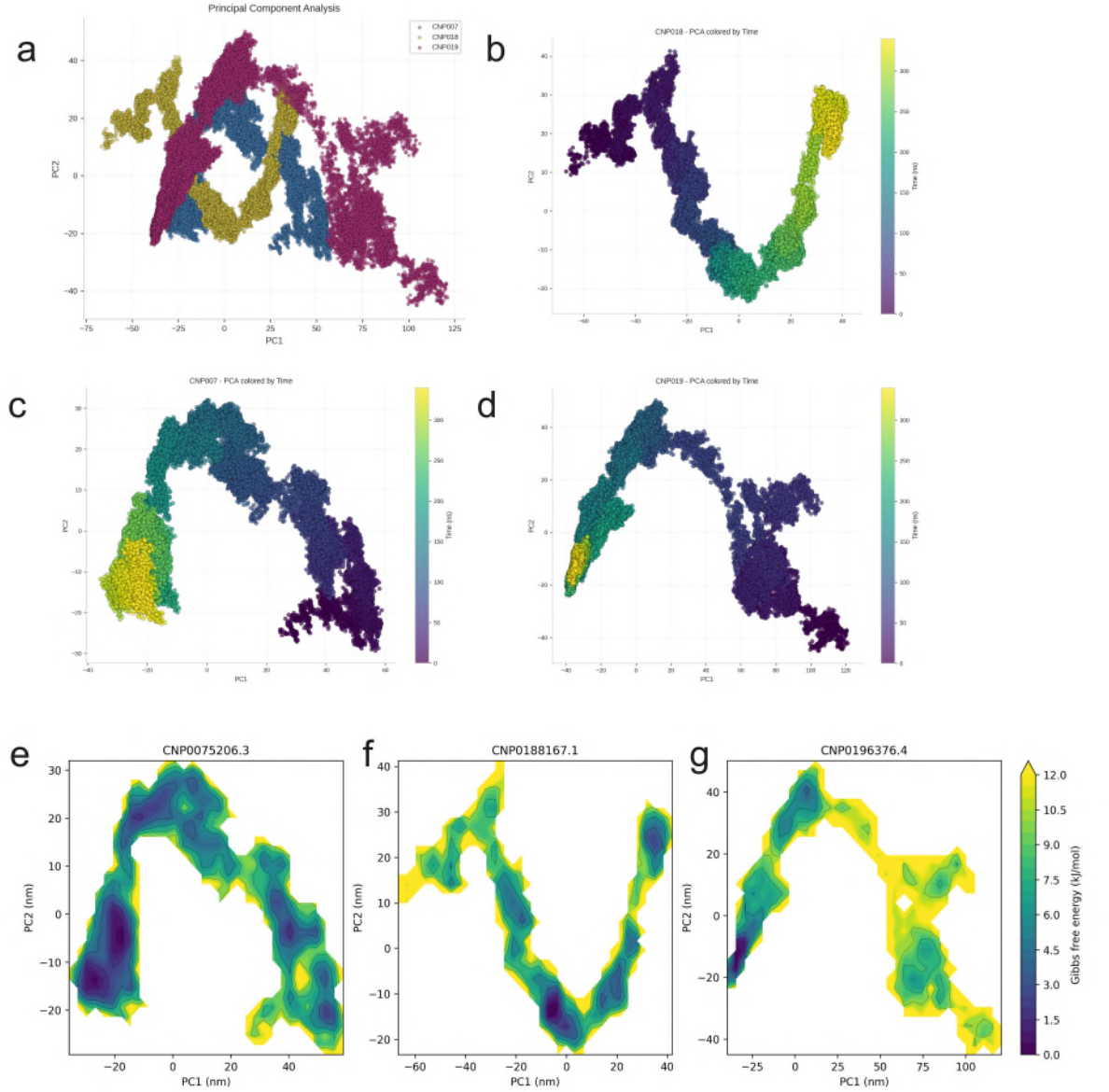


Figure 6. Principal component analysis and free energy landscapes. (a) Projection of the three trajectories onto the first two principal components. (b to d) Individual projections coloured by simulation time, showing the temporal evolution of the conformational regions explored by each complex. (e to g) Gibbs free energy landscapes constructed from the joint probability distribution of the first two principal components. Free energy is expressed relative to the most populated state. Regions left in white correspond to areas of conformational space never sampled during the simulation.

Discussion

One secondary result is worth noting. The screening was conducted without restricting the search space to a predefined pocket, and yet the best poses of the top ten compounds converged on a single cavity, whose centroids lie within 4.4 Å of their average position, while 17% of the lower-ranked poses were distributed elsewhere on the receptor surface. This cavity corresponds to the interface independently defined by crystallographic structure analysis, topological pocket prediction, and data from Tsutsumi et al. (2014). The convergence of an unbiased screening toward an interface identified by three independent approaches provides further support for the relevance of Site I as a viable binding site on IL-18Rα.

The three selected compounds are natural products retrieved from the COCONUT database. **CNP0188167.1** is a securinane-type alkaloid related to norsecurinine and the fluevirosinines, reported in *Flueggea virosa*, whose extracts have been shown to inhibit pro-inflammatory cytokines (Diarra et al. 2024). **CNP0075206.3** is a C-glycosylated anthraquinone–anthrone dimer reported in the genus *Asphodelus* (Khalfaoui et al. 2021), a Mediterranean group whose anthraquinones have been reported to modulate inflammatory signalling pathways (Bouyahya et al. 2022). **CNP0196376.4** is a triglycosylated quercetin derivative (C33H40O20) reported in several plant species, including *Cupressus sempervirens*, which was the entry point for this compound in the present library; quercetin and its glycosides are polyphenolic flavonoids with documented anti-inflammatory properties (Bouyahya et al. 2022).

CNP0188167.1 showed the strongest binding of the three candidates, with a docking score of -11.50 kcal/mol, placing it 3.65 standard deviations below the mean of the screened library. It engages Thr29 through a conventional hydrogen bond, Gln124 through a carbon-hydrogen bond and Lys200 through an alkyl contact, two of which, Thr29 and Gln124, belong to the consensus interface defined in Table 2. Among the three compounds it also satisfies both the Lipinski and Pfizer filters and shows the most favourable predicted absorption profile, with a low probability of P-glycoprotein efflux and no hydrogen bond donors.

Molecular dynamics gave a more nuanced picture. The ligand itself behaved as a conformationally rigid molecule, with the lowest internal RMSD of the three (0.044 nm), and the receptor remained compact, with a radius of gyration between 2.25 and 2.35 nm and the lowest per-residue fluctuations of the three complexes. Its engagement with the receptor was nevertheless discontinuous: hydrogen bond occupancy fell to zero between 68 and 136 ns before rising above 98 % from 238 ns onward, giving an overall occupancy of 32.2 % across the trajectory. Since a ligand released into bulk solvent would not re establish a persistent hydrogen bond network,

this profile is consistent with repositioning within the binding site rather than with dissociation, but it indicates that the initial docked pose was not the one retained over the second half of the simulation.

However, these favorable binding and stability properties are offset by a predicted unfavorable safety profile. **CNP0188167.1** is flagged on eleven toxicological endpoints, including genotoxicity (1.000), nephrotoxicity (0.973), neurotoxicity (0.967), acute oral toxicity (0.957), skin sensitization (0.962), human hepatotoxicity (0.906), and carcinogenicity (0.820), and carries structural alerts for genotoxic carcinogenicity and skin sensitization. Securinine, the parent compound of this class, contains an α,β-unsaturated lactone and acts as a selective GABA receptor antagonist, causing severe convulsions in mice at 20 mg/kg (Klochkov and Neganova 2021). The Michael acceptor motif present in this lactone may account for the genotoxicity and sensitization warnings, while the documented GABAergic activity of the class is consistent with the predicted neurotoxicity. This compound should therefore be considered a starting skeleton requiring structural optimization rather than a candidate for direct experimental evaluation.

CNP0075206.3 bound with a docking score of -10.00 kcal/mol and showed the most extensive interaction network of the three, engaging Arg25, His27, Thr29, Phe36, Thr199 and Lys200 through a combination of hydrogen bonds, a π-cation interaction with Arg25 and a π-π T-shaped contact with Phe36.

In molecular dynamics, it showed intermediate behaviour, with a ligand RMSD of 0.12 nm, occasional conformational fluctuations and a variable number of hydrogen bonds (0 to 4). Its free energy landscape displayed two adjacent low-energy regions together with several shallower minima, consistent with more than one accessible binding geometry rather than a single dominant pose (Qiang et al. 2025). Optimization of this scaffold would need to address its very low predicted aqueous solubility (LogS -6.34) and its high plasma protein binding (95.07 %), in addition to the redox-active core discussed below.

This compound showed the least favourable predicted safety profile of the three, with thirteen unfavourable endpoints including genotoxicity (0.997), drug-induced liver injury (0.996), AMES mutagenicity (0.989) and nephrotoxicity (0.988), together with predicted activation of the aryl hydrocarbon and oestrogen receptor pathways. It also carries a pan-assay interference substructure. Polyhydroxylated anthraquinones can undergo redox cycling that generates reactive oxygen species, a well-documented mechanism of quinone toxicity (Bolton and Dunlap 2017), which may account for part of this profile and would require specific control experiments in any subsequent evaluation.

CNP0196376.4 behaved differently from the other two candidates. It showed the weakest predicted binding of the three (-9.80 kcal/mol) and two unfavourable donor-donor contacts with Ser24 and Thr29, yet it maintained the most continuous hydrogen bond network of the three, present in 91.8 % of frames. This combination of weak initial complementarity and extensive hydrogen bonding was accompanied by the most pronounced conformational rearrangement of the receptor: the highest mean backbone RMSD (2.16 nm), the broadest conformational sampling in the PCA projection **(Figure 6.a, 6.d)**, and the longest adaptation phase, with solvent accessible surface area decreasing from 237.8 nm² over the first 50 ns to 199.4 nm² after 200 ns and the radius of gyration converging to approximately 2.30 nm over the same period.

Taken together, these observations are consistent with a slow accommodation process in which the ligand and the binding site adjust progressively to one another, rather than with the behavior of **CNP0188167.1**, which combined conformational rigidity with a discontinuous hydrogen bond network. Whether this reflects an induced-fit binding mode or simply poor steric complementarity cannot be determined from these simulations alone. Distinguishing between the two would require binding kinetics measurements and longer or replicate trajectories (Bernetti et al. 2017).

CNP0196376.4 has the most favorable predicted toxicological profile of the three compounds, with only four unfavorable endpoints (skin sensitization 1.000, AMES mutagenicity 0.936, drug-induced liver injury 0.925, and genotoxicity 0.834), and favorable predictions for carcinogenicity (0.020), hERG channel blockade (0.005), acute oral toxicity (0.036), and neurotoxicity (0.001). Its main limitation is pharmacokinetic: its triglycosylation, its polar surface area of 328.35 Å², and its twelve hydrogen-bond donors result in a high probability of poor intestinal absorption (0.859) and efflux via P-glycoprotein (0.987), which rules out oral administration and suggests a parenteral route for this compound. Its structural analysis also identifies a pan-assay interference substructure (PAINS), a chelation alert, and eight substructures associated with skin sensitization. The catechol group in ring B, derived from quercetin, can oxidize to ortho-quinone (Bolton and Dunlap 2017), which would account for both these structural alerts and the sensitization signal.

These results come in a context where the relevance of IL-18Rα as a target in acute GvHD has already been established experimentally. Neutralization of IL-18Rα by a monoclonal antibody confers protection in a mouse model of acute GvHD (Li et al. 2015), and overexpression of the receptor on T cells has been documented in both humans and mice (Fujimori et al. 2002; Itoi et al. 2004), particularly under the influence of the IL-33/ST2 axis (Reichenbach et al. 2015). The present study explores whether this interface, which has thus far been

584 targeted by biologics, could be engaged by small molecules. However, it is important to note a difference in
585 scale: Site I covers an interaction surface area of approximately 1,850 Å², which far exceeds the surface area
586 occupied by a low-molecular-weight ligand. Complete inhibition of such a protein-protein interface by a small
587 molecule remains a challenging goal, and the compounds identified here should be considered as initial chemical
588 probes rather than as complete antagonists.

589 **Conclusion**

590 Through unbiased screening of a library of 2,611 phytochemicals derived from 50 Moroccan medicinal species,
591 this study identified three compounds that bind to Site I of IL-18Rα, the interface through which IL-18 binds to
592 its receptor. The best conformations of the ten highest-ranked compounds converged on this cavity without it
593 having been imposed, which supports its relevance as a binding site.

594 Among the three candidates selected for molecular dynamics simulations, **CNP0188167.1** showed the strongest
595 predicted affinity and behaved as a conformationally rigid molecule, but its hydrogen bond network with the
596 receptor was interrupted for more than 60 ns before being reestablished. **CNP0075206.3** displayed an
597 intermediate internal flexibility, with hydrogen bond occupancy decreasing steadily throughout the trajectory.
598 **CNP0196376.4**, despite the weakest predicted affinity of the three, maintained the most continuous interaction
599 with the receptor, with hydrogen bonds present in 91.8 % of frames and the most populated free energy basin.
600 The predicted ADMET profiles were unfavourable for all three compounds, along a gradient ranging from
601 **CNP0196376.4,** the least problematic, to **CNP0075206.3**, which carried the most alerts. These molecules should
602 therefore be regarded as structural starting points requiring optimisation rather than as therapeutic candidates.

603 Future work should focus on three areas. First, modifying the motifs responsible for toxicological alerts: the α,β-
604 unsaturated lactone in **CNP0188167.1**, the redox-active anthraquinone core in **CNP0075206.3**, and the catechol
605 group in **CNP0196376.4**. The analogs thus obtained could be run through the same pipeline to verify whether the
606 removal of these motifs preserves affinity for Site I. Next, computational validation should be strengthened
607 through cross-validation of docking, replicated simulations, and calculation of binding free energies. Finally,
608 experimental validation should be conducted through direct affinity measurement, cellular assays assessing IL-
609 18-induced IFN-γ production, and evaluation in the murine model of acute GvHD previously used for the anti-
610 IL-18Rα antibody.

611

Limitations

This study is entirely computational. The reported affinities, dynamic profiles and ADMET properties are predictions, and no experimental validation was performed in vitro or in vivo.

Since no IL-18Rα structure co-crystallised with a low molecular weight ligand was available, the docking protocol could not be validated by redocking or by a positive control, and the scores were therefore interpreted relative to the distribution of the screened library. No cross validation with a second docking program was performed, the default exhaustiveness value was retained for a search space covering the entire ectodomain, and the receptor was treated as rigid. In addition, residues Ile49 to Thr51 and the C-terminal segment Lys316 to Arg329 were not resolved in the crystallographic structure and were not modelled. The resulting chain discontinuity, which was preserved throughout the simulations, probably contributes to the overall backbone deviations observed.

A single trajectory was produced per complex, as the available computing resources did not allow independent replicates. Production was scheduled for 400 ns, but repeated interruptions of the HPC cluster stopped the three runs at different times, and all trajectories were truncated to 340 ns, the duration reached by the shortest run. This duration falls within the range commonly used for receptor-ligand complexes, but it does not rule out conformational rearrangements occurring on longer time scales, and the absence of replicates prevents any assessment of the reproducibility of the observed trajectories.

The extracellular domain of IL-18Rα was simulated in isolation, without IL-18 and without membrane anchoring. Under these conditions the three immunoglobulin-like domains reorient and the receptor compacts, as reflected by the decrease in the radius of gyration and in the solvent accessible surface area. The conformations sampled at the end of the simulations therefore do not necessarily correspond to the state of the receptor within the physiological ternary complex. Ligand RMSD was calculated after superimposing each ligand onto its own reference structure and reflects internal conformational flexibility rather than retention at the binding site, which was assessed here through hydrogen bond occupancy. For CNP0075206.3 this occupancy decreases continuously and becomes negligible by the end of the trajectory. As this compound is a polyaromatic dimer capable of interacting through π stacking, the loss of hydrogen bonding does not establish that it left the site, but its retention could not be demonstrated.

No binding free energy calculation could be completed. The energy ranking of the three compounds therefore relies exclusively on docking scores, whose accuracy is limited, and on interaction descriptors extracted from the trajectories. Finally, ADMET predictions are based on models, some of which are trained on limited datasets, and the compounds examined here, which are structurally complex natural products, may fall outside their applicability domain. The alerts identified for pan-assay interference and colloidal aggregation would require specific experimental controls. The screened library is also restricted to the phytochemicals documented in COCONUT for the selected species.

Acknowledgements

The authors gratefully acknowledge the High-Performance Computing platform Marwan, operated by the National Center for Scientific and Technical Research (CNRST), Rabat, Morocco, for providing the computational resources used to perform the molecular dynamics simulations in this study.

Declaration of generative AI use

During the preparation of this work, the authors used Deepseek AI and ChatGPT to assist with the development of Python analysis scripts, the refinement of GROMACS and AutoDock Vina command syntax, and the improvement of the English language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Conflict of interest

The authors reported no potential conflict of interest.

Ethics declaration

Not applicable

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Authors' contribution

LA, HC, and SR conceived the idea and designed the objectives. LA and SR performed the computational study, LA wrote the paper, HC participated in the data analysis, and HC, FG, and MB performed the corrections of the manuscript.

References

Bernetti M, Cavalli A, Mollica L (2017) Protein-ligand (un)binding kinetics as a new paradigm for drug discovery at the crossroad between experiments and modelling. Medchemcomm 8:534–550. https://doi.org/10.1039/c6md00581k

692 Bickerton GR, Paolini GV, Besnard J, et al (2012) Quantifying the chemical beauty of drugs. Nature
693 Chem 4:90–98. https://doi.org/10.1038/nchem.1243

694 Bolton JL, Dunlap T (2017) Formation and Biological Targets of Quinones: Cytotoxic versus
695 Cytoprotective Effects. Chem Res Toxicol 30:13–37.
696 https://doi.org/10.1021/acs.chemrestox.6b00256

697 Bouyahya A, Guaouguaou F-E, El Omari N, et al (2022) Anti-inflammatory and analgesic properties of
698 Moroccan medicinal plants: Phytochemistry, in vitro and in vivo investigations, mechanism
699 insights, clinical evidences and perspectives. Journal of Pharmaceutical Analysis 12:35–57.
700 https://doi.org/10.1016/j.jpha.2021.07.004

701 Chakravorty A, Hussain A, Cervantes LF, et al (2024) Exploring the Limits of the Generalized CHARMM
702 and AMBER Force Fields through Predictions of Hydration Free Energy of Small Molecules. J
703 Chem Inf Model 64:4089–4101. https://doi.org/10.1021/acs.jcim.4c00126

704 Chandrasekhar V, Rajan K, Kanakam SRS, et al (2025) COCONUT 2.0: a comprehensive overhaul and
705 curation of the collection of open natural products database. Nucleic Acids Research
706 53:D634–D643. https://doi.org/10.1093/nar/gkae1063

707 D. Segall M (2012) Multi-Parameter Optimization: Identifying High Quality Compounds with a Balance
708 of Properties. CPD 18:1292–1310. https://doi.org/10.2174/138161212799436430

709 Diarra A, Agossa K, Youl ENH (2024) The Potential of Cochlospermum tinctorium, Flueggea virosa,
710 and Waltheria indica Traditional Plants From Burkina Faso in Treating Periodontitis: A
711 Systematic Review. Cureus. https://doi.org/10.7759/cureus.52471

712 Dinarello CA (2001) Novel targets for interleukin 18 binding protein. Annals of the Rheumatic
713 Diseases 60:iii18–iii24. https://doi.org/10.1136/ard.60.90003.iii18

714 Dinarello CA (1999) IL-18: A TH1 -inducing, proinflammatory cytokine and new member of the IL-1
715 family. Journal of Allergy and Clinical Immunology 103:11–24.
716 https://doi.org/10.1016/S0091-6749(99)70518-X

717 Fujimori Y, Yoshimoto T, Matsui K, et al (2002) Increased Expression of Interleukin-18 Receptor on T
718 Lymphocytes in Patients with Acute Graft-versus-Host Disease After Allogeneic Bone Marrow
719 Transplantation. Journal of Interferon & Cytokine Research 22:751–754.
720 https://doi.org/10.1089/107999002320271332

721 Goodsell DS, Sanner MF, Olson AJ, Forli S (2021) The AutoDock suite at 30. Protein Science 30:31–43.
722 https://doi.org/10.1002/pro.3934

723 Itoi H, Fujimori Y, Tsutsui H, et al (2004) Differential Upregulation of Interleukin-18 Receptor α Chain
724 Between $CD4^+$ and $CD8^+$ T Cells During Acute Graft-Versus-Host Disease in Mice. Journal of
725 Interferon & Cytokine Research 24:291–296. https://doi.org/10.1089/107999004323065075

726 Kanate AS, Majhail NS, Savani BN, et al (2020) Indications for Hematopoietic Cell Transplantation and
727 Immune Effector Cell Therapy: Guidelines from the American Society for Transplantation and
728 Cellular Therapy. Biology of Blood and Marrow Transplantation 26:1247–1256.
729 https://doi.org/10.1016/j.bbmt.2020.03.002

730 Kaplan W (2001) Swiss-PDB Viewer (Deep View). Briefings in Bioinformatics 2:195–197.
731 https://doi.org/10.1093/bib/2.2.195

732 Khalfaoui A, Noumi E, Belaabed S, et al (2021) LC-ESI/MS-Phytochemical Profiling with Antioxidant,
733 Antibacterial, Antifungal, Antiviral and In Silico Pharmacological Properties of Algerian
734 Asphodelus tenuifolius (Cav.) Organic Extracts. Antioxidants 10:628.
735 https://doi.org/10.3390/antiox10040628

736 Klochkov S, Neganova M (2021) Unique indolizidine alkaloid securinine is a promising scaffold for the
737 development of neuroprotective and antitumor drugs. RSC Adv 11:19185–19195.
738 https://doi.org/10.1039/D1RA02558A

739 Li X, Zhang C, Chen W, et al (2015) Protective effect of neutralizing anti-IL-18α monoclonal antibody
740 on a mouse model of acute graft-versus-host disease. Oncology Reports 34:2031–2039.
741 https://doi.org/10.3892/or.2015.4176

742 Malard F, Holler E, Sandmaier BM, et al (2023) Acute graft-versus-host disease. Nat Rev Dis Primers
743 9:27. https://doi.org/10.1038/s41572-023-00438-1

744 Morris GM, Huey R, Lindstrom W, et al (2009) AutoDock4 and AutoDockTools4: Automated docking
745 with selective receptor flexibility. J Comput Chem 30:2785–2791.
746 https://doi.org/10.1002/jcc.21256

747 Nakamura K, Okamura H, Wada M, et al (1989) Endotoxin-induced serum factor that stimulates
748 gamma interferon production. Infect Immun 57:590–595.
749 https://doi.org/10.1128/iai.57.2.590-595.1989

750 Pusic I, Pavletic SZ (2019) Challenges in Conducting Studies in Chronic Graft-versus-Host Disease. CHI
751 1:36. https://doi.org/10.2991/chi.d.190314.001

752 Qiang Y, Wu P, Liu X, et al (2025) Structural basis of the multiple ligand binding mechanisms of the
753 P2X1 receptor. Acta Pharmacol Sin 46:2564–2573. https://doi.org/10.1038/s41401-025-
754 01512-y

755 Reddy P, Davies S, Majhail N, et al (2020) In response to “American Society of Hematology 2020
756 guidelines for treating newly diagnosed acute myeloid leukemia in older adults.” Blood
757 Advances 4:5431–5432. https://doi.org/10.1182/bloodadvances.2020003413

Reichenbach DK, Schwarze V, Matta BM, et al (2015) The IL-33/ST2 axis augments effector T-cell responses during acute GVHD. Blood 125:3183–3192. https://doi.org/10.1182/blood-2014-10-606830

Rex DAB, Agarwal N, Prasad TSK, et al (2020) A comprehensive pathway map of IL-18-mediated signalling. J Cell Commun Signal 14:257–266. https://doi.org/10.1007/s12079-019-00544-4

Sorokina M, Merseburger P, Rajan K, et al (2021) COCONUT online: Collection of Open Natural Products database. J Cheminform 13:2. https://doi.org/10.1186/s13321-020-00478-9

Talley S, Rademacher DJ, Campbell EM (2023) Inflammasome activation occurs in CD4+ and CD8+ T cells during graft-versus-host disease. Cell Death Dis 14:632. https://doi.org/10.1038/s41419-023-06138-8

Teshima T, Reddy P, Zeiser R (2016) Acute Graft-versus-Host Disease: Novel Biological Insights. Biology of Blood and Marrow Transplantation 22:11–16. https://doi.org/10.1016/j.bbmt.2015.10.001

Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455–461. https://doi.org/10.1002/jcc.21334

Tsutsumi N, Kimura T, Arita K, et al (2014) The structural basis for receptor recognition of human interleukin-18. Nat Commun 5:5340. https://doi.org/10.1038/ncomms6340

Zhang L, Chu J, Yu J, Wei W (2016) Cellular and molecular mechanisms in graft-versus-host disease. Journal of Leukocyte Biology 99:279–287. https://doi.org/10.1189/jlb.4RU0615-254RR

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DataSupp.zip