POLITECNICO DI TORINO
Repository ISTITUZIONALE

Pore-level quantitative structure-activity relationship (QSAR) for water permeation rate in aquaporins

Original
Pore-level quantitative structure-activity relationship (QSAR) for water permeation rate in aquaporins / Galano-Frutos, J.J., Bergamasco, L., Vigo, P., Morciano, M., Fasano, M., Pirolli, D., Chiavazzo, E., De Rosa, M.C.. - In: BIOCHIMICA ET BIOPHYSICA ACTA-BIOMEMBRANES. - ISSN 0005-2736. - 1868:3(2026). [10.1016/j.bbamem.2026.184529]

Availability:
This version is available at: 11583/3009871 since: 2026-04-14T15:44:57Z

Publisher:
Elsevier

Published
DOI:10.1016/j.bbamem.2026.184529

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

19 November 2026

BBA - Biomembranes 1868 (2026) 184529

Contents lists available at ScienceDirect

BBA - Biomembranes

journal homepage: www.elsevier.com/locate/bbamem

Research paper

Pore-level quantitative structure-activity relationship (QSAR) for water permeation rate in aquaporins

Juan José Galano-Frutos [a,*,1,2], Luca Bergamasco [b,*,2], Paolo Vigo [b], Matteo Morciano [b], Matteo Fasano [b], Davide Pirolli [a], Eliodoro Chiavazzo [b], Maria Cristina De Rosa [a]

[a] Istituto di Scienze e Tecnologie Chimiche "Giulio Natta" (SCITEC) - National Research Council (CNR), Via Largo Francesco Vito 1, 00168, Rome, Italy
[b] Department of Energy, Politecnico di Torino, C.so Duca degli Abruzzi 24, 10129, Turin, Italy

ARTICLE INFO

Keywords:
QSAR
Aquaporins
Pore permeation rate
Pore selectivity
Pore descriptors

ABSTRACT

Aquaporins (AQPs) and aquaglyceroporins (AQGPs) play a crucial role in regulating water transport and solute selectivity across biological membranes. Besides their biological relevance, AQPs have attracted growing interest as models for the design of next-generation biomimetic membranes for water filtration. In this work, we present a pore-level Quantitative Structure–Activity Relationship (QSAR) approach that relates structural and physicochemical pore descriptors with experimentally reported water permeation rates across a set of AQ(G)Ps with high-resolution 3D structures. This data-driven methodology, presented here as a proof of concept, introduces a multi-feature framework for determining pore descriptors associated with water transport efficiency in AQ(G)Ps. Applied to two compiled permeation rate datasets, this framework recapitulates determinants previously reported in single-feature studies, while also highlighting additional pore descriptors that emerge as relevant in a multi-variable context. The insights gained through this approach may, in perspective, contribute to advancing the rational design of AQP-based filtration devices and to deepening the molecular understanding of the function of these valuable macromolecules in health and disease.

1. Introduction

The maintenance of ionic homeostasis and water levels within the cell and within different organisms' tissues and compartments is a function that has been linked to a protein family called aquaporins (AQPs), which includes aquaglyceroporins (AQGPs). Initially referred to as Channel-forming Integral Protein (CHIP) [1–3] or Major Intrinsic Proteins (MIPs), AQ(G)Ps are 28–35 kDa integral membrane protein channels, which enable the transport of water across cells with a high efficiency and a unique selectivity. Typically assembled as tetramers (Fig. 1a), each subunit forms an integral membrane pore, characterized by six transmembrane spanning domains, with cytosolic amino and carboxy termini (Fig. 1b and c). Only pure water and small uncharged solutes, e.g. glycerol, urea, metalloids, carbon dioxide, oxygen, nitrogen, hydrogen peroxide and ammonia, are allowed to pass through the pore [4–7]. This selectivity is due to an electrostatic barrier located at the center of the channel, which prevents the passage of protons and other charged molecules [8] (Fig. 1b). Phenomena like the survival of some plants in extreme environments, characterized by saline-alkaline soils, high or low temperatures, heavy metal toxicity and drought are owed to adaptations developed by these organisms thanks to the presence of AQ(G)Ps [9–12]. AQ(G)Ps have also been associated with cell volume regulation, transepithelial fluid transport, cell migration and proliferation, neuroexcitation, among other physiological events [13–15], and thus linked to various non-infectious and infectious diseases in humans [3,13,14,16,17], highlighting the importance of their study across a variety of research and development fields.

There are more than 350 known aquaporins, with 13 subfamilies found in mammalian, namely AQP0–12. Mammalian AQ(G)Ps are usually divided into three sub-classes, one containing proteins that allow only water molecules flow, namely AQP0, 1, 2, 4, 5, 6 and 8, which are referred to as orthodox aquaporins, another one that allows the transport of water and other small solutes, like glycerol or urea, that includes AQP3, 7, 9 and 10, referred to as aquaglyceroporins (AQGPs), and a

* Corresponding authors.
E-mail addresses: jjgalano@certest.es (J.J. Galano-Frutos), luca.bergamasco@polito.it (L. Bergamasco).
[1] Present address: Certest Biotec S.L., 50840 San Mateo de Gallego, Zaragoza, Spain.
[2] These authors contributed equally to this work.

https://doi.org/10.1016/j.bbamem.2026.184529
Received 13 August 2025; Received in revised form 17 February 2026; Accepted 3 April 2026
Available online 6 April 2026
0005-2736/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

more recently discovered class that has comparatively lower sequence homology (about 20%) with the orthodox aquaporins, that includes AQP11 and AQP12, referred to as unorthodox or superaquaporins [17,19,20]. A higher AQPs diversity is present in plants, where there are five homologous AQPs subfamilies, namely plasma membrane intrinsic proteins (PIP), tonoplast intrinsic protein (TIP), small intrinsic protein (SIP), nodulin-26 like intrinsic protein (NIP) and X intrinsic proteins (XIP) [10]. Amphibians [21], bacteria [22] and fungi [23] have also been reported to carry AQPs.

Thanks to their inherent functional characteristics, aquaporins are used for developing next-generation ultra-permeable membranes [24,25]. Aquaporin-based biomimetic membranes (ABMs) [26] have demonstrated remarkable promise both in water desalination [27,28] (see reviews in Ref. [29–33]) and wastewater treatment [34–36]. Their application for these and other related purposes has proven to be a valuable solution [37–39] as an alternative to conventional reverse osmosis (RO) [40,41] or thermal desalination techniques [42,43], offering potential energy and cost savings [39]. Hélix-Nielsen [31] and Beratto-Ramos et al. [36] emphasized the necessity of adopting novel optimization and design strategies to enhance ABM technology further. In close alignment with this imperative, the past 15 years have witnessed a significant surge in the development of bio-inspired artificial water channels (AWCs), which has been driven by the ambition to fabricate synthetic counterparts of AQ(G)Ps (see Ref. [44–46], and recent applications in Ref. [47, 48]), marking a notable frontier in scientific exploration.

In connection with these challenges, significant efforts—spanning both experimental and computational domains—have been dedicated to unraveling the working principles governing the permeation and selective capabilities of AQ(G)Ps [19,49–60]. Most of them, however, have

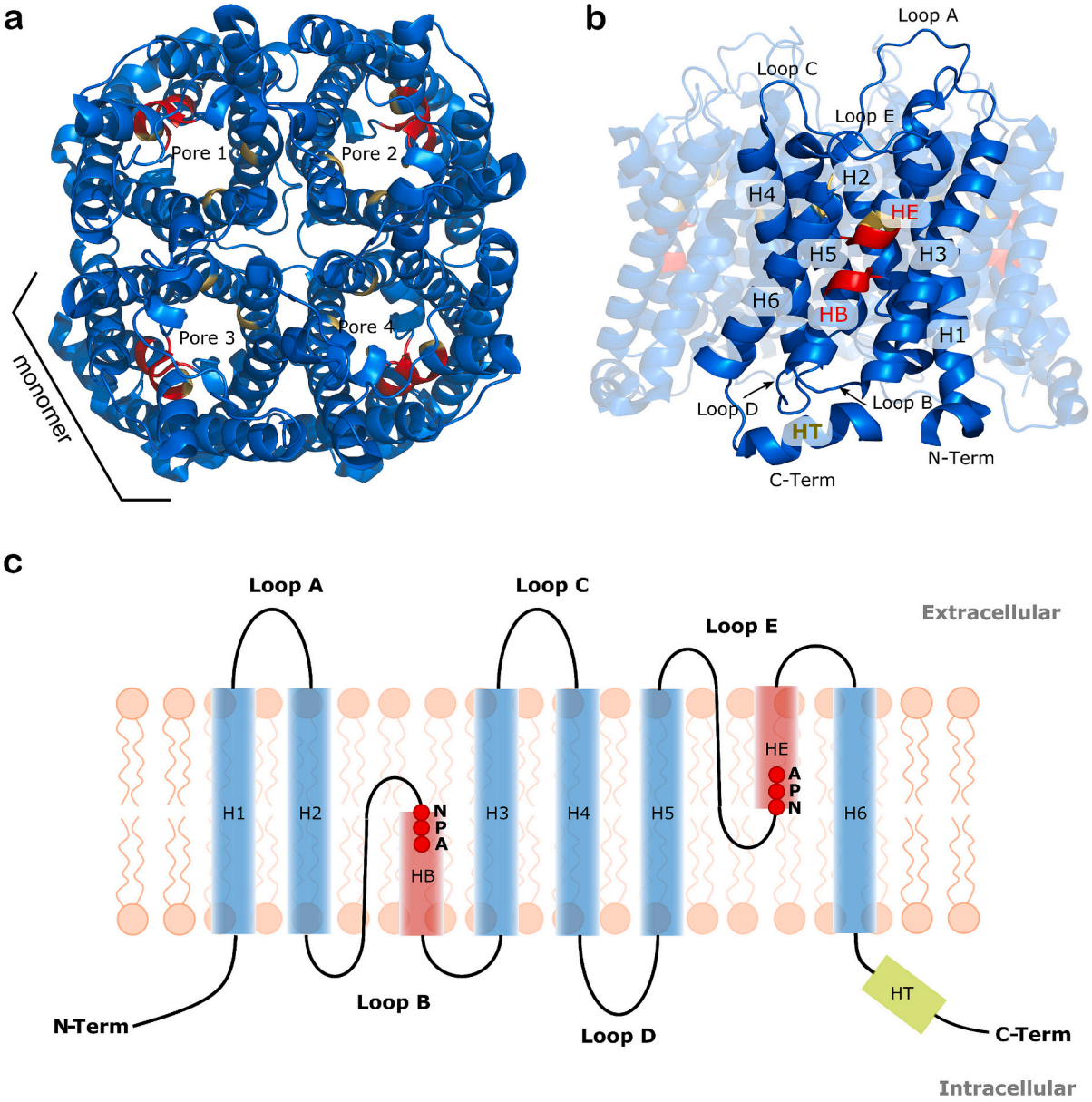


Fig. 1. Structural features of mammalian AQ(G)Ps. (a) Tetrameric structure of hAQP1 (PDB ID 7UZE [18]), viewed from the extracellular side, with the pores indicated on each monomer. (b) Side view of hAQP1 highlighting the main features of the pore for one monomer. The two NPA motifs are highlighted in red, and the selectivity filter (SF) motif is highlighted in yellow. (c) Schematic representation of the mammalian AQ(G)P channel topology. Each monomer comprises six transmembrane helices (H1–H6) connected by five loops (A–E) and an additional C-terminal short amphipathic helix (HT). Loops B and E fold back into the membrane, forming two half-membrane spanning helices (HB and HE) that harbor the two asparagine–proline–alanine (NPA) motifs (red colored circles). The N- and C-termini are located in the cytoplasm, and the amphipathic helix (HT) is a common protein–protein interaction site.

predominantly focused on individual features of these proteins [61], as for instance, the amino acid compositions of the NPA and SF motifs, the pore gating [59], the pore shape and size [60,61], the putative number of hydrogen bonds formed along the pore lining [57], or the number of positive charges at the pore mouth [58], to name a few. For instance, it has been shown that AQ(G)Ps' water permeability and high solute rejection occur through a narrow hydrophilic constriction, typically around ~3 Angstroms, which facilitates the formation of a water single-file wire [62]. This process is characterized by hydrogen bonding among water molecules [50,63], alongside bonding and polarizing interactions with the pore lining [57,58,62]. Also, amino acids positions determinative of AQP vs. AQGPs selectivity, and those involved in auto-regulatory events linked to post-translational modifications have been unveiled through a variety of experimental methods, including site-directed mutagenesis, chimeric domain swaps, membrane permeability assays, electron crystallography, X-ray crystallography, and molecular dynamics simulations [19,64,65]. Despite such progress, a precise and comprehensive understanding of the permeation properties, selectivity and mechanisms exhibited by natural AQ(G)Ps remains incomplete [66].

Quantitative Structure-Activity Relationship (QSAR) (or Quantitative Structure-Property Relationship, QSPR) modeling, is one of the most extended in silico tools employed in medicinal chemistry [67]. In QSAR, features such as solubility, acidity, and polarity are computed as molecular descriptors and then applied to determine which of them can be used to predict measurable responses, like toxicity, drug-likeness, or enzyme binding. The application of QSAR to model the activities or properties of peptides and proteins is an emerging field that is not yet as well established as it is for small molecules, due to the added complexity of these systems. Nevertheless, this approach has proven useful for uncovering hidden relationships that can be exploited to guide the design and optimization of the properties of systems of interest [68]. For valuable insights into the current state, development, and optimization of molecular descriptors to be used in QSAR studies on peptides and proteins, readers may refer to the following works: Ref. [69] (a review paper) and Ref. [70] (a research article).

Although well-established, QSAR methods appear to be still unexplored in the context of AQ(G)Ps. Thus, here we present—as a methodological proof of concept—a pore-level QSAR approach that combines available structural and permeability data to reveal key, and potentially novel, determinants of the filtering capabilities of these remarkable biomolecules. The proposed methodology could provide deeper insights at the molecular level, offering a better understanding of the AQ(G)Ps' functioning, and creating opportunities for the rational improvement of their properties and for targeted interventions in disease-related contexts.

2. Materials and methods

2.1. AQPs with available 3D structure

A search for proteins of the MIP family was performed on two different databases, namely PROSITE [71,72] and InterPro [73,74]. Proteins with experimental 3D structures available were subsequently retrieved from the RCSB PDB archive [75].

The PROSITE database [71,72], returned 285 entries (with accession code PS00221 related to MIP family). Out of the true positive sequences (264), 53 structure entries were available on the PDB archive according to this database. The InterPro database [73,74], which contained all the structure entries reported in PROSITE, returned (with accession code IPR000425 related to MIP family) 73 different PDB structures. This number of structures represented 22 individual aquaporins. In cases where multiple structures were available for a protein, the one with the highest structural resolution among those with canonical sequences (no mutations) was selected as its representative structure for the present study. An overview of these 22 AQ(G)Ps and their respective PDB IDs is reported in Table 1.

2.2. Training set and AQ(G)Ps' permeation rate

We reviewed the literature for osmotic unitary (single-channel) water permeation rates (p_f)—encompassing both experimental and molecular dynamics (MD)-based approaches—reported for the 22 AQ(G)Ps with available experimental structure (see Table S1 and Fig. 2). Permeation data for other AQ(G)Ps lacking resolved 3D structures, and therefore not included in Table 1, are also provided in Tables S1 and S2 and in Fig. 2 (see details below).

Experimental data was not found for some AQ(G)Ps listed in Table 1, for which only in silico data is reported. For some AQ(G)Ps, we could not find any data, or the reported permeation rates were not provided as

Table 1
Overview of AQPs with available 3D structure and selected PDB entries.

AQ(G)P (short name)	PDB ID	Source organism	Ref.[a]
Aquaporin 0 (bAQP0)	1YMG	*Bos taurus*	[76]
Aquaporin 0 (oAQP0)	2B6O	*Ovis aries*	[77]
Aquaporin 1 (bAQP1)	1J4N	*Bos Taurus*	[78]
Aquaporin 1 (hAQP1)	7UZE	*Homo sapiens*	[18]
Aquaporin 1 (aAqp1)	7W7S	*Anabas testudineus*	[79]
Aquaporin 1 (kAqp1)	3ZOJ	*Komagataella pastoris*	[80]
Aquaporin 2 (hAQP2)	4NEF	*Homo sapiens*	[81]
Aquaporin 4 (hAQP4)	3GD8	*Homo sapiens*	[82]
Aquaporin 4 (rAQP4)	2D57	*Rattus norvegicus*	[83]
Aquaporin 5 (hAQP5)	3D9S	*Homo sapiens*	[84]
Aquaporin 7 (hAQP7)	6QZI	*Homo sapiens*	[85]
Aquaporin 10 (hAQP10)	6F7H	*Homo sapiens*	[86]
Aquaporin (SoPIP2;1)	1Z98	*Spinacia oleracea*	[87]
Aquaporin M (mAqpM)	2F2B	*Methanothermobacter marburgensis*	[88]
Probable aquaporin M (aAqpM)	3NE2	*Archaeoglobus fulgidus*	[89]
Aquaporin NIP2;1 (OsNIP2;1)	7CJS	*Oryza sativa japonica*	[90]
Aquaporin TIP2;1 (AtTIP2;1)	5I32	*Arabidopsis thaliana*	[91]
Probable aquaporin PIP2;4 (AtPIP2;4)	6QIM	*Arabidopsis thaliana*	[6]
Aquaporin Z (eAqpZ)	1RC2	*Escherichia coli*	[92]
Aquaporin Z 2 (aAqpZ2)	3LLQ	*Agrobacterium fabrum*	[93]
Aquaglyceroporin (pAqgp)	3C02	*Plasmodium falciparum*	[94]
Glycerol uptake facilitator protein (eGlpF)	1FX8	*Escherichia coli*	[62]

[a] Reference articles for structures 3LLQ and 3NE2 are not yet available. In these cases the references provided are the links to the entries in the RCSB Protein Data Bank [75].

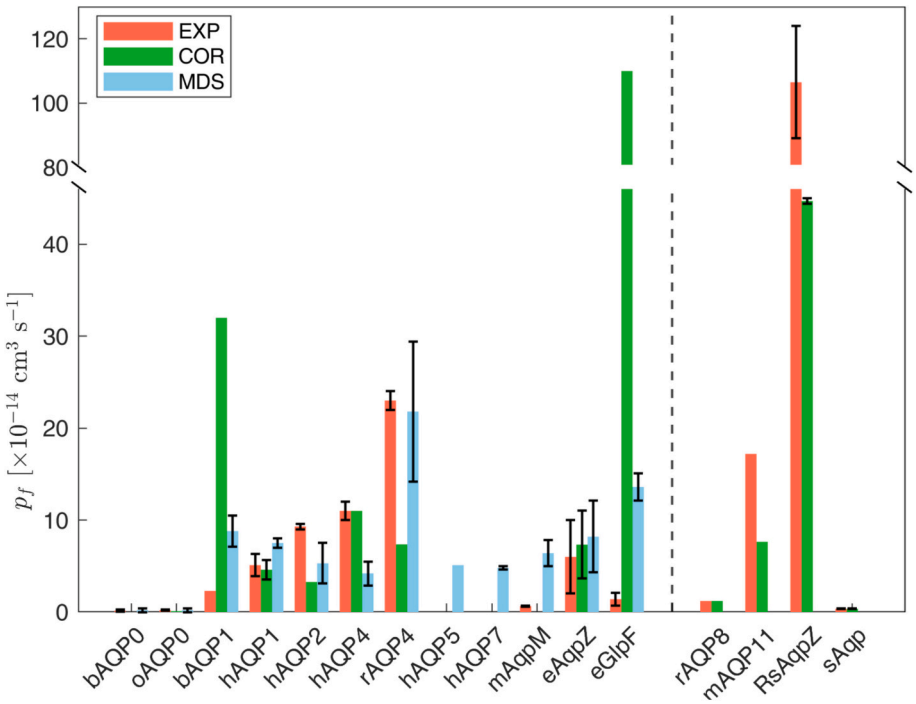


Fig. 2. Osmotic unitary water permeation rates ($\times$ 10^{-14} cm^3•s^{-1}). Red bars correspond to individual p_f measurements or averages, as compiled in Table S1. Green bars correspond to corrected unitary water permeabilities (p_f^{corr} [56]), as compiled in Table S2. Cyan bars correspond to p_f values (individual or averages) obtained from molecular dynamics simulations (MDS; as compiled in Table S1). AQPs and AQGPs to the left of the vertical dashed line correspond to those with experimentally resolved 3D structures (listed in Table 1). Within this subset, experimental p_f values could not be found or derived for hAQP5 and hAQP7, and the compiled p_f^{corr} values correspond to corrections or averages derived strictly from low-temperature ($\leq$15 °C) measurements (Table S2). AQPs to the right of the dashed line lack resolved 3D structures and were used as an external validation set in the pore-level QSAR approach presented here. The experimental methods employed, along with the main conditions and simulation setups used to derive the compiled values, are summarized in Table S1. When more than one value is available—regardless of the method used—the bar represents the mean value and the black vertical lines indicate the standard error of the mean. Otherwise, the black vertical line denotes the experimental or computational uncertainty—when reported—associated with the individual measurement or calculation.

single-channel water permeabilities (i.e., p_f) but rather as global osmotic permeability coefficient (P_f) or other metrics. In cases such as hAQP7 [95–97] and hAQP10 [86], we were unable to convert P_f into p_f due to the lack of required data for such a transformation. In the case of hAQP5, the functional analysis carried out by Horsefield et al. [84] relied on a comparative light-scattering–based approach, for which only rate constants (k, in s^{-1}) were reported; consequently, we were not able to retrieve the p_f values.

Importantly, Wachlmayr et al. [56] recently revisited the stopped-flow methodology underlying many historical measurements of osmotic water permeability (P_f) and unitary water permeability (p_f). They argued that a significant subset of older reports is systematically biased because P_f was often obtained by fitting light-scattering traces to a single exponential and inserting the resulting time constant into approximate formulas that are only valid under restrictive conditions. To address this, they compare the main analysis strategies and promote an empirically corrected analytical-solution-based exponential approach (E-ASF) that retains the practical simplicity of exponential fitting while reproducing permeabilities consistent with the proper analytical solution. Crucially for legacy datasets, they provide osmolarity-based correction factors that can be applied even when the original traces are unavailable, and show that many published p_f values should be revised downward (typically to ~0.25–0.5 of the originally reported value), in agreement with independent numerical re-fits [56]. Here, we compiled—from Ref. [56]—the corrected unitary water permeabilities (hereafter p_f^{corr}) for 8 of AQ(G)Ps listed in Table 1 (i.e., those with experimentally resolved 3D structures), as well as for 4 additional AQPs lacking 3D structures, which are used as an external validation set in the pore-level QSAR approach presented here (see Table S2 and Fig. 2).

Since equilibrium MDS-based approaches, mostly used in earlier modeling of AQ(G)Ps, have well-recognized limitations [98], and because MDS-derived osmotic unitary water permeabilities do not appear to align well with either non-corrected or corrected experimental estimates (see Fig. 2), we did not consider MDS-derived permeation rates for the pore-level QSAR framework presented here. We therefore carried out the proposed modeling separately for two subsets of structurally resolved AQ(G)Ps listed in Table 1: (i) those for which experimental p_f values are available (individual measurements or averages thereof, as compiled in Table S1), comprising 10 AQ(G)Ps: bAQP0, oAQP0, bAQP1, hAQP1, hAQP2, hAQP4, rAQP4, mAqpM, eAqpZ and eGlpF (see Table S1 and Fig. 2); and (ii) those for which corrected unitary permeabilities, p_f^{corr}, are reported by Wachlmayr et al. [56], comprising 8 AQ(G)Ps: oAQP0, bAQP1, hAQP1, hAQP2, hAQP4, rAQP4, eAqpZ and eGlpF (see Table S2 and Fig. 2). This second set of AQ(G)Ps with available p_f^{corr} values is indeed a subset of the first one.

When compiling the p_f^{corr} data from Ref. [56], we followed the authors' recommendation that osmotic permeabilities should preferably be measured at low temperature (<10 °C) to minimize systematic errors. Thus, we applied a temperature-based filtering to maximize internal consistency of the corrected permeability dataset: measurements performed above 10 °C, as well as entries with unreported temperature were discarded (with the sole exception of sheep AQP0, for which only one value at 15 °C was available and was retained). In this context, in cases where multiple low-temperature permeabilities are reported for the same AQ(G)P, we averaged their p_f^{corr} values to obtain a single representative corrected permeability per aquaporin (Table S2).

For QSAR modeling, unitary permeabilities were expressed on a common numerical scale ($\times$ 10^{-14} cm^3•s^{-1}, after which the 10^{-14} factor

was removed prior to log-transformation. We thus used $ln\left(p_f\right)$ or $ln\left(p_f^{corr}\right)$ as the response variable in the QSAR modeling. This log transformation was necessary because p_f and p_f^{corr} spans approximately two and four orders of magnitude, respectively, and their values are not evenly distributed (Table S2). The log transformation compresses the dynamic range, reduces the influence of extreme values, and yields a distribution closer to normality, thereby facilitating model fitting and interpretation.

The scenario outlined above, further enables us to assess how the use of uncorrected versus corrected permeability values impacts predictive performance and, critically, the physicochemical interpretability of the resulting QSAR models. Importantly, the analysis and discussion presented below are intended in a constructive spirit—aimed at contributing additional evidence to the ongoing effort to refine experimental permeability measurements in narrow water channels such as AQ(G)Ps—rather than to endorse or disparage any particular experimental or analytical approach.

2.3. Pore descriptors

Different computational resources were used to obtain or derive pore features (hereafter 'pore descriptors', PDs). Geometric PDs (including averaged pore diameter, minimum pore diameter and regularity) and pore-lining amino acid composition were obtained using the PoreWalker server [99]. Per-residue hydropathy (Kyte–Doolittle index [100]) was computed in Discovery Studio [101]. The Kyte-Doolittle hydrophobicity scale ranges from −4.5, the most hydrophilic value for Arg, to 4.5, the most hydrophobic value for Ile, with Gly being the most amphipathic residue (−0.4). PDs providing an indirect proxy for local mobility/flexibility were derived from crystallographic atomic displacements (B-factors) extracted from the PDB files. Finally, H-bonding- and electrostatic/solvation-related PDs were calculated following procedures analogous to those described by Horner et al. in Ref. [57] and Ref. [58], respectively (see Fig. 3).

When using the PoreWalker tool [99], the initial chain of each AQ(G)P structure was uploaded to the server for calculation. The PoreWalker output (a summary .TXT file is provided in the **Supplementary Information** for every AQ(G)P studied) includes the pore diameter profile along the pore z-coordinate, the 3D shape regularity percentage, 2D projection of the pore shape, cross-sectional images of the pore, as well as the list of amino acids lining the pore.

To carry out the present pore-level QSAR modeling, we first equalized the pore z-coordinate across the AQ(G)Ps with p_f and/or p_f^{corr} data compiled (Fig. 2 and Tables S1-S3). This was done by aligning all the pores along the z-axis, using as a common reference point the β-carbon (βC) of the asparagine residue in the second NPA motif. The βC coordinate was thus set to z = 0 (Table S3, Fig. 4a, and left panels in Fig. S1).

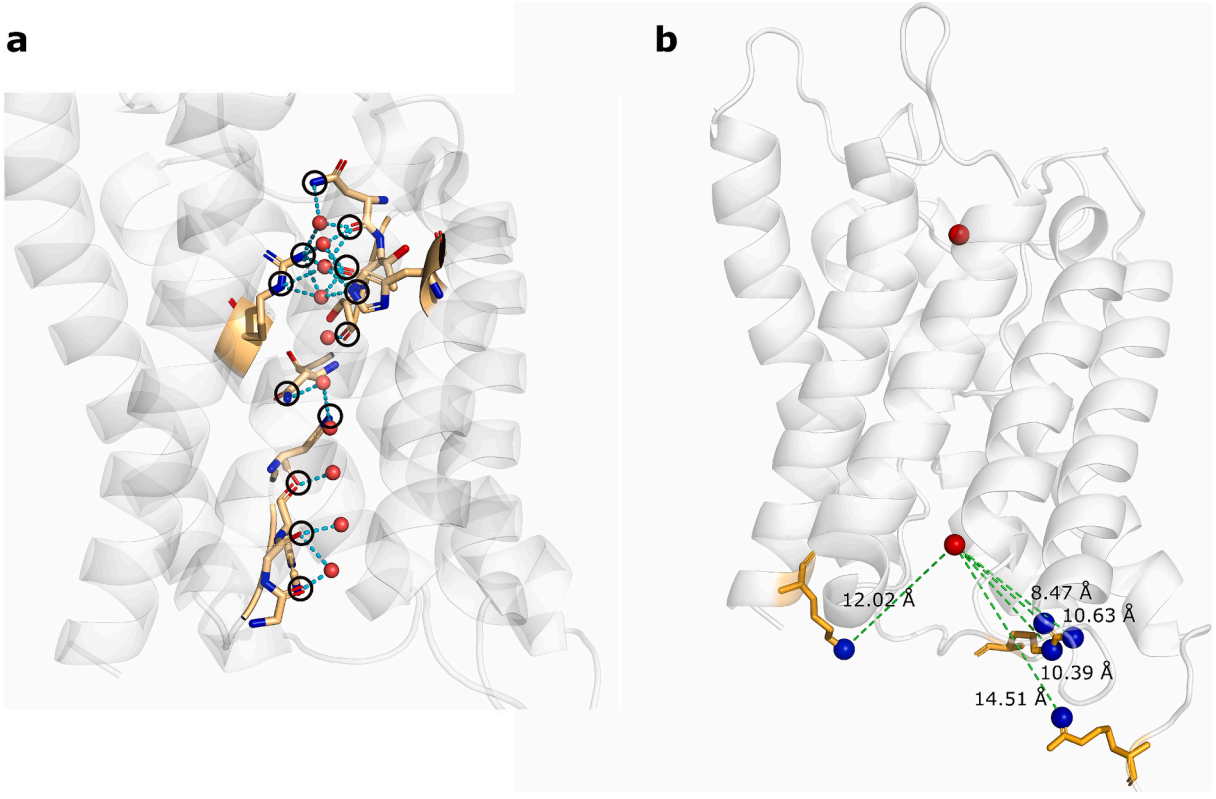
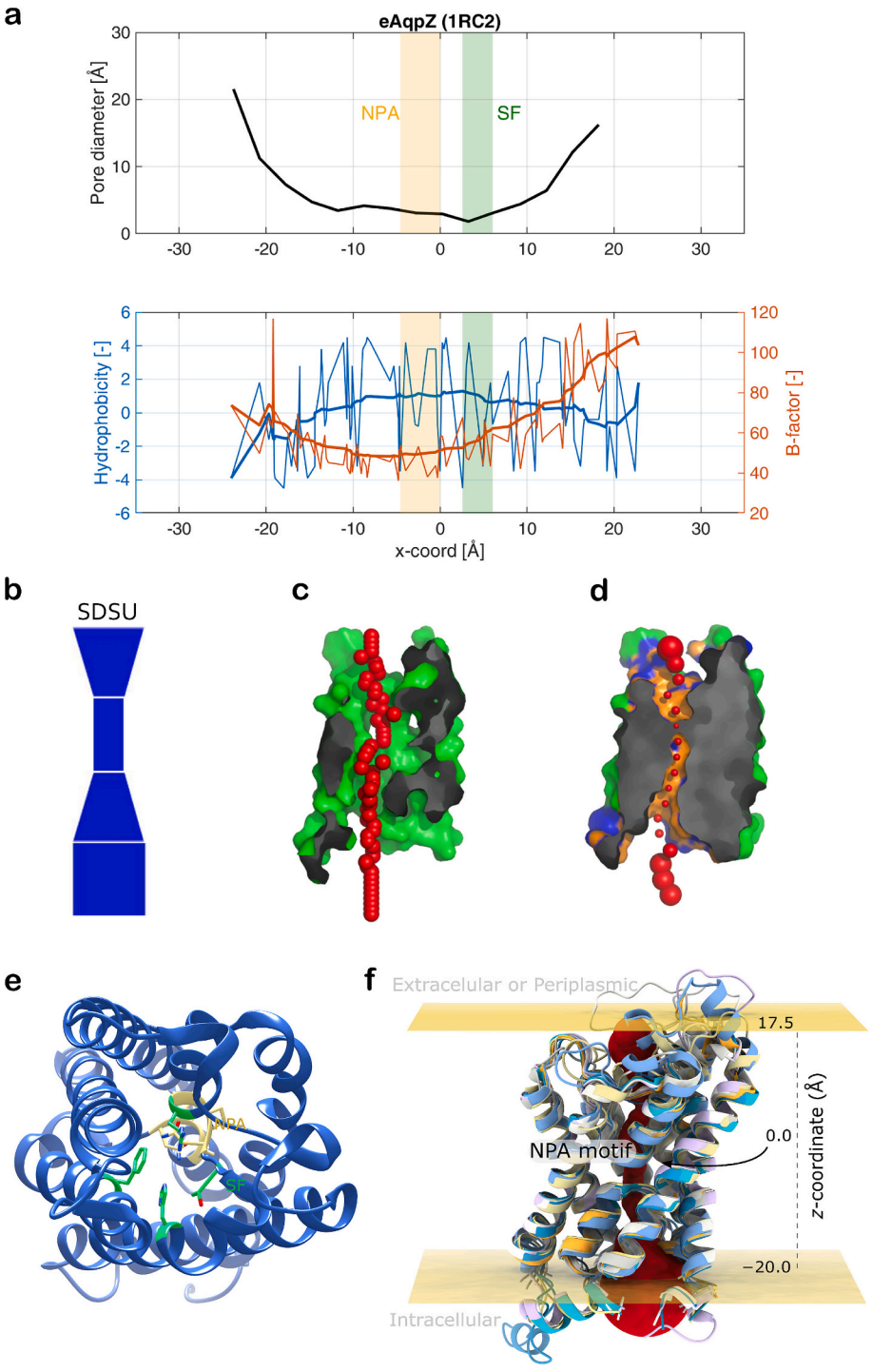


Fig. 3. Schematic illustration for eAqpZ (PDB ID: 1RC2) showing the computed number of potential pore-lining hydrogen-bonds ($\boldsymbol{N_H}$), charged residues and charged sites located in the periplasmic and cytoplasmic pore vestibules. (a) Calculated potential H-bonding sites (N_H = 12, black circles) along the single-file region of eAqpZ. The single-file water column (red spheres) was defined following exactly the procedure described by Horner et al. [57]: water coordinates were taken from the 0.88 Å-resolution yeast AQP1 structure (PDB ID: 3ZOJ) [80] after structural superposition onto eAqpZ. Putative hydrogen bonds (cyan dashed lines) were identified using a 3.8 Å distance cut-off. (b) Positively charged Lys and Arg residues (Res^+ = 3, tan-colored sticks), and associated charged side-chain nitrogen atoms ($Sites^+$ = 5, blue spheres) counted within the pore cytoplasmic vestibule in eAqpZ, according to the established 15 Å cut-off to the reference waters (red spheres) represented on both the periplasmic and cytoplasmic sides. The red spheres mark the oxygen atoms of the first water molecules flanking the single-file water column, i.e., waters that are no longer part of the continuous single-file chain traversing the pore. Their coordinates were also taken from the above-mentioned yeast AQP1 structure (PDB ID 3ZOJ) [80] after structural superposition onto eAqpZ. The numbers next to the green dashed lines indicate the distances (Å)—all below the 15 Å cut-off (same defined by Horner et al. [58])—computed between each water oxygen atom (red spheres) and the charged side-chain nitrogen atoms (blue spheres) of the identified Lys and Arg residues within the same outer vestibule. Custom Python scripts were developed to compute these PDs and the rest of related pore features (see Table 2), including the number of negatively charged residues (Res^-, Glu and Asp) and charged sites/atoms ($Sites^-$) located in the pore vestibules.



(caption on next page)

Fig. 4. Pore properties profiles and geometrical shape for *Escherichia coli* aquaporin Z (eAqpZ, PDB ID 1RC2), and scheme showing the *z*-coordinates equalization of the AQ(G)Ps modeled in this work. (a) Top panel: diameter profile along the pore principal axis (*z*-coordinate) at 3 Å step as calculated by the PoreWalker tool [99]; bottom panel: per-residue hydrophobicity index (blue) and B-factor (orange) of amino acids in the pore lining along the pore z-coordinate, as calculated by Discovery Studio (BIOVIA, Dassault Systèmes) [101] and as obtained from the crystal PDB file (1RC2), respectively. The thicker lines are smooth curves (noise removed). Vertical bands show the *z*-coordinates ranges where the NPA (orange) and SF (green) motifs are located. (b) Funnel-like form representing a 2D projection of the predicted pore shape as calculated by variations of diameter values along the pore: the bottom of the pore axis (i.e. the lowest *z*-coordinate) corresponds to the bottom of the stack. In the letter code on the top of the form, 'D' indicates a conical frustum generated by decreasing pore diameter values, i.e. the diameter of the lower base of the frustum is bigger than the diameter of its upper circle, 'U' indicates the opposite conical frustum, i.e. lower base diameter smaller than upper base diameter, and 'S' indicates a cylinder (the code ordering from bottom to top). (c) and (d) Longitudinal cross-sectional images of the pore issued by PoreWalker, where the pore regularity is shown through red spheres located at the centre of the pore every 1 Å step (c), whereas the pore dimension is represented by red spheres—at the center of the pore at 3 Å step—whose sizes are proportional to the pore diameter calculated at that point (d). In the latter, orange and blue areas show pore-lining atoms and corresponding residues, respectively. (e) Cartoon representation of a single monomer of *E. coli* aquaporin Z (PDB ID: 1RC2), viewed from the extracellular side. The NPA motifs are shown as sticks in pale yellow (residues 63–65 for NPA-1 and 186–188 for NPA-2), and the selectivity filter residues (F43, H174, T183, and R189) are shown as green sticks. (f) Scheme showing the *z*-coordinates equalization performed on the modeled AQ(G)Ps. The central red yarn ball-like form approximately represents the pore diameter of one of the aligned AQ(G)Ps. **Notes:** Amino acids in the pore lining are reported in the PoreWalker's output files (**Supplementary Files**). The *z*-coordinates in the properties profiles (a) were shifted from the original ones issued by PoreWalker, to place the beta carbon (βC) of the asparagine (N) in the second NPA motif at the center of the *z*-coordinate system ($z = 0$). This ensures that all the AQ(G)Ps are aligned to have the same reference *z*-coordinate, as showed in panel f.

After alignment, z-coordinates outside the range of −20 and 17.5 Å were not considered (Fig. 4f). This sectional "trimming" was motivated by the observation that pore diameter profiles for several AQ(G)Ps (obtained by PoreWalker [99]) showed artifactually elongated z-coordinates at either the intracellular or extracellular end, or both (see Fig. 4 and Fig. S1). In these regions, the pore shape is less well defined, making it difficult to unambiguously determine the residues lining the channel.

Trimming the pore ends therefore helps to reduce the uncertainty in identifying the pore-lining residues and allows the application of a uniform framework across all the AQ(G)Ps, thereby enabling a more consistent calculation of pore properties, including average diameter, hydrophobicity, and B-factor–related pore descriptors.

All in all, thirteen (13) pore descriptors (PDs) were selected to conduct the pore-level QSAR modeling proposed in this study. Table 2 contains an overview of these PDs. A subset (*Reg*, *SF_Ave_Bfac*, *SF_Ave_Bfac_BB* and *SF_Ave_Bfac_SC*) was first scaled to the 0–10 range. Then, standardization was applied to selected descriptors included in the QSAR models. These preprocessing steps were implemented to minimize scale-related bias during model training.

2.4. Pore-level QSAR modeling and performance assessment

Given the limited number of AQ(G)Ps with compiled values of p_f or p_f^{corr} (n = 10 and n = 8, respectively) among those listed in Table 1, we use these entire sets to train the models. These two training datasets—containing the unitary log-transformed water permeabilities ($ln\left(p_f\right)$ or $ln\left(p_f^{corr}\right)$) together with the calculated PDs—are provided as **Supplementary Files** in CSV format.

The iterative regularization technique LASSO (Least Absolute Shrinkage and Selection Operator) [102] was used, along with the Ordinary Least Squares (OLS) multiple linear regression method (MLR), to select optimal combinations of PDs for QSAR modeling. LASSO repeatedly fits a regression model by applying an L1 penalty so that small coefficients are shrunk to zero, thereby removing irrelevant variables and helping to prevent overfitting and multicollinearity.

Statistical parameters obtained to assess the performance of the QSAR models include: R-squared (R^2, internal coefficient of determination), regression p-value (from a Fisher F-test), statistical significance (from a Student t-test, $\alpha = 0.05$) for each independent variable (PD) in the models and the independent term, as well as the predictive parameters: SPRESS: Standard Prediction Error Sum of Squares, and loo-Q^2 (leave-one-out cross-validation coefficient).

Regarding the parameters mentioned above, R^2 is a measure of how well the model fits the data used to train the model (desirable R^2 is close to 1, whereas R^2 close to 0 denotes bad performance). The loo-Q^2 parameter is a measure of the internal predictive ability, and not just the fit. This metric is obtained after: (i) removing one item from the dataset, (ii) training again the model with the rest of items in the dataset, (iii) predicting the left-out item using the re-trained model. This process is then repeated until every item has been left out once. Then, all the prediction errors are combined into Q^2. Thus, higher loo-Q^2 (crucially, >0.5 up to 1) means that the model predicts well data internally, whereas low loo-Q^2 means it predicts poorly (basically noise).

To evaluate collinearity among the retained predictors, we inspected the pairwise correlation matrix and calculated variance inflation factors (VIFs) for each final descriptor set. Candidate models displaying pairwise intercorrelations $|r| > 0.5$ and/or $VIF > 5$ were excluded. In addition, the condition number (CN) was calculated to assess the sensitivity of the regression coefficients to perturbations in the data; models with $CN > 5$ were considered numerically unstable and were not pursued further (additional details are provided in the **Supplementary Information**).

A y-scrambling analysis was also conducted. The y-scrambling (also called y-randomization) is a method to assess chance correlation issues. In QSAR, an unscrambled model is the natural relationship that is observed between the dependent Y variable(s) and the independent X variable(s). A y-scrambled model, instead, is obtained by deliberately randomizing (shuffling) the Y values in the dataset while keeping X unchanged across the analyzed items. This procedure is done several times so that allows to verify that none of the "randomized" (scrambled) models perform better than the observed (trained) model, thus excluding the possibility that the proposed model is a result of accidentally fitting noise.

For model assessment, we took advantage of the availability of four aquaporins that, despite lacking experimentally resolved 3D structures, have p_f^{corr} permeability values reported in the literature [56] (Table S2 and Fig. 2). These proteins were used as an external test set to evaluate one of the best-performing models developed in this study for the p_f^{corr} dataset (see **Results and Discussion**). We used AlphaFold2-predicted 3D structures [103,104] for this purpose, and computed the pore descriptors included in the model in the same manner as for the rest of the dataset. Since AlphaFold2 does not provide B-factor information, and its confidence metrics (pLDDT and PAE) have been shown not to be equivalent to experimental B-factors [105,106], we restricted this validation only to the QSAR model that do not rely on B-factor features.

A dedicated Python script was developed to perform all regression, statistical, and y-scrambling analyses. This script is also supplied, together with the AQ(G)P PDB or AlphaFold structure files, as **Supplementary Data**. Further details on the pore-level QSAR modeling implemented in this study can be found in the **Supplementary Information** document.

Table 2
Overview of the pore descriptors computed and analyzed in this work.

Pore Descriptor (PD)	PD Short Name	PD Type	Description & Computation
Averaged diameter	*Ave_diam*	Geometric	Diameter average computed from the diameter versus z-coordinate data provided by PoreWalker server [99], along the pore z-coordinate-'trimming' analyzed. Computed by a custom script.
Minimum diameter	*Min_diam*	Geometric	Minimum diameter identified from the diameter versus z-coordinate data provided by the PoreWalker server [99].
3D shape regularity	*Reg*	Geometric	Measure of how uniformly the pore geometry (cross-sectional symmetry) is maintained along the channel axis (z-coordinate). Reported by the PoreWalker server [99] as a percentage.
Averaged hydrophobicity indexes on pore-lining residues	*Ave_hydropho*	Hydropathy	Average of per-residue hydrophobicity along the pore z-coordinate-'trimming' analyzed. Hydropathy indexes from the Kyte-Doolittle scale [100], computed by Discovery Studio [98]. *Ave_hydropho* accounts for strictly hydrophobic amino acids (i.e., those with positive hydrophobicity).
Averaged hydrophilicity indexed on pore-lining residues	*Ave_hydrophi*	Hydropathy	Average of per-residue hydrophilicity along the pore z-coordinate-'trimming' analyzed. Hydropathy indexes from the Kyte-Doolittle scale [100], computed by Discovery Studio [98]. *Ave_hydrophi* accounts for strictly hydrophilic amino acids (i.e., those with negative Kyte-Doolittle hydrophobicity). These negative indexes were converted into positive values—by multiplying by −1—to favor better interpretability.
Averaged per-residue B-factor on pore-lining residues within the selectivity filter region	*SF_Ave_Bfac*	Mobility/ flexibility	Average of per-residue displacement (B-factor) on those residues located within the pore *z*-coordinate 2–6 Å (see Fig. 4a and Fig. S1). B-factor values extracted from the PDB files using a custom script.
Average side chain B-factor on pore-lining residues within the selectivity filter region	*SF_Ave_Bfac_SC*	Mobility/ flexibility	Average of per-residue side-chain displacement on those residues located within the pore *z*-coordinate 2–6 Å (see Fig. 4a and Fig. S1). Extracted from the PDB file by a custom script.
Averaged backbone B-factor on pore-lining residues within the selectivity filter region	*SF_Ave_Bfac_BB*	Mobility/ flexibility	Average of per-residue backbone displacement on those residues located within the pore *z*-coordinate 2–6 Å (see Fig. 4a and Fig. S1). Extracted from the PDB file by a custom script.
Averaged B-factor on pore-lining residues	*Ave_Bfac*	Mobility/ flexibility	Average of per-residue displacement on all the residues lining the pore. B-factor values extracted from the PDB files using a custom script.
Averaged side chain B-factor on the pore-lining residues	*Ave_Bfac_SC*	Mobility/ flexibility	Average of per-residue side chain displacement on all the residues lining the pore. B-factor values extracted from the PDB files using a custom script.
Averaged backbone B-factor on the pore-lining residues	*Ave_Bfac_BB*	Mobility/ flexibility	Average of per-residue backbone displacement on all the residues lining the pore. B-factor values extracted from the PDB files using a custom script.
Number of potential H-bonds with pore-lining residues along the single-file region	N_H	H-bonding	Number of potential H-bonding sites along the single-file region. The single-file water column defined following exactly the procedure described by Horner et al. [93] (see the caption of Fig. 3, panel a).
Positively charged Lys and Arg residues in the pore vestibules	Res^+	Electrostatic/ solvation	Number of positively charged Lys and Arg residues counted within the pore outer vestibules (with side-chain nitrogen atoms within 15 Å distance cut-off to the reference waters in their own vestibule, see Fig. 3, panel b). It is a measure of enrichment of positively charged sites at the pore mouths, computed by a custom script.

(*continued on next page*)

Table 2 (*continued*)

Pore Descriptor (PD)	PD Short Name	PD Type	Description & Computation
Positively charged side-chain nitrogen atoms from Lys and Arg in the pore vestibules	$Sites^{+}$	Electrostatic/ solvation	Number of positively charged side-chain nitrogen atoms from Lys and Arg residues counted within the pore outer vestibules (within 15 Å distance cut-off to the above-mentioned reference waters, see Fig. 3, panel b). It is another measure of enrichment of positively charged sites at the pore mouths, computed by a custom script.
Negatively charged Glu and Asp residues in the pore vestibules	Res^{-}	Electrostatic/ solvation	Number of negatively charged Asp and Glu residues counted within the pore outer vestibules (with side-chain oxygen atoms within 15 Å distance cut-off to the above-mentioned reference waters. It is a measure of enrichment of negatively charged sites at the pore mouths, computed by a custom script.
Positively charged side-chain nitrogen atoms from Lys and Arg in the pore vestibules	$Sites^{-}$	Electrostatic/ solvation	Number of negatively charged side-chain oxygen atoms from Asp and Glu residues counted within the pore outer vestibules (within 15 Å distance cut-off to the above-mentioned reference waters). It is another measure of enrichment of negatively charged sites at the pore mouths, computed by a custom script.

3. Results and discussion

3.1. On AQ(G)Ps' permeation rate

Two key functional elements of AQ(G)Ps are their selectivity and permeation properties. In the **Supplementary Information**, we provide a systematic description of the main structural motifs that determine these properties, highlighting the similarities and differences across the representative set of AQ(G)Ps investigated in this study (see Table S4, Fig. S2 and Fig. S3). AQ(G)Ps' pore shape (also analyzed in detail in the **Supplementary Information**), whose conical design—both in the cytoplasmic and the extracellular entrances (Fig. 4b, Fig. 4f and Fig. S1)—ensures the creation of highly permeable water channels [60] that can transport ~10^9 water molecules per second (per monomer) in a single-file manner [107–110]. Both experimental [1,56,65,107,108,110–112] and in silico [8,63,98,109,113] approaches have been devised to measure or simulate the water permeation rate of AQ(G)Ps.

Access to reliable p_f data of AQ(G)Ps is required to unravel the key structure–function relationships to a molecular level, as we intend here. Back in the early 90's, in vitro functional assays were developed to measure the osmotic water permeability of membrane water channels. Classical—though still used—swelling assays based either on the microinjection and expression of a transcribed protein sequence (typically the mRNA of the protein channel of interest) in Xenopus laevis oocytes (oocytes swelling assay: OSA) [1] or the reconstitution of the protein in liposomes (proteoliposome swelling assay: PSA) [107], have been indispensable tools for the estimation of the water permeability (P_f) and, by inference, unitary water permeabilities (p_f) for many AQ(G)Ps.

Despite the advances achieved in methodologies for measuring the water permeabilities of AQ(G)Ps [114]—particularly improvements in the physical techniques employed [65]—, the vast majority of unitary permeability measurements (p_f) have been obtained by indirect methods that heavily rely on the accurate estimation of the number of channels present in the sample. As stated by Gonzalez and collaborators, innovative methodological efforts are still needed to enable direct experimental measurements of the events occurring in water molecules inside AQ(G)P channels [51]. In the last decade, although still insufficient, some efforts have been made to obtain more direct permeation measurements [97] and to apply improved experimental strategies [115].

Notably, recent re-analyses have suggested that a non-negligible fraction of legacy stopped-flow–derived osmotic and unitary permeability estimates may be systematically overstated (up to 75%), owing to methodological assumptions that were common in earlier studies [56]. This has two practical implications for comparative structure–function analyses: first, absolute p_f values across the literature should be interpreted with caution unless they derive from modern analysis pipelines or include explicit corrections; and second, relative rankings and broad trends are generally more robust than the original numerical magnitudes.

On the computational side, most approaches have relied on molecular dynamics simulations (MDS), including simulations under equilibrium conditions [63,98,116–123], pressure-induced [98], and steered molecular dynamics [113]. However, Zhu et al. [98] reported that equilibrium MDS is not a reliable method for computing osmotic unitary water permeabilities and suggested instead imposing a hydrostatic pressure difference across the membrane to better approximate experimental conditions [98]. Despite these limitations, MDS-based approaches—unlike experimental methods—enable the direct computation of p_f under fully controlled conditions and have provided invaluable atom– and residue–level insights into filtration mechanisms, selectivity, and proton exclusion [50,63,85,98,117]. Moreover, they hold the potential to deepen our understanding of the behavior of protein channels in lipid membranes and to enable detailed thermodynamic and stability analyses [124,125]. Such analyses may help rationalize, and ultimately mitigate, the stability issues reported for several AQ(G)Ps [36].

3.2. Pore-level QSAR study as a methodological proof of concept

3.2.1. Pore descriptors (PDs)

Several mechanistic studies have already attempted to quantitatively dissect how specific pore features modulate the unitary water permeability of narrow AQ(G)Ps, albeit typically by interrogating one factor at a time and/or using limited protein sets.

In particular, Horner et al. [57] proposed that single-file water mobility is governed by the number of potential hydrogen-bonds (N_H) between permeating waters and pore-lining residues, with transport limited by the kinetics of breaking and reforming these interactions during translocation. Closely related work further highlighted the role of entrance/exit dehydration–rehydration energetics, showing that the electrostatic environment at the pore mouth (e.g., enrichment of positively charged residues) can measurably boost p_f even when the single-file H-bonding capacity is held constant, based on comparisons across a

small panel of AQPs [58].

Complementary MDS-based analyses likewise indicated that permeability is not dictated by pore radius alone, but reflects a competition between channel geometry and the degree of single-file character along the permeation pathway [117]. Finally, experimental mutagenesis combined with in silico pore profiling has reinforced that subtle variations in constriction geometry and local physico-chemistry can differentially impact permeation [126], underscoring that "pore size" is rarely a sufficient standalone descriptor.

Notwithstanding these important advances—including the broader recognition that gating/regulatory phenomena can reversibly switch AQ(G)Ps between more and less permeable states under specific modulators (e.g., pH, ions, phosphorylation), adding another structural layer to consider—only a small number of pore descriptors have been rigorously benchmarked against extensive, harmonized experimental p_f datasets to date. This motivates a dedicated descriptor framework for the pore-lining environment, in which geometric, H-bonding, electrostatic/solvation-related features, and other features are systematically encoded and evaluated jointly within a unified quantitative modeling context.

In the proposed pore-level QSAR modeling, we consider all pore features described above and additionally include descriptors that, to our knowledge, have not been systematically examined in the literature on determinants of water permeation, such as pore regularity and hydrophobicity- and B-factor–related pore descriptors (see Table 2).

Computed hydrophobicity profiles (Fig. 4a and Fig. S1) show significantly more hydrophobic residues (positive indexes) lining the pore than hydrophilic ones (negative indexes), in agreement with previous reports [127,128]. These differences are less pronounced near and at the intra- and extracellular pore vestibules, where the pore diameter and water contact surface increase, and there is a higher concentration of hydrophilic residues. The hydrophobicity profiles, thus, predominantly show a downward bell-shaped smoothed curve, with the peak located around the central part of the pore, where the NPA and SF motifs are situated. It is important to note that, although hydrophobic residues predominate—particularly in the central region of the pore—the presence of hydrophilic residues interspersed among them could be essential for water passage [62,117]. In relation to this, a pore lining composition analysis of the AQ(G)Ps listed in Table 1 is provided in the **Supplementary Information** (see Fig. S4) for the reader's reference.

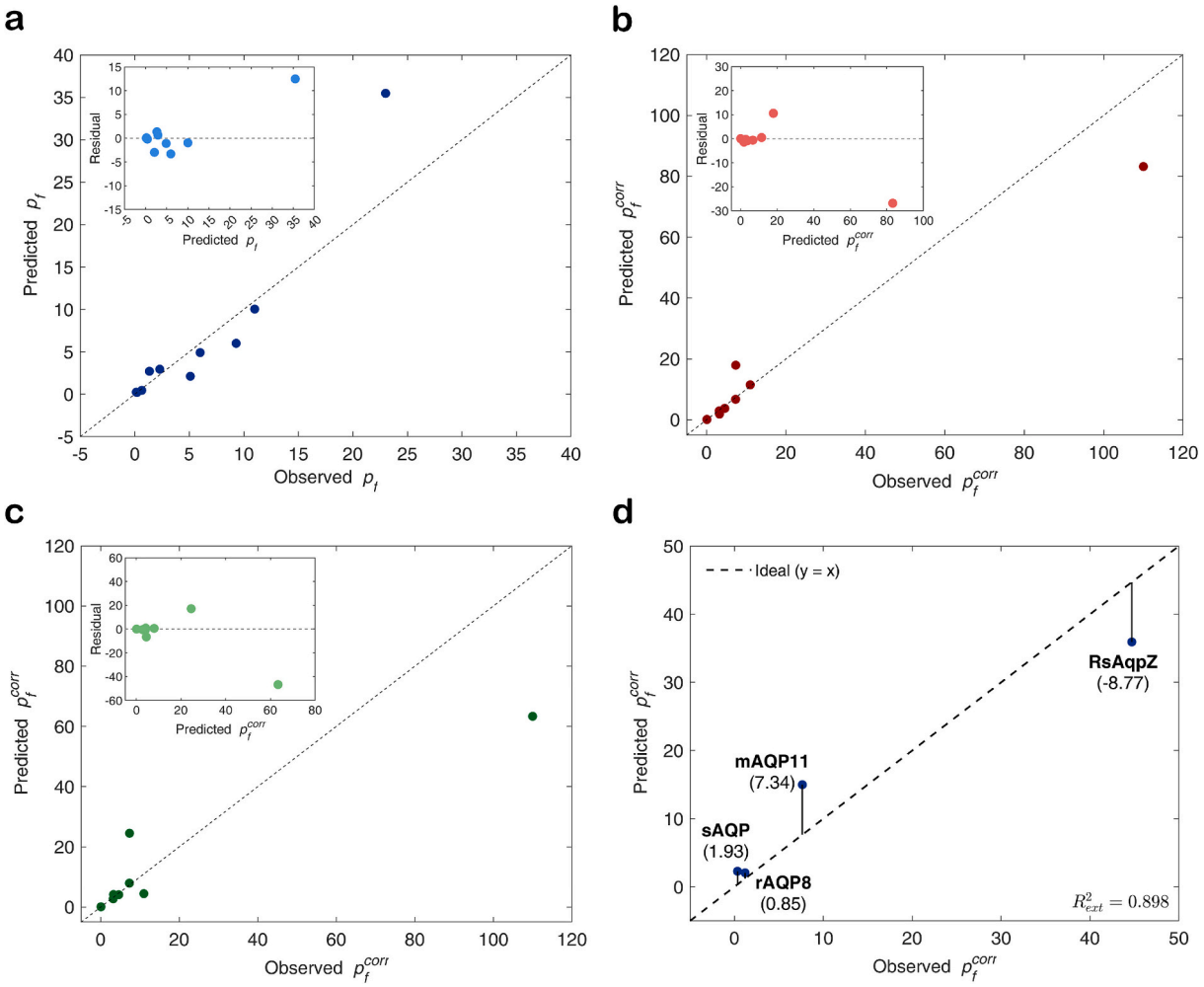


Fig. 5. Predictive performance and evaluation of the proposed pore-level QSAR models. (a) Predicted versus observed non-corrected permeability (p_f) values obtained from Model 1.3 (Table 4). (b) Predicted versus observed corrected permeability (p_f^{corr}) values obtained from Model 2.3a (Table 4). (c) Predicted versus observed corrected permeability (p_f^{corr}) values obtained from Model 2.3b (Table 4). In (a-c) the diagonal dashed lines show the 1:1 relationship, and the insets display residuals versus predicted p_f^{corr}. (d) Predicted versus observed p_f^{corr} values obtained with Model 2.3b for an independent set (rAQP8, mAQP11, RsAqpZ and sAqp; Table S6) used for external validation. The diagonal dashed line corresponds to the ideal prediction ($y = x$). The external coefficient of determination ($R^2_{Ext} = 0.898$) indicates a strong overall agreement between observed and predicted values. For each data point, the deviation from the diagonal (Δ = predicted – observed) is indicated in parentheses, highlighting over– and under–predictions by the model.

Moreover, B-factor profiles display the opposite shape, that is, an upward bell smoothed curve (Fig. 4a and Fig. S1). The B-factor values, which are typically derived from fitting X-ray diffraction data, measure the atomic displacement in molecular structures, indicating how much atoms deviate from their average positions due to thermal motion or structural disorder. According to these graphs, residues in the pore with the lowest B-factors are located in the region encompassing the pore constriction sites, namely the SF and NPA motifs. This is likely due to the proximity of the pore walls, which facilitates the formation of interaction networks that are crucial for the transport and selectivity exhibited by these proteins. Accordingly, within our QSAR framework, we evaluated B-factor–derived descriptors at the whole-residue, backbone, and side-chain levels, computed as averages over residues in the selectivity filter region as well as over the all pore-lining residues for the AQ(G)Ps analyzed.

We additionally computed B-factor descriptors at the whole-protein level and for the extra-pore region (i.e., non–pore-lining residues) using the same structure alignment and pore-trimming workflow. However, none of these descriptors was statistically significant and, accordingly, they provided no informative signal in our tests. These descriptors are therefore not included in Table 2, which focuses on pore-related descriptors. Instead, their computed values are provided in the **Supplementary Information** as a separate file for transparency.

3.2.2. *Pore-level QSAR models: Performance and mechanistic insights*

We present here the best-performing models ($R^2 > 0.9$ and loo-$Q^2 > 0.6$) obtained for the two permeability datasets, following the criteria defined in Section 2.4: (1) AQP(G)Ps with compiled p_f (n = 10), and (2) AQP(G)Ps with compiled p_f^{corr} (n = 8). Candidate QSAR models containing four or more PDs exhibited substantial multicollinearity and/or instability and were therefore discarded, even when their apparent goodness of fit was high (data not shown).

Model 1.3 *(dataset 1 of uncorrected permeabilities; 3 PDs; n = 10):*

$$ln\left(p_f\right) = -1.38(\pm 0.21)^{*}N_H + 0.98(\pm 0.20)^{*}SF_Ave_Bfac_SC - 1.18(\pm 0.22)^{*}Sites^{-} + 0.83(\pm 0.19)$$

This model exhibits an excellent fit ($R^2 = 0.92$; Adj. $R^2 = 0.88$; Prob (F) = 0.001) together with moderate-to-good internal predictivity (loo-$Q^2 = 0.71$). In real-scale terms, predicted p_f values remained within the same order of magnitude as the experimental (observed) values across all cases. The largest absolute deviation (12.5×10^{-14} $cm^3{\cdot}s^{-1}$) was observed for the highest permeability channel (rat AQP4; Table 4 and Fig. 5a), which is an expected feature of small-n QSAR models where high-leverage points can dominate.

The strongest effect in Model 1.3 is the negative coefficient for N_H, defined as the number of potential H-bonding sites along the single-file region. This indicates that increasing the density of H-bond-capable pore-lining residues is associated with reduced water permeability. The direction of this effect is consistent with the single-file transport framework, in which wall–water H-bonding introduces kinetic friction: permeating waters must repeatedly break and reform interactions with pore-lining residues, slowing collective motion through the constriction. The dependence on N_H in Model 1.3 is also consistent with prior work by Horner et al. [57], who reported an inverse relationship between p_f and N_H across a small set of AQ(G)Ps, protein channels, and nanotubes.

The positive coefficient for *SF_Ave_Bfac_SC* (average side-chain displacement/B-factor for residues within z = 2–6 Å, i.e., around the selectivity filter) suggests that greater local side-chain mobility in the constriction region correlates with increased water permeability in this dataset. A plausible mechanistic interpretation is that moderate flexibility could transiently relax steric or interaction constraints, smoothing the free-energy landscape for single-file water translation (e.g., by reducing the persistence of "sticky" configurations). However, because B-factors conflate genuine dynamics with crystal packing, resolution, and refinement procedures, this descriptor should be interpreted cautiously as a proxy for flexibility rather than a direct dynamical observable.

The negative coefficient for $Sites^{-}$ indicates that a higher density of negatively charged sites in the outer vestibules is associated with reduced p_f in this dataset. This observation is compatible with, but also extends, previous mechanistic analyses of aquaporin entrances. In particular, Horner et al. reported a clear positive association between positively charged residues at the channel mouth and single-file water flow, whereas the contribution of negatively charged residues appeared comparatively marginal in the small set of channels examined [58]. In Model 1.3, however, the LASSO procedure retained $Sites^{-}$—defined not as the number of negatively charged residues, but as the number of negatively charged oxygen atoms from Asp/Glu side chains within a defined vestibular proximity to reference waters (see Table 2)—suggesting that a more fine-grained representation of vestibular anionic functionality may capture effects that are not resolved by residue counts alone.

A plausible interpretation is that vestibular anionic groups can oppose water entry and throughput by increasing the energetic cost of reorganizing hydration upon approach to the pore mouth, thereby counteracting the permeability-enhancing effect of cationic sites (with a lower desolvation penalty, see discussion below for Model 2.3b) [58]. In addition, an increased abundance of Asp/Glu oxygen atoms provides additional opportunities for transient water–protein hydrogen bonding in the vestibular regions, which could slow water exchange into the single-file pathway by raising the local residence time and effectively increasing friction before and/or after the principal constriction. These mechanisms are not mutually exclusive and may act in concert. Given the limited sample size (n = 10), the contribution of $Sites^{-}$ is best interpreted as a mechanistically plausible association that merits targeted validation, rather than as a definitive causal determinant of p_f.

Model 2.3a *(dataset 2 of corrected permeabilities; 3 PDs; n = 8)*:

$$ln\left(p_f^{corr}\right) = -0.53(\pm 0.21)^{*}SF_Ave_Bfac_SC - 0.77(\pm 0.21)^{*}Reg - 1.55(\pm 0.21)^{*}N_H + 1.57(\pm 0.20)$$

This model shows an excellent goodness-of-fit ($R^2 = 0.95$; Adj. $R^2 = 0.92$; Prob(F) = 0.004) but a more modest internal predictability (loo-$Q^2 = 0.63$). In real-scale terms, predicted p_f^{corr} values remained within the same order of magnitude as the compiled (observed) permeabilities across all AQ(*G*)Ps. The largest absolute deviation was observed for *E. coli* GlpF (26.8×10^{-14} $cm^3{\cdot}s^{-1}$; Table 4 and Fig. 5b).

Notably, N_H remains a dominant negative term, reinforcing the robustness of the H-bond friction concept across datasets. The negative coefficient for *Reg* (PoreWalker-derived regularity [100]) suggests that a more uniform pore geometry along the pore axis is associated with lower permeability in the corrected dataset. At this stage, *Reg* is best viewed as a global structural correlate of permeability rather than a direct mechanistic determinant. Clarifying its physical meaning will likely require a more detailed analysis that resolves local radius variations along the pore and quantifies how these features couple to the free-energy barriers governing single-file water flow.

In contrast to Model 1.3, the coefficient associated with *SF_Ave_Bfac_SC* is negative in Model 2.3a. However, this term also displays the smallest effect size among the selected predictors (−0.53 in standardized units) and the weakest statistical support within the model ($P > |t| = 0.064$, whereas the remaining terms show P-values ≤0.01). Taken together, the modest magnitude, marginal significance, and sign reversal relative to the non-corrected-permeability model suggest that B-factor–derived descriptors could be particularly sensitive to changes in the response definition (i.e., corrected versus non-corrected permeabilities) and to dataset composition. In addition, the training set comprises structures obtained under heterogeneous experimental conditions (including resolution, temperature, pH, crystallization/buffer

conditions, and structure-determination method; see Table S6), all of which can influence crystallographic displacement parameters and limit the direct comparability of B-factors across AQ(G)Ps. Accordingly, at this stage, the *SF_Ave_Bfac_SC* term should be interpreted cautiously and revisited once a larger dataset becomes available.

Model 2.3b (*dataset 2 of corrected permeabilities; 3 PDs; n = 8*):

$$ln\left(p_f^{corr}\right) = 0.39(\pm 0.39)*Sites^{+} - 0.95(\pm 0.33)*Reg - 1.59(\pm 0.35)*N_H + 1.57(\pm 0.29)$$

This model also provides a strong fit ($R^2 = 0.90$; Adj. $R^2 = 0.826$; Prob(F) = 0.018) together with similar moderate internal predictivity (loo-$Q^2 = 0.65$). In real-scale terms, predicted p_f^{corr} values remained within the same order of magnitude as the compiled (observed) permeabilities for all channels, with the exception of *E. coli* GlpF, for which the absolute discrepancy reached 46.6 × 10^{-14} $cm^3 \cdot s^{-1}$ (Table 4 and Fig. 5c). Relative to Model 2.3a, Model 2.3b replaces the *SF_Ave_Bfac_SC* descriptor with $Sites^{+}$, defined as the number of positively charged Lys/Arg side-chain nitrogen atoms within the pore vestibules (Table 2). The positive coefficient for $Sites^{+}$ is consistent with previous evidence that positive charges at the pore mouth (entrances/exits) can boost water flow [58]. Accordingly, Model 2.3b supports the interpretation that vestibular cationic groups may facilitate water entry and/or reduce the energetic penalty associated with reorganization of hydration at the pore mouth, thereby promoting the initiation and maintenance of the single-file water column.

Importantly, Model 2.3b is the most amenable to external testing among the QSAR models presented here, as it relies on descriptors that can be computed from AlphaFold2 [103,104] structures (in contrast to B-factor–based terms). External evaluation on four AQPs (rAQP8, mAQP11, RsAqpZ and sAQP) shows good overall agreement between predicted and compiled p_f^{corr} values (Fig. 5d), with $R^2_{Ext} = 0.898$. In real-scale terms, absolute deviations ranged from 0.85 to 8.77 × 10^{-14} $cm^3 \cdot s^{-1}$; in relative terms, three AQPs were predicted within ~2-fold (rAQP8, mAQP11 and RsAqpZ), whereas the lowest-permeability case (sAQP) was overestimated (~6.5-fold). Given the very small external set (n = 4) and the wide range in p_f^{corr} values across this subset of AQPs, these results should be interpreted as preliminary evidence of transferability rather than as a definitive validation.

Overall, N_H displays a consistent negative association with permeability across all models, supporting the view that an increased density of potential wall–water hydrogen-bonding sites in the single-file region imposes an effective friction that limits water throughput. By contrast, descriptors derived from crystallographic B-factors show dataset-dependent behavior, including a reversal in coefficient sign between the non-corrected and corrected permeability datasets, reinforcing their interpretation as empirical proxies rather than robust mechanistic determinants. Importantly, vestibular electrostatics emerge as a complementary modulator of permeability: the positive association of $Sites^{+}$ with p_f^{corr}, together with the negative trend observed for $Sites^{-}$ in the non-corrected dataset, is consistent with the notion that charge distribution at the pore mouth can influence water entry/exit by altering hydration reorganization and the associated desolvation penalty.

From a statistical standpoint, multicollinearity was negligible (VIF ≈ 1–1.76; condition number ≤ 2.21) across all models, supporting coefficient stability and interpretability. Moreover, y-scrambling analysis (100 permutations; Fig. S5, panels b, d and f) yielded no scrambled model with superior R^2/loo-Q^2 compared with the corresponding original model, which argues against chance correlations being responsible for the reported fits. All statistical parameters for the QSAR models are summarized in Table 3, and the per-descriptor LASSO coefficients resulted after the application of the feature-selection procedure are shown in the bar plots in Fig. S5 (panels a, c and e).

3.2.3. *Challenges and main limitations of the proposed methodology*

Modeling water permeation through narrow channels such as AQPs and AQGPs remains intrinsically challenging. This complexity stems, on the one hand, from the coupling between local and global phenomena occurring both within and around the pore, including auto-regulatory processes (e.g., pore gating; discussed in more detail in the **Supplementary Information**, Fig. S6) and post-translational modifications [13,87,129–132]. On the other hand, defining descriptors that faithfully capture the structural and physicochemical determinants of permeation is not straightforward a priori.

In addition, the still limited availability of experimentally determined osmotic permeability coefficients (p_f) across AQ(G)Ps precludes the development of more robust predictive models based on extended training sets and, at the same time, restricts the possibility of reserving independent external test sets for rigorous validation.

Beyond data availability, substantial room remains to devise and validate new pore features / descriptors that encode the physical determinants of pore-mediated filtration and selectivity, an area where

Table 3
Best-performing pore-level QSAR models and their associated statistical fit and predictive performance metrics.

Training Dataset[a]	No. PDs in model[c]	Model ID	PD1 (coeff. ± SE)[d]	PD2 (coeff. ± SE)[d]	PD3 (coeff. ± SE)[d]	Indep. term (± SE)[d]	R^2	P-value (F-test)	CN[e]	VIF[f]	loo-Q^2[g]	SPRESS[h]
1 $ln\left(p_f\right)$	3	1.3	N_H (−1.38 ± 0.21)***	*SF_Ave_Bfac_SC* (0.98 ± 0.20)***	$Sites^{-}$ (−1.18 ± 0.22)***	0.84 ± 0.19***	0.92	0.001	1.68	1.20; 1.08; 1.29	0.71	0.87
2 $ln\left(p_f^{corr}\right)$	3	2.3a	*Ave_Bfac_SC* (−0.53 ± 0.21)*	*Reg* (−0.77 ± 0.21)**	N_H (−1.55 ± 0.21)***	1.57 ± 0.20***	0.95	0.004	1.38	1.07; 1.03; 1.11	0.63	1.12
		2.3b	$Sites^{+}$ (0.39 ± 0.38)*	*Reg* (−0.95 ± 0.33)**	N_H (−1.59 ± 0.35)***	1.57 ± 0.29***	0.90	0.018	2.21	1.30; 1.41; 1.76	0.65	1.09

[a] Water permeability datasets used to train the QSAR models presented in the rows: 1) dataset of individual or averaged p_f measurements—without correction—for: bAQP0, oAQP0, bAQP1, hAQP1, hAQP2, hAQP4, rAQP4, mAqpM, eAqpZ and eGlpF (as reported in Table S1 and depicted in Fig. 2); 2) dataset of corrected water permeabilities for: oAQP0, bAQP1, hAQP1, hAQP2, hAQP4, rAQP4, eAqpZ and eGlpF (as reported in Table S2 and depicted in Fig. 2).

[c] Number of PDs included in the QSAR models presented in the rows.

[d] S.E. stands for standard error, and the number of asterisks shown over the values indicates the statistical significance (*p*-value from a Student *t*-test) obtained for the PDs or the independent term of the models, namely: * if $p < 0.1$, ** if $p < 0.05$, and *** if $p < 0.01$.

[e] The condition number (CN) obtained for each model. A summarized description of CN is given in the **Supplementary Information**.

[f] The Variance Inflation Factor (VIF) for all the variables (PDs) in the models. A summarized description of VIF is also provided in the **Supplementary Information**.

[g] loo-Q^2 is the leave-one-out cross-validation coefficient used to assess the internal predictability of the models.

[h] SPRESS stands for Standard Prediction Error Sum of Squares.

Table 4
Measured or corrected water permeation rates (observed), predicted values, and residuals obtained from Models 1.3, 2.3a and 2.3b in Table 3.[a]

AQ(G)P (PDB)	Observed p_f ($\times 10^{-14}$ cm^3•s^{-1})[b]	Predicted p_f ($\times 10^{-14}$ cm^3•s^{-1})	Residual p_f (Pred. minus Obs.)	Observed p_f^{corr} ($\times 10^{-14}$ cm^3•s^{-1})[c]	Predicted p_f^{corr} ($\times 10^{-14}$ cm^3•s^{-1})	Residual p_f^{corr} (Pred. minus Obs.)	Observed p_f^{corr} ($\times 10^{-14}$ cm^3•s^{-1})[c]	Predicted p_f^{corr} ($\times 10^{-14}$ cm^3•s^{-1})	Residual p_f^{corr} (Pred. minus Obs.)
	Model 1.3			Model 2.3a			Model 2.3b		
eGlpF (1FX8)	1.35	2.71	1.36	110.0	83.2	−26.8	110.0	63.4	−46.6
bAQP1 (1J4N)	2.3	2.95	0.65	3.2	2.93	−0.27	3.2	2.75	−0.45
eAqpZ (1RC2)	6.0	4.91	−1.09	7.33	6.75	−0.58	7.33	7.95	0.62
bAQP0 (1YMG)	0.15	0.25	0.1	–	–	–	–	–	–
oAQP0 (2B6O)	0.2	0.21	0.01	0.09	0.11	0.02	0.09	0.11	0.02
rAQP4 (2D57)	23.0	35.5	12.5	7.36	17.94	10.58	7.36	24.52	17.16
mAqpM (2F2B)	0.65	0.46	−0.19	–	–	–	–	–	–
hAQP4 (3GD8)	11.0	10.0	−1.0	11.0	11.5	0.5	11.0	4.45	−6.55
hAQP2 (4NEF)	9.3	6.0	−3.3	3.27	1.89	−1.38	3.27	4.17	0.9
hAQP1 (7UZE)	5.1	2.12	−2.98	4.6	3.77	−0.83	4.6	4.12	−0.48

[a] All unitary permeation rates and their associated residuals are presented on the original (non-log) scale, obtained by back-transforming the modeled log-transformed data.
[b] Experimental water permeation rates (averages if more than one is reported for a given AQ(G)P), as listed in Table S1 (original p_f values, i.e., without the correction proposed by Wachlmayr et al. [56]).
[c] Corrected water permeation rates [56] compiled in Table S2 and Fig. 2.

computational approaches could play a particularly important role. In this context, MDS-based methodologies could be leveraged to derive dynamic descriptors that go beyond static structural proxies. This perspective is increasingly feasible given the availability of high-quality structure predictors such as AlphaFold [103,104], together with growing access to high-performance computing resources. In particular, AlphaFold-derived structures of AQ(G)Ps can enable descriptor calculation and, consequently, support broader model development and validation—as illustrated here by the external set used to validate Model 2.3b (Table S6)—particularly as larger datasets of AQ(G)Ps with experimentally measured osmotic permeability become available.

It should also be noted that crystal structures, such as those used here, are not fully representative of proteins embedded in a lipid membrane. However, because our modeling framework focuses on the pore region, it is reasonable to assume that pore geometry and key constriction features are comparatively less sensitive to the surrounding membrane environment than other protein regions.

For this study, AQ(G)P structures solved at the highest available resolution were selected. In addition, to avoid biases and minimize inter-structure differences, all B-factor-based descriptors were computed after structural alignment of all AQ(G)Ps and applying identical pore segment definitions (z-coordinate pore trimming in the range of −20 to 17.5 Å). Nonetheless, modest differences remain across structures in terms of resolution, crystallization conditions (e.g., temperature, pH, and solvent), and data-collection protocols across the AQ(G)Ps analyzed here (Table S5). This issue limits direct comparisons between AQ(G)Ps and, particularly for models relying on B-factor features, complicate interpretation in terms of "true" dynamics or flexibility.

Despite these limitations, the good statistical metrics and predictive performance achieved by the selected QSAR models, together with their consistency with previously reported experimental observations (see previous section), render the proposed approach particularly encouraging. Importantly, this framework enables the joint assessment of multiple pore features and their combined effects on permeability. In our view, the results presented here may also motivate continued improvements in experimental procedures for determining osmotic water permeability, as well as the development of richer and more physically grounded pore descriptors, ultimately extending the applicability and robustness of this QSAR framework in future studies. Nonetheless, experimental validation and prospective design tests will be required before any predictive or engineering capability can be firmly established.

4. Conclusions

A pore-level QSAR approach is proposed to help identify key characteristics influencing water permeation through molecular pores. The best QSAR models achieved, while showing good correlation and moderate-to-good internal predictive performance, still present some limitations due to the scarcity of experimental permeation rate data for AQ(G)Ps, which hinders the use of broader training sets. Nevertheless, this constraint does not diminish the potential of the approach for gaining a deeper residue-level understanding of AQ(G)Ps' filtering properties, particularly by integrating multiple pore descriptors within a single quantitative framework rather than interrogating them individually. We believe that this methodology holds significant promise and, as additional permeation data become available, could yield a more comprehensive view of the pore permeation process. Ultimately, it may also support the rational design of highly efficient AQ(G)Ps or other biomimetic molecules with tailored filtration properties and provide further insights into the molecular mechanisms underlying AQ(G)P-related diseases.

CRediT authorship contribution statement

Juan José Galano-Frutos: Writing – review & editing, Writing –

original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Luca Bergamasco:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Paolo Vigo:** Writing – original draft, Investigation, Data curation. **Matteo Morciano:** Writing – review & editing. **Matteo Fasano:** Writing – review & editing. **Davide Pirolli:** Writing – review & editing. **Eliodoro Chiavazzo:** Writing – review & editing. **Maria Cristina De Rosa:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbamem.2026.184529.

Data availability

Files with the raw and processed data, and the Python script prepared for the pore-level QSAR performed can be accessed from: https://zenodo.org/records/18675157.

References

[1] G.M. Preston, T.P. Carroll, W.B. Guggino, P. Agre, Appearance of water channels in Xenopus oocytes expressing red cell CHIP28 protein, Science 256 (1992) 385–387, https://doi.org/10.1126/science.256.5055.385.

[2] P. Agre, S. Sasaki, M.J. Chrispeels, Aquaporins: a family of water channel proteins, American Journal of Physiology-Renal Physiology (1993), https://doi.org/10.1152/ajprenal.1993.265.3.F461.

[3] P. Agre, M.J. Borgnia, M. Yasui, J.D. Neely, J. Carbrey, D. Kozono, et al., Discovery of the aquaporins and their impact on basic and clinical physiology, Curr. Top. Membr. 51 (2001) 1–38, https://doi.org/10.1016/S1063-5823(01)51003-0.

[4] J.J. Zwiazek, H. Xu, X. Tan, A. Navarro-Ródenas, A. Morte, Significance of oxygen transport through aquaporins, Sci. Rep. 7 (2017) 40411, https://doi.org/10.1038/srep40411.

[5] V. Lindahl, P. Gourdon, M. Andersson, B. Hess, Permeability and ammonia selectivity in aquaporin TIP2;1: linking structure to function, Sci. Rep. 8 (2018) 2995, https://doi.org/10.1038/s41598-018-21357-2.

[6] H. Wang, S. Schoebel, F. Schmitz, H. Dong, K. Hedfalk, Characterization of aquaporin-driven hydrogen peroxide transport, Biochim. Biophys. Acta Biomembr. 1862 (2020) 183065, https://doi.org/10.1016/j.bbamem.2019.183065.

[7] B. Yang, A. Verkman, Water and glycerol permeabilities of aquaporins 1–5 and MIP determined quantitatively by expression of epitope-tagged constructs inXenopus oocytes, J. Biol. Chem. 272 (1997) 16140–16146, https://doi.org/10.1074/jbc.272.26.16140.

[8] E. Tajkhorshid, P. Nollert, M.Ø. Jensen, L.J. Miercke, J. O'Connell, R.M. Stroud, et al., Control of the selectivity of the aquaporin water channel family by global orientational tuning, Science (New York, N.Y.) 296 (2002) 525–530, https://doi.org/10.1126/science.1067778.

[9] S.G. Goto, R.E. Lee, D.L. Denlinger, Aquaporins in the Antarctic midge, an extremophile that relies on dehydration for cold survival, Biol. Bull. 229 (2015) 47–57, https://doi.org/10.1086/BBLv229n1p47.

[10] A. Gautam, A.K. Pandey, Aquaporins responses under challenging environmental conditions and abiotic stress tolerance in plants, Bot. Rev. 87 (2021) 467–495, https://doi.org/10.1007/s12229-021-09249-z.

[11] R. Kapilan, M. Vaziri, J.J. Zwiazek, Regulation of aquaporins in plants under stress, Biol. Res. 51 (2018) 4, https://doi.org/10.1186/s40659-018-0152-0.

[12] B.N. Philip, A.J. Kiss, R.E. Lee, The protective role of aquaporins in the freeze-tolerant insect *Eurosta solidaginis* : functional characterization and tissue abundance of EsAQP1, J. Exp. Biol. 214 (2011) 848–857, https://doi.org/10.1242/jeb.051276.

[13] Yang B, Editor. Aquaporins. Vol. 1398. Second Edition. Singapore: Springer Nature; 2023. doi:https://doi.org/10.1007/978-981-19-7415-1.

[14] A.S. Verkman, Aquaporins in Clinical Medicine, Annu. Rev. Med. 63 (2012) 303–316, https://doi.org/10.1146/annurev-med-043010-193843.

[15] A.S. Verkman, Aquaporins, Curr. Biol. 23 (2013) R52–R55, https://doi.org/10.1016/j.cub.2012.11.025.

[16] A.K. Azad, T. Raihan, J. Ahmed, A. Hakim, T.H. Emon, P.A. Chowdhury, Human Aquaporins: functional diversity and potential roles in infectious and non-infectious diseases, Front. Genet. 12 (2021) 654865, https://doi.org/10.3389/fgene.2021.654865.

[17] A. Dutta, M. Das, Deciphering the role of aquaporins in metabolic diseases: a mini review, Am J Med Sci 364 (2022) 148–162, https://doi.org/10.1016/j.amjms.2021.10.029.

[18] F. Vallese, K. Kim, L.Y. Yen, J.D. Johnston, A.J. Noble, T. Calì, et al., Architecture of the human erythrocyte ankyrin-1 complex, Nat. Struct. Mol. Biol. 29 (2022) 706–718, https://doi.org/10.1038/s41594-022-00792-w.

[19] C.M. Krane, D.L. Goldstein, Comparative functional analysis of aquaporins/glyceroporins in mammals and anurans, Mammalian Genome : Official Journal of the International Mammalian Genome Society 18 (2007) 452–462, https://doi.org/10.1007/s00335-007-9041-5.

[20] L. Calvanese, M. Pellegrini-Calace, R. Oliva, *In silico* study of human aquaporin AQP11 and AQP12 channels, Protein Sci. 22 (2013) 455–466, https://doi.org/10.1002/pro.2227.

[21] M. Suzuki, Y. Shibata, Y. Ogushi, R. Okada, Molecular machinery for Vasotocin-dependent Transepithelial water movement in amphibians: Aquaporins and evolution, Biol. Bull. 229 (2015) 109–119, https://doi.org/10.1086/BBLv229n1p109.

[22] H. Tong, Q. Hu, L. Zhu, X. Dong, Prokaryotic aquaporins, Cells 8 (2019) 1316, https://doi.org/10.3390/cells8111316.

[23] N. Pettersson, C. Filipsson, E. Becit, L. Brive, S. Hohmann, Aquaporins in yeasts and filamentous fungi, Biol. Cell. 97 (2005) 487–500, https://doi.org/10.1042/BC20040144.

[24] M. Kumar, M. Grzelakowski, J. Zilles, M. Clark, W. Meier, Highly permeable polymeric membranes based on the incorporation of the functional water channel protein aquaporin Z, Proc. Natl. Acad. Sci. USA 104 (2007) 20719–20724, https://doi.org/10.1073/pnas.0708762104.

[25] Z. Yang, C. Wu, C.Y. Tang, Making waves: why do we need ultra-permeable nanofiltration membranes for water treatment? Water Res X 19 (2023) 100172 https://doi.org/10.1016/j.wroa.2023.100172.

[26] C.H. Nielsen, Biomimetic membranes for sensor and separation applications, Anal. Bioanal. Chem. 395 (2009) 697–718, https://doi.org/10.1007/s00216-009-2960-0.

[27] Y. Li, S. Qi, M. Tian, W. Widjajanti, R. Wang, Fabrication of aquaporin-based biomimetic membrane for seawater desalination, Desalination 467 (2019) 103–112, https://doi.org/10.1016/j.desal.2019.06.005.

[28] A. Güvensoy-Morkoyun, S. Velioğlu, M.A. Göktuğ, S.B. Tantekin-Ersolmaz, Desalination potential of aquaporin-inspired functionalization of carbon nanotubes: bridging between simulation and experiment, ACS Appl. Mater. Interfaces 14 (2022) 28174–28185, https://doi.org/10.1021/acsami.2c03700.

[29] C. Tang, Z. Wang, I. Petrinić, A.G. Fane, C. Hélix-Nielsen, Biomimetic aquaporin membranes coming of age, Desalination 368 (2015) 89–105, https://doi.org/10.1016/j.desal.2015.04.026.

[30] C.Y. Tang, Y. Zhao, R. Wang, C. Hélix-Nielsen, A.G. Fane, Desalination by biomimetic aquaporin membranes: review of status and prospects, Desalination 308 (2013) 34–40, https://doi.org/10.1016/j.desal.2012.07.007.

[31] C. Hélix-Nielsen, Biomimetic membranes as a technology platform: challenges and opportunities, Membranes 8 (2018), https://doi.org/10.3390/membranes8030044.

[32] A. Azarafza, M.A. Islam, Y. Golpazirsorkheh, I. Efteghar, M. Sadrzadeh, M. Kamkar, et al., Aquaporin-based biomimetic membranes for low energy water desalination and separation applications, Adv. Funct. Mater. 33 (2023) 2213326, https://doi.org/10.1002/adfm.202213326.

[33] M.A. Shannon, P.W. Bohn, M. Elimelech, J.G. Georgiadis, B.J. Mariñas, A. M. Mayes, Science and technology for water purification in the coming decades, Nature 452 (2008) 301–310, https://doi.org/10.1038/nature06599.

[34] Space-Age Water Conservation | NASA Spinoff n.d. https://spinoff.nasa.gov/page/space-age-water-conservation-nasa (accessed April 27, 2024).

[35] S. Atkinson, Aquaporin's membranes play a vital role in effectively cleaning both produced and oil and gas flow-back water, Membr. Technol. 2020 (2020) 5–6, https://doi.org/10.1016/S0958-2118(20)30107-5.

[36] A. Beratto-Ramos, J. Dagnino-Leone, J. Martínez-Oyanedel, M. Aranda, R. Bórquez, Fabrication and filtration performance of aquaporin biomimetic membranes for water treatment, Sep. Purif. Rev. 51 (2022) 340–357, https://doi.org/10.1080/15422119.2021.1948865.

[37] M. Grzelakowski, M.F. Cherenet, Y.X. Shen, M. Kumar, A framework for accurate evaluation of the promise of aquaporin based biomimetic membranes, J. Membr. Sci. 479 (2015) 223–231, https://doi.org/10.1016/j.memsci.2015.01.023.

[38] A. Abdelrasoul, H. Doan, A. Lohi, C.-H. Cheng, Aquaporin-based biomimetic and bioinspired membranes for new Frontiers in sustainable water treatment technology: approaches and challenges, Polym. Sci., Ser. A 60 (2018) 429–450, https://doi.org/10.1134/S0965545X18040016.

[39] D. Cohen-Tanugi, R.K. McGovern, S.H. Dave, J.H. Lienhard, J.C. Grossman, Quantifying the potential of ultra-permeable membranes for water desalination, Energy Environ. Sci. 7 (2014) 1134–1141, https://doi.org/10.1039/C3EE43221A.

[40] M. Fasano, M. Morciano, L. Bergamasco, E. Chiavazzo, M. Zampato, S. Carminati, et al., Deep-sea reverse osmosis desalination for energy efficient low salinity enhanced oil recovery, Appl. Energy 304 (2021) 117661, https://doi.org/10.1016/j.apenergy.2021.117661.

[41] M. Qasim, M. Badrelzaman, N.N. Darwish, N.A. Darwish, N. Hilal, Reverse osmosis desalination: a state-of-the-art review, Desalination 459 (2019) 59–104, https://doi.org/10.1016/j.desal.2019.02.008.

[42] M. Morciano, M. Fasano, L. Bergamasco, A. Albiero, M.L. Curzio, P. Asinari, et al., Sustainable freshwater production using passive membrane distillation and waste heat recovery from portable generator sets, Appl. Energy 258 (2020) 114086, https://doi.org/10.1016/j.apenergy.2019.114086.
[43] G. Antonetto, M. Morciano, M. Alberghini, G. Malgaroli, A. Ciocia, L. Bergamasco, et al., Synergistic freshwater and electricity production using passive membrane distillation and waste heat recovered from camouflaged photovoltaic modules, J. Clean. Prod. 318 (2021) 128464, https://doi.org/10.1016/j.jclepro.2021.128464.
[44] I. Kocsis, Z. Sun, Y.M. Legrand, M. Barboiu, Artificial water channels—deconvolution of natural Aquaporins through synthetic design, Npj Clean Water 1 (2018) 1–11, https://doi.org/10.1038/s41545-018-0013-y.
[45] W. Song, M. Kumar, Artificial water channels: toward and beyond desalination, Curr. Opin. Chem. Eng. 25 (2019) 9–17, https://doi.org/10.1016/j.coche.2019.06.007.
[46] M. Barboiu, Y. Le Duc, A. Gilles, P.-A. Cazade, M. Michau, Y. Marie Legrand, et al., An artificial primitive mimic of the gramicidin-a channel, Nat. Commun. 5 (2014) 4142, https://doi.org/10.1038/ncomms5142.
[47] I.M. Andrei, A. Chaix, B.T. Benkhaled, R. Dupuis, C. Gomri, E. Petit, et al., Selective water pore recognition and transport through self-assembled alkyl-Ureido-Trianglamine artificial water channels, J. Am. Chem. Soc. 145 (2023) 21213–21221, https://doi.org/10.1021/jacs.3c02815.
[48] M.D. Vincenzo, A. Tiraferri, V.E. Musteata, S. Chisca, R. Sougrat, L.B. Huang, et al., Biomimetic artificial water channel membranes for enhanced desalination, Nat. Nanotechnol. 16 (2021) 190–196, https://doi.org/10.1038/s41565-020-00796-x.
[49] E. Kruse, N. Uehlein, R. Kaldenhoff, The aquaporins, Genome Biol. 7 (2006) 206, https://doi.org/10.1186/gb-2006-7-2-206.
[50] Groot B.L. de, H. Grubmüller, The dynamics and energetics of water permeation and proton exclusion in aquaporins, Curr. Opin. Struct. Biol. 15 (2005) 176–183, https://doi.org/10.1016/j.sbi.2005.02.003.
[51] M. Ozu, J.J. Alvear-Arias, M. Fernandez, A. Caviglia, A. Peña-Pichicoi, C. Carrillo, et al., Aquaporin gating: a new twist to unravel permeation through water channels, Int. J. Mol. Sci. 23 (2022) 12317, https://doi.org/10.3390/ijms232012317.
[52] M. Ozu, L. Galizia, C. Acuña, G. Amodeo, Aquaporins: more than functional monomers in a tetrameric arrangement, Cells 7 (2018) 209, https://doi.org/10.3390/cells7110209.
[53] Structure, Dynamics, and Function of Aquaporins n.d. https://www.ks.uiuc.edu/Research/aquaporins/ (accessed April 28, 2024).
[54] A. Froger, D. Thomas, C. Delamarche, B. Tallur, Prediction of functional residues in water channels and related proteins, Protein Sci. 7 (1998) 1458–1468, https://doi.org/10.1002/pro.5560070623.
[55] D. Fu, M. Lu, The structural basis of water permeation and proton exclusion in aquaporins (review), Mol. Membr. Biol. (2007), https://doi.org/10.1080/09687680701446965.
[56] J. Wachlmayr, L. Samineni, D.G. Knyazev, T. Barta, A. Speletz, C. Yao, et al., Biophysical quantification of unitary solute and solvent permeabilities to enable translation to membrane science, J. Membr. Sci. 675 (2023) 121308, https://doi.org/10.1016/j.memsci.2022.121308.
[57] A. Horner, F. Zocher, J. Preiner, N. Ollinger, C. Siligan, S.A. Akimov, 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. 1 (2015) e1400083, https://doi.org/10.1126/sciadv.1400083.
[58] A. Horner, C. Siligan, A. Cornean, P. Pohl, Positively charged residues at the channel mouth boost single-file water flow, Faraday Discuss. 209 (2018) 55–65, https://doi.org/10.1039/C8FD00050F.
[59] A. Horner, P. Pohl, Single-file transport of water through membrane channels, Faraday Discuss. 209 (2018) 9–33, https://doi.org/10.1039/C8FD00122G.
[60] S. Gravelle, L. Joly, F. Detcheverry, C. Ybert, C. Cottin-Bizonne, L. Bocquet, Optimizing water permeability through the hourglass shape of aquaporins, Proc. Natl. Acad. Sci. USA 110 (2013) 16367–16372, https://doi.org/10.1073/pnas.1306447110.
[61] N. Gössweiner-Mohr, C. Siligan, K. Pluhackova, L. Umlandt, S. Koefler, N. Trajkovska, et al., The hidden intricacies of Aquaporins: remarkable details in a common structural scaffold, Small 18 (2022) 2202056, https://doi.org/10.1002/smll.202202056.
[62] D. Fu, A. Lisbon, L.J. Miercke, C. Weitzman, P. Nollert, J. Krucinski, et al., Structure of a glycerol-conducting channel and the basis for its selectivity, Science 290 (2000) 481–486, https://doi.org/10.1126/science.290.5491.481.
[63] B.L. de Groot, H. Grubmüller, Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF, Science 294 (2001) 2353–2357, https://doi.org/10.1126/science.1066115.
[64] T. Gonen, T. Walz, The structure of aquaporins, Q. Rev. Biophys. 39 (2006) 361–396, https://doi.org/10.1017/S0033583506004458.
[65] E.I. Solenov, G.S. Baturina, L.E. Katkova, S.G. Zarogiannis, Methods to measure water permeability, in: B. Yang (Ed.), Aquaporins, Springer Netherlands, Dordrecht, 2017, pp. 263–276, https://doi.org/10.1007/978-94-024-1057-0_18.
[66] M. Xiong, C. Li, W. Wang, B. Yang, Protein Structure and Modification of Aquaporins, in: B. Yang (Ed.), Aquaporins, Springer Nature, Singapore, 2023, pp. 15–38, https://doi.org/10.1007/978-981-19-7415-1_2.
[67] A. Cherkasov, E.N. Muratov, D. Fourches, A. Varnek, I.I. Baskin, M. Cronin, et al., QSAR modeling: where have you been? Where are you going to? J. Med. Chem. 57 (2014) 4977–5010, https://doi.org/10.1021/jm4004285.
[68] E.N. Muratov, J. Bajorath, R.P. Sheridan, I.V. Tetko, D. Filimonov, V. Poroikov, et al., QSAR without borders, Chem. Soc. Rev. 49 (2020) 3525–3564, https://doi.org/10.1039/D0CS00098A.
[69] J. Emonts, J.F. Buyel, An overview of descriptors to capture protein properties – tools and perspectives in the context of QSAR modeling, Comput. Struct. Biotechnol. J. 21 (2023) 3234–3247, https://doi.org/10.1016/j.csbj.2023.05.022.
[70] M.H. Barley, N.J. Turner, R. Goodacre, Improved descriptors for the quantitative structure–activity relationship modeling of peptides and proteins, J. Chem. Inf. Model. 58 (2018) 234–243, https://doi.org/10.1021/acs.jcim.7b00488.
[71] PROSITE: Database of Protein Domains, Families and Functional Sites n.d. https://prosite.expasy.org/ (accessed October 16, 2023).
[72] C.J. Sigrist, Castro E. de, L. Cerutti, others, New and continuing developments at PROSITE, Nucleic Acids Res. 41 (2013) D344–D347.
[73] T. Paysan-Lafosse, M. Blum, S. Chuguransky, T. Grego, B.L. Pinto, G.A. Salazar, et al., InterPro in 2022, Nucleic Acids Res. 51 (2023) D418–D427, https://doi.org/10.1093/nar/gkac993.
[74] InterPro: Classification of protein families n.d. https://www.ebi.ac.uk/interpro/ (accessed April 26, 2024).
[75] H.M. Berman, K. Mitsuoka, J. Westbrook, Z.F. andG Gilliland, T.N. Bhat, H. Weissig, et al., The protein data bank, Nucleic Acids Res. 28 (2000) 235–242, https://doi.org/10.1093/nar/28.1.235.
[76] W.E. Harries, D. Akhavan, L.J. Miercke, S. Khademi, R.M. Stroud, The channel architecture of aquaporin 0 at a 2.2-A resolution, Proc. Natl. Acad. Sci. USA 101 (2004) 14045–14050, https://doi.org/10.1073/pnas.0405274101.
[77] T. Gonen, Y. Cheng, P. Sliz, Y. Hiroaki, Y. Fujiyoshi, S.C. Harrison, et al., Lipid-protein interactions in double-layered two-dimensional AQP0 crystals, Nature 438 (2005) 633–638, https://doi.org/10.1038/nature04321.
[78] H. Sui, B.G. Han, J.K. Lee, P. Walian, B.K. Jap, Structural basis of water-specific transport through the AQP1 water channel, Nature 41 (2001) 872–878, https://doi.org/10.1038/414872a.
[79] J. Zeng, F. Schmitz, S. Isaksson, J. Glas, O. Arbab, M. Andersson, et al., High-resolution structure of a fish aquaporin reveals a novel extracellular fold, Life Sci. Alliance 5 (2022), https://doi.org/10.26508/lsa.202201491.
[80] U. Kosinska Eriksson, G. Fischer, R. Friemann, G. Enkavi, E. Tajkhorshid, R. Neutze, Subangstrom resolution X-ray structure details aquaporin-water interactions, Science 340 (2013) 1346–1349, https://doi.org/10.1126/science.1234306.
[81] A. Frick, U.K. Eriksson, Mattia F. de, F. Oberg, K. Hedfalk, R. Neutze, et al., X-ray structure of human aquaporin 2 and its implications for nephrogenic diabetes insipidus and trafficking, Proc. Natl. Acad. Sci. USA 111 (2014) 6305–6310, https://doi.org/10.1073/pnas.1321406111.
[82] J.D. Ho, R. Yeh, A. Sandstrom, I. Chorny, W.E. Harries, R.A. Robbins, et al., Crystal structure of human aquaporin 4 at 1.8 A and its mechanism of conductance, Proc. Natl. Acad. Sci. USA 106 (2009) 7437–7442, https://doi.org/10.1073/pnas.0902725106.
[83] Y. Hiroaki, K. Tani, A. Kamegawa, N. Gyobu, K. Nishikawa, H. Suzuki, et al., Implications of the aquaporin-4 structure on array formation and cell adhesion, J. Mol. Biol. 355 (2006) 628–639, https://doi.org/10.1016/j.jmb.2005.10.081.
[84] R. Horsefield, K. Nordén, M. Fellert, A. Backmark, S. Törnroth-Horsefield, Scheltinga A.C.T. van, et al., High-resolution x-ray structure of human aquaporin 5, Proc. Natl. Acad. Sci. USA 105 (2008) 13327–13332, https://doi.org/10.1073/pnas.0801466105.
[85] S.W. de Mare, R. Venskutonytė, S. Eltschkner, B.L. de Groot, K. Lindkvist-Petersson, Structural basis for glycerol efflux and selectivity of human aquaporin 7, Structure 28 (2020) 215–222, https://doi.org/10.1016/j.str.2019.11.011.
[86] K. Gotfryd, A.F. Mósca, J.W. Missel, S.F. Truelsen, K. Wang, M. Spulber, et al., Human adipose glycerol flux is regulated by a pH gate in AQP10, Nat. Commun. 9 (2018) 1–11, https://doi.org/10.1038/s41467-018-07176-z.
[87] S. Törnroth-Horsefield, Y. Wang, K. Hedfalk, U. Johanson, M. Karlsson, E. Tajkhorshid, et al., Structural mechanism of plant aquaporin gating, Nature 439 (2006) 688–694, https://doi.org/10.1038/nature04316.
[88] J.K. Lee, D. Kozono, J. Remis, Y. Kitagawa, P. Agre, R.M. Stroud, Structural basis for conductance by the archaeal aquaporin AqpM at 1.68 Å, Proc. Natl. Acad. Sci. 102 (2005) 18932–18937, https://doi.org/10.1073/pnas.0509469102.
[89] Lee JK, Finer-Moore JS, Stroud RM. RCSB Protein Data Bank, 3NE2: *Archaeoglobus fulgidus* Aquaporin n.d. https://www.rcsb.org/structure/3ne2 (accessed October 6, 2023).
[90] Y. Saitoh, N. Mitani-Ueno, K. Saito, K. Matsuki, S. Huang, L. Yang, et al., Structural basis for high selectivity of a rice silicon channel Lsi1, Nat. Commun. 12 (2021) 6236, https://doi.org/10.1038/s41467-021-26535-x.
[91] A. Kirscht, S.S. Kaptan, G.P. Bienert, F. Chaumont, P. Nissen, Goot B.L. de, et al., Crystal structure of an ammonia-permeable aquaporin, PLoS Biol. 14 (2016) e1002411, https://doi.org/10.1371/journal.pbio.1002411.
[92] D.F. Savage, P.F. Egea, Y. Robles-Colmenares, J.D. O'Connell 3rd, R.M. Stroud, Architecture and selectivity in aquaporins: 2.5 a X-ray structure of aquaporin Z, PLoS Biol. 1 (2003) E72, https://doi.org/10.1371/journal.pbio.0000072.
[93] Liu Q, Hillerich B, Love J, NYCOMPS. RCSB Protein Data Bank, 3LLQ: Aquaporin structure from plant pathogen Agrobacterium Tumerfaciens n.d. https://www.rcsb.org/structure/3llq (accessed October 6, 2023).
[94] Z.E. Newby, J. O'Connell Iii, Y. Robles-Colmenares, S. Khademi, L.J. Miercke, R. M. Stroud, Crystal structure of the aquaglyceroporin PfAQP from the malarial parasite plasmodium falciparum, Nat. Struct. Mol. Biol. 15 (2008) 619–625, https://doi.org/10.1038/nsmb.1431.
[95] F.J. Moss, P. Mahinthichaichan, D.T. Lodowski, T. Kowatz, E. Tajkhorshid, A. Engel, et al., Aquaporin-7: a dynamic aquaglyceroporin with greater water and

glycerol permeability than its bacterial homolog GlpF, Front. Physiol. 11 (2020) 728, https://doi.org/10.3389/fphys.2020.00728.
[96] T. Katano, Y. Ito, K. Ohta, T. Yasujima, K. Inoue, H. Yuasa, Functional characteristics of aquaporin 7 as a facilitative glycerol carrier, Drug Metab. Pharmacokinet. 29 (2014) 244–248, https://doi.org/10.2133/dmpk.DMPK-13-RG-121.
[97] M. Palmgren, M. Hernebring, S. Eriksson, K. Elbing, C. Geijer, S. Lasič, et al., Quantification of the intracellular life time of water molecules to measure transport rates of human Aquaglyceroporins, J. Membr. Biol. 250 (2017) 629–639, https://doi.org/10.1007/s00232-017-9988-4.
[98] F. Zhu, E. Tajkhorshid, K. Schulten, Theory and simulation of water permeation in aquaporin-1, Biophys. J. 86 (2004) 50–57, https://doi.org/10.1016/S0006-3495(04)74082-5.
[99] M. Pellegrini-Calace, T. Maiwald, J.M. Thornton, PoreWalker: a novel tool for the identification and characterization of channels in transmembrane proteins from their three-dimensional Structure, PLoS Comput. Biol. 5 (2009) e1000440, https://doi.org/10.1371/journal.pcbi.1000440.
[100] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105–132, https://doi.org/10.1016/0022-2836(82)90515-0.
[101] Discovery Studio, BIOVIA, Dassault Systèmes, 2022.
[102] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. B. Methodol. 58 (1996) 267–288, https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
[103] O. Kovalevskiy, J. Mateos-Garcia, K. Tunyasuvunakool, AlphaFold two years on: validation and impact, Proc. Natl. Acad. Sci. 121 (2024) e2315002121, https://doi.org/10.1073/pnas.2315002121.
[104] J. Jumper, R. Evans, A. Pritzel, others, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589, https://doi.org/10.1038/s41586-021-03819-2.
[105] Y. Vander Meersche, J. Diharce, J.-C. Gelly, T. Galochkina, Flexibility or uncertainty? A critical assessment of AlphaFold 2 pLDDT, Structure 33 (2025) 2157–2163.e2, https://doi.org/10.1016/j.str.2025.09.001.
[106] O. Carugo, pLDDT values in AlphaFold2 protein models are unrelated to globular protein local flexibility, Crystals 13 (2023) 1560, https://doi.org/10.3390/cryst13111560.
[107] M.L. Zeidel, S.V. Ambudkar, B.L. Smith, P. Agre, Reconstitution of functional water channels in liposomes containing purified red cell CHIP28 protein, Biochemistry 31 (1992) 7436–7440, https://doi.org/10.1021/bi00148a002.
[108] M.L. Zeidel, S. Nielsen, B.L. Smith, S.V. Ambudkar, A.B. Maunsbach, P. Agre, Ultrastructure, pharmacologic inhibition, and transport selectivity of aquaporin CHIP in proteoliposomes, Biochemistry 33 (1994) 1606–1615, https://doi.org/10.1021/bi00172a042.
[109] M.Ø. Jensen, O.G. Mouritsen, Single-channel water permeabilities of Escherichia coli aquaporins AqpZ and GlpF, Biophys. J. 90 (2006) 2270–2284, https://doi.org/10.1529/biophysj.105.073965.
[110] T. Walz, B.L. Smith, M.L. Zeidel, A. Engel, P. Agre, Biologically active two-dimensional crystals of aquaporin CHIP, J. Biol. Chem. 269 (1994) 1583–1586, https://doi.org/10.1016/S0021-9258(17)42062-X.
[111] L.A. Coury, J.C. Mathai, G.V. Prasad, J.L. Brodsky, P. Agre, M.L. Zeidel, Reconstitution of water channel function of aquaporins 1 and 2 by expression in yeast secretory vesicles, Am. J. Phys. 274 (1998) F34–F42, https://doi.org/10.1152/ajprenal.1998.274.1.F34.
[112] P. Pohl, S.M. Saparov, M.J. Borgnia, P. Agre, Highly selective water channel activity measured by voltage clamp: analysis of planar lipid bilayers reconstituted with purified AqpZ, Proc. Natl. Acad. Sci. 98 (2001) 9624–9629, https://doi.org/10.1073/pnas.161299398.
[113] M.Ø. Jensen, S. Park, E. Tajkhorshid, K. Schulten, Energetics of glycerol conduction through aquaglyceroporin GlpF, Proc. Natl. Acad. Sci. USA 99 (2002) 6731–6736, https://doi.org/10.1073/pnas.102649299.
[114] D. Boytsov, C. Hannesschlaeger, A. Horner, C. Siligan, P. Pohl, Micropipette aspiration-based assessment of Single Channel water permeability, Biotechnol. J. 15 (2020) e1900450, https://doi.org/10.1002/biot.201900450.
[115] K. Yue, J. Jiang, P. Zhang, L. Kai, Functional analysis of aquaporin water permeability using an *Escherichia coli*-based cell-free protein synthesis system, Front. Bioeng. Biotechnol. 8 (2020), https://doi.org/10.3389/fbioe.2020.01000.
[116] R. Araya-Secchi, J.A. Garate, D.S. Holmes, T. Perez-Acle, Molecular dynamics study of the archaeal aquaporin AqpM, BMC Genomics 12 (Suppl. 4) (2011) S8, https://doi.org/10.1186/1471-2164-12-S4-S8.
[117] M. Hashido, M. Ikeguchi, A. Kidera, Comparative simulations of aquaporin family: AQP1, AQPZ, AQP0 and GlpF, FEBS Lett. 579 (2005) 5549–5552, https://doi.org/10.1016/j.febslet.2005.09.018.
[118] M. Hashido, A. Kidera, M. Ikeguchi, Water transport in aquaporins: osmotic permeability matrix analysis of molecular dynamics simulations, Biophys. J. 93 (2007) 373–385, https://doi.org/10.1529/biophysj.106.101170.
[119] A.R. Binesh, R. Kamali, Molecular dynamics insights into human aquaporin 2 water channel, Biophys. Chem. 207 (2015) 107–113, https://doi.org/10.1016/j.bpc.2015.10.002.
[120] D. Alberga, O. Nicolotti, G. Lattanzi, G.P. Nicchia, A. Frigeri, F. Pisani, et al., A new gating site in human aquaporin-4: insights from molecular dynamics simulations, Biochimica et Biophysica Acta (BBA)-Biomembranes 1838 (2014) 3052–3060, https://doi.org/10.1016/j.bbamem.2014.08.015.
[121] J.-A. Garate, N.J. English, J.M.D. MacElroy, Human aquaporin 4 gating dynamics in dc and ac electric fields: a molecular dynamics study, J. Chem. Phys. 134 (2011) 055110, https://doi.org/10.1063/1.3529428.
[122] J.S. Jung, R.V. Bhat, G.M. Preston, W.B. Guggino, J.M. Baraban, P. Agre, Molecular characterization of an aquaporin cDNA from brain: candidate osmoreceptor and regulator of water balance, Proc. Natl. Acad. Sci. USA 91 (1994) 13052–13056, https://doi.org/10.1073/pnas.91.26.13052.
[123] M. Alishahi, R. Kamali, Forced diffusion of water molecules through aquaporin-5 biomembrane; a molecular dynamics study, Biophysics and Physicobiology 15 (2018) 255–262, https://doi.org/10.2142/biophysico.15.0_255.
[124] J.J. Galano-Frutos, J. Sancho, Energy, water, and protein folding: a molecular dynamics-based quantitative inventory of molecular interactions and forces that make proteins stable, Protein Sci. 33 (2024) e4905, https://doi.org/10.1002/pro.4905.
[125] J.J. Galano-Frutos, F. Nerín-Fonz, J. Sancho, Calculation of protein folding thermodynamics using molecular dynamics simulations, J. Chem. Inf. Model. 63 (2023) 7791–7806, https://doi.org/10.1021/acs.jcim.3c01107.
[126] P.O. Saboe, C. Rapisarda, S. Kaptan, Y.-S. Hsiao, S.R. Summers, R. De Zorzi, et al., Role of pore-lining residues in defining the rate of water conduction by Aquaporin-0, Biophys. J. 112 (2017) 953–965, https://doi.org/10.1016/j.bpj.2017.01.026.
[127] D. Kozono, M. Yasui, L.S. King, P. Agre, Aquaporin water channels: atomic structure molecular dynamics meet clinical medicine, J. Clin. Invest. 109 (2002) 1395–1399, https://doi.org/10.1172/JCI15851.
[128] A.S. Verkman, A.K. Mitra, Structure and function of aquaporin water channels, Am. J. Physiol. Ren. Physiol. 278 (2000) F13–F28, https://doi.org/10.1152/ajprenal.2000.278.1.F13.
[129] A. Frick, M. Järvå, S. Törnroth-Horsefield, Structural basis for pH gating of plant aquaporins, FEBS Lett. 587 (2013) 989–993, https://doi.org/10.1016/j.febslet.2013.02.038.
[130] C. Tournaire-Roux, M. Sutka, H. Javot, E. Gout, P. Gerbeau, D.-T. Luu, et al., Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins, Nature 425 (2003) 393–397, https://doi.org/10.1038/nature01853.
[131] A. Yaneff, L. Sigaut, M. Marquez, K. Alleva, L.I. Pietrasanta, G. Amodeo, Heteromerization of PIP aquaporins affects their intrinsic permeability, Proc. Natl. Acad. Sci. 111 (2014) 231–236, https://doi.org/10.1073/pnas.1316537111.
[132] A. Yaneff, L. Sigaut, N. Gómez, C. Aliaga Fandiño, K. Alleva, L.I. Pietrasanta, et al., Loop B serine of a plasma membrane aquaporin type PIP2 but not PIP1 plays a key role in pH sensing, Biochim. Biophys. Acta Biomembr. 1858 (2016) 2778–2787, https://doi.org/10.1016/j.bbamem.2016.08.002.