

## Computational Modeling on Aquaporin-3 as Skin Cancer Target: A Virtual Screening and Molecular Dynamic Simulation Study

Dharmendra Kumar Yadav<sup>a, \$\*</sup>, Surendra Kumar<sup>a, \$</sup>, Eun-Ha Choi<sup>b</sup>, Sandeep Chaudhary<sup>c</sup>,  
Mi-Hyun Kim<sup>a, \*</sup>

<sup>a</sup>Gachon Institute of Pharmaceutical Science & Department of Pharmacy, College of Pharmacy,  
Gachon University, Incheon, 21936, South Korea.

<sup>b</sup>Plasma Bioscience Research Center/PDP Research Center, Kwangwoon University, Nowon-Gu,  
Seoul, 139-791, South Korea

<sup>c</sup>Laboratory of Organic & Medicinal Chemistry, Department of Chemistry, Malaviya National  
Institute of Technology, Jawaharlal Nehru Marg, Jaipur-302017 (India)

*Email: [dharmendra30oct@gmail.com](mailto:dharmendra30oct@gmail.com), [kmh0515@gachon.ac.kr](mailto:kmh0515@gachon.ac.kr)*

\*corresponding author

**Dr.Dharmendra Kumar Yadav, Ph.D.**

Assistant Professor

Office: +82-32-820-4947

Email: [dharmendra30oct@gmail.com](mailto:dharmendra30oct@gmail.com)

<sup>\$</sup>These authors contributed equally to this work.

**Abstract:** Aquaporin-3 (AQP3) is one of the aquaglyceroporins, which is expressed in the basolateral layer of the skin membrane. Studies have reported that human skin squamous cell carcinoma overexpresses AQP3 and inhibition of its function may alleviate skin tumorigenesis. In the present study, we have applied a virtual screening method that encompasses filters for physicochemical properties and molecular docking to select potential hit compounds that bind to the Aquaporin-3 protein. Based on molecular docking results, the top 20 hit compounds were analyzed for stability in the binding pocket using unconstrained molecular dynamics simulations and further evaluated for binding free energy. Furthermore, examined the ligand-unbinding pathway of the inhibitor from its bound form to explore possible routes for inhibitor approach to the ligand-binding site. With a good docking score, stability in the binding pocket, and free energy of binding, these hit compounds can be developed as Aquaporin-3 inhibitors in the near future.

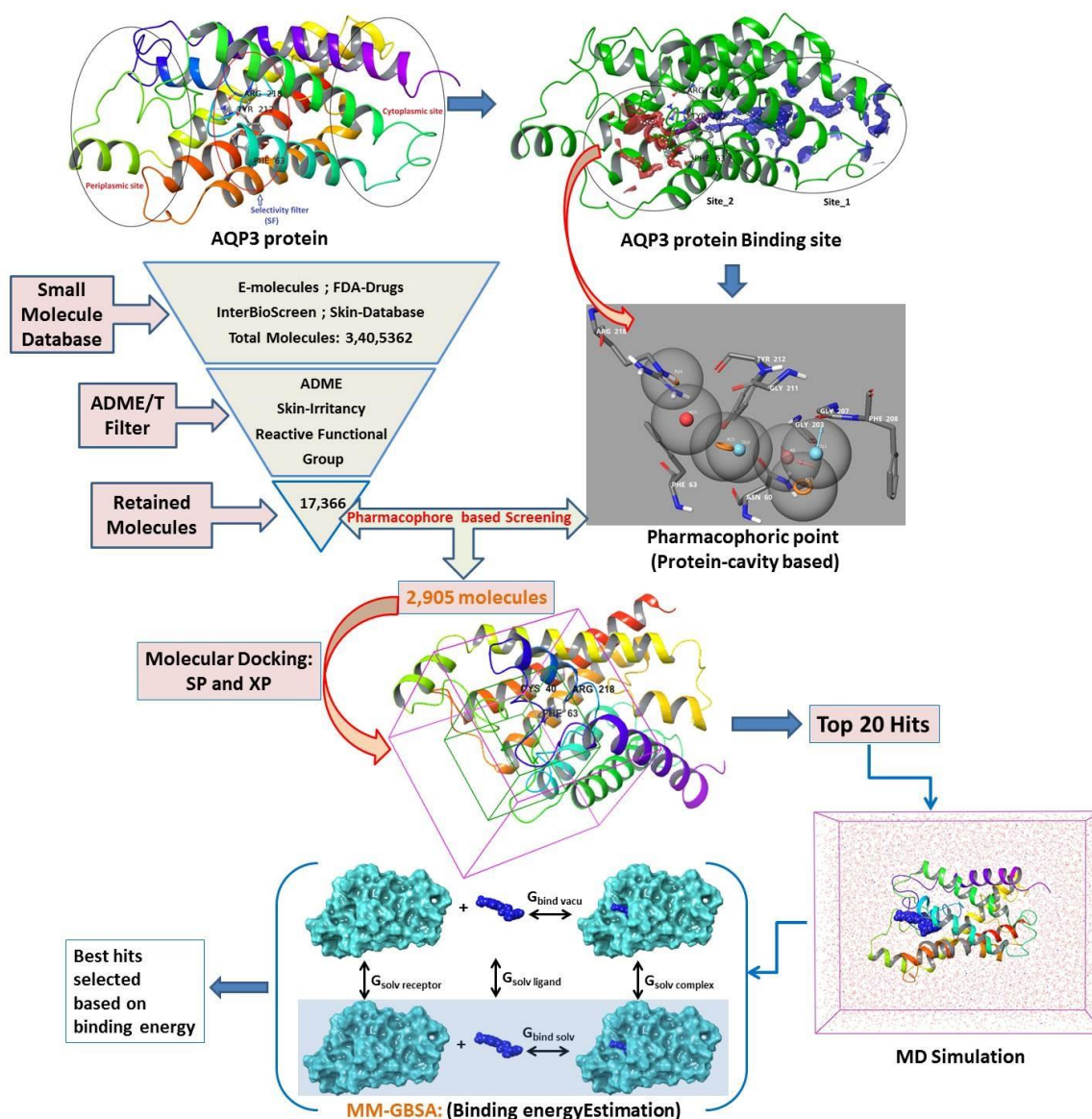
**Keywords:** AQP3 protein, molecular docking, molecular dynamics, MM-GBSA analysis, pharmacophore-based filter

## 1. Introduction

Skin cancer represents a major and growing public health problem, accounting for ~40 % of all newly diagnosed cancer cases [1]. Skin cancer includes basal cell carcinomas (BCC), squamous cell carcinomas (SCC), and malignant melanomas [2]. BCC and SCC each constitute non-melanoma skin cancer, originate from epidermal keratinocytes, and have been associated with chronic sun exposure, whereas melanoma skin cancer originates from melanocytes and has been associated with intermittent sun exposure [3-5]. The stratum corneum (SC) is the epidermal layer of the skin, consisting of terminally differentiated keratinocytes and containing a lipid extracellular matrix. The appearance and physical properties of the skin depend on a number of factors, including the lipid/protein composition, barrier properties, and the concentration of water-retaining osmolytes or “natural moisturizing factors” such as free amino acids, ions, and other small solutes [6]. The aquaporins (AQPs) are ubiquitous family of small, hydrophobic, and highly conserved membrane proteins involved in the transport of water and small solutes such as glycerol, nitrates, and urea [7,8]. To date, 13 human AQP isoforms (AQP0-12) have been identified and differentially expressed in many types of cells and tissues in the body [9]. The AQPs are broadly classified into two groups: orthodox aquaporins (selective for water) and aquaglyceroporins (permeable to small solutes including glycerol) [10]. Genotype and phenotype studies of AQPs identified their role in refractory edema, brain swelling, neuroinflammation, glaucoma, epilepsy, cancer, pain, and obesity [11]. The involvement of AQPs in cell migration has implicated them in tumor angiogenesis, local invasion, and metastasis [12]. Among all the identified AQP isoforms, AQP1 (expressed in endothelial cells) and AQP3 (expressed in plasma membranes in the basal layer of keratinocytes in human skin) are of particular interest for the study of cancer model [13,14]. The functions of AQPs in the skin have not been thoroughly investigated; however, AQP3 has gained attention over the last few decades, as it is abundantly expressed in the skin [15,16]. A human keratocarcinoma cell line was found to express AQP3. AQP3 (aquaglyceroporins) transport water, glycerol, urea, and hydrogen peroxide, playing an important role in SC hydration, skin elasticity, barrier recovery, wound healing, cell proliferation, tumorigenesis, and cell migration [14]. Previous studies have reported that Aquaporin-3 (AQP3) deficient mice may have dry skin and delayed barrier recovery due to the absence of AQP3 facilitated glycerol transport [17,18]. Chikuma et al. have studied a multistage skin tumor model in mice and report that AQP3 is overexpressed in skin cancer, while AQP-null mice show complete resistance to the formation of skin cancer [15]. The

study also found that disrupting the AQP3 controlled gene leads to reduced cell proliferation and skin tumorigenesis. Thus, resistance to skin tumorigenesis as a function of AQP3 deficiency provides a rationale for evaluating AQP3 inhibitors in the prevention and therapy of melanoma, as well as other tumors associated with AQP3 overexpression.

There are compelling opportunities, yet so far little progress, in the search for AQP-based therapeutics. There are a few reported AQP inhibitors suitable for clinical trials; however, none of them has shown any specificity for AQP3 [19-21]. AQP1, a close congener of AQP3 in terms of protein sequence, reported to be inhibited by tetraethylammonium salts [22], acetazolamide [23], bumetanide [24], and DMSO [25]. Further, in search of AQP inhibitors, Preston et al. [26] and Niemietz et al. [19] found mercurial ( $\text{HgCl}_2$ ), silver (Ag), and gold (Au) containing inorganic compounds act as AQP inhibitors; however, they are non-selective and extremely toxic. Later, Martins et al. evaluated metal-based drugs already known to possess different therapeutic properties (such as anticancer, antirheumatic, and antibacterial agents) for AQP1 and AQP3 inhibition and found promising results [27]. Moreover, several authors have synthesized and reported gold-based compounds for AQP3 inhibitors and elucidated the mechanism of inhibition [28-30]. Here for the first time, we explore small molecule inhibitors (hits) of AQP3 using a series of virtual screening tools, followed by molecular dynamics and binding free energy calculations (**Fig. 1**). Virtual screening tools comprising molecular docking and pharmacophore-based methods may reduce false positives in potential hits.



**Fig. 1** Designed workflow for the virtual screening of AQP3 inhibitors.

## 2. Materials and Methods

### 2.1. Chemical datasets and virtual screening methods

For a molecule to be an effective drug, it must reach its target site in the body and stay there in a bioactive form long enough for the expected biological events to occur [31]. The assessment of molecules for skin permeability is always a major challenge, as molecules must cross the heterogeneous nature of the SC of the skin membrane to reach the target site. The majority of drugs available are administered orally and their permeability across various membranes depends on their size, shape, and physicochemical properties [32,33]. Permeation of drug molecules

across the various membranes is governed by the Biopharmaceutical Classification System that has classified the permeability of molecules based on their solubility (class I: high solubility, high permeability; class II: low solubility, high permeability; class III: high solubility, low permeability; and class IV: low solubility, low permeability) [34,35]. Thus, in order to evaluate the permeability of drug molecules across the skin (selecting those having no toxicological [skin irritancy] properties) a series of filters comprising physicochemical (QikProp) [36,37], toxicity (skinirritancy) [38,39], and reactive functional group [40,41] were applied to narrow the list of available drug molecules that may efficiently bind to the target receptor. The following filter conditions were set: Molecular Weight: 20 to 300; LogS (Predicted Solubility): -9.0 to 1.0; LogKp (Predicted Skin Permeability): -8.0 to 1.0; Jm (Predicted Transdermal Transport Rate): less than 10; Reactive Functional Groups: 0 to 2; Skinirritancy: low or none [42-44]. A total of 3,379,981 small molecules obtained from the e-molecules database [45], IBS database (synthetic and natural compounds) [46], the dataset from Braga et al. [47], and US-FDA approved drugs (**Table 1**) [48] were passed through the series of filters. Furthermore, pharmacophore-based (protein cavity)[49] screening and molecular docking-based (Standard Precision and Extra Precision) [50] screening were performed to find good scoring compounds. From these good scoring compounds, the top 20 hits were selected for molecular dynamics [51] and binding free energy prediction (MM-GBSA) [52].

**Table 1.** Small molecules considered under present study.

Database/Literature	Total number of molecules
E-molecules	3,32,8465
InterBioScreen (IBS)	67,609
Small molecules from <a href="http://chembench.mml.unc.edu">http://chembench.mml.unc.edu</a>	87
US-FDA approved drugs	9101

## 2.2. Ligand preparation

The structures were prepared using the LigPrep module in Schrodinger suite [53]. The LigPrep generates energy minimized 3D structures using an OPLS3 force field. The correct Lewis structure, tautomer, and ionization states (pH 7.0  $\pm$ 2.0) for each structure were generated, optimized, and energy minimized with default settings.

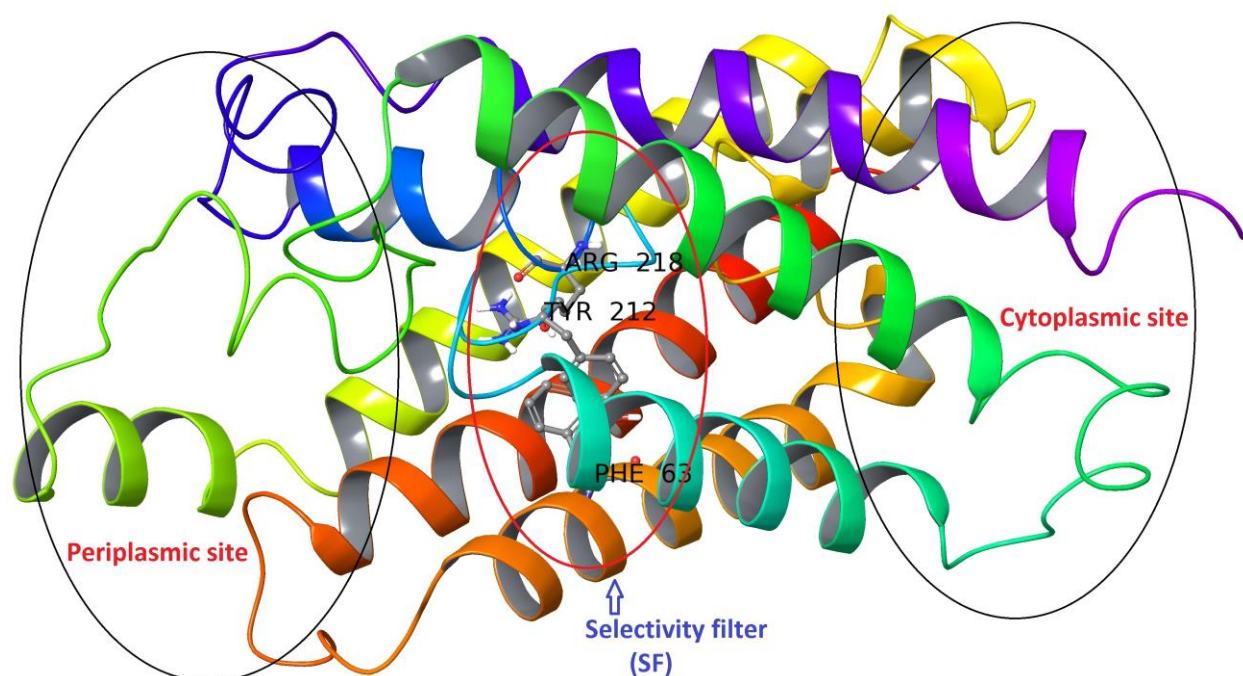
## 2.3. Protein preparations, active site prediction, and receptor grid generation

The crystal structure of AQP3 has not been solved yet and thus obtained from the reported work [27]. Martins et al. [27] have built an AQP3 structure via homology modeling. The modeled structure was imported into the maestro workspace and the multistep Protein Preparation Wizard was used to prepare the protein, including addition of H-atoms, bond order correction, and H-bond network optimization, followed by energy minimization using Impref module of Impact with an OPLS3 force field [54].

The AQP3 structure comprises six membrane-spanning helices and two half-helices with their positive and N-terminal ends located at the center of the protein and C-terminal ends pointing toward either side of the membrane. The helices surround the 20 Å long and 3-4 Å wide amphipathic AQP channel. A structural study on aquaporins has identified two asparagine-proline-alanine (NPA) sequence motifs located at the ends of the two quasi 2-fold related half-spanning helices. AQP3 has a tunnel-like structure with periplasmic and cytoplasmic pore sites on opposite ends, separated by a selectivity filter (SF) domain comprising conserved amino acid residues (Phe63, Arg218 and Tyr212), the domain being a distinguishing features that identify subfamilies of AQPs (**Fig. 2**) [27,55]. Martins et al. reported that a periplasmic pocket of AQP3 has an extended hydrophobic region in the proximity of the SF domain, which is absent in AQP1 (another member of aquaglyceroporins), providing the latter with higher hydrophilic character [56].

The prepared structure lacks any bound ligand, thus the binding sites were defined using SiteMap tools of Schrodinger [57] with default settings. The AQP3 structure with identified sites is shown in **Fig. S1**. The SiteMap tool has identified two probable binding sites; based on the previous binding site evaluation criteria, Site\_2 was selected as the binding site. A receptor grid was generated on Site\_2 with the default setting for molecular docking.



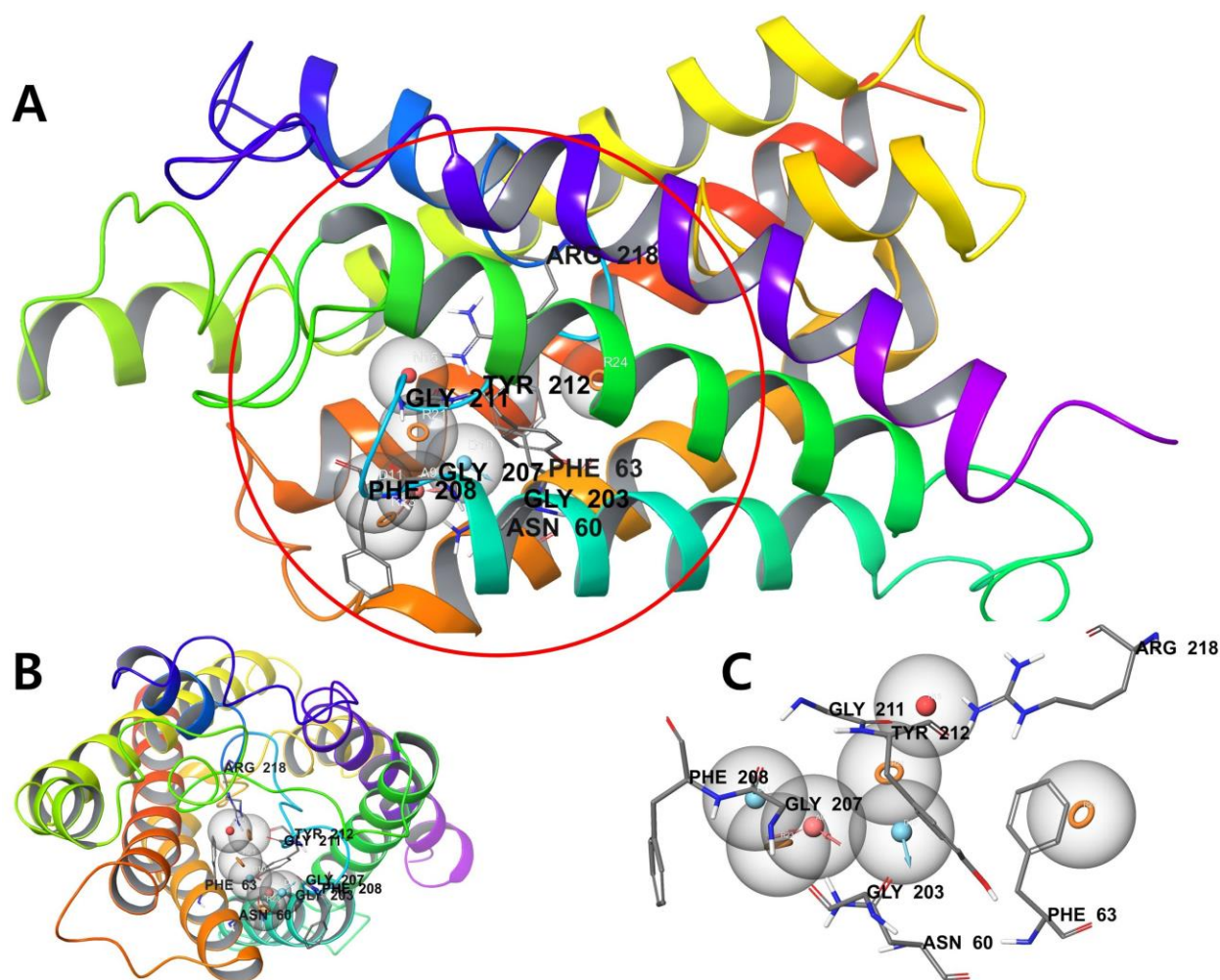


**Fig. 2** AQP3 protein with periplasmic and cytoplasmic sites. The SF region comprising key amino acid residues (Phe63, Tyr212, and Arg218) are displayed as ball and stick representations.

#### 2.4. Pharmacophore-based (protein cavity) filtering

We have employed the pharmacophore filtering methods to further screen the small molecules. Numerous approaches have been available and described in detail elsewhere [58-60]. However, due to lack of a bound ligand in the present study, we used a protein cavity-based pharmacophore point enumeration method. As Site\_2 was previously defined from the SiteMap tool, it was further utilized to enumerate the pharmacophoric point for the pharmacophore-based filtering method with default settings. The pharmacophore filtering methodology is based on docking of fragments to a protein receptor (e-pharmacophore model), followed by a selection of fragment features that maximize the binding interaction. Finally, common features identified by pharmacophoric points were chosen to satisfy criteria for their positions and directions (**Fig. 3**).





**Fig. 3** (A) Whole protein showing cavity-based pharmacophoric point; (B) top view with pharmacophoric point; (C) pharmacophoric point with label residues in the binding pocket.

### 2.5. Molecular docking

The filtered small molecules from the pharmacophore-based filtering method were analyzed by molecular docking. The molecular docking was performed with glide v7.8 in the SP (Standard Precision) and XP (Xtra Precision) protocol of the Schrodinger Suite with default settings [50]. The filtered molecules docked into the binding pocket (Site\_2) and the final hits were selected based on the threshold criteria (docking score > -6.0) in each step.

### 2.6. Molecular dynamics

All the MD simulations were carried out using the Desmond 5.3 MD package [51]. The protein-ligand complex system was inserted into the pre-equilibrated POPC (300°K) lipid bilayer membrane using the Set up Membrane option of system builder module of Desmond. The upper and lower lipid bilayer region of system was filled with water model (TIP3P) [61] as the solvent in

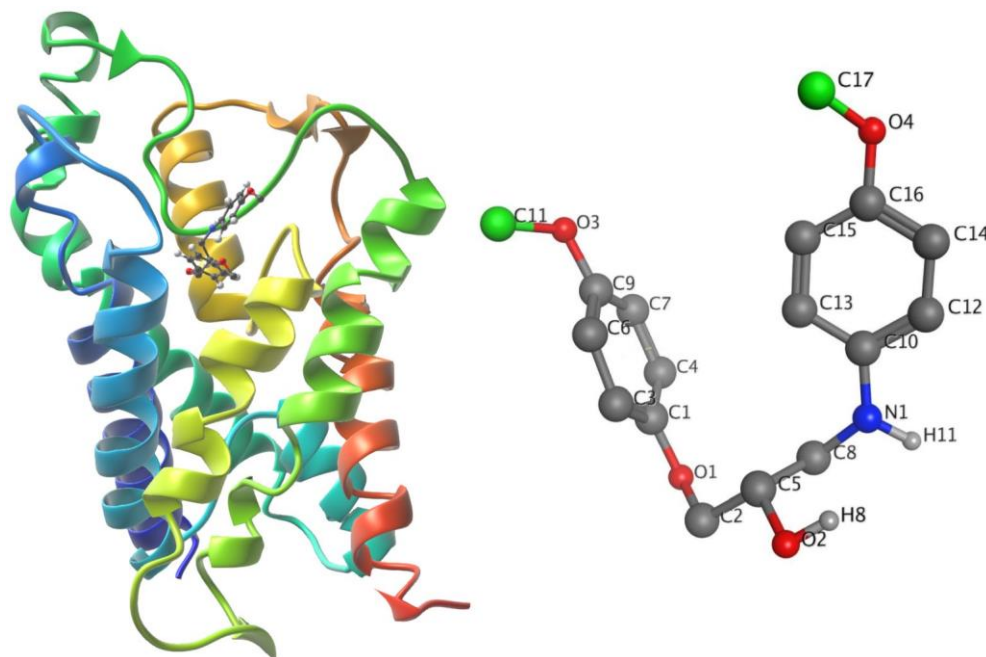
orthorhombic box with OPLS\_2005 force-field. The shape and size was set at 10 Å buffered distance. The desired neutral system was built with the addition of 0.15 M NaCl in the system. A built system is shown in **Fig. S2**. The system was relaxed by implementing Steepest Descent and the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithms in a hybrid manner [62]. The simulation was performed under NPT ensemble using the Nose-Hoover thermostat [63] and Martyna-Tobias-Klein barostat methods [64] applying a constant temperature of 300°K and 1.01325 atm of pressure, respectively. The short-range coulombic interactions were analyzed using a cut-off value of 9.0 Å and the short-range method. The smooth particle mesh ewald method [65] was used for handling long-range coulombic interactions and RESPA-based constraints allowing 2 fs time steps. The MD simulation was carried out for 100 ns and the trajectory sampling was done at an interval of 10 ps.

### 2.7. Binding free energy analysis

The interaction energies between the protein and the selected top poses were computed using the MM-GBSA (generalized-born/surface area) method implemented in Schrodinger [66,67]. The average binding free energy ( $\Delta G_{\text{bind}}$ ) based on MM-GBSA was calculated using the `thermal_mmgsa.py` script. During the MM-GBSA calculation, the last 10 ns MD simulation trajectory (every snap shot) was used as input to compute the average binding free energy.

### 2.8. Ligand-unbinding pathway

We have further explored the ligand-unbinding pathway using the ART-RRT method [68]. We adopted the same methodology that was described in detail. The ligand-unbinding pathways were searched on the last frame from dynamic simulations of the protein-ligand complex. In this study, we used the compound 5633879 (1-(4-methoxyphenoxy)-3-((4-methoxyphenyl)amino)propan-2-ol) to explore its unbinding path from the binding pocket. The gromos53a6 force field[69] assigned to the protein and ligand parameters were obtained from the ATB (Automated Topology Builder) server [70]. The ligand atom shown in green was assigned to active ARAP atoms and a box with a dimension of 100 Å X 100 Å X 120 Å was assigned during the pathway search (**Fig. 4**). We performed 10 runs with default settings for the ligand-unbinding pathway search.



**Fig. 4** Left: the system in ribbons with bound ligand in ball and stick form, Right: The ligand with active atoms in green.

### 3. Results and Discussion

#### 3.1. Chemical database curation and ligand preparation

In the present study the chosen databases had varied ranges of physicochemical and toxicological properties (**Fig. S3**). Thus, in order to find the molecules that could easily permeate across the membrane with no toxicological (skin irritancy) properties, the chosen database molecules were filtered based on a pre-condition filter (mentioned in the Materials and Methods section). This two-stage filter (ADME and Toxicity) greatly reduced the number of molecules, which were prepared using the LigPrep module of Schrodinger for further processing.

#### 3.2. Pharmacophore-based (protein cavity) filtering

The pharmacophore-based (protein cavity) hypothesis, shown in **Fig. 3**, depicts seven chemical features comprising three aromatic rings (R), two hydrogen bond donors (D), one hydrogen bond acceptor (A), and one negative ion (N). Among the features generated, hydrogen bond donor and acceptor features are vector properties possessing a vectorial nature, which indicates the direction of electron sharing [71]. The features in the hypothesis were superimposed on the Site<sub>2</sub> binding site, which showed that the chemical features were complementary to key amino acid residues, including H-bonding interactions with Asn60, Gly203, Gly207, Gly211, and Arg218, corresponding to A, D, and N features. Likewise, Phe63, Phe208, and Tyr212 corresponded to R

features. Such complementarity between amino acid residues and chemical features reveal the importance of binding to small molecules. The reduced dataset (after application of the initial physicochemical and toxicity filter) was further screened against the generated pharmacophore-based (protein cavity) hypothesis as a query and molecules were ranked according to the PhaseScreenScore. The PhaseScreenScore measures the quality and quantity of features matching the hypothesis, where molecules with a score above 0.8 are selected for the next stage.

### 3.3. Molecular docking

The identified binding site (Site\_02) encompasses the hydrophilic/hydrophobic areas comprising amino acid residues within 3 Å of the binding pocket, namely, Val43, Val46, Phe56, Ile59, Asn60, Phe63, Phe147, Ala148, Thr149, Tyr150, Gly207, Gly211, Tyr212, and Arg218. Martins et al. [27] reported that among all the binding site residues, the triad amino acid residues (Phe63, Arg218, Tyr212) near the periplasmic gate are involved in the binding of small molecules and may modulate the function of the AQP3 protein. Hence, Site\_02 was chosen in this study as the binding site and used for docking of small molecules obtained after application of the pharmacophore-based (protein cavity) filter. The molecular docking comprised two filters. In the first we have employed an SP mode, where dock poses with a docking score above -6.0 were selected for an Extra Precision (XP) mode docking with default settings. After the second filter (XP docking), the dock poses were ranked according to the docking score and the top 20 poses were selected as potential hits that could modulate the function of AQP3 protein (**Table S1**). The ligand-binding amino acid residues from the top 20 hits are summarized in **Table S2**.

From all the docked poses (XP mode), the top 20 poses were selected and analyzed, where all had docking scores ranging from -7.550 kcal/mol to -6.747 kcal/mol. The comparison of binding poses reveals that despite the diverse scaffold of these poses, all interact with common binding site residues (**Fig. S4**). All the docked poses displayed multiple direct hydrogen bond interactions with key amino acid residues, such as Asn60 and Arg218, whereas the backbone atoms of residues Gly145, Ala148, Gly207, Gly211, and Phe208 were involved in hydrogen bond interactions. Similarly, most of the compounds displayed  $\pi$ - $\pi$  stacking interactions with the aromatic rings of the Tyr150 and Phe208 amino acid residues. Studies reported that, the amino acid residues from the periplasmic gate and selectivity filter (SF) region, (Phe63, Arg218, Tyr212) of AQP3 has important role in ligand binding [27,72]. Thus, the analysis of the top scoring docked poses revealed

that, the aromatic fragments of the molecules are juxtaposed into a group of aromatic amino acid residues (including Phe63, Tyr150, Phe208, and Tyr212) located at the opening of periplasmic gate and near the SF region, revealing their importance in modulating the function of the target protein.

### 3.4. Molecular dynamics

In order to access the stability, dynamics, reliability, and underlying molecular interactions of the docked poses at the atomic level, the top 20 docked poses were submitted for molecular dynamics simulation. The quality of the molecular dynamics simulation was measured by analyzing the total energy, potential energy, temperature, pressure, and volume of the protein-ligand complexes (**Fig. S5-S24**). The potential energy includes the sum of the bond, angle, torsion, and non-bonded terms represent the system stability. Thus, the plots of potential energy clearly indicated that the system was well equilibrated and remained stable throughout the simulations. The other measured structural parameters (i.e. RMSDs, Root Mean Square Fluctuations (RMSFs), and interaction histograms) are shown in **Fig. S25-S48**. The overall structural fluctuation was evaluated by measuring the RMSDs of C $\alpha$ -atoms against simulation time. In all simulations the protein attains equilibrium within 10ns of simulation, and then oscillates afterward with RMSDs below 3.5 Å, indicating that the system had evolved into stable states and was reasonably converged. In general, the trajectories in each pose during the simulation produced a stable protein with an average RMSD value that ranged from 1.75 Å to 3.50 Å, where the large RMSDs found may occur due to motion of the loop region of the protein. Furthermore, it was observed from the RMSDs plot that protein-ligand complexes did not dissociate and remained bound throughout the simulations (**Fig. S25 – S28**).

Further, in order to access the movement of residues during the simulation we plotted the RMSFs for C $\alpha$ -atoms of all residues (see supplementary file). The RMSFs plot displayed the helix (pink-colored bar region) and ligand contacts region (green-colored vertical bar) during the simulation. The RMSFs measure the average atomic mobility of the C $\alpha$ -atoms and it was observed from residue analysis that residues 49-55, 75-85, 94-100, 123-125, 134-156, 178-190, 211-220, 229-248, and 265-269 are part of the loop region and may be flexible during the simulations. The molecular docking analysis shows that top scored poses have interacted with loop region residues



(i.e. 55, 141, 142, 145-152, 211, 212, and 218) of the target protein. However, the ligand contact analysis of the loop region from the RMSFs plot revealed that, despite the flexible loop region, interactions from poses (STOCKIN-03432, 2801237, 16694164, 5633879, 13477729, 36716128, 42888719, 31879059, 25665268) were stable during the simulation (**Fig. S29(a) – S48(a)**).

It is noteworthy to mention here that, all the selected 20 poses are structurally diverse with multiple rotatable bonds. The ligand RMSDs, as shown in **Fig. S49-S52**, revealed that the compounds were relatively stable despite such structural flexibility. For example, CMPD17 (31966421) (**Fig. S52**) contains a seven rotatable group and its RMSD, despite initial higher fluctuations, was stable  $\sim 1.5$  Å. Likewise, CMPD09 (16694164) (**Fig. S50**) with a eight rotatable group was stable around  $\sim 2.25$  Å. Thus, it is assumed that compounds repositioned their binding mode during the initial simulation and then subsequently acquired stability.

From molecular docking analysis, it was shown that all the docked poses make multiple interactions with the binding residues; thus, we further investigated its stability over the course of the simulation. **Fig. S29(b&c) – S48(b&c)** display the histograms (interaction fraction) and 2D interactions between ligands and binding amino acid residues throughout the simulation. The interaction fraction in the histogram shows different color bars, each representing the contribution of respective interactions with amino acid residues. The green, purple, red, and blue colored bars correspond to H-bonds, Hydrophobic, Ionic, and water bridge interactions, respectively. The 2D interaction diagram shows interacting residues that have been retained over 20% of the simulation time. In addition, the total number of hydrogen bond established between protein and ligand during the simulation were also calculated and shown in **Fig. S29(d)-S48(d)**. The total number of hydrogen bond observed ranges from 1-6 and in most of the compounds, these hydrogen bonds were retained throughout the simulation, further stabilizing the compounds in the binding pocket of the protein. Furthermore the protein secondary structure elements (SSE) like  $\alpha$ -helix and  $\beta$ -strands were monitored during the simulation in order to understand the stability of secondary structures. The **Fig. S29(e)-S48(e)** shows the SSE composition for each trajectory frame and each residue during the course of the simulation. The SSE result shows that, secondary structure elements were stable during the molecular dynamic simulations. In any dynamic simulation, multiple forces act on protein-ligand complexes resulting in the establishment of different types of



molecular interactions. The dynamic interactions between the protein and the ligands during simulation are summarized and further compared with the interaction analysis from the molecular docking results (**Table 2**). Comparative analysis of poses from dynamics and docking simulation interactions revealed that during dynamics, either previously formed interactions (during docking) were retained or new interactions with amino acid residues were formed. Among all the 20 dynamic simulations, compounds C8C, 42888719, and 31879059 failed, while other compounds were able to retain the interactions observed during the docking simulation, though some new interactions were seen with all the poses. Moreover, we were also interested to observe the interactions with the key triad amino acid residues (Phe63, Tyr212, Arg218) during the simulation; ligand interaction analysis from trajectory revealed that all compounds except 25284644, 32927247, 36657947, 36716128, 42888719, IP6, and 31879059 make interactions with these key amino acid residues. Thus, the molecular dynamic simulation results support the docking simulation and suggest these molecules could be important modulators of AQP3 function.

- 1 **Table 2.** Comparison of amino acid residues interactions observed during Molecular Dynamic simulation and Molecular Docking (Bold faces  
 2 represent the common binding amino acid residues).

Compound ID	Amino Acid Residues Interaction observed during simulation	Amino Acid Residues Interaction observed during docking
<b>C8C</b>	Val43 (H-Bond with Backbone), Phe56 ( $\pi$ - $\pi$ ), Gly211 (H-Bond), Tyr212 (H-Bond with Backbone)	Asn60 (H-Bond), Tyr150 (H-Bond with Backbone), Phe208 (H-Bond with Backbone)
<b>STOCKIN-03432</b>	Cys40 (H-Bond), <b>Phe56 (<math>\pi</math>-<math>\pi</math>)</b> , Ile146 (H-Bond with Backbone), <b>Ala148 (H-Bond with Backbone)</b> , Gly207 (Water Bridge H-Bond Network with Backbone), Arg218 (H-Bond)	Gly145 (H-Bond with Backbone), <b>Ala148 (H-Bond with Backbone)</b> , Gly211 (H-Bond with Backbone), <b>Phe208 (<math>\pi</math>-<math>\pi</math>)</b>
<b>25284644</b>	<b>Phe208 (<math>\pi</math>-<math>\pi</math>)</b>	Asn60 (H-Bond), <b>Phe208 (<math>\pi</math>-<math>\pi</math>)</b> , Gly211 (H-Bond with Backbone), Arg218 (H-Bond)
<b>32927247</b>	Gly211 (Water Bridge H-Bond Network with Backbone), <b>Arg218 (Water Bridge H-Bond Network)</b>	<b>Asn60 (H-Bond)</b> , <b>Arg218 (H-Bond)</b>
<b>27371521</b>	Gly145 (Water Bridge H-Bond with Backbone), Tyr150 (H-Bond with Backbone), Tyr212 ( $\pi$ - $\pi$ and H-Bond with Backbone)	Asn60 (H-Bond), Tyr150, Phe208 (H-Bond with Backbone), Phe208 ( $\pi$ - $\pi$ )
<b>2801237</b>	Ile59 (Water Bridge H-Bond Network with Backbone), Ile146 (Water Bridge H-Bond Network with Backbone), <b>Arg218 (Water Bridge H-Bond Network)</b>	Asn60 (H-Bond), Ala148 (H-Bond with Backbone), Gly211 (H-Bond with Backbone), Phe208 ( $\pi$ - $\pi$ ), <b>Arg218 (H-Bond)</b>
<b>5633879</b>	<b>Gly145 (H-Bond with Backbone)</b> , <b>Ala148 (H-Bond with Backbone)</b> , <b>Tyr212 (H-Bond with Backbone)</b> , <b>Arg218 (H-Bond)</b>	<b>Asn60 (H-Bond)</b> , <b>Ala148 (H-Bond with Backbone)</b> , <b>Arg218 (H-Bond)</b> ,
<b>36994203</b>	<b>Gly211 (H-Bond with Backbone)</b> , <b>Arg218 (H-Bond and <math>\pi</math>-cation)</b>	Asn60 (H-Bond), Gly145 (H-Bond with Backbone), <b>Gly211 (H-Bond with Backbone)</b> , <b>Arg218 (H-Bond)</b>
<b>16694164</b>	<b>Gly145 (H-Bond with Backbone)</b> , Gly207 (Water Bridge H-Bond Network with Backbone), Arg218 (H-Bond)	Asn60 (H-bond), Gly211 (H-Bond with Backbone), <b>Gly145 (H-Bond with Backbone)</b> , Ala148 (H-Bond with Backbone)
<b>13477729</b>	Arg218 (Water Bridge Network H-Bond)	Asn60 (H-Bond), Gly145 (H-Bond with Backbone), Gly211 (H-Bond with Backbone)
<b>36657947</b>	Ala148 (Water Bridge H-Bond Network with Backbone), Tyr150 (H-Bond with Backbone), Phe208 ( $\pi$ - $\pi$ )	Asn60 (H-bond), Gly207 (H-Bond with Backbone)

<b>36716128</b>	Phe63 ( $\pi$ - $\pi$ ), Tyr212 ( $\pi$ - $\pi$ )	Asn60 (H-bond), Gly211, <b>Ala148 (H-Bond with Backbone)</b>
<b>3325122</b>	Val43 (Water Bridge H-Bond Network with Backbone), Val47 (Water Bridge H-Bond Network with Backbone), Leu48 (Water Bridge H-Bond Network with Backbone), <b>Phe208 (<math>\pi</math>-<math>\pi</math>)</b>	Asn60 (H-Bond), <b>Gly145 (H-Bond with Backbone)</b> , Gly211 (H-Bond with Backbone), Tyr150, <b>Phe208 (<math>\pi</math>-<math>\pi</math>)</b>
<b>42888719</b>	Gly145 (Water Bridge H-Bond Network with Backbone), <b>Phe208 (<math>\pi</math>-<math>\pi</math>)</b> , Tyr212 (H-Bond with Backbone), <b>Arg218 (H-Bond and Water Bridge H-Bond Network)</b>	Phe208 ( $\pi$ - $\pi$ ), Gly211 (H-Bond with Backbone), <b>Arg218 (H-Bond)</b>
<b>IP6</b>	Asn60 (H-Bond), <b>Tyr150 (H-Bond with Backbone)</b> , Pro151 (H-Bond with Backbone), Ser152 (H-Bond with Backbone), <b>Phe208, (H-Bond with Backbone)</b>	<b>Tyr150 (H-Bond with Backbone)</b> , Phe208 (H-Bond with Backbone), Phe208 ( $\pi$ - $\pi$ )
<b>31879059</b>	Tyr150 (Water Bridge H-Bond Network with Backbone), Ser210 (Water Bridge H-Bond Network with Backbone), <b>Arg218 (H-Bond and <math>\pi</math>-cation)</b> , Trp242 (Water Bridge H-Bond Network with Backbone)	Gly142 (H-Bond with Backbone), <b>Arg218 (H-Bond)</b>
<b>31966421</b>	Phe63 ( $\pi$ - $\pi$ ), <b>Gly145 (H-Bond with Backbone)</b> , Ala148 (H-Bond with Backbone), <b>Arg218 (H-Bond)</b>	Asn60 (H-Bond), <b>Gly145 (H-Bond with Backbone)</b> , Phe208 ( $\pi$ - $\pi$ ), <b>Arg218 (H-Bond)</b>
<b>7658775</b>	<b>Gly145 (H-Bond with Backbone)</b> , <b>Ala148 (H-Bond with Backbone)</b> , Tyr150 (H-Bond with Backbone), <b>Arg218 (H-Bond)</b>	<b>Gly142 (H-Bond with Backbone)</b> , Gly145 (H-Bond with Backbone), Ala148 (H-Bond with Backbone), <b>Arg218 (H-Bond)</b>
<b>25665268</b>	Gly145 (Water Bridge H-Bond Network with Backbone), Ala148 (Water Bridge H-Bond Network with Backbone), <b>Phe208 (H-Bond with Backbone)</b> , Tyr212 ( $\pi$ - $\pi$ ), Arg218 ( $\pi$ -cation)	Asn60 (H-Bond), Tyr150 (H-Bond with Backbone), <b>Phe208 (H-Bond with Backbone)</b> , Phe208 ( $\pi$ - $\pi$ )
<b>37101119</b>	Ile146 (H-Bond with Backbone), <b>Gly211 (Water Bridge H-Bond Network with Backbone)</b> , <b>Arg218 (H-Bond and Water Bridge H-Bond Network)</b>	Asn60 (H-Bond), Phe208 ( $\pi$ - $\pi$ ), <b>Gly211 (H-Bond with Backbone)</b> , <b>Arg218 (H-Bond)</b>

1 \*The name of the compounds for these given compound ID is mentioned in the supplementary file

### 3.5. Binding free energy analysis

To estimate the binding association between the AQP3 protein and the selected 20 poses obtained from dynamic simulation, an MM-GBSA analysis was carried out on the last 10ns of each dynamic simulation trajectory. The MM-GBSA analysis results are shown in **Table 3**. Compounds 5633879, IP6, 25284644, 36994203, and 27371521 displayed higher binding free energy values ( $\Delta G_{\text{Bind}}$ ) of -74.01 kcal/mol, -68.48 kcal/mol, -63.87 kcal/mol, -62.94 kcal/mol, and -61.30 kcal/mol, respectively, among all 20 dynamic poses. Further, comparison of the molecular docking scores and binding free energies of all top 20 poses revealed that compound C8C (6-(2-(1H-indol-6-yl)ethyl)-2-aminopyrimidin-4(3H)-one), despite having the highest docking score of -7.55 kcal/mol, shows the lowest binding free energy of -27.28 kcal/mol. Similarly, some marginally low scored compounds in the molecular docking emerged as better inhibitors based on binding free energy. In this context, compounds that were sorted and selected based on docking scores alone, without considering binding free energy, might show a different order of priority compared to compounds that were selected based on binding free energy.

The contributions of all parameters to the binding free energy are shown in **Table 3**, which clearly indicate that  $\Delta G_{\text{Bind}}$  is greatly influenced by the total contributions of coulombic, H-bond, Lipo, and vdW interactions. Furthermore, vdW and H-bond interactions are important contributors to ligand binding in all cases; however, it seems  $\Delta G_{\text{Lipo}}$  may also significantly affect the binding free energy of compounds with the AQP3 protein.

Among all the binding free energies observed, compound 5633879 (1-(4-methoxyphenoxy)-3-((4-methoxyphenyl)amino)propan-2-ol) had the highest  $\Delta G_{\text{Bind}}$  (-74.01 kcal/mol) with a docking score of -7.152 kcal/mol, forming H-bond with Ala148, Tyr212, and Arg218 residues in a dynamic trajectory that was similar to the interaction observed in the docking pose (Arg60, Ala148, Arg218). In the docked pose, the oxygen atoms of the  $-\text{OCH}_3$  and  $-\text{OH}$  groups formed H-bonds with the  $-\text{NH}-$  group of Asn60 and the guanidino group of Arg218 at a distance of 2.02 Å and 2.24 Å, respectively. Similarly, H-bond network with the backbone carbonyl of Ala148 has been observed with the  $-\text{NH}-$  and  $-\text{OH}$  groups of the linker region of the compound. Furthermore, we observed some additional interactions in dynamic poses, such as H-bonds with the backbone atoms of the Tyr212 residues. The interaction fractions of the compound 5633879 showed it in the vicinity of Phe63, Tyr150, Phe208, and Tyr212, imparting hydrophobic interactions during simulation (**Fig. S35 (b&c)**).

**Table 3.** MM-GBSA calculation for selected hit compounds (MM-GBSA was performed on last 5ns of simulation trajectory and mean values are shown with standard error.

S. No.	Compound ID	$\Delta G_{\text{Bind}}$	$\Delta G_{\text{Coulomb}}$	$\Delta G_{\text{Covalent}}$	$\Delta G_{\text{Hbond}}$	$\Delta G_{\text{Lipo}}$	$\Delta G_{\text{vdW}}$	$\Delta G_{\text{Packing}}$	$\Delta G_{\text{SolGB}}$
1	C8C	-29.63±8.60	-12.68±5.69	1.86±0.96	-0.85±0.50	-8.22±1.66	-30.47±3.01	-1.97±0.94	22.71±2.86
2	STOCKIN-03432	-58.62±5.60	-17.36±6.02	1.85±1.57	-2.21±0.31	-18.49±1.71	-44.37±2.12	-0.62±0.43	22.57±2.41
3	25284644	-38.89±4.94	-13.49±9.87	0.73±1.61	-0.04±0.14	-14.36±1.55	-30.36±3.32	-0.03±0.04	18.66±9.06
4	32927247	-35.29±8.61	-8.08±4.40	0.38±1.62	-0.46±0.31	-12.04±3.14	-31.02±4.76	-2.15±1.80	18.08±3.04
5	27371521	-48.58±6.90	-12.60±9.56	0.98±1.65	-0.89±0.52	-13.55±1.89	-39.97±2.87	-2.32±0.56	19.77±6.11
6	2801237	-41.97±6.05	-11.06±3.13	2.02±1.51	-5.18±0.29	-14.36±2.26	-36.66±2.83	-0.65±0.49	19.25±2.58
7	5633879	-61.39±6.54	-16.16±3.24	0.82±1.54	-2.06±0.38	-19.47±4.20	-45.61±3.96	-2.04±0.83	23.13±1.90
8	36994203	-59.60±4.08	-29.37±4.66	0.36±1.43	-1.79±0.22	-13.59±1.24	-44.60±2.36	-1.67±0.60	31.05±2.66
9	16694164	-59.51±5.19	-22.51±5.12	1.78±1.50	-1.68±0.25	-19.87±2.77	-49.12±3.06	-1.47±0.81	33.37±3.03
10	13477729	-38.22±5.11	-8.22±4.20	1.50±1.74	-0.20±0.31	-17.15±2.38	-34.86±3.35	-1.20±0.89	21.90±3.16
11	36657947	-46.65±6.63	-6.91±2.97	0.70±1.04	-0.59±0.21	-18.49±2.37	-36.78±3.85	-0.81±0.40	16.22±2.07
12	36716128	-45.02±4.49	-4.67±2.86	0.98±1.17	-0.10±0.18	-16.89±1.85	-36.18±2.68	-2.09±0.63	13.93±3.12
13	3325122	-48.08±4.07	-8.37±4.67	1.81±2.04	-0.28±0.35	-18.06±2.13	-38.74±3.63	-0.80±0.35	16.38±3.44
14	42888719	-43.41±5.27	-11.96±5.93	1.75±1.21	-0.77±0.42	-14.95±2.09	-36.30±2.94	-1.23±0.53	20.06±3.43
15	IP6	-56.96±5.91	-38.84±11.98	4.39±2.35	-1.83±0.42	-16.13±1.64	-36.57±2.54	-2.24±1.20	34.26±10.93
16	31879059	-52.83±4.02	-26.36±4.86	5.16±1.97	-1.91±0.44	-16.09±1.04	-48.97±2.62	-2.47±0.57	37.80±3.65
17	31966421	-60.74±3.93	-19.59±2.92	1.86±1.73	-1.53±0.18	-17.98±1.65	-43.42±3.45	-2.96±0.92	22.88±1.65
18	7658775	-56.92±5.02	-20.51±3.35	2.15±1.32	-1.94±0.42	-18.14±2.10	-41.92±2.95	-0.90±0.51	24.35±1.53
19	25665268	-57.84±3.48	-15.72±4.61	3.36±0.83	-0.63±0.35	-18.53±0.75	-47.57±1.65	-1.44±0.36	22.67±3.00
20	37101119	-41.65±6.63	-5.43±5.81	-0.29±1.47	-0.32±0.26	16.03±1.74	-35.70±3.82	-0.94±0.73	17.08±3.81

\*The name of the compounds for these given compound ID is mentioned in the supplementary file.

Similarly, compound 31966421 (3-(1H-benzo[d]imidazol-2-yl)-N-(2-methyl-3-(1H-pyrazol-1-yl)propyl)propanamide) had an intermediary  $\Delta G_{\text{Bind}}$  (-56.35 kcal/mol) with a docking score of -6.747 kcal/mol. It forms H-bonds with Asn60, Gly145, and Arg218, and  $\pi$ - $\pi$  interactions with Phe208 residues during docking, while in dynamic simulation the compound retains H-bonds with Asn60, and Arg218, as well as forming new H-bond and  $\pi$ - $\pi$  interactions with the Ile146, Gly211, Tyr212 and Phe56 residues, respectively. In the docking pose, the -OH group of the linker and the ring nitrogen of imidazole form H-bonds with Arg218 and Asn60 at a distance of 2.19 Å and 1.96 Å, respectively. The -NH- group of the benzimidazole group also forms a H-bond with the backbone carbonyl of Gly145. In dynamic trajectories, the Asn60 and Arg218 residues retain the H-bond interactions for more than 30% of the simulation time and the imidazole group engages into a  $\pi$ - $\pi$  interaction with Phe56.

Compound C8C, which was identified as the top scorer in molecular docking, shows the lowest binding free energy among all the compounds analyzed. In the docking pose, the compound has H-bonds with the carbonyls of Asn60, Tyr150, and Phe208 with distances of 2.07Å, 2.06Å, and 2.0Å, respectively; while in the dynamic trajectory the compound failed to display any H-bonds for 20% of the simulation while compound showed a  $\pi$ - $\pi$  interaction with Phe63. Despite all the individual components contributing to the binding free energy in the MM-GBSA analysis, the insignificant contribution of the  $\Delta G_{\text{Lipo}}$  resulting from higher ligand strain energy might lead to its lower binding free energy.

Furthermore, the energy of protein receptor and in complex with compounds were also measured during the simulation (**Table S4**). The result shows that the energy of the complex is lowered as compared to protein alone and suggested the stability of the pose.

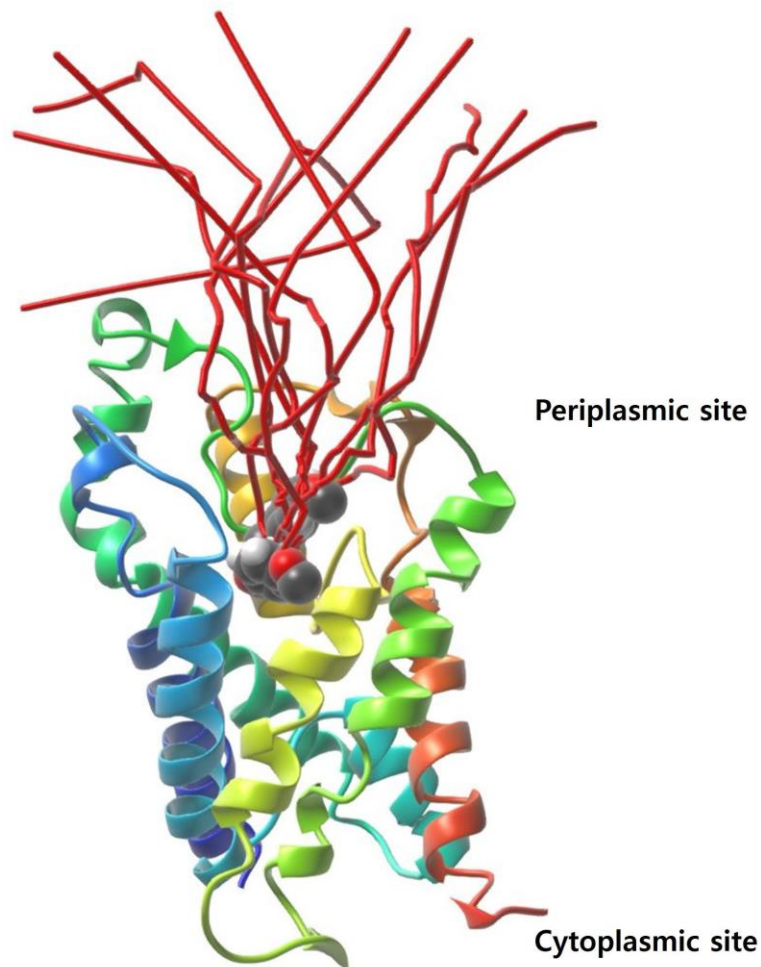
Thus, based on binding free energy values, the order of top 10 compounds is 5633879 > IP6 > 25284644 > 36994203 > 27371521 > 31879059 > 16694164 > 7658775 > 31966421 > 36716128.

### 3.6. Exploration of the ligand-unbinding pathways

The initial protein-ligand complex and the active atoms on compound 5633879 (1-(4-methoxyphenoxy)-3-((4-methoxyphenyl)amino)propan-2-ol) are shown in **Fig. 4**, illustrating no direction-bias sampling-domain. From all 10 runs, we found that the ligand-unbinding process followed a common pathway (**Fig. 5 and supplementary movie-01**). The ligand-unbinding path statistics are shown in **Table S3**. The AQP3 membrane protein has two sites, one that is periplasmic and one that is cytoplasmic. To permeate the cell, solute molecules must follow the periplasmic site. Our ligand-unbinding pathway search revealed that despite no



direction-bias sampling, ligand unbinding and exit from the protein followed the periplasmic site.

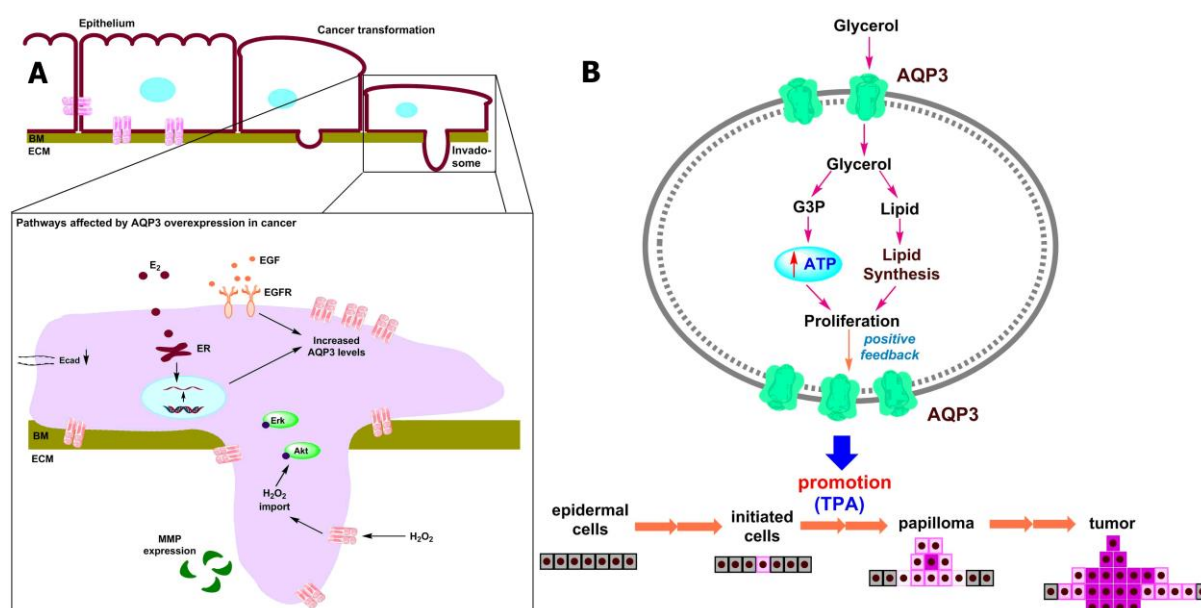


**Fig. 5** The ligand-unbinding path (in red color). The protein is represented with a ribbon and the ligand in VdW sphere.

### 3.7. Biological network

AQP3 was overexpressed in the plasma membrane of keratinocytes in the basal and spinous layers of epidermis, where it contributes to metastasis, proliferation, and the epithelial-to-mesenchymal transition (EMT) [73,74]. Here AQP3 serves as a water and glycerol transporter, facilitating skin hydration and possibly playing a role in cell migration. A study of an AQP3 knockout mouse model revealed these mice did not develop skin tumors, even after exposure to a tumor initiator [73]. Epidermal growth factor and estrogen both contribute to cancer development and are upstream regulators of AQP3 expression [75]. Since cancer cells have increased levels of  $H_2O_2$ , AQP3 mediated  $H_2O_2$  transport plays an important role in cancer progression [76]. It has also been suggested that AQP3-mediated  $H_2O_2$  transport increased phosphorylation of the protein kinase B (Akt) and the extracellular signal-regulated kinase

(Erk) 1/2. Similarly, overexpressed AQP3 increases the MMPs (matrix-metalloproteases), which further promote the cancer cell invasiveness (**Figure 6A**) [77]. Similarly, Verkman et al. have reported that AQP3-facilitated glycerol transport is involved in ATP generation and regulation of epidermal cell proliferation. Cells with high levels of intracellular glycerol due to overexpressed AQP3 are in states of epidermal hyperproliferation, such as psoriasis, ichthyosis, atopic dermatitis, wound healing, and tumorigenesis. Thus, AQP3-facilitated glycerol transport generates ATP and mediates the growth and survival of tumor cells (**Figure 6B**) [15]. Therefore, targeting AQP3 expression reduces several intracellular signalling pathways, resulting in reduced cell proliferation, migration, and invasion.



**Fig. 6** Schematic diagram showing the effect of overexpressed AQP3 protein in the basolateral layer of the skin membrane.

#### 4. Conclusions

The diverse roles AQPs in physiology and its involvement in prognosis of a variety of disease states have been well established, which necessitate the discovery of selective modulators or inhibitors as therapeutic agents. AQP3 is widely distributed in epithelial cells of the kidneys, airways, and skin, playing a role in water reabsorption, mucosal secretions, skin hydration, and cell volume regulation. However, a previous study has reported the aberrant expression of AQP3 in melanoma cells, suggesting its inhibition will lead to a new therapeutic treatment. In the present study, we performed a virtual screening to identify novel hits for inhibitors of the AQP3 target protein. A total of 20 hits with good binding affinity were obtained from a combination of pharmacophore and docking-based screening strategies. The physicochemical properties of the selected compounds comply with skin permeability properties. The hit

compounds obtained bind to key amino acid residues (Phe63, Tyr212, and Arg218) to inhibit the activity of the AQP3 protein. The molecular dynamics and MM-GBSA analysis revealed the compound 5633879 (1-(4-methoxyphenoxy)-3-((4-methoxyphenyl)amino)propan-2-ol) has good free energy of binding. Since the AQP3 is a channel protein embedded into the basolateral layer of the skin membrane, it has two relevant sites: a periplasmic site and a cytoplasmic site. The inhibitor must approach the AQP3 protein via the periplasmic site to modulate its function. This led us to further explore the ligand-unbinding pathway of the bound protein-ligand complex system. The ligand-unbinding pathway also revealed that the inhibitor approached the binding site through the periplasmic site. The hit compounds obtained from the present study promise good docking scores and binding free energy for the AQP3 protein; however, further research is required for hit optimization and biological screening.

**Supplementary Materials:** Figure S1. The AQP3 protein structure with identified Site\_1 and Site\_2 (Circled region) from SiteMap tool of Schrodinger suite. Figure S2. The representative Lipid-protein-ligand system built for molecular dynamic simulation. The lipids proteins and ligands are represented in green colour, cartoon and yellow colour with vdW sphere respectively.

Figure S3. The physicochemical properties based on predefined filter. Table S1. The top 20 hit compound selected after docking simulation (XP mode). Table S2. Docking Score, and amino acid residues in the binding site within 3Å for top 20 hits. Table S3. Ligand unbinding path result for compound (5633879). Figure S4 Binding site residues interactions of docked poses with AQP-3 protein. (Amino acid residues displayed interactions within 3Å are shown). Figure S5 to Figure S24 shows the Total energy (E) (kcal/mol); Potential energy (E\_P) (kcal/mol); Pressure (P) (in bar), Temperature (T) (in K), and Volume (V) (in Å<sup>3</sup>) during the molecular simulation for compounds in complex with AQP-3. Figure S25 to Figure S28 shows the RMSD of C $\alpha$ -atoms of AQP-3 in complex compounds. Figure S29 to Figure S48 shows the RMSD, RMSF, protein-ligand histogram and diagram, H-bond interaction, and secondary structure elements for compounds in AQP3 protein during the simulation. Figure S49 to Figure S52 shows the RMSD of ligands in binding pocket of the AQP3 protein. Table S4. The energy of protein alone and in complex with compounds from dynamic simulation.

**Author contributions:** D.K.Y. conceived and designed the project and collected data from the literature and databases. S.K. and D.K.Y. performed the experiments, analyzed the data, and wrote the manuscript. E.H.C. SC. and M.H.K provided the lab facility. All authors contributed

to the interpretation and discussion of the results. All authors have read and approved the final version of the manuscript.

**Acknowledgments:** Author D.K.Y. is thankful to the Basic Science Research Program of the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science, and Technology, who supported this study (No.2017R1C1B2003380). The authors are also thankful to Gachon Institute of Pharmaceutical Science and the Department of Pharmacy, College of Pharmacy, Gachon University of Medicine and Science, Incheon, Korea for providing the computational modeling facility. This research was also partially supported by Leading Foreign Research Institute Recruitment Program through the National Research Foundation of Korea (NRF) funded by the Korea government (MSIP) (NRF-2016K1A4A3914113).

**Conflicts of interest:** The author(s) confirm that this article content has no conflicts of interest.

## References

1. Jemal, A.; Murray, T.; Samuels, A.; Ghafoor, A.; Ward, E.; Thun, M.J. Cancer statistics, 2003. *CA: a cancer journal for clinicians* **2003**, *53*, 5-26.
2. Einspahr, J.G.; Stratton, S.P.; Bowden, G.T.; Alberts, D.S. Chemoprevention of human skin cancer. *Critical reviews in oncology/hematology* **2002**, *41*, 269-285.
3. Katalinic, A.; Kunze, U.; Schäfer, T. Epidemiology of cutaneous melanoma and non-melanoma skin cancer in Schleswig-Holstein, Germany: incidence, clinical subtypes, tumour stages and localization (epidemiology of skin cancer). *British Journal of Dermatology* **2003**, *149*, 1200-1206.
4. Elwood, J.M.; Gallagher, R.P.; Hill, G.; Pearson, J. Cutaneous melanoma in relation to intermittent and constant sun exposure—the Western Canada Melanoma Study. *International journal of cancer* **1985**, *35*, 427-433.
5. Lauth, M.; Unden, A.B.; Toftgård, R. Non-melanoma skin cancer: pathogenesis and mechanisms. *Drug Discovery Today: Disease Mechanisms* **2004**, *1*, 267-272.
6. Elias, P.M. Structure and function of the stratum corneum extracellular matrix. *The Journal of investigative dermatology* **2012**, *132*, 2131-2133, doi:10.1038/jid.2012.246.
7. Verkman, A.S.; Mitra, A.K. Structure and function of aquaporin water channels. *American journal of physiology. Renal physiology* **2000**, *278*, F13-28, doi:10.1152/ajprenal.2000.278.1.F13.
8. Fujiyoshi, Y.; Mitsuoka, K.; de Groot, B.L.; Philippsen, A.; Grubmüller, H.; Agre, P.; Engel, A. Structure and function of water channels. *Current opinion in structural biology* **2002**, *12*, 509-515.
9. Hara-Chikuma, M.; Verkman, A. Roles of aquaporin-3 in the epidermis. *Journal of Investigative Dermatology* **2008**, *128*, 2145-2151.
10. Takata, K.; Matsuzaki, T.; Tajika, Y. Aquaporins: water channel proteins of the cell membrane. *Progress in histochemistry and cytochemistry* **2004**, *39*, 1-83.
11. Verkman, A.S. Aquaporins in clinical medicine. *Annual review of medicine* **2012**, *63*, 303-316, doi:10.1146/annurev-med-043010-193843.
12. Verkman, A.S.; Hara-Chikuma, M.; Papadopoulos, M.C. Aquaporins--new players in cancer biology. *Journal of molecular medicine* **2008**, *86*, 523-529, doi:10.1007/s00109-008-0303-9.
13. Verkman, A.S.; Anderson, M.O.; Papadopoulos, M.C. Aquaporins: important but elusive drug targets. *Nature reviews. Drug discovery* **2014**, *13*, 259-277, doi:10.1038/nrd4226.
14. Hara-Chikuma, M.; Verkman, A. Aquaporin-3 facilitates epidermal cell migration and proliferation during wound healing. *Journal of molecular medicine* **2008**, *86*, 221-231.
15. Hara-Chikuma, M.; Verkman, A. Prevention of skin tumorigenesis and impairment of epidermal cell proliferation by targeted aquaporin-3 gene disruption. *Molecular and cellular biology* **2008**, *28*, 326-332.
16. Sugiyama, Y.; Ota, Y.; Hara, M.; Inoue, S. Osmotic stress up-regulates aquaporin-3 gene expression in cultured human keratinocytes. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* **2001**, *1522*, 82-88.
17. Hara, M.; Ma, T.; Verkman, A. Selectively reduced glycerol in skin of aquaporin-3 deficient mice may account for impaired skin hydration, elasticity and barrier recovery. *Journal of Biological Chemistry* **2002**.
18. Hara, M.; Verkman, A. Glycerol replacement corrects defective skin hydration, elasticity, and barrier function in aquaporin-3-deficient mice. *Proceedings of the National Academy of Sciences* **2003**, *100*, 7360-7365.



19. Niemietz, C.M.; Tyerman, S.D. New potent inhibitors of aquaporins: silver and gold compounds inhibit aquaporins of plant and human origin. *FEBS letters* **2002**, *531*, 443-447.
20. Migliati, E.; Meurice, N.; DuBois, P.; Fang, J.S.; Somasekharan, S.; Beckett, E.; Flynn, G.; Yool, A.J. Inhibition of Aquaporin-1 and Aquaporin-4 water permeability by a derivative of the loop diuretic bumetanide acting at a internal pore-occluding binding site. *Molecular pharmacology* **2009**.
21. Brown, D.; Lu, H.A.J. Aquaporin-2 inhibitors: fishing in the chemical pool. *Journal of the American Society of Nephrology* **2013**, *24*, 685-686.
22. Brooks, H.L.; Regan, J.W.; Yool, A.J. Inhibition of aquaporin-1 water permeability by tetraethylammonium: involvement of the loop E pore region. *Molecular pharmacology* **2000**, *57*, 1021-1026.
23. Bing, M.; XIANG, Y.; Sheng-mei, M.; Tao, L.; He-ming, Y.; Xue-jun, L. Effects of acetazolamide and anordiol on osmotic water permeability in AQP1-cRNA injected *Xenopus* oocyte. *Acta Pharmacol Sin* **2004**, *25*, 90-97.
24. Kourghi, M.; Pei, J.V.; De Ieso, M.L.; Flynn, G.; Yool, A.J. Bumetanide derivatives AqB007 and AqB011 selectively block the aquaporin-1 ion channel conductance and slow cancer cell migration. *Molecular pharmacology* **2016**, *89*, 133-140.
25. Vanhoek, A.; Vanos, C. Inhibition of water channels in isolated rat renal cortical plasma-membranes by dimethylsulfoxide and mercuric sulfhydryl-reagents. In *Proceedings of Journal of Physiology-London*; pp. P142-P142.
26. Preston, G.M.; Jung, J.S.; Guggino, W.B.; Agre, P. The mercury-sensitive residue at cysteine 189 in the CHIP28 water channel. *Journal of Biological Chemistry* **1993**, *268*, 17-20.
27. Martins, A.P.; Marrone, A.; Ciancetta, A.; Cobo, A.G.; Echevarría, M.; Moura, T.F.; Re, N.; Casini, A.; Soveral, G. Targeting aquaporin function: potent inhibition of aquaglyceroporin-3 by a gold-based compound. *PloS one* **2012**, *7*, e37435.
28. Serna, A.; Galán-Cobo, A.; Rodrigues, C.; Sánchez-Gomar, I.; Toledo-Aral, J.J.; Moura, T.F.; Casini, A.; Soveral, G.; Echevarría, M. Functional Inhibition of Aquaporin-3 With a Gold-Based Compound Induces Blockage of Cell Proliferation. *Journal of cellular physiology* **2014**, *229*, 1787-1801.
29. Martins, A.P.; Ciancetta, A.; de Almeida, A.; Marrone, A.; Re, N.; Soveral, G.; Casini, A. Aquaporin inhibition by gold (III) compounds: new insights. *ChemMedChem* **2013**, *8*, 1086-1092.
30. De Almeida, A.; Soveral, G.; Casini, A. Gold compounds as aquaporin inhibitors: new opportunities for therapy and imaging. *MedChemComm* **2014**, *5*, 1444-1453.
31. Daina, A.; Michielin, O.; Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports* **2017**, *7*, 42717, doi:10.1038/srep42717.
32. van de Waterbeemd, H.; Camenisch, G.; Folkers, G.; Chretien, J.R.; Raevsky, O.A. Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors. *Journal of drug targeting* **1998**, *6*, 151-165, doi:10.3109/10611869808997889.
33. Potts, R.O.; Guy, R.H. A predictive algorithm for skin permeability: the effects of molecular size and hydrogen bond activity. *Pharmaceutical research* **1995**, *12*, 1628-1633.
34. Ku, M.S. Use of the Biopharmaceutical Classification System in early drug development. *The AAPS journal* **2008**, *10*, 208-212, doi:10.1208/s12248-008-9020-0.



35. Golovenko, N.; Borisiuk, I. [Biopharmaceutical classification system--experimental model of the prediction of drug bioavailability]. *Biomeditsinskaia khimiia* **2008**, *54*, 392-407.
36. Yadav, D.K.; Kumar, S.; Saloni; Misra, S.; Yadav, L.; Teli, M.; Sharma, P.; Chaudhary, S.; Kumar, N.; Choi, E.H., et al. Molecular Insights into the Interaction of RONS and Thieno[3,2-c]pyran Analogs with SIRT6/COX-2: A Molecular Dynamics Study. *Scientific reports* **2018**, *8*, 4777-4777, doi:10.1038/s41598-018-22972-9.
37. Yadav, D.K.; Kumar, S.; Saloni; Singh, H.; Kim, M.-H.; Sharma, P.; Misra, S.; Khan, F. Molecular docking, QSAR and ADMET studies of withanolide analogs against breast cancer. *Drug design, development and therapy* **2017**, *11*, 1859-1870, doi:10.2147/DDDT.S130601.
38. Verma, S.; Kumar, S.; Kumar, S. Design, synthesis, computational and biological evaluation of new benzodiazepines as CNS agents. *Arabian Journal of Chemistry* **2017**, <https://doi.org/10.1016/j.arabjc.2017.08.005>, doi:<https://doi.org/10.1016/j.arabjc.2017.08.005>.
39. Saluja, A.K.; Tiwari, M.; Kumar, S. Ligand-based design, virtual screening, and ADME/T-based profiling on a dataset of 1,3,5-triazine-substituted benzene sulfonamides as carbonic anhydrase inhibitors. *Journal of Chemometrics* **2014**, *28*, 108-115, doi:doi:10.1002/cem.2579.
40. Huggins, D.J.; Venkitaraman, A.R.; Spring, D.R. Rational methods for the selection of diverse screening compounds. *ACS chemical biology* **2011**, *6*, 208-217, doi:10.1021/cb100420r.
41. Tsaion, K., and Steven A. Kates, eds. *ADMET for medicinal chemists: a practical guide*; John Wiley & Sons: 2011.
42. Lian, G.; Chen, L.; Han, L. An evaluation of mathematical models for predicting skin permeability. *Journal of pharmaceutical sciences* **2008**, *97*, 584-598, doi:10.1002/jps.21074.
43. Mitragotri, S.; Anissimov, Y.G.; Bunge, A.L.; Frisch, H.F.; Guy, R.H.; Hadgraft, J.; Kasting, G.B.; Lane, M.E.; Roberts, M.S. Mathematical models of skin permeability: an overview. *International journal of pharmaceutics* **2011**, *418*, 115-129, doi:10.1016/j.ijpharm.2011.02.023.
44. Potts, R.O.; Guy, R.H. Predicting skin permeability. *Pharmaceutical research* **1992**, *9*, 663-669.
45. <https://www.emolecules.com>. Eolecules Database: La Jolla, CA, .
46. <https://www.ibscreen.com/bases>. InterBioScreen Database. Availabe online: (accessed on
47. Braga, R.C.; Alves, V.M.; Muratov, E.N.; Strickland, J.; Kleinstreuer, N.; Tropsha, A.; Andrade, C.H. Pred-Skin: A Fast and Reliable Web Application to Assess Skin Sensitization Effect of Chemicals. *Journal of Chemical Information and Modeling* **2017**, *57*, 1013-1017, doi:10.1021/acs.jcim.7b00194.
48. <https://www.fda.gov/Drugs/default.htm>. FDA Approved Drugs. Availabe online: (accessed on
49. Loving, K.; Salam, N.K.; Sherman, W. Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation. *Journal of computer-aided molecular design* **2009**, *23*, 541-554, doi:10.1007/s10822-009-9268-1.
50. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K., et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry* **2004**, *47*, 1739-1749, doi:10.1021/jm0306430.

51. Bowers, K.J.; Chow, D.E.; Xu, H.; Dror, R.O.; Eastwood, M.P.; Gregersen, B.A.; Klepeis, J.L.; Kolossvary, I.; Moraes, M.A.; Sacerdoti, F.D., et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In Proceedings of SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, 11-17 Nov. 2006; pp. 43-43.
52. Greenidge, P.A.; Kramer, C.; Mozziconacci, J.C.; Wolf, R.M. MM/GBSA binding energy prediction on the PDBbind data set: successes, failures, and directions for further improvement. *J Chem Inf Model* **2013**, *53*, 201-209, doi:10.1021/ci300425v.
53. Maestro11.6. Schrödinger Release 2018-2: LigPrep, Schrödinger, LLC, New York, NY, 2018.
54. Sastry, G.M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided molecular design* **2013**, *27*, 221-234, doi:10.1007/s10822-013-9644-8.
55. Savage, D.F.; O'Connell, J.D., 3rd; Miercke, L.J.; Finer-Moore, J.; Stroud, R.M. Structural context shapes the aquaporin selectivity filter. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107*, 17164-17169, doi:10.1073/pnas.1009864107.
56. Park, J.H.; Saier, M.H., Jr. Phylogenetic characterization of the MIP family of transmembrane channel proteins. *The Journal of membrane biology* **1996**, *153*, 171-180.
57. Halgren, T.A. Identifying and characterizing binding sites and assessing druggability. *J Chem Inf Model* **2009**, *49*, 377-389, doi:10.1021/ci800324m.
58. Peach, M.L.; Nicklaus, M.C. Combining docking with pharmacophore filtering for improved virtual screening. *Journal of cheminformatics* **2009**, *1*, 6-6, doi:10.1186/1758-2946-1-6.
59. Planesas, J.M.; Claramunt, R.M.; Teixido, J.; Borrell, J.I.; Perez-Nueno, V.I. Improving VEGFR-2 docking-based screening by pharmacophore postfiltering and similarity search postprocessing. *J Chem Inf Model* **2011**, *51*, 777-787, doi:10.1021/ci1002763.
60. Chang, C.; Ekins, S.; Bahadduri, P.; Swaan, P.W. Pharmacophore-based discovery of ligands for drug transporters. *Advanced drug delivery reviews* **2006**, *58*, 1431-1450, doi:10.1016/j.addr.2006.09.006.
61. Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *The Journal of Physical Chemistry A* **2001**, *105*, 9954-9960, doi:10.1021/jp003020w.
62. Saputro, D.R.S.; Widyaningsih, P. Limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method for the parameter estimation on geographically weighted ordinal logistic regression model (GWOLR). *AIP Conference Proceedings* **2017**, *1868*, 040009, doi:10.1063/1.4995124.
63. Evans, D.J.; Holian, B.L. The nose-hoover thermostat. *The Journal of chemical physics* **1985**, *83*, 4069-4074.
64. Cho, K.; Joannopoulos, J.D.; Kleinman, L. Constant-temperature molecular dynamics with momentum conservation. *Physical Review E* **1993**, *47*, 3145-3151, doi:10.1103/PhysRevE.47.3145.
65. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **1995**, *103*, 8577-8593, doi:10.1063/1.470117.
66. Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy

- Perturbation and the OPLS Force Field. *Journal of Chemical Theory and Computation* **2010**, 6, 1509-1519, doi:10.1021/ct900587b.
67. Li, J.; Abel, R.; Zhu, K.; Cao, Y.; Zhao, S.; Friesner, R.A. The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins: Structure, Function, and Bioinformatics* **2011**, 79, 2794-2812.
  68. Nguyen, M.K.; Jaillet, L.; Redon, S. ART-RRT: As-Rigid-As-Possible exploration of ligand unbinding pathways. *Journal of computational chemistry* **2018**, 39, 665-678, doi:10.1002/jcc.25132.
  69. Oostenbrink, C.; Villa, A.; Mark, A.E.; van Gunsteren, W.F. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of computational chemistry* **2004**, 25, 1656-1676, doi:10.1002/jcc.20090.
  70. Malde, A.K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P.C.; Oostenbrink, C.; Mark, A.E. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *J Chem Theory Comput* **2011**, 7, 4026-4037, doi:10.1021/ct200196m.
  71. Dixon, S.L.; Smondyrev, A.M.; Rao, S.N. PHASE: a novel approach to pharmacophore modeling and 3D database searching. *Chemical biology & drug design* **2006**, 67, 370-372, doi:10.1111/j.1747-0285.2006.00384.x.
  72. De Almeida, A.; Mósca, A.F.; Wragg, D.; Wenzel, M.; Kavanagh, P.; Barone, G.; Leoni, S.; Soveral, G.; Casini, A. The mechanism of aquaporin inhibition by gold compounds elucidated by biophysical and computational methods. *Chemical communications* **2017**, 53, 3830-3833, doi:10.1039/C7CC00318H.
  73. Hara-Chikuma, M.; Verkman, A.S. Prevention of skin tumorigenesis and impairment of epidermal cell proliferation by targeted aquaporin-3 gene disruption. *Mol Cell Biol* **2008**, 28, 326-332, doi:10.1128/mcb.01482-07.
  74. Chen, J.; Wang, T.; Zhou, Y.C.; Gao, F.; Zhang, Z.H.; Xu, H.; Wang, S.L.; Shen, L.Z. Aquaporin 3 promotes epithelial-mesenchymal transition in gastric cancer. *Journal of experimental & clinical cancer research : CR* **2014**, 33, 38, doi:10.1186/1756-9966-33-38.
  75. Huang, Y.-T.; Zhou, J.; Shi, S.; Xu, H.-Y.; Qu, F.; Zhang, D.; Chen, Y.-D.; Yang, J.; Huang, H.-F.; Sheng, J.-Z. Identification of Estrogen Response Element in Aquaporin-3 Gene that Mediates Estrogen-induced Cell Migration and Invasion in Estrogen Receptor-positive Breast Cancer. *Scientific reports* **2015**, 5, 12484-12484, doi:10.1038/srep12484.
  76. Lennicke, C.; Rahn, J.; Lichtenfels, R.; Wessjohann, L.A.; Seliger, B. Hydrogen peroxide - production, fate and role in redox signaling of tumor cells. *Cell communication and signaling : CCS* **2015**, 13, 39, doi:10.1186/s12964-015-0118-6.
  77. Marlar, S.; Jensen, H.H.; Login, F.H.; Nejsun, L.N. Aquaporin-3 in Cancer. *International journal of molecular sciences* **2017**, 18, 2106, doi:10.3390/ijms18102106.

## **TABLE CAPTIONS**

**Table 1** Small molecules considered under the present study

**Table 2** Comparison of amino acid residue interactions observed during Molecular Dynamic simulations and Molecular Docking (Bold faces represent the common binding amino acid residues).

**Table 3** MM-GBSA calculation for selected hit compounds (MM-GBSA was performed on the last 10ns of the simulation trajectory; mean values are shown with standard error).

## **FIGURE CAPTIONS**

**Fig. 1** Designed workflow for the virtual screening of AQP3 inhibitors.

**Fig. 2** AQP3 protein with periplasmic and cytoplasmic sites. The SF region comprising key amino acid residues (Phe63, Tyr212, and Arg218) are displayed as ball and stick representations.

**Fig. 3** (A) Whole protein showing cavity-based pharmacophoric point; (B) top view with pharmacophoric point; (C) pharmacophoric point with label residues in the binding pocket.

**Fig. 4** Left: the system in ribbons with bound ligand in ball and stick form, Right: The ligand with active atoms in green.

**Fig. 5** The ligand-unbinding path (in red color). The protein is represented with a ribbon and the ligand in VdW sphere.

**Fig. 6** Schematic diagram showing the effect of overexpressed AQP3 protein in the basolateral layer of the skin membrane.