

## RESEARCH ARTICLE

# Prediction of LncRNA-encoded small peptides in glioma and oligomer channel functional analysis using in silico approaches

Yipeng Cao<sup>1,2</sup>, Rui Yang<sup>3</sup>, Imshik Lee<sup>4</sup>, Wenwen Zhang<sup>1</sup>, Jiana Sun<sup>1</sup>, Xiangfei Meng<sup>2\*</sup>, Wei Wang<sup>1\*</sup>

**1** National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Hexi District, Tianjin, P. R. China, **2** National Supercomputer Center in Tianjin, TEDA – Tianjin Economic-Technological Development Area, Tianjin, P. R. China, **3** Department of Infection and Immunity, Tianjin Union Medical Center, Nankai University Affiliated Hospital, Hongqiao District, Tianjin, P. R. China, **4** College of Physics, Nankai University, Nankai District, Tianjin, P. R. China

\* [mengxf@nscj.cn](mailto:mengxf@nscj.cn) (XM); [weiwang\\_2@126.com](mailto:weiwang_2@126.com) (WW)



## OPEN ACCESS

**Citation:** Cao Y, Yang R, Lee I, Zhang W, Sun J, Meng X, et al. (2021) Prediction of LncRNA-encoded small peptides in glioma and oligomer channel functional analysis using in silico approaches. *PLoS ONE* 16(3): e0248634. <https://doi.org/10.1371/journal.pone.0248634>

**Editor:** Shailesh Kumar, National Institute of Plant Genome Research (NIPGR), INDIA

**Received:** May 26, 2020

**Accepted:** March 3, 2021

**Published:** March 18, 2021

**Copyright:** © 2021 Cao et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its [Supporting information](#) files.

**Funding:** The manuscript is supported by Science & Technology Development Fund of Tianjin Education Commission for Higher Education (2018KJ067).

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Glioma is a lethal malignant brain cancer, and many reports have shown that abnormalities in the behavior of water and ion channels play an important role in regulating tumor proliferation, migration, apoptosis, and differentiation. Recently, new studies have suggested that some long noncoding RNAs containing small open reading frames can encode small peptides and form oligomers for water or ion regulation. However, because the peptides are difficult to identify, their functional mechanisms are far from being clearly understood. In this study, we used bioinformatics methods to identify and evaluate lncRNAs, which may encode small transmembrane peptides in gliomas. Combining ab initio homology modeling, molecular dynamics simulations, and free energy calculations, we constructed a predictive model and predicted the oligomer channel activity of peptides by identifying the lncRNA ORFs. We found that one key hub lncRNA, namely, DLEU1, which contains two smORFs (ORF1 and ORF8), encodes small peptides that form pentameric channels. The mechanics of water and ion ( $\text{Na}^+$  and  $\text{Cl}^-$ ) transport through this pentameric channel were simulated. The potential mean force of the  $\text{H}_2\text{O}$  molecules along the two ORF-encoded peptide channels indicated that the energy barrier was different between ORF1 and ORF8. The ORF1-encoded peptide pentamer acted as a self-assembled water channel but not as an ion channel, and the ORF8 permeated neither ions nor water. This work provides new methods and theoretical support for further elucidation of the function of lncRNA-encoded small peptides and their role in cancer. Additionally, this study provides a theoretical basis for drug development.

## 1. Introduction

Glioma is one of the most prevalent types of primary intracranial carcinomas. Glioma has varying malignancy grades and histological subtypes [1], and it remains a highly lethal

malignancy worldwide. Glioma is characterized by rapid cell proliferation and angiogenesis [2]. Traditional treatments have limited effectiveness in the majority of gliomas because glioblastoma (GBM) stem-like cells (GSCs) are highly recrudescence [3]. Consequently, investigations exploring the accurate molecular mechanisms of and reliable therapeutic targets for gliomas have attracted extensive attention.

Many studies have shown that long noncoding RNAs (lncRNAs) are involved in many diseases, including glioma, and participate in gene regulation via various mechanisms [4, 5]. lncRNAs are a class of RNAs > 200 nucleotides in length that lack coding potential [6, 7]. Based on the increasing amount of functional lncRNAs aberrantly expressed in glioma tissues and cell lines [8–10], they are closely related to glioma occurrence by influencing tumor development, invasion, and metastasis. Recently, some studies have reported the discovery of open reading frames (ORFs) that encode functional small peptides, particularly lncRNAs [11, 12]. For example, Huang et al. found that the lncRNA HOXB-AS3 encodes micro peptides that can bind to heterogeneous nuclear ribonucleoproteins (hnRNPs) and suppress glucose metabolism reprogramming in colon cancer [13]. SLN, PLB, MLN, and DWORF [14–16] are involved in some physiological processes of diseases by regulating skeletal muscle activity via interaction with sarcoplasmic reticulum  $\text{Ca}^{2+}$ -ATPase (SERCA) [17]. Surprisingly, Ji et al. found that more than 40% of lncRNA fragments encode proteins and that more than half of these fragments are in the form of small peptides (median number of amino acids < 46) [18].

Previous reports have shown that glioma proliferation, apoptosis, migration, and invasion are driven by water permeability mostly through aquaporins (AQPs) [19, 20], and this information has often been used as a marker of expected survival in cancer patients. Water regulation is strongly correlated with glioma angiogenesis, brain edema, tumor migration, and tumor invasion [21, 22], indicating that the study of the water permeability mechanism could lead to the discovery of pharmaceutical targets for glioma management. In addition to and more importantly than AQP water channels, lncRNA-encoded small peptides, such as SLN and PLB, can form pentameric channels in the membrane through the leucine/isoleucine zipper in their transmembrane (TM) regions [23]. The pentameric channels formed by these peptides act as TM water and ion transport channels as well as hydrophobic gate channels.

It is interesting to note that the AQPs in gliomas are not sufficiently permeable to cause glioma progression [24–26]. Some reports have indicated that water molecules permeate through the lncRNA-encoded small peptide oligomer pores. Therefore, we considered similar mechanisms for the water permeation in glioma. Because experimental identification of lncRNA-encoded small peptides is difficult, the present bioinformatics methods were implemented using algorithms, such as support vector machine (SVM) algorithms, depending on the length of the amino acid chain and the homology to known sequences [12]. However, the encoded small TM peptides contain so few amino acids ( $\leq 46$  AA) that most are not homologous with other proteins and thus cannot be identified accurately by the current algorithms.

In this study, we identified the key hub lncRNAs in gliomas by multiple bioinformatics methods, including survival analysis, lncRNA functional enrichment analysis, bioinformatics predictions, and ab initio modeling, to improve the predictive reliability of lncRNA-encoded small peptides. Subsequent structural analysis evaluated the possible encoding of small TM peptides. Molecular dynamics (MD) simulation was applied to investigate whether the small TM peptides could stabilize the formation of a pentameric structure acting as a water or ion channel. Our results showed that the peptides encoded by a hub lncRNA formed a pentamer with water and ion permeability, supplementing our understanding of the permeability mechanism of water in gliomas. Thus, our results suggested that the hub lncRNAs in glioma may encode small TM peptides and that their self-assembled oligomers function as water channels. Additionally, our study provided new methods and theoretical support for further elucidation

of the functions of lncRNA-encoded small peptides and their role in cancer, and it provided a theoretical basis for drug development.

## 2. Materials and methods

### 2.1 Screening and analysis of glioma hub lncRNAs

**1) Computational analysis of RNA sequences.** All data were downloaded from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov>) [27] with complete lncRNA, mRNA, and miRNA expression profiles. The total glioma RNA sequences of 700 tumor samples and 5 adjacent nontumorous brain tissues were included in the subsequent analysis. The patients had an overall survival time of > 10 years. Because this study followed TCGA Research Network guidelines (<http://cancergenome.nih.gov/publications/publicationguidelines>), no further ethical approval was required.

**2) Statistical analysis.** Differential expression analysis was performed to identify differentially expressed lncRNAs, mRNAs, and miRNAs (DElncRNAs, DEmRNAs, and DEmiRNAs, respectively) by the edgeR package in Bioconductor 3.1 [28] with the cutoff criteria of  $|\log_2$  fold change (FC)| > 2 and false discovery rate (FDR) < 0.01. The *Homo sapiens* GTF file for gene annotation was downloaded from the National Center for Biotechnology Information (NCBI) website ([http://ftp.ncbi.nlm.nih.gov/genomes/Homo\\_sapiens/](http://ftp.ncbi.nlm.nih.gov/genomes/Homo_sapiens/)).

**3) Gene Ontology (GO), pathway analysis, and ceRNA network construction.** Significant DElncRNAs, DEmRNAs, and DEmiRNAs were examined in the GO database (<http://www.geneontology.org>) [29], in which significantly enriched GO terms were identified to analyze their biological function.

A ceRNA network was built, including lncRNA-miRNA-mRNA interactions. Human lncRNA-miRNA and mRNA-miRNA interactions were obtained from StarBase 2.0 (<http://starbase.sysu.edu.cn/starbase2/index.php>) [30]. According to the lncRNA-miRNA and miRNA-mRNA interactions, we utilized Cytoscape version 3.7.0 [31] to visualize the RNA network.

**4) Survival differences according to hub lncRNAs.** From the obtained ceRNA network, the hub lncRNAs were analyzed by using the cytoHubba [32] plugin in Cytoscape 3.7.0. We defined the hub lncRNA threshold as edges  $\geq 3$ . The Cox proportional hazards regression model was employed to analyze the hub lncRNAs in the ceRNA network. The survival periods of the glioma patients were obtained from TCGA. Statistically significant lncRNAs affecting the survival period ( $P < 0.05$ ) were determined by Cox regression univariate analysis to construct Kaplan-Meier survival curves for patients with glioma.

### 2.2 Prediction and evaluation of small open reading frames (smORFs) encoding TM peptides

The complete sequences of the lncRNAs were obtained from the NCBI Nucleotide database (<https://www.ncbi.nlm.nih.gov>). A combination of the above ceRNA network and the following four bioinformatics methods was used to predict the encoding ability of TM peptides in lncRNAs: 1) the candidate lncRNA sequences were predicted by Open Reading Frame Finder in NCBI (<https://www.ncbi.nlm.nih.gov/orffinder/>) [33] by using the nucleotide sequences above to find smORFs; 2) the Coding Potential Calculator (CPC) and CPC2 [34, 35] were used to assess the quality of ORFs; 3) the BLAST tool in NCBI was used to assess the conservation of lncRNA sequences between different species; and 4) the translated amino acid sequences of the smORFs were then used to identify the potential TM regions via TMHMM Server v2.0 [36] based on the hidden Markov model.

### 2.3 Peptide modeling and MD simulation

**1) Preparation of hub lncRNA-encoded candidate small peptide models.** We prepared candidate small peptide models by using I-TASSER [37] online software. The I-TASSER server was developed for protein structure predictions with homology modeling based on the ab initio algorithm, and it was ranked as the best automated method for protein structure prediction in the Critical Assessment of Structure Prediction (CASP) experiments [38, 39]. Because all of the small peptides contained potential  $\alpha$ -helix TM regions, we marked the TM regions and used the nuclear magnetic resonance (NMR) models of SLN and PLB as the reference structures for modeling.

**2) Preoptimization of peptide models.** To select the most accurate small peptides for further simulation, the models predicted by I-TASSER to have the highest scores were chosen. The small peptides were embedded in a pre-equilibrium palmitoyl-oleoyl-phosphatidylcholine (POPC) membrane using the DESMOND software package [40]. Then, the peptides were optimized by a 20 ns MD simulation. All of the parameters used were the same in subsequent simulations. Finally, the Rampage server was used for evaluating the models and the structure with the best scores.

**3) Preparation of the peptide pentamer-membrane models and MD simulation.** Two peptide structures were obtained from the prediction described above. To build the structure of self-assembled pentapeptides in the membrane, the PLB pentamer in membranes [41] (PDB code: 1ZLL) was used as the template due to its high similarity in the TM region. The equilibrated POPC membrane was used to mimic the sarcoplasmic reticulum (SR) membrane environment. The simulation was performed using a  $14 \times 14 \times 10$  nm<sup>3</sup> box with three-dimensional periodic boundaries. The box was filled with a TIP3P [42] water model and used as water molecules. The size of the entire system was ~140,000 atoms.

**4) MD simulation.** To obtain the equilibrated initial pentameric peptides in the POPC membrane, the Charmm36 [43] force field was chosen for MD simulation. The particle mesh Ewald (PME) algorithm was applied for electrostatic interactions with a nonbonded interaction at 1.2 nm and a cutoff at 1.2 nm. The linear constraint solver (LINCS) algorithm was used to constrain the bond lengths. The system temperature was maintained at 310 K by a Nose-Hoover thermostat, corresponding to body temperature. The pressure was maintained semi-isotropically at 1 bar on the x- and y-axes by using a Parrinello-Rahman barostat, and the time step for simulation was 2 fs. A 20 ns equilibration process was applied under both constant volume (NVT) and constant pressure (NPT). Then, both peptides were simulated using a 1  $\mu$ s MD simulation.

**5) Umbrella sampling and the potential of mean force (PMF).** The PMF of the two peptide systems was calculated by using the method described by Zhu et al [44, 45]. For all of the pentameric peptide systems, the H<sub>2</sub>O molecules were located in the cytoplasm bulk water region. The PMFs of water molecules were calculated from the two peptide models. One H<sub>2</sub>O molecule was pulled smoothly along the z-coordinate axis (the bilayer orientation was normalized to the z-axis), which corresponds to the direction from the N- to the C-terminal. An approximate 5 ns simulation was performed to obtain snapshots of the umbrella sampling windows. Seventy umbrella sampling windows for H<sub>2</sub>O molecules at an interval of 0.1 nm along the z-coordinate axis were produced to calculate the PMFs from 3.5 to -3.5 nm (0 nm indicates the center of the membrane thickness). The H<sub>2</sub>O molecules were harmonically restrained by a force constant of 3000 kJ·mol/nm<sup>2</sup> on the z-axis. The simulation time for each window was 3 ns; the first 1 ns was used for system equilibration, and 2 ns was applied for analysis. The weighted histogram analysis method (WHAM) was used to compute the PMFs. The profile was generated by the GROMACS protocol 'g\_wham' [46]. Bootstrap analysis (N = 50) was used to estimate the statistical error.

Both VMD 1.9.3 and PyMOL 2.2 software packages [47, 48] were used to visualize all simulation processes. The GROMACS 5.05 software package [49] was used on the NVIDIA CUDA acceleration workstation and the Tianhe Supercomputer (Tianhe I HPC) at the National Supercomputer Center in Tianjin with Xeon processors. The total simulation time was  $\sim 4 \mu\text{s}$ .

### 3. Results

#### 3.1 ceRNA network and hub lncRNA selection

**1) Identification of DEmRNAs, DE miRNAs, and DE lncRNAs.** We analyzed lncRNA expression profiles in glioma patient tissues ( $n = 700$ ) and adjacent nontumor tissues ( $n = 5$ ). There were 420 downregulated and 128 upregulated lncRNAs within 548 aberrantly expressed lncRNAs (absolute FC  $> 2$ ,  $P < 0.01$ ; the volcano plot is shown in [S1 Fig](#)) as well as 2370 downregulated and 1615 upregulated mRNAs within 3985 aberrantly expressed mRNAs. Filtering analysis with the same criteria also identified 62 miRNAs differentially expressed between glioma and normal tissues with 20 downregulated miRNAs and 42 upregulated miRNAs (the full list of the RNAs is included in [S1–S3 Files](#)).

**2) Glioma ceRNA network.** Based on the ceRNA theory, the potential interactions among the above dysregulated genes were predicted to understand the DE lncRNA functions by bioinformatics analysis. Eight specific DE miRNAs were identified to interact with 32 DE lncRNAs through miRNA response elements. The miRTarBase and StarBase databases [29, 49, 50] were used to predict the candidate mRNA targets of the 6 DE miRNAs and 41 mRNAs within the 3985 DEmRNAs that were used to build the ceRNA network. The ceRNA network involved 32 lncRNAs, 8 miRNAs, and 41 mRNAs. [Fig 1](#) shows the ceRNA network that was visualized using Cytoscape.

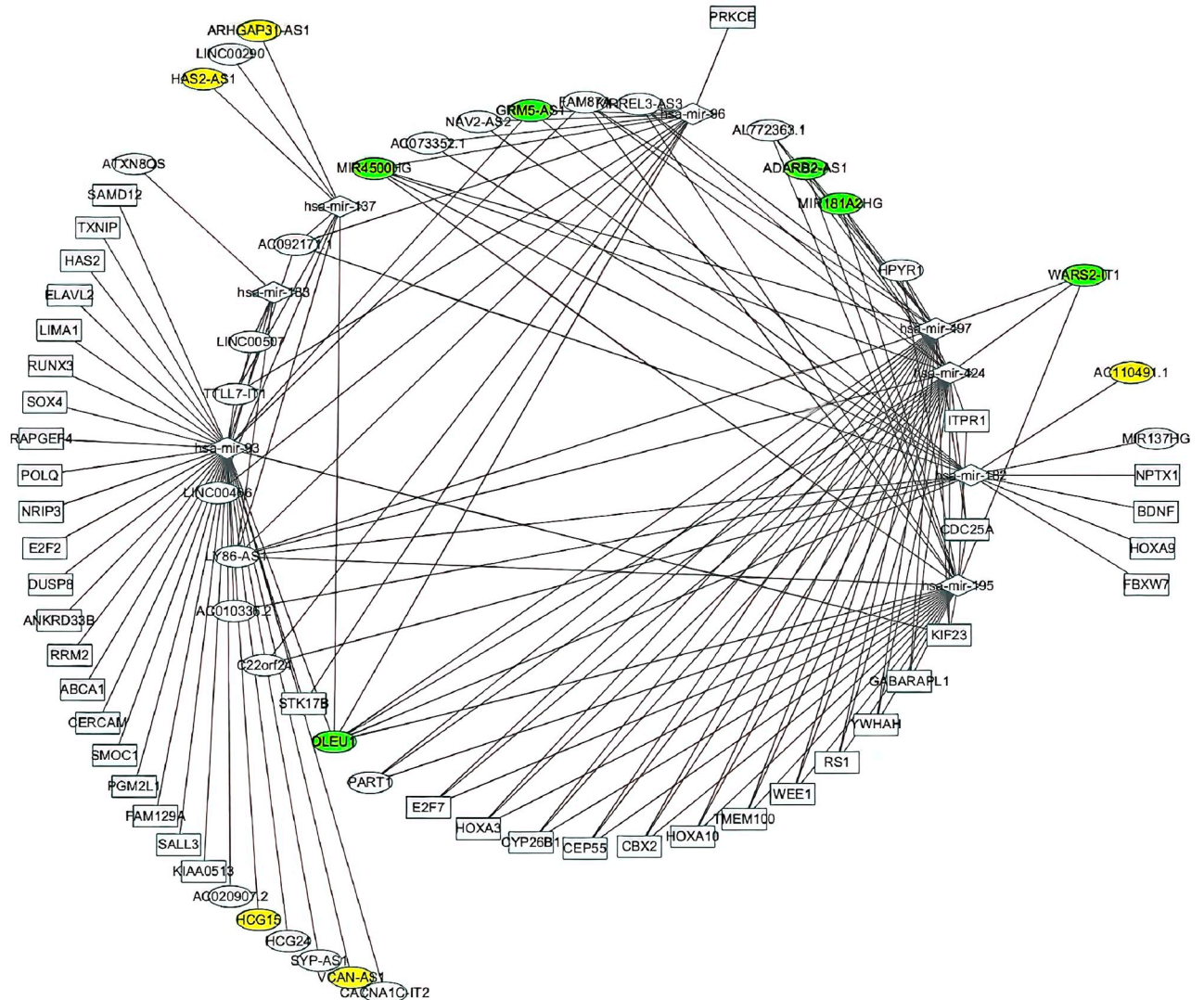
**3) DE lncRNAs and their associated clinical features.** In analyzing the association between DE lncRNAs and glioma patient survival periods, all DE lncRNAs in the ceRNA network were chosen according to the bioinformatics analysis. The significant effects on survival were evaluated with  $P < 0.05$  to identify the key hub lncRNAs with prognostic characteristics. In the ceRNA network, 11 lncRNAs ([Table 1](#)) were found to be associated with the overall survival of patients with glioma by univariate Cox regression analysis.

Node lncRNAs were defined as edges of the ceRNA network  $\geq 3$ . Six node lncRNAs (MIR4500HG, GRM5-AS1 ([Fig 1](#)), ADARB2-AS1, MIR181A2HG, WARS2-IT1, and DLEU1) were identified. Kaplan-Meier survival curves indicated that four node lncRNAs (MIR181A2HG, MIR4500HG, GRM5-AS1, and ADARB2-AS1) positively correlated with overall survival but that two node lncRNAs (IT1 and DLEU1) were negatively associated with overall survival. The Kaplan-Meier survival curves of the 11 lncRNAs are shown in [S4 File](#).

**4) Functional analysis of lncRNAs negatively associated with overall survival.** The functional analysis revealed a negative correlation with overall survival. Two lncRNAs, DLEU1 and IT1, in the above ceRNA network were enriched in 103 GO biological process categories. The GO biological processes of the above two lncRNAs were involved in ion channel activity (GO:0005216) and ion gated channel activity (GO:0022839). This finding suggested that high expression of these two lncRNAs leads to a low survival rate in glioma patients, probably due to their abnormal biological processes involving ion or water permeability, which was also our main focus in this study.

#### 3.2 ORF and TM region identification

Using the methods described in the Methods section, the candidate ORFs in hub lncRNAs were searched by ORF FINDER online software. We found that DLEU1 and IT1 contain ORFs. The ORFs were translated into amino acid sequences and analyzed by TMHMM 2.0 to



**Fig 1. The ceRNA network.** The colored circles represent key lncRNAs (with both green and yellow) and key hub lncRNAs (with green) ( $n \geq 3$ ).

<https://doi.org/10.1371/journal.pone.0248634.g001>

**Table 1. P-values of the lncRNAs associated with the survival rate of glioma patients.**

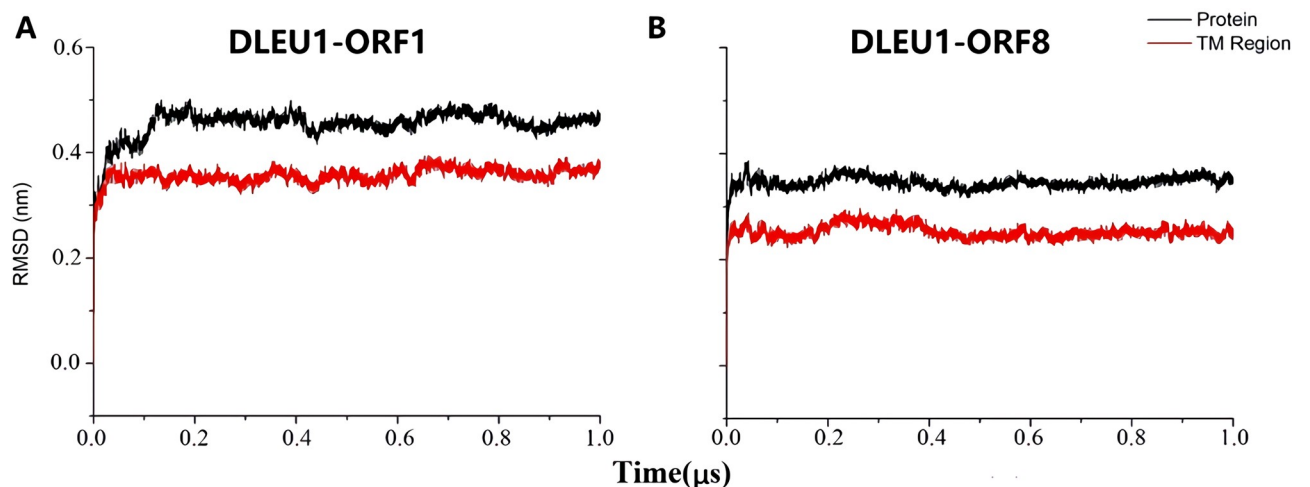
lncRNA	P-value
VCAN.AS1	9.99E-16
ARHGAP31.AS1	3.00E-15
AC110491.1	9.85E-14
MIR4500HG	2.04E-09
HCG15	5.43E-05
DLEU1	6.19E-05
IT1	3.02E-04
HAS2.AS1	6.77E-04
GRM5.AS1	1.14E-03
ADARB2.AS1	3.05E-02
MIR181A2HG	4.69E-02

<https://doi.org/10.1371/journal.pone.0248634.t001>

predict the TM regions. The CPC and CPC2 web tools were used to assess the quality of the lncRNA ORFs. Then, the BLAST tool in NCBI was used to compare the homology of DLEU1 and IT1 to determine whether they are conserved between different species. The following criteria were used: 1) the sequence must be a continuous ORF in a lncRNA; 2) the amino acid sequence, including the TM region, has to achieve a 100% probability (S3 Fig); 3) the potential TM region has to be conserved and compared with those in other species by multiple sequence alignment in the NCBI database; 4) given the length of the identified small peptides (SLN, PLB, MLN, and DWORF), the number of amino acids in the predicted peptide must be  $\leq 45$ ; 5) the ORF quality must be  $> 95\%$  as assessed by both CPC and CPC2; and 6) to ensure reliable results, we repeated the above steps using the known lncRNA-encoded small TM peptides, namely, MLN, DWORF, SLN, and PLB. Through these steps, we identified two possible TM peptides, encoded by ORF1 and ORF8, in the lncRNA DLEU1 according to the conditions above. These peptides were suitable for subsequent functional study, and the results are shown in S2–S4 Figs.

### 3.3 Model construction and MD simulation

Our results showed that the lncRNA DLEU1 was closely related to tumor proliferation and migration, while IT1 had no candidate peptides encoding ORFs. The subsequent molecular modeling and MD simulation focused on the two ORFs of DLEU1, which contained the candidates of the small TM peptides. We built the secondary structure of the ORFs using the I-TASSER online software package based on the ab initio algorithm. Five candidate small peptides were predicted for each ORF. The peptides chosen for MD simulation had higher C-Scores, and the TM region appeared to have an  $\alpha$ -helical configuration. The modeled peptide was embedded in a pre-equilibrated POPC membrane, and a 50 ns MD simulation for optimization of the pentameric assemblies was then performed until the RMSD became smooth. The optimized structure was then extracted. Homology modeling for the pentameric candidates of the small peptides was performed as we described previously for the SLN and PLB pentamer structures. Two small peptide pentamers were embedded into the membrane, and 1  $\mu$ s (1000 ns) of a MD simulation was then performed for optimization. The stability of the pentameric assemblies during the optimization process was then evaluated using the C- $\alpha$ 's RMSD and the Rampage online software. As shown in Fig 2A and 2B, after approximately 100 ns and 30 ns,



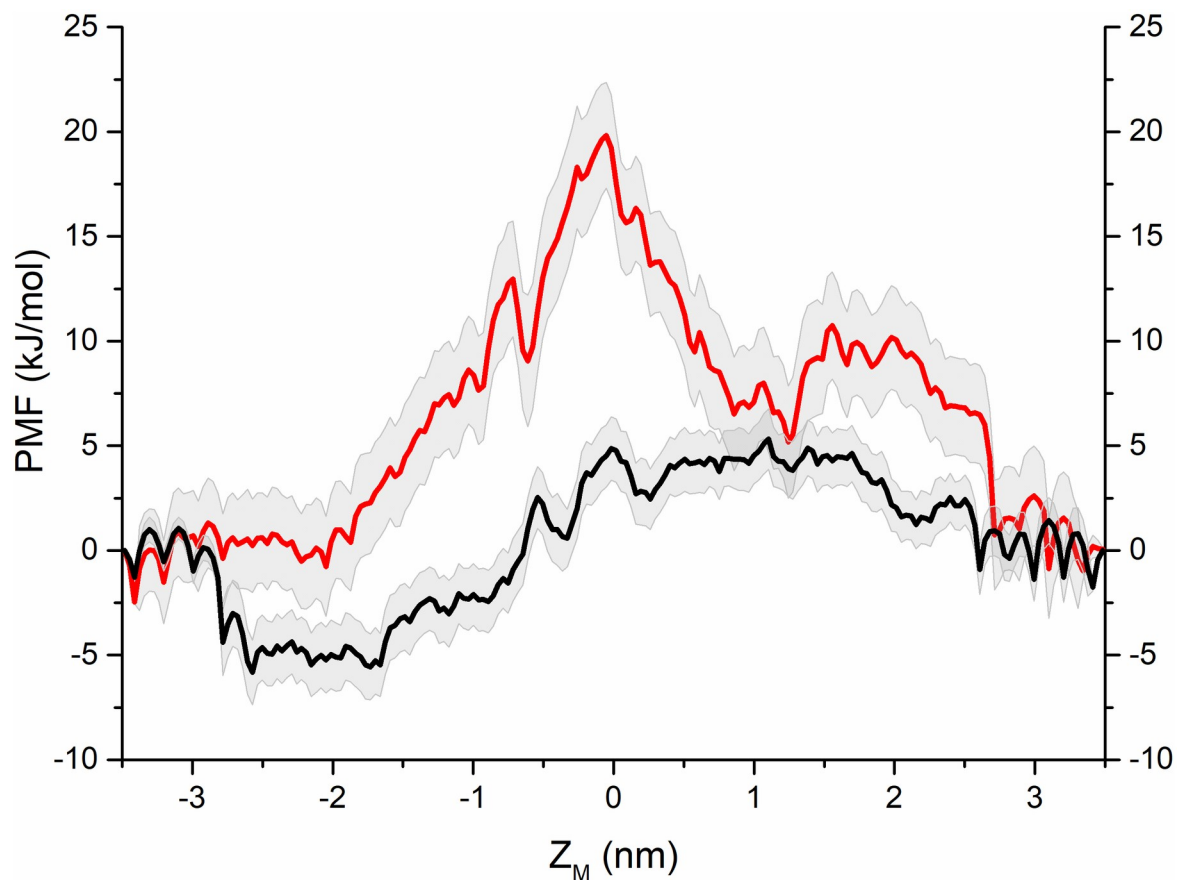
**Fig 2. Structural stability of pentameric DLEU1 ORF1 (A) and ORF8 (B) peptides.** The black and red solid curves represent the root mean square deviation (RMSD) of the whole protein and TM region, respectively.

<https://doi.org/10.1371/journal.pone.0248634.g002>

the RMSD curves for ORF1 and ORF8 were flattened, respectively, indicating that the two models were stabilized within 1  $\mu$ s (Fig 2). The RMSD of both peptide pentameric assemblies showed that the RMSD of the TM region was  $\sim$ 0.1 nm less than that of the backbones, indicating that the TMs have a higher stability. Comparison of the two pentamers showed that ORF1 had a more flexible backbone than ORF8 and that the RMSD of the ORF8 pentamer was  $\sim$ 0.1 nm less than that of the ORF1 pentamer.

### 3.4 Computation of H<sub>2</sub>O and ion permeability

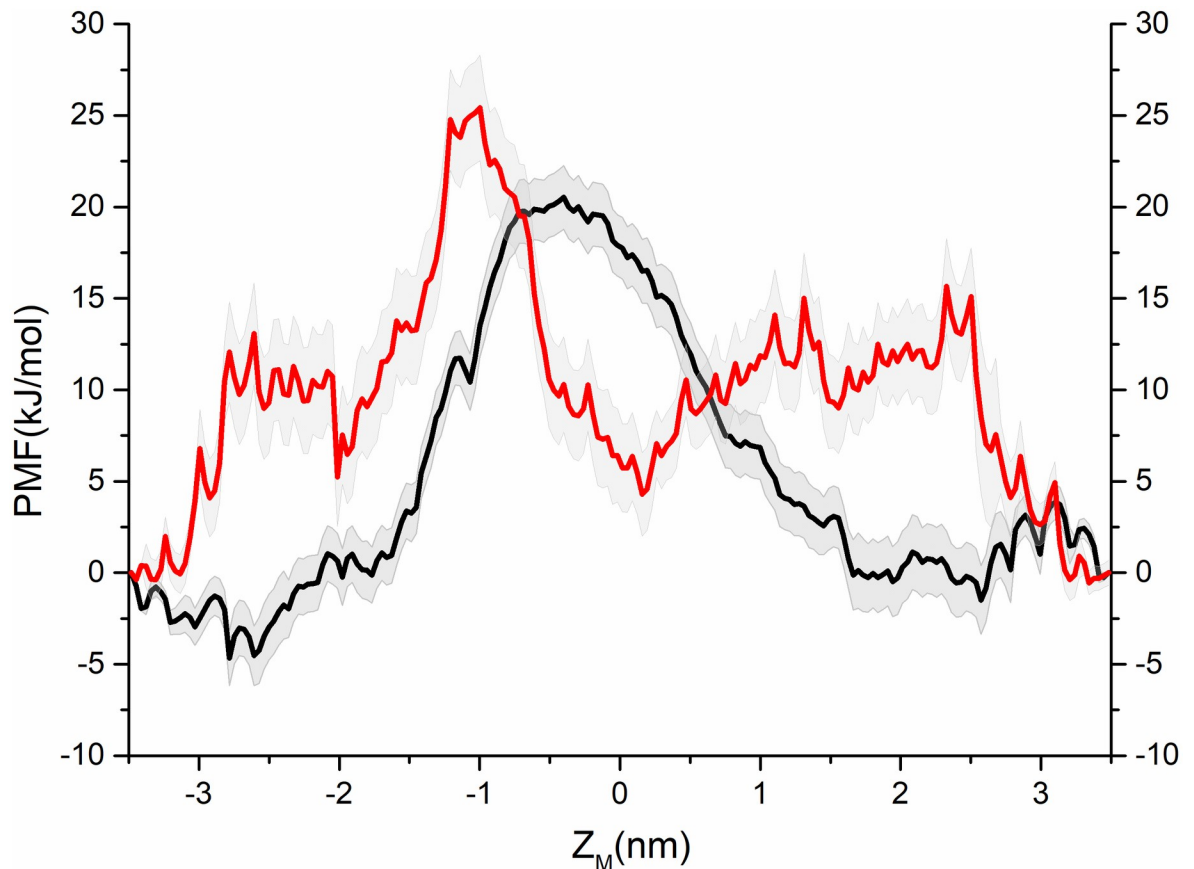
To quantify the energy barrier for an individual H<sub>2</sub>O molecule through the DLEU1 candidate peptide channels, free energy profiles were used to identify the individual water molecular permeation from the cytoplasm (-z valued side) to the lumen (+z valued side) region of each pentamer. Fig 3 shows the PMF of H<sub>2</sub>O permeating through the two pores. The z-coordinate was defined as a function of the length of the pore. A H<sub>2</sub>O molecule was pulled into the pentamer channel, and the energy barrier (PMF) decreased sharply from 0 to -5 kJ/mol between -3 and -2.5 nm along the z-coordinate. Then, the curve ascended slowly from -2 to 0 nm. After a stable PMF barrier from 0 to 1.5 nm, there was another descent. These profiles are attributed to the energetic cost of moving H<sub>2</sub>O through the pore. The maximum free energy barrier for H<sub>2</sub>O crossing the ORF1 pentamer pore was  $\sim$ 5 kJ/mol at  $Z_M = 0$  to 1.5 nm. For the ORF8 pentamer, when a H<sub>2</sub>O molecule permeated through the pore from -3.5 to 0 nm, the energy cost rapidly increased from 0 to 20 kJ/mol, which corresponded to the maximum energy barrier



**Fig 3. PMF associated with water permeation through the channel pore of ORF1 (black curve) and ORF8 (red curve) pentamer.**

<https://doi.org/10.1371/journal.pone.0248634.g003>





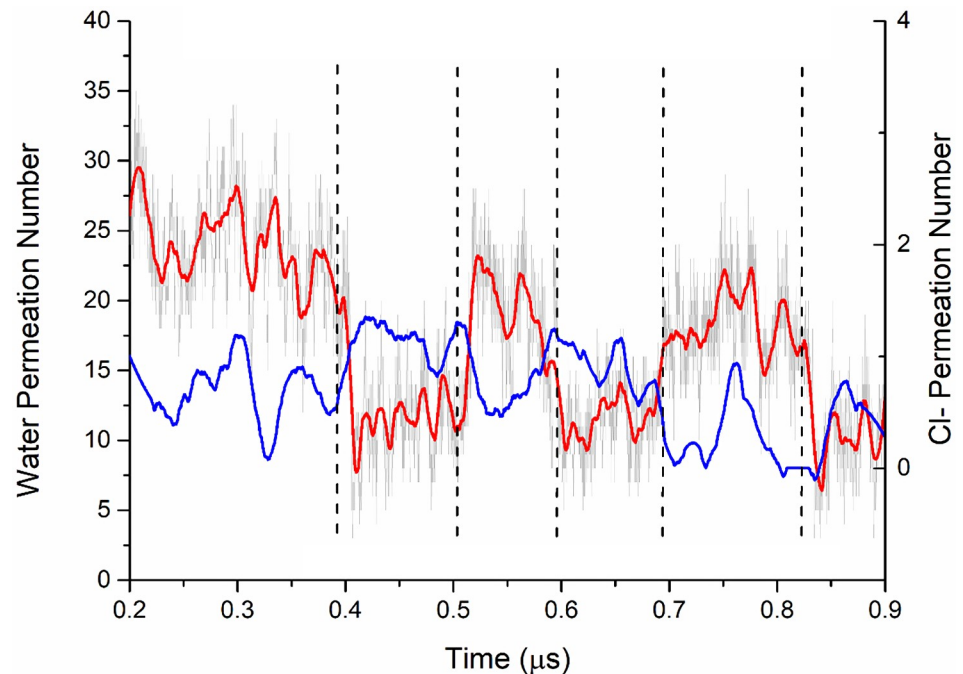
**Fig 4. PMF associated with Na<sup>+</sup> (black curve) and Cl<sup>-</sup> (red curve) permeation through the channel pore of the ORF1 pentamer.**

<https://doi.org/10.1371/journal.pone.0248634.g004>

where three leucine residues were placed. The location of the leucine residue showed the strongest hydrophobicity of the pore. Subsequently, the free energy profile dropped rapidly from 0 to 3 nm and eventually to zero. The simulation results indicated that the H<sub>2</sub>O molecule penetrated more easily through the pentamer pore of DLEU1 ORF1 than that of ORF8. The lower energy barrier for water transportation from the cytoplasm to the lumen suggested that the H<sub>2</sub>O molecules had a favorable flow in the ORF1 pore. The amino acid residues of ORF1 and ORF8 at  $Z_M = 0$  and  $-1$  nm are serine and leucine, respectively, indicating that the serine and leucine side chains have a greater influence on water permeability.

Due to the substantial energy barrier difference in water permeation between the pentamer pores of ORF1 and ORF8, it is likely that a gate will also be impermeable to ions if there is an energetic barrier to water permeation. Therefore, we focused on the ability of the ORF1 pore to permeate ions. To determine whether the ORF1 pentamer plays a role as an ion channel, the PMFs of Na<sup>+</sup> and Cl<sup>-</sup> through the ORF1 pore were calculated. Both Na<sup>+</sup> and Cl<sup>-</sup> had to overcome a large energy barrier in the pore. As shown in Fig 4, the maximum energy barrier reached 20 kJ/mol at  $Z_M = 0$  nm for Na<sup>+</sup>. The maximum energy barrier of Cl<sup>-</sup> was 25 kJ/mol at  $Z_M = -1$  nm, which was slightly larger than that of Na<sup>+</sup>. The energy barriers of Na<sup>+</sup> and Cl<sup>-</sup> revealed that the ions were not allowed to freely permeate through the ORF1 pentameric channel.

The number of water molecules in the ORF1 pore showed complete water permeation and periodic variation during the 1- $\mu$ s MD simulation (see Fig 5, red and gray curve). The quantity



**Fig 5. The change in ORF1 with water and chloride ions during the 1  $\mu$ s simulation.** The gray curve represents the absolute quantity of water molecules. The red and blue curves represent the smooth line of water and Cl<sup>-</sup>, respectively (200 pts).

<https://doi.org/10.1371/journal.pone.0248634.g005>

of water molecules changed between  $\sim 7$  and  $\sim 20$  kJ/mol at  $Z_M = -1-1$  nm, and the time interval was  $\sim 0.1$   $\mu$ s. The permeated number of Cl<sup>-</sup> ions also fluctuated periodically as water molecules changed at the 0–2 scale, which indicated that the ORF1 pore presented a higher energy barrier for Cl<sup>-</sup> permeation. This result was consistent with the PMF profiles. Both the PMF and water molecular changes indicated that the ORF1 channel mainly functioned as a water channel rather than as an ion channel.

#### 4. Discussion

Glioma is the most common intracranial tumor and is characterized by various abnormal gene and protein regulatory events that drive tumorigenesis [51]. Currently, many efforts have been made to elucidate the role of lncRNAs and the related membrane proteins involved in tumorigenesis, progression, and cerebral edema [52, 53]. According to a previous study, lncRNAs can participate in the regulation of cellular ions and water by encoding small TM peptides [54–56], but the mechanisms of small TM peptides is still unknown.

There are no accurate methods for predicting lncRNA-encoded small peptides. In this study, ceRNA network, survival, and GO enrichment analyses were conducted to identify hub lncRNAs with channel functions. Subsequently, multiple bioinformatics methods were used to predict and validate the small TM peptides encoded in ORFs of lncRNAs in glioma. Finally, the ORFs were translated into amino acid sequences. A combination of ab initio modeling, MD simulations and other methods was used to determine the pore characteristics of the pentamers.

To establish the ceRNA network among mRNAs, lncRNAs, and miRNAs in gliomas, 705 sequencing data points in the TCGA database from a cohort of tissue samples (700 cancer tissues and 5 adjacent tissues) were characterized to establish a ceRNA network. We identified a

total of 4,595 aberrantly expressed lncRNAs, mRNAs, and miRNAs, and we successfully constructed a prognosis- and lncRNA-related ceRNA network in gliomas by biological prediction. Eleven lncRNAs with significant prognostic differences were detected (Table 1). Two of the lncRNAs were negatively correlated with the potential prognosis of glioma patients. The high expression of DLEU1 and IT1 significantly reduces the survival rate of glioma patients. In subsequent GO enrichment analysis, we found that the functions of the two lncRNAs were related to channel activity, especially in particle transport and TM transport [57]. This finding indicated that their effect on prognosis may be related to the dynamics of intracellular ions and water, but there is no evidence that the lncRNAs are directly involved in cellular ion transport. Therefore, the ORFs of the lncRNAs were identified by a bioinformatics method and could be translated into small TM peptides.

Many studies have shown that the permeability and movement of water and ions play a key role in regulating the migration/invasion of cancer cells [58, 59], which has an important relationship with the cerebral edema caused by glioma. The movement of water may accelerate tumor migration, making tumors more prone to metastasis [20]. Although water infiltration is often regulated by the classical AQP water channel family [60], lncRNA-encoded small TM peptide oligomers, such as PLB and SLN, may also play an important role. The water permeability of the small TM peptide oligomer channel works in the same manner as PLB pentamers in muscle cells. To further confirm the potential permeability of these peptide oligomer functions in gliomas, we chose DLEU1 for subsequent simulations. DLEU1 activates KPNA3, which is related to the increased proliferation and migration of cancer cells [61]. Some reports have indicated that DLEU1 increases cell proliferation. Our results showed that ORF1 and ORF8 in DLEU1 encode small TM peptides. Through molecular modeling, MD simulations, PMF computation, and structural analysis, we found that there were different water and ion behaviors between the ORF1 and ORF8 pentamers.

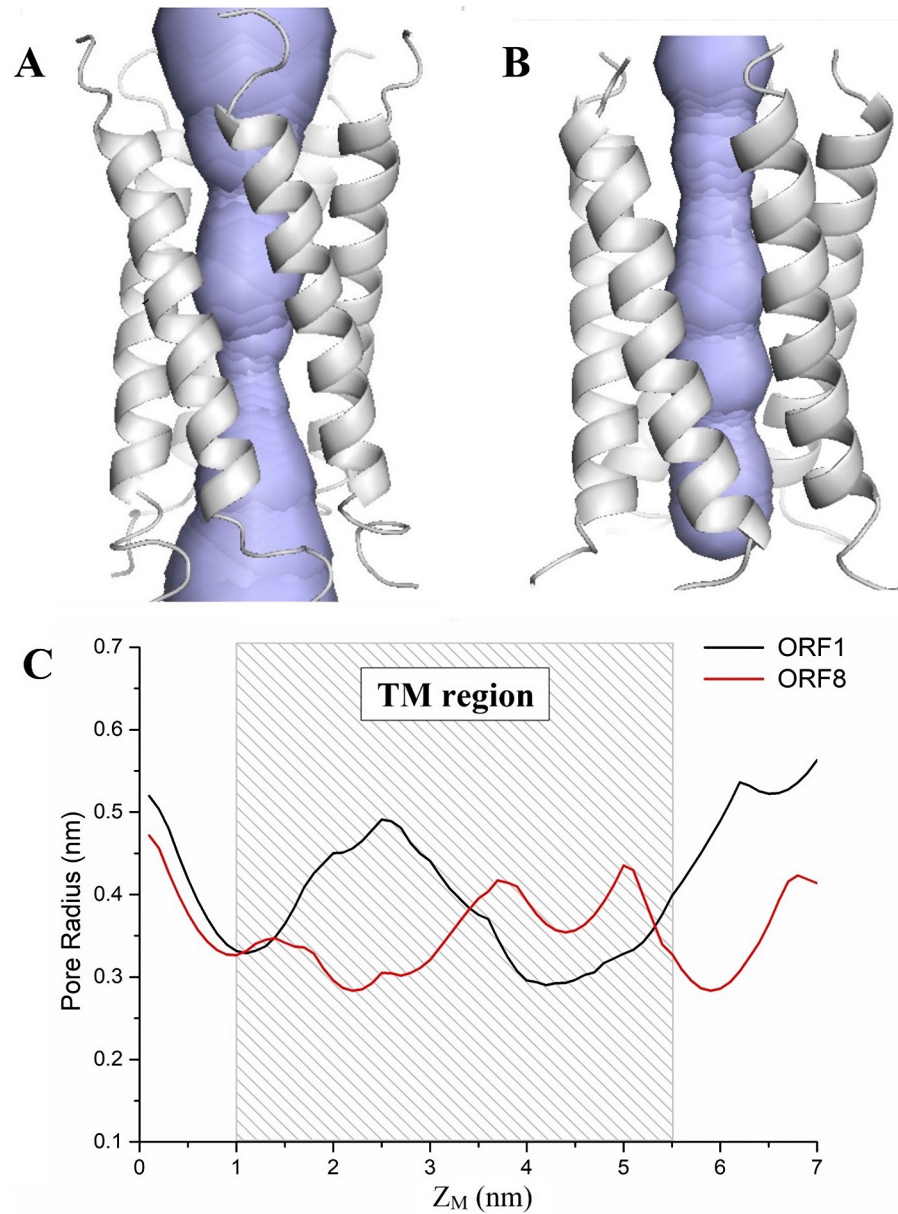
The 1  $\mu$ s MD simulation results of the two peptide pentamers (ORF1 and ORF8) showed that the RMSD was stabilized within 0.1  $\mu$ s. The RMSD of the whole protein and TM region fluctuations were maintained within 0.5 nm. Both peptide RMSD and simulation trajectories showed that there were no structural fractures or depolymerizations, indicating that the peptide pentamer channels remained stable in the simulation environment. This result was consistent with our previous simulation results for the SLN and PLB pentamers [62]. Comparison of the ORF1 and ORF8 pentamers showed that the RMSD of the ORF1-encoded peptide pentamer was  $\sim$ 0.1 nm higher than that of the ORF8-encoded peptide pentamer, suggesting that the ORF1 pentamer was more flexible than the ORF8 pentamer.

The water permeability of the ORF1 pentamer channel (black curve in Fig 3) can be characterized as a three-stage process as follows: i) the spontaneous water molecule entry stage with -5 kJ/mol free energy; ii) the overcoming stage of an  $\sim$ 8 kJ/mol energy barrier at the central region of the channel pore at approximately  $Z_M = -2-0$  nm; and iii) the overcoming stage of the decreasing energy barrier from  $\sim$ 3 kJ/mol. The energy barrier of the ORF8 channel was completely different from that of the ORF1 channel. When water molecules entered the ORF8 pores, the energy barrier steeply increased until the maximum energy barrier was  $\sim$ 20 kJ/mol, which was nearly 15 kJ/mol higher than that of ORF1. Then, the curve steeply decreased, exhibiting a flattened barrier region until the water molecule permeated through the pore. Thus, water molecules spontaneously permeated through the ORF1 pentamer channel pores but not through the ORF8 pentamer channel due to the higher-energy barrier. The count diagram (Fig 5) of water molecules showed that the water permeation through the ORF1 channel was a spontaneous process. The quantity variation in water molecules showed an obvious hydrated-dehydrated alteration, suggesting that the pentameric ORF1 channel is a hydrophobic channel. However, the counting diagram of Cl<sup>-</sup> showed no spontaneous TM activity

through the pore. Interestingly, the distribution of water and  $\text{Cl}^-$  along the channel pore showed competing relationships with each other as  $\text{Cl}^-$  ion occupation in the pore blocked water permeation. To further validate the ionic function in the ORF1 pentamer, we calculated the PMF of the  $\text{Na}^+$  and  $\text{Cl}^-$  permeation through the ORF1 pentameric channel. The PMF results suggested that the highest energy barriers for  $\text{Na}^+$  and  $\text{Cl}^-$  were 20 kJ/mol and 25 kJ/mol, respectively. The MD trajectories showed that neither  $\text{Na}^+$  nor  $\text{Cl}^-$  passed through the pore. Intriguingly,  $\text{Na}^+$  ions exhibited Brownian motion in water without entering the pore, but  $\text{Cl}^-$  ions appeared at the entrance of the pore, appearing to be ‘stuck’ at leucine amino acids and blocked by the channel. This finding was consistent with the results of the PMF and water distribution. The ORF1 pentamer channel showed water transport capacity but no  $\text{Cl}^-$  transport process. According to the permeation of water and ion simulation results, we reduced the energy barrier for water and ion transport by placing the charge group at the mouth of the channel pore, such as that in the ORF1 peptide. This finding was also consistent with previous studies on water transport through narrow pores or nanopores [63–65]. In addition, the hydrogen bond distribution between the water molecules and amino acids inside the pores of the ORF1 and ORF8 pentamers were different. The water molecules in the ORF1 pore had fewer hydrogen bonds than those in the ORF8 pore, which may have led to the difference in water permeation.

To reveal the difference in the water and ionic permeability of the small peptide pentamer channels of the two ORFs, the radii of the two pores were calculated as shown in Fig 6. These two ORF channels had similar pore radius profiles. A minimum radius of  $\sim 0.3$  nm was larger than the van der Waals radius of the water molecule (0.19 nm) as shown in Fig 6C. However, the energy barriers of water permeation were different between the ORF1 and ORF8 pentamers (Fig 3). The structure and sequence of the small peptides of the two ORFs were analyzed. The results showed that the water permeability difference between the two ORF pentamers was mainly affected by three factors. (1) Although both the ORF1 and ORF8 small peptides have a valine/leucine/isoleucine zipper domain, ORF1 lacks one valine/leucine/isoleucine residue at the TM region (Fig 7B), which weakens the intermolecular interaction among the small peptides of the ORF1 pentamer in the formation of the pentameric assembly. Consequently, these weakened interactions lead to a more flexible pentameric structure and easier permeation of water molecules. (2) There are four hydrophilic amino acids in ORF1 located in the N-terminus of the peptide (cytoplasmic region), but these amino acids in ORF8 are located in the C-terminus (lumen region), resulting in a significant absolute energy difference between the water permeation of the ORF1 and ORF8 pores. (3) There are two positively charged groups (Arg2 and Arg3) at the N-terminus of the ORF1 peptides (Fig 7A). These electric charges in arginine attract water molecules, supporting the explanation of water accessibility on the N-terminus of the pentamer. Some water molecules spontaneously enter into the pentamer channel pore, while the  $\text{Cl}^-$  ions are blocked from the pore. In contrast, the N-terminus of ORF8 has no charged amino acid residues, which decreases the amount of water molecule aggregation, reducing the probability of spontaneous permeability.

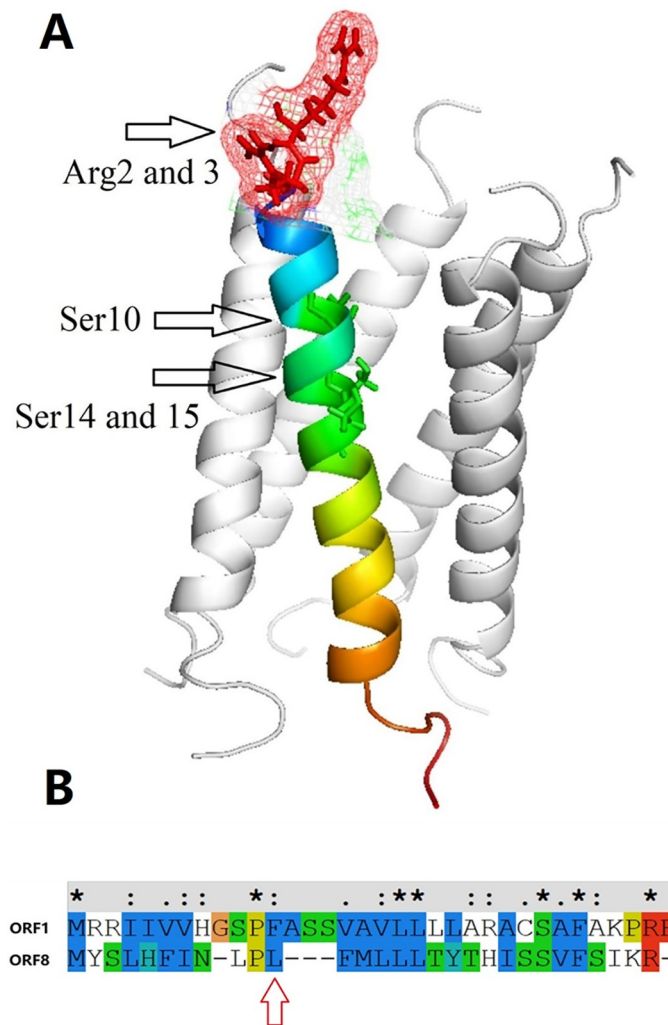
As some ORFs in lncRNAs can encode small peptides with biological functions [18], our results suggested that the ORFs of the hub lncRNAs with survival differences, including ORFs encoding small TM peptides, may affect the prognosis of glioma patients. The ORF1-encoded peptides can interact with each other on the membrane to form stable oligomeric hydrophobic channel pores. The ORF channel enhanced water influx, leading to water and ion leakage. The leaked ions and water would affect the microenvironment of the cell or interact with the known water and ion channel permeation. The intracellular and extracellular water and ion exchange causes tumor cell proliferation and migration, aggravating the



**Fig 6.** Channel structures in the ORF1 (A) and ORF8 (B) with the pore's lining surfaces. The variation in the inner pore radius of ORF1 (black) and ORF8 (red) is shown in (C). The  $Z_M$  shows the distance from the center of mass, and the rectangle with a sparse pattern in (C) is the TM region of the pore. The calculated channel was measured using CAVER 3.01 [66].

<https://doi.org/10.1371/journal.pone.0248634.g006>

occurrence of cerebral edema. Many small TM peptides have a general hydrophobicity that allows water and ions to influx. In this study, the simulations of the ORF1 and ORF8 pentamer channels showed significant differences in water and ion permeability, indicating that the hydrophobicity of the lncRNA-encoded small peptides was different. In addition, this difference explained why lncRNAs had different effects on cancer prognosis even though they had similar ORFs that could encode small TM peptides. The survival differences caused by lncRNA-encoded small TM peptides are worthy of further study and corresponding experimental confirmation.



**Fig 7.** (A) Pentameric ORF1 channel. The arrows indicate two charged amino acid groups (Arg2 and 3) and three hydrophilic groups (Ser10, 14, and 15). (B) Sequence alignment of ORF1 and ORF8. One leucine is present in ORF8 but not in ORF1.

<https://doi.org/10.1371/journal.pone.0248634.g007>

## 5. Conclusion

A model was built from a glioma ceRNA network to screen for differentially expressed hub lncRNAs and to identify ORF fragments encoding small TM peptides by bioinformatics methods and ab initio homology modeling combined with MD simulation. Our established models revealed that the small peptides of the two encodable ORFs (ORF1 and ORF8) of the DLEU1 lncRNA could form pentameric channels. The simulations of the water behavior along the pentameric channel showed that the ORF1-encoded channel was possibly a hydrophobic gate. These findings were consistent with our two previous studies, which suggested that lncRNAs encode the small peptide pentameric structures of SLN and PLB. We proposed that DLEU1 can encode small peptides with ion channel activity, leading to an increase in glioma permeability, which in turn increases brain edema and even increases the risk of cancer cell development, invasion and metastasis. This mechanism could be used for prospective studies on the functions of lncRNAs encoding small TM peptides and for the development of new glioma treatment strategies.

## Supporting information

**S1 Fig. The volcano of lncRNA, mRNA and miRNA.** The red and green dots represent up-regulate and downregulated RNAs respectively.

(DOCX)

**S2 Fig. The predicted ORFs in WARS2.IT1 and DLEU1 lncRNA.** The DLEU1 contain ORF1 and ORF8 can encoding small peptides.

(DOCX)

**S3 Fig. Two ORFs second structure, the red lines represent a predicted transmembrane  $\alpha$ -helix configuration.**

(DOCX)

**S4 Fig. Comparison of two ORFs sequence conservation of DLEU1.**

(DOCX)

**S1 File.**

(CSV)

**S2 File.**

(CSV)

**S3 File.**

(CSV)

**S4 File.**

(DOCX)

## Acknowledgments

We thank the National Supercomputer Center in Tianjin (Tianhe HPC) for providing computing power and support.

## Author Contributions

**Conceptualization:** Yipeng Cao.

**Data curation:** Yipeng Cao, Rui Yang, Wenwen Zhang.

**Formal analysis:** Yipeng Cao, Rui Yang, Wenwen Zhang, Jiana Sun.

**Funding acquisition:** Yipeng Cao, Wei Wang.

**Methodology:** Yipeng Cao, Imshik Lee, Jiana Sun.

**Resources:** Xiangfei Meng.

**Software:** Xiangfei Meng.

**Writing – original draft:** Yipeng Cao.

**Writing – review & editing:** Imshik Lee, Wei Wang.

## References

1. Omuro A, DeAngelis LM. Glioblastoma and other malignant gliomas: a clinical review. *Jama*. 2013; 310: 1842–1850. <https://doi.org/10.1001/jama.2013.280319> PMID: 24193082
2. Altieri R, Agnoletti A, Quattrucci F, Garbossa D, Specchia FMC, Bozzaro M, et al. Molecular biology of gliomas: present and future challenges. *Transl Med UniSa*. 2014; 10: 29–37. PMID: 25147764

3. Peng Z, Liu C, Wu M. New insights into long noncoding RNAs and their roles in glioma. *Mol Cancer*. 2018; 17: 61. <https://doi.org/10.1186/s12943-018-0812-2> PMID: 29458374
4. Xing Z, Lin A, Li C, Liang K, Wang S, Liu Y, et al. lncRNA Directs Cooperative Epigenetic Regulation Downstream of Chemokine Signals. *Cell*. 2014; 159: 1110–1125. <https://doi.org/10.1016/j.cell.2014.10.013> PMID: 25416949
5. Luo Y, Liang M, Yao W, Liu J, Niu Q, Chen J, et al. Functional role of lncRNA LOC101927497 in N-methyl-N'-nitro-N-nitrosoguanidine-induced malignantly transformed human gastric epithelial cells. *Life Sci*. 2018; 193: 93–103. <https://doi.org/10.1016/j.lfs.2017.12.007> PMID: 29223541
6. Guttman M, Russell P, Ingolia NT, Weissman JS, Lander ES. Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell*. 2013; 154: 240–251. <https://doi.org/10.1016/j.cell.2013.06.009> PMID: 23810193
7. Kung JTY, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics*. 2013; 193: 651–669. <https://doi.org/10.1534/genetics.112.146704> PMID: 23463798
8. Jing SY, Lu YY, Yang JK, Deng WY, Zhou Q, Jiao BH. Expression of long non-coding RNA CRNDE in glioma and its correlation with tumor progression and patient survival. *Epilepsia*. 2016; 20: 3992–3996. PMID: 27775801
9. Wu F, Zhang C, Cai J, Yang F, Jiang T. Upregulation of long noncoding RNA HOXA-AS3 promotes tumor progression and predicts poor prognosis in glioma. *Oncotarget*. 2017; 8: 53110–53123. <https://doi.org/10.18632/oncotarget.18162> PMID: 28881797
10. Chen X, Sun Y-Z, Guan N-N, Qu J, Huang Z-A, Zhu Z-X, et al. Computational models for lncRNA function prediction and functional similarity calculation. *Brief Funct Genomics*. 2018; 18: 58–82.
11. Andrews SJ, Rothnagel JA. Emerging evidence for functional peptides encoded by short open reading frames. *Nat Rev Genet*. 2014; 15: 193–204. <https://doi.org/10.1038/nrg3520> PMID: 24514441
12. Choi S-W, Kim H-W, Nam J-W. The small peptide world in long noncoding RNAs. *Brief Bioinform*. 2019; 20: 1853–1864. <https://doi.org/10.1093/bib/bby055> PMID: 30010717
13. Huang JZ, Chen M, Chen, Gao XC, Zhu S, Huang H, et al. A Peptide Encoded by a Putative lncRNA HOXB-AS3 Suppresses Colon Cancer Growth. *Mol Cell*. 2017; 68: 171–184. <https://doi.org/10.1016/j.molcel.2017.09.015> PMID: 28985503
14. Anderson D, Anderson K, Chang CL, Makarewicz C, Nelson B, Mcanally J, et al. A Micropeptide Encoded by a Putative Long Noncoding RNA Regulates Muscle Performance. *Cell*. 2015; 160: 595–606. <https://doi.org/10.1016/j.cell.2015.01.009> PMID: 25640239
15. Nelson BR, Makarewicz CA, Anderson DM, Winders BR, Troupes CD, Wu F, et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science*. 2016; 351: 271–275. <https://doi.org/10.1126/science.aad4076> PMID: 26816378
16. Cao Y, Yang R, Sun J, Zhang W, Lee I, Wang W, et al. Effects of amino acid modifications on the permeability of the pentameric sarcolipin channel. *Proteins Struct Funct Bioinforma*. 2020; 1–9. <https://doi.org/10.1002/prot.26028> PMID: 33244801
17. Asahi M, Sugita Y, Kurzydowski K, De LS, Tada M, Toyoshima C, et al. Sarcolipin regulates sarco (endo)plasmic reticulum Ca<sup>2+</sup>-ATPase (SERCA) by binding to transmembrane helices alone or in association with phospholamban. *Proc Natl Acad Sci*. 2003; 100: 5040–5045. <https://doi.org/10.1073/pnas.0330962100> PMID: 12692302
18. Ji Z, Song R, Regev A, Struhl K. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *Elife*. 2015; 4: e08890. <https://doi.org/10.7554/eLife.08890> PMID: 26687005
19. Papadopoulos MC, Saadoun S, Verkman AS. Aquaporins and cell migration. *Pflügers Arch J Physiol*. 2008; 456: 693–700. <https://doi.org/10.1007/s00424-007-0357-5> PMID: 17968585
20. Stroka KM, Jiang H, Chen S-H, Tong Z, Wirtz D, Sun SX, et al. Water permeation drives tumor cell migration in confined microenvironments. *Cell*. 2014; 157: 611–623. <https://doi.org/10.1016/j.cell.2014.02.052> PMID: 24726433
21. Betz AL, Iannotti F, Hoff JT. Brain edema: a classification based on blood-brain barrier integrity. *Cerebrovasc Brain Metab Rev*. 1989; 1: 133–154. PMID: 2701373
22. Ding T, Gu F, Fu L, Ma Y-J. Aquaporin-4 in glioma invasion and an analysis of molecular mechanisms. *J Clin Neurosci*. 2010; 17: 1359–1361. <https://doi.org/10.1016/j.jocn.2010.02.014> PMID: 20685122
23. Autry JM, Rubin JE, Pietrini SD, Winters DL, Robia SL, Thomas DD. Oligomeric interactions of sarcolipin and the Ca-ATPase. *J Biol Chem*. 2011; 286: 31697–31706. <https://doi.org/10.1074/jbc.M111.246843> PMID: 21737843
24. Umans RA, Sontheimer H. Combating malignant astrocytes: Strategies mitigating tumor invasion. *Neurosci Res*. 2018; 126: 22–30. <https://doi.org/10.1016/j.neures.2017.09.010> PMID: 29054465



25. Thompson EG, Sontheimer H. A role for ion channels in perivascular glioma invasion. *Eur Biophys J*. 2016; 45: 635–648. <https://doi.org/10.1007/s00249-016-1154-x> PMID: 27424110
26. Cuddapah VA, Robel S, Watkins S, Sontheimer H. A neurocentric perspective on glioma invasion. *Nat Rev Neurosci*. 2014; 15: 455–465. <https://doi.org/10.1038/nrn3765> PMID: 24946761
27. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol*. 2015; 19: A68. <https://doi.org/10.5114/wo.2014.47136> PMID: 25691825
28. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; 26: 139–140. <https://doi.org/10.1093/bioinformatics/btp616> PMID: 19910308
29. Consortium GO. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*. 2004; 32: D258–D261. <https://doi.org/10.1093/nar/gkh036> PMID: 14681407
30. Li J-H, Liu S, Zhou H, Qu L-H, Yang J-H. starBase v2. 0: decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 2013; 42: D92–D97. <https://doi.org/10.1093/nar/gkt1248> PMID: 24297251
31. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003; 13: 2498–2504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658
32. Chin C-H, Chen S-H, Wu H-H, Ho C-W, Ko M-T, Lin C-Y. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014; 8: S11. <https://doi.org/10.1186/1752-0509-8-S4-S11> PMID: 25521941
33. Rombel IT, Sykes KF, Rayner S, Johnston SA. ORF-FINDER: a vector for high-throughput gene identification. *Gene*. 2002; 282: 33–41. [https://doi.org/10.1016/s0378-1119\(01\)00819-8](https://doi.org/10.1016/s0378-1119(01)00819-8) PMID: 11814675
34. Kong L, Zhang Y, Ye Z-Q, Liu X-Q, Zhao S-Q, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res*. 2007; 35: W345–W349. <https://doi.org/10.1093/nar/gkm391> PMID: 17631615
35. Kang Y-J, Yang D-C, Kong L, Hou M, Meng Y-Q, Wei L, et al. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res*. 2017; 45: W12–W16. <https://doi.org/10.1093/nar/gkx428> PMID: 28521017
36. Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol*. 2001; 305: 567–580. <https://doi.org/10.1006/jmbi.2000.4315> PMID: 11152613
37. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods*. 2015; 12: 7–8. <https://doi.org/10.1038/nmeth.3213> PMID: 25549265
38. Zhang W, Yang J, He B, Walker SE, Zhang H, Govindarajoo B, et al. Integration of QUARK and I-TASSER for ab initio protein structure prediction in CASP11. *Proteins Struct Funct Bioinforma*. 2016; 84: 76–86. <https://doi.org/10.1002/prot.24930> PMID: 26370505
39. Yang J, Zhang W, He B, Walker SE, Zhang H, Govindarajoo B, et al. Template-based protein structure prediction in CASP11 and retrospect of I-TASSER in the last decade. *Proteins Struct Funct Bioinforma*. 2016; 84: 233–246. <https://doi.org/10.1002/prot.24918> PMID: 26343917
40. Bowers KJ, Chow E, Xu H, Dror RO, Eastwood MP, Gregersen BA, et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. *SC 2006 Conference, Proceedings of the ACM/IEEE*. 2006. p. 43–43.
41. Oxenoid K, Chou JJ. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc Natl Acad Sci U S A*. 2005; 102: 10870–10875. <https://doi.org/10.1073/pnas.0504920102> PMID: 16043693
42. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983; 79: 926–935.
43. Huang J, MacKerell AD. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J Comput Chem*. 2013; 34: 2135–2145. <https://doi.org/10.1002/jcc.23354> PMID: 23832629
44. Zhu F, Hummer G. Pore opening and closing of a pentameric ligand-gated ion channel. *Proc Natl Acad Sci*. 2010; 107: 19814–19819. <https://doi.org/10.1073/pnas.1009313107> PMID: 21041674
45. Zhu F, Hummer G. Drying transition in the hydrophobic gate of the GLIC channel blocks ion conduction. *Biophys J*. 2012; 103: 219–227. <https://doi.org/10.1016/j.bpj.2012.06.003> PMID: 22853899
46. Hub JS, De Groot BL, Van Der Spoel D. g\_wham: A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J Chem Theory Comput*. 2010; 6: 3713–3720.

47. DeLano WL. Pymol: An open-source molecular graphics tool. *CCP4 Newsl protein Crystallogr.* 2002; 40: 82–92.
48. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph.* 1996; 14: 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) PMID: 8744570
49. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* 2015; 1: 19–25.
50. Chou C-H, Shrestha S, Yang C-D, Chang N-W, Lin Y-L, Liao K-W, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 2017; 46: D296–D302.
51. Gao F, Cui Y, Jiang H, Sui D, Wang Y, Jiang Z, et al. Circulating tumor cell is a common property of brain glioma and promotes the monitoring system. *Oncotarget.* 2016; 7: 71330. <https://doi.org/10.18632/oncotarget.11114> PMID: 27517490
52. Papadopoulos MC, Verkman AS. Aquaporin-4 and brain edema. *Pediatr Nephrol.* 2007; 22: 778–784. <https://doi.org/10.1007/s00467-006-0411-0> PMID: 17347837
53. Griesdale DEG, Honey CR. Aquaporins and brain edema. *Surg Neurol.* 2004; 61: 418–421. <https://doi.org/10.1016/j.surneu.2003.10.047> PMID: 15120208
54. Payre F, Desplan C. Small peptides control heart activity. *Science.* 2016; 351: 226–227. <https://doi.org/10.1126/science.aad9873> PMID: 26816363
55. Winther AML, Bublitz M, Karlsen JL, Miller J V, Hansen JB, Nissen P, et al. The sarcolipin-bound calcium pump stabilizes calcium sites exposed to the cytoplasm. *Nature.* 2013; 495: 265–269. <https://doi.org/10.1038/nature11900> PMID: 23455424
56. Singh DR, Dalton MP, Cho EE, Pribadi MP, Zak TJ, Šeflová J, et al. Newly Discovered Micropeptide Regulators of SERCA Form Oligomers but Bind to the Pump as Monomers. *J Mol Biol.* 2019; 431: 4429–4443. <https://doi.org/10.1016/j.jmb.2019.07.037> PMID: 31449798
57. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. *Nat Genet.* 2000; 25: 25–29. <https://doi.org/10.1038/75556> PMID: 10802651
58. Schwab A, Fabian A, Hanley PJ, Stock C. Role of ion channels and transporters in cell migration. *Physiol Rev.* 2012; 92: 1865–1913. <https://doi.org/10.1152/physrev.00018.2011> PMID: 23073633
59. Pardo LA, Stühmer W. The roles of K<sup>+</sup> channels in cancer. *Nat Rev Cancer.* 2014; 14: 39–48. <https://doi.org/10.1038/nrc3635> PMID: 24336491
60. Papadopoulos MC, Verkman AS. Aquaporin water channels in the nervous system. *Nat Rev Neurosci.* 2013; 14: 265–277. <https://doi.org/10.1038/nrn3468> PMID: 23481483
61. Liu T, Han Z, Li H, Zhu Y, Sun Z, Zhu A. LncRNA DLEU1 contributes to colorectal cancer progression via activation of KPNA3. *Mol Cancer.* 2018; 17: 118. <https://doi.org/10.1186/s12943-018-0873-2> PMID: 30098595
62. Cao Y, Wu X, Yang R, Wang X, Sun H, Lee I. Self-assembling study of sarcolipin and its mutants in multiple molecular dynamic simulations. *Proteins-structure Funct Bioinforma.* 2017; 85: 1065–1077 <https://doi.org/10.1002/prot.25273> PMID: 28241400
63. Tunuguntla RH, Henley RY, Yao Y-C, Pham TA, Wanunu M, Noy A. Enhanced water permeability and tunable ion selectivity in subnanometer carbon nanotube porins. *Science.* 2017; 357: 792–796. <https://doi.org/10.1126/science.aan2438> PMID: 28839070
64. Horner A, Zocher F, Preiner J, Ollinger N, Siligan C, Akimov SA, et al. The mobility of single-file water molecules is governed by the number of H-bonds they may form with channel-lining residues. *Sci Adv.* 2015; 1: e1400083. <https://doi.org/10.1126/sciadv.1400083> PMID: 26167541
65. Cao Y, Wu X, Wang X, Sun H, Lee I. Transmembrane dynamics of the Thr-5 phosphorylated sarcolipin pentameric channel. *Arch Biochem Biophys.* 2016; 604: 143–151. <https://doi.org/10.1016/j.abb.2016.06.016> PMID: 27378083
66. Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, et al. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol.* 2012; 8: e1002708. <https://doi.org/10.1371/journal.pcbi.1002708> PMID: 23093919