Bioinformatics Advances, 2025, **00**, vbaf102
https://doi.org/10.1093/bioadv/vbaf102
Application Note

Software

MDTAP: a tool to analyze permeation events across membrane proteins

Sruthi Sundaresan[1,†], Palur Venkata Raghuvamsi[1,†], Thenmalarchelvi Rathinavelan[1,*]

[1]Department of Biotechnology, Indian Institute of Technology Hyderabad, Kandi, Telangana 502284, India

*Corresponding author. Department of Biotechnology, Indian Institute of Technology Hyderabad, Kandi, Telangana 502284, India. E-mail: tr@bt.iith.ac.in.

†= equal contribution.

Associate Editor: Michael Gromiha

Abstract

Motivation: Molecular dynamics (MD) simulations provide critical insights into the transport of solutes, solvents, and drug molecules across protein channels embedded in a membrane bilayer. However, identifying and analyzing the permeation events from complex simulation data remains as a challenging and laborious task. Thus, an automated tool that facilitates the capture of permeation events of any molecular type across any channel is essential to streamline MD trajectory analysis and enhance the understanding of biological processes in a timely manner.

Results: Molecular Dynamics Trajectory Analysis of Permeation (MDTAP) is a Linux/Mac-based software that automatically detects permeation events across membrane-embedded protein and nucleic acid channels. The tool accepts trajectories in DCD (CHARMM/NAMD) and PDB format (obtained from any MD simulation package) and employs bash scripts to analyze the input trajectories to characterize the molecular permeation. The efficiency of MDTAP is demonstrated using MD trajectories of *Escherichia coli* outer membrane protein Wzi and *E. coli* Aquaporin Z. MDTAP can also analyze permeation across heterogeneous lipid membranes and artificial nucleic acid channels, addressing their growing importance. Thus, MDTAP simplifies trajectory analysis and also reduces the need for manual inspection.

Availability and implementation: MDTAP is open-source and is freely available on GitHub (https://github.com/MBL-lab/MDTAP), including source code, installation instructions, and usage documentation.

1 Introduction

Molecular dynamics (MD) simulation has emerged as a computational microscope to capture the time-dependent conformational dynamics of biomolecules (Dror *et al.* 2012, Feng *et al.* 2019). Since its flourishing in the 1970s, the efficacy of MD simulation techniques in providing structural insights about biological phenomena has been well-demonstrated (Gao *et al.* 1989, Marszalek *et al.* 1999, Zhang *et al.* 2010, Cheatham and Case 2013, Littlefield *et al.* 2014, Schlick and Portillo-Ledesma 2021, Schlick *et al.* 2021). One such example is the translocation of small molecules like water, ions, or other solute molecules, which plays a pivotal role in maintaining cellular homeostasis across membrane-embedded protein channels (Lynch *et al.* 2020). For instance, the mechanism of water conduction by aquaporins, a water-conducting membrane protein that regulates the amount of water inside cells across all the kingdoms of life (Kruse *et al.* 2006), was elucidated in greater detail by MD simulations (de Groot and Grubmuller 2001, Tajkhorshid *et al.* 2002). Similarly, MD simulations helped understand the gating mechanisms of mechanosensitive (Gullingsrud and Schulten 2004), voltage-dependent anion channel (VDAC; Rui *et al.* 2011), potassium channels (Noskov *et al.* 2004), and so on (Beckstein *et al.* 2003, Carnevale *et al.* 2021, Niitsu and Sugita 2023, Freites and Tobias 2024, Acharya and Kleinekathofer 2025). MD simulations were also helpful in understanding the solute conduction mechanism across artificial nucleic acid channels (Joshi *et al.* 2023).

There are many tools to map the transmembrane channels, like CHAP (Klesse *et al.* 2019), HOLE (Smart *et al.* 1996), CAVER (Petrek *et al.* 2006), and MOLEonline (Pravda *et al.* 2018). Interestingly, the PerMM tool differs from the others (Smart *et al.* 1996, Petrek *et al.* 2006, Pravda *et al.* 2018, Klesse *et al.* 2019) as it provides the permeation coefficient and the translocation pathway of drug-like molecules through artificial and natural membranes (Lomize *et al.* 2019). However, to the best of our knowledge, there is no tool available to capture the solute/solvent conduction events across a membrane protein from the vast volume of MD trajectory in a simple fashion irrespective of the nature of the pore shape, pore size, and directionality of transportation [viz., unidirectional (extracellular to intracellular or vice-a-versa) or bidirectional]. To this end, a trajectory analysis tool, *Molecular Dynamics Trajectory Analysis of Permeation* (MDTAP), is developed here to capture and quantify the permeation events across protein and nucleic acid channels. MDTAP offers flexibility for the user to define the molecule of interest (water/ions/solute/small molecules) whose permeation across a protein or nucleic acid channel can be tracked using the PDB files or CHARMM/NAMD DCD files generated from the MD trajectory irrespective of the mode of conduction (viz., a single molecule in narrow channels or bulk

Received: January 31, 2025; **Revised:** April 8, 2025; **Editorial Decision:** April 18, 2025

molecules in wider channels which may either be unidirectional or bidirectional). MDTAP uses a distinct approach to facilitate the capture of molecules transported across a channel having a straight or tortuous diffusion path. To the best of our knowledge, there is no single platform available that can comprehensively characterize molecular permeation across a channel using MD trajectories. Thus, MDTAP is developed to seamlessly extract detailed information about the permeation path, the number of permeating events, the time taken by a molecule to permeate across the channel, the capacity of a channel to accommodate the permeating molecule of interest at any instance, and the net flux. The source code, installation, and usage instructions for MDTAP are freely available to the public on GitHub.

2 Methods

2.1 Development of a methodology to calculate permeation events

A molecule (e.g. water, ion, solute, or small molecule) is considered to be permeating only if it moves from one side of the channel to the other. However, the permeation path may not always follow a single-file trajectory (Sachdeva *et al.* 2016, Sachdeva *et al.* 2017), as seen in the outer membrane protein Wzi of *Escherichia coli* (Bushell *et al.* 2013) (Fig. 1A). Due to multiple entry and exit points in this protein, a water molecule entering the Wzi channel does not necessarily have to cross the channel during the MD simulation time. Thus, this complicates the understanding of water permeation mechanism(s). This study introduces an automated method to capture molecular permeation events across any channel type (viz., straight or tortuous with single or multiple entries/exits) from the MD simulation trajectories (Fig. 1B).

First, the channel's dimensions are defined by identifying the minimum (X_{min}, Y_{min}, and Z_{min}) and maximum (X_{max}, Y_{max}, and Z_{max}) coordinates along the X, Y, and Z axes. These coordinates can be extracted from the first frame of the trajectory or obtained from prior knowledge of the channel's geometry. Note that the minimum and maximum coordinates are identified by considering all the atoms (including hydrogens) of the channel based on the user-defined chain ID and/or seg ID. Exceptionally, for the Z-density profile submodule, while the minimum and maximum X and Y coordinates (of the channel) are determined as mentioned above, the Z_{min} and Z_{max} values are identified by considering all the atoms of both the channel as well as the reservoir to capture the permeation path of the molecule along the Z-direction (along which the channel is aligned). However, the user has the freedom to define the dimensions manually. Soon after obtaining the channel's dimensions, the permeating molecule(s) of interest (for instance, water, as depicted in Fig. 1A) falling within the defined channel dimensions in the given MD trajectory are identified using a representative atom of the molecule. For example, in the case of a water molecule, the oxygen atom can serve as the representative.

To improve the computational efficiency, a frequency cut-off defining the minimum number of occurrences of a molecule inside the channel is subsequently defined. For every molecule above the frequency threshold, the trajectories where the Z-coordinate falls within the range of Z_{min} to Z_{max} are considered for further analysis. The Z-coordinates of the first and last frames satisfying the frequency cut-off are designated as Z_1 and Z_2 to determine the permeation events. Additionally, the Z-coordinates of the molecule before entering the channel (frame before Z_1) and after exiting the channel (frame after Z_2) are checked to ensure the accurate identification of permeating or non-permeating molecule(s). This is achieved by comparing the Z-coordinates of the molecule (Z_1 and Z_2) with those of the channel (Z_{min} and Z_{max}), as described in the following. Non-permeation events can be of five different categories: (i) the molecule enters from Z_{max} direction but does not cross the channel (viz., $Z_1 > Z_{max}$ and $Z_{min} < Z_2 < Z_{max}$; indicated as 1 in Fig. 1B), (ii) the molecule enters from Z_{min} direction but does not cross the channel (viz., $Z_1 < Z_{min}$ and $Z_{min} < Z_2 < Z_{max}$), (iii) the molecule is already within the channel at the start ($Z_{min} < Z_1 < Z_{max}$) but exits from either side of the channel ($Z_2 > Z_{max}$ or $Z_2 < Z_{min}$; indicated as 2 in Fig. 1B), (iv) the molecule remains within the channel throughout ($Z_{min} < Z_1 < Z_{max}$ and $Z_{min} < Z_2 < Z_{max}$; indicated as 3 in Fig. 1B), or (v) the molecule enters and exits from the same side of the channel ($|Z_2 - Z_1| < |Z_{max} - Z_{min}|$) (indicated as 4 and 5 in Fig. 1B). If a molecule does not meet these criteria, it is likely to permeate the channel. Permeation events are divided into two types based on the direction: (i) the molecule crosses the channel from the Z_{max} to Z_{min} direction (direction 1; indicated as 6 in Fig. 1B) and (ii) the molecule crosses the channel from the Z_{min} to Z_{max} direction (direction 2; indicated as 7 in Fig. 1B). The direction of permeation is determined by checking the proximity of Z_1 to Z_{max} or Z_{min} at the entry frame and the proximity of Z_2 to Z_{max} or Z_{min} at the exit frame. However, some of the molecules might still not satisfy any of the above criteria due to periodic boundary conditions and are classified as non-permeating.

2.2 MDTAP software development: implementation, prerequisites, functionality, and validation

The above methodology is implemented as a Linux/Mac-based software named "MDTAP: Molecular Dynamics Trajectory Analysis of Permeation." The software executes this methodology in a Bourne Again Shell (bash) environment and utilizes a Fortran 90 script to read CHARMM or NAMD trajectories in DCD format, as DCD files are written in unformatted Fortran format. The Fortran 90 file is executed using a GNU Fortran compiler and is called as "gfortran." The software has multiple interactive modules, as detailed in the following sections, along with the input and output information. Since some modules generate output graphs for the user's preliminary inspection, the system executing the software should have gnuplot installed and can be accessed by the name "gnuplot."

Note that MDTAP is not limited to capturing permeation across transmembrane protein or nucleic acid channels; it also facilitates the characterization of molecular permeation across heterogeneous lipid membranes.

3 Results and discussion

3.1 Starting and exiting MDTAP

3.1.1 Rename module

The interface of the MDTAP software is given in Supplementary Table S1, wherein the user has to first select whether the trajectories are in PDB or DCD format. Subsequently, the "Rename" module (invoked by "A" or "a") has to be executed to rename the PDBs (any MD simulation package) or DCDs (CHARMM/NAMD) to make them accessible to MDTAP (Supplementary

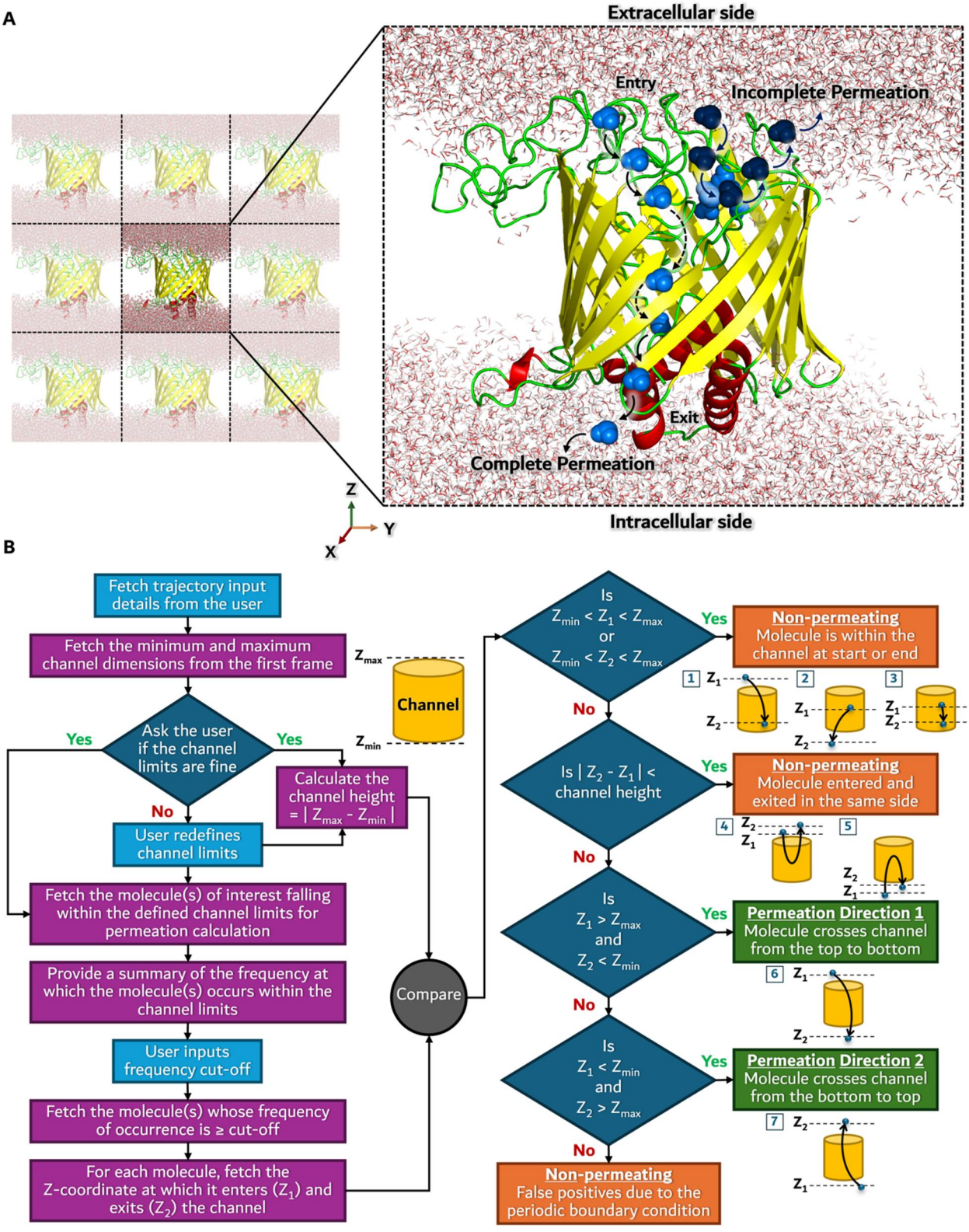

Figure 1. Methodological development of MDTAP. (A) A schematic representation describing complete and incomplete permeation events by considering the water-conducting *E. coli* Wzi protein channel (cartoon representation, PDB ID: 2YNK) as an example. Wzi in a periodic boundary condition is shown on the left. For the sake of clarity, POPE and POPG (3:1) membrane molecules are not shown. The water reservoir on either side of the channel is shown in line representation. While the light blue colored water molecule shows the complete permeation event, the dark blue colored water molecule shows an incomplete permeation. (B) Flowchart describing the methodology implemented in MDTAP to identify the permeation events in an automated fashion. See text for details. The yellow cylinder indicates the channel through which the molecule(s) is expected to permeate.

Table S2). This module renames the PDBs/DCDs in sequential order with the prefix "step_" and extension ".pdb" or ".dcd" (viz., step_1.pdb, step_2.pdb, step_3.pdb, etc., or step_1.dcd, step_2.dcd, step_3.dcd, etc).

3.1.2 Exit module

The user can leave the software interface using options "C" or "c" of the main module. It is important to note that while the examples discussed here have used PDB as the input trajectory format, MDTAP accepts both PDB and DCD files as input.

3.2 Analyze permeation module

Since parameters such as permeation coefficient, pore dimensions, and diffusion events are crucial for characterizing the permeation of a solvent/solute/small molecule(s) across a channel, the second module of the software, "Analyze Permeation Module" (APM), which is invoked by typing "B" or "b," includes nine submodules, as described in Table 1 and Supplementary Fig. S1. Notably, all the APM submodules assume that the channel is aligned with respect to the *Z*-axis of the Cartesian coordinate system. Figure 2 summarizes the outputs generated by the submodules of the "APM" by considering the previously published *E. coli* Wzi trajectory (Sachdeva *et al.* 2016) as a test case.

MDTAP automatically recommends the channel dimensions based on the minimum (X_{min}, Y_{min}, and Z_{min}) and maximum (X_{max}, Y_{max}, and Z_{max}) coordinates of the protein/nucleic acid channel along the *X*, *Y*, and *Z* axes. This helps capture the permeation event(s) even if the channel extends beyond the inner/outer membranes (extracellular, periplasmic, or cytoplasmic sides). One such example is Wza, an *E. coli* protein that spans beyond the outer membrane bilayer (Dong *et al.* 2006, Supplementary Fig. S3).

To facilitate the multichannel analysis, the software also allows the user to enter the segment and/or chain IDs of the channel of interest. For example, in a system like aquaporin with four channels labeled with seg IDs as "segA," "segB," "segC," and "segD" in the input PDB files, the user can specify "segA" to analyze the permeation across the first channel. In the absence of a seg ID in the PDB files (like in a single channel), the user may skip this option by pressing the enter key. Similarly, the user can use the chain ID as an input; in the absence of which in the PDB, the user may skip this option by pressing the enter key.

3.2.1 *Z*-density profile (APM 1)

This submodule provides information about the molecule's conduction path along the channel axis. After obtaining the necessary information from the user (Supplementary Table S3), the positions of the transporting molecule(s) along the *Z*-axis (viz., the channel axis) are determined. This is achieved by counting the number of molecules in each 1 Å block. The submodule generates individual plots and text files for each input PDB file and a time-averaged plot (with standard deviations) showing the "Number of molecules in every 1 Å block versus Position along *Z*-coordinate (Å)" (Fig. 3A, Supplementary Fig. S4A). Note that, unlike the other submodules, this submodule uses the maximum and minimum values of *Z*-coordinates of the entire system to fetch the *Z*-axis dimension of the channel. However, the user can also manually define the dimensions.

3.2.2 *XY* area profile (APM 2)

This submodule provides insight into the localization of the permeating molecules within the *XY* cross-sectional area of the channel. Its input format is similar to the previous submodule (Supplementary Table S4). However, while specifying the channel dimensions in the input, the user can refer to the plots generated by the previous submodule to identify appropriate upper and lower *Z*-axis limits. This submodule generates a series of two-dimensional (2D) plots to give the user a clue about the localization of the transporting molecule in the *XY* plane with respect to time (Supplementary Table S4). Along with the 2D plots, the submodule also generates a series of 3D plots (Supplementary Fig. S5) with respect to time by incorporating the *Z*-axis coordinate of the permeating molecule in the third dimension (for each input PDB). The submodule further generates a series of density plots (using all the input PDBs), which is the projection of the transporting molecule onto the respective *XY* plane for each 1 Å block defined along the *Z*-axis to precisely obtain the pore size (in the *XY* plane) of the channel (Fig. 3B, Supplementary Fig. S4B). An animated GIF file is also created by combining all the density plots (Supplementary Movie S1).

Table 1. List of MDTAP modules and their functionalities.

APM number	Submodule name	Function
1	*Z*-density profile	Identifies the variation in the population of permeating molecules along the channel axis (viz., *Z*-axis)
2	*XY* area profile	Identifies the aerial and spatial distribution of the molecule of interest projected to the *XY* plane
3	Rate of change of molecules	Calculates the number of molecules present inside the channel with respect to time
4	Permeation	Identifies and lists molecules that undergo complete and incomplete permeation through the protein
5	Net flux and permeability coefficient (P_d)	Provides the net flux and the diffusion permeability coefficient (P_d) of the molecule of interest across the channel
6	Residence time	Calculates the time (in ps) that the permeating molecule resides in the conduction path
7	Track molecule	Tracks the path followed by the permeating molecule along the pore axis (*Z*-axis) as it passes through the channel
8	Diffusion entry/exit	Calculates the number of molecules that enter or exit through a diffusion plane defined by three or more amino acids
9	Distance calculation	Calculates the distance between electronegative atom pairs of the molecule(s) of interest and the channel residues and reports atoms that are within a distance of 3.5 Å, along with the frequency of occurrence

Notes: The submodule number to be used in the software interface (Supplementary Table S1) to invoke the appropriate submodule is given in column 1. Note that these submodules do not need to be executed in a sequential order.

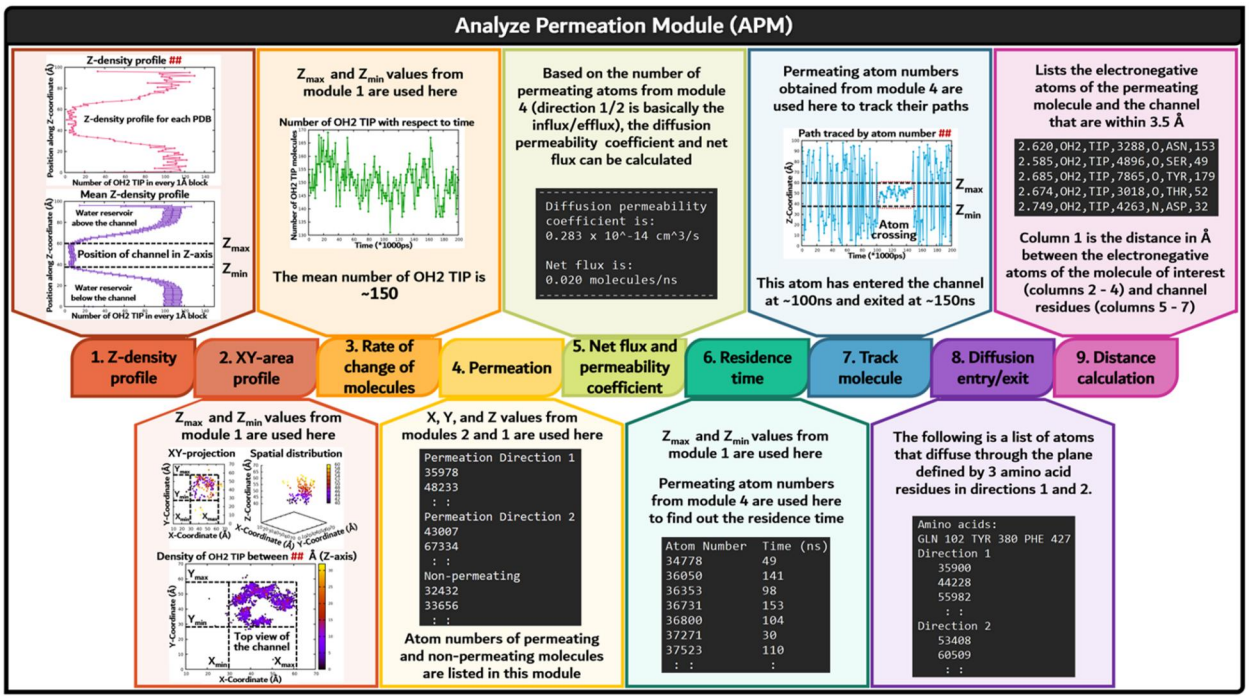

Figure 2. Schematic representation describing the outputs generated by the nine different submodules of the "APM" (refer to Table 1 for the functionality details and Supplementary Fig. S2 for the interlink between these submodules) by considering *E. coli* Wzi trajectory (200 ns) (Sachdeva *et al.* 2016) as a case in point is shown. The channel's position with respect to the *Z*-axis is obtained from the "*Z*-density profile" submodule and is used as the Z_{max} and Z_{min} values for the subsequent submodules. Similarly, the "*XY* area profile" submodule gives a clue about the channel's position with respect to the *XY* plane. The list of atom numbers that are found to permeate across the channel is obtained from the "Permeation" submodule (viz., the atom numbers of those molecules that undergo complete or true permeation) and is used in the subsequent submodules (refer to Supplementary Fig. S2). The red-colored double hashtags (##) in the graph titles indicate that it is a representative output of the respective submodule.

3.2.3 Rate of change of molecules (APM 3)

This submodule calculates the number of molecules inside the channel over the simulation time, providing insights into the channel's capacity to accommodate the molecule of interest. Defining the *X*-, *Y*-, and *Z*-axis dimensions based on the plots generated by the previous submodules is recommended. Once the required data is obtained (Supplementary Table S5), the submodule outputs a text file and a corresponding "Time (ps) versus Number of Molecules" graph, illustrating the distribution of the molecule inside the channel (Fig. 3C, Supplementary Fig. S4C). Additionally, the submodule generates a text file with the volume occupied by the molecule within the specified *X*, *Y*, and *Z* dimensions with respect to time in picoseconds (ps) by multiplying the number of molecules by the molar volume of the molecule of interest.

3.2.4 Permeation (APM 4)

This submodule calculates the number of molecules permeating through the channel, regardless of directionality (Supplementary Table S6). A "permeation event" is defined as the complete translocation or transport of water, ions, or small molecules through the channel from one reservoir to another (Supplementary Fig. S6). The submodule generates three output text files: Permeation-dir1.dat, Permeation-dir2.dat, and Non-permeating.dat, listing molecules that permeate in two distinct directions and those that do not permeate, respectively. Note that for unidirectional channels, only molecules crossing in one specific direction will be listed. To further visualize the path traced by the permeating molecule (s), the submodule also generates individual PDBs that consist of the positions of the permeating molecule during its residence inside the channel. Each PDB file corresponds to a specific atom number.

3.2.5 Net flux and permeability coefficient (P_d) calculation (APM 5)

Insights obtained from the "Permeation" submodule can be used for calculating the net flux and permeability coefficients using the "Net flux and permeability coefficient" submodule (Supplementary Table S7). The net flux is calculated by subtracting the total number of permeating molecules in one direction from the opposite direction (for bidirectional channels) during the simulation time frame. Such a net flux calculation from the simulation trajectories can be utilized to observe the channel conduction mechanism and determine the attainment of equilibrium, wherein the number of molecules translocating from both sides should be equal. This is done by counting the number of water molecules crossing the channel in two different directions by considering the total number of permeating molecules listed as direction 1 and direction 2 in the "Permeation" submodule. The resultant number will be the net flux. For the unidirectional channels, the number of permeating molecules will be zero in one of the directions. The submodule also calculates the diffusion permeability coefficient (P_d) as described elsewhere (Equation (1); Zhu *et al.* 2004):

$$P_d = \left(\frac{V_w}{N_A}\right) q_o \quad (1)$$

Where V_w is the molar volume of the molecule of interest, which is 18.07 cm^3 for water, N_A is Avogadro's number, and q_o is the number of permeation events in one direction per

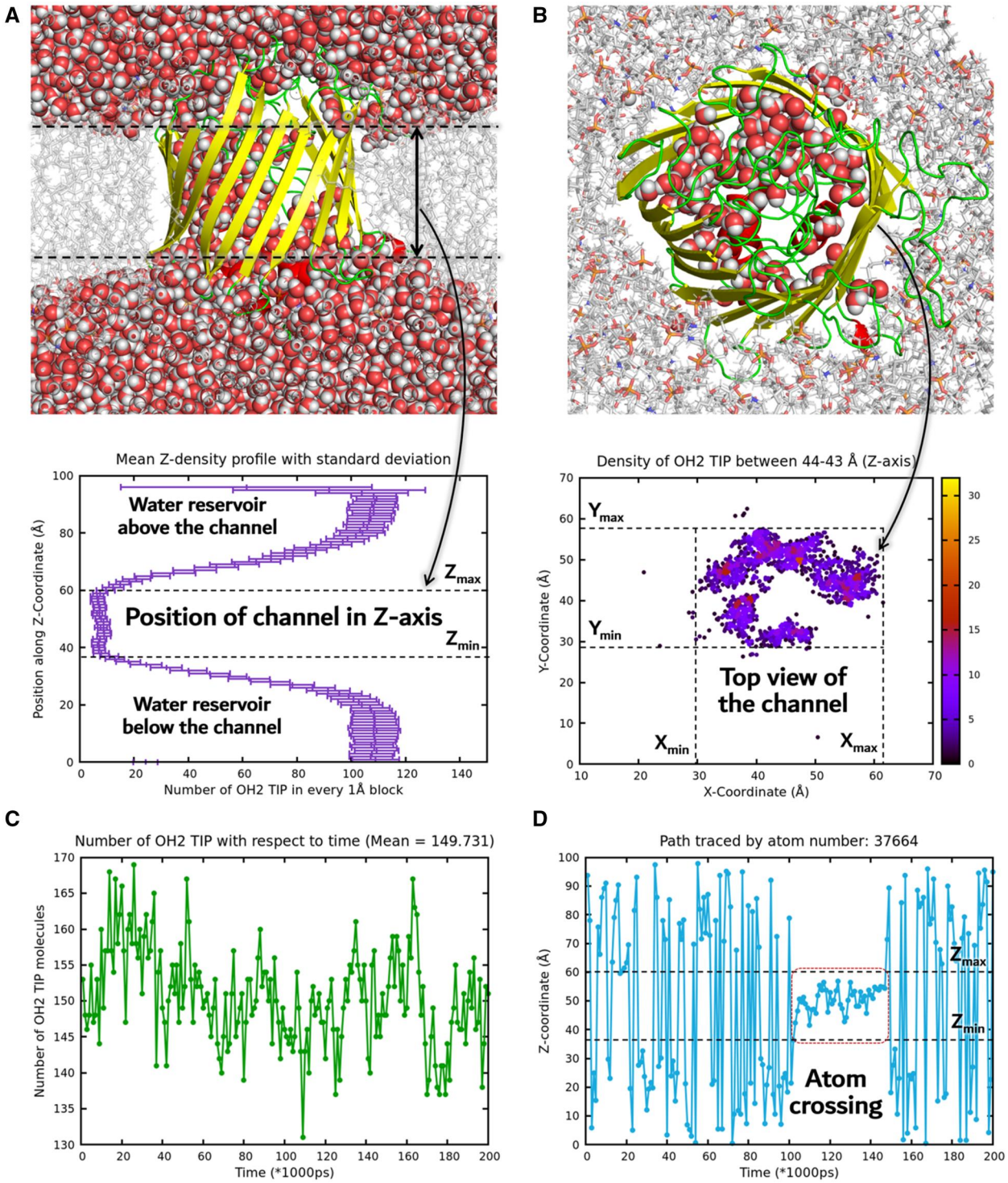

Figure 3. Graphical outputs generated by submodules 1, 2, 3, and 7 of APM: (A) A snapshot of the MD trajectory showing the bulk water (represented by spheres) translocation across the *E. coli* Wzi (cartoon representation) channel: (top) front view of the channel and (bottom) the concomitant "Number of water molecules in every 1 Å block versus Position along *Z*-coordinate (Å)" plot generated by the submodule "*Z*-density profile." The dotted lines indicate the position of the channel in the *Z*-dimension. (B) (top) Top view of the Wzi channel and (bottom) the corresponding position of the channel in the *XY* plane projected onto the 1 Å block defined along the *Z*-axis (representative plot between 44–43 Å) generated by the "*XY* area profile" module. (C) The number of molecules within the channel (*Z*-coordinates between 40 and 60 Å as defined by Figure (A) (bottom)) is shown as a "Time (ps) versus Number of molecules" plot generated by the "Rate of change of molecules" submodule. A mean of ∼150 molecules is seen in the channel in accordance with the previous study (Sachdeva *et al.* 2016). (D) The path traced by the water molecule with atom number 37664 (oxygen) is shown as a "Time (ps) versus *Z*-coordinate (Å)" plot generated by the "Track molecule" submodule. This molecule takes about ∼50 ns (viz., enters and exits between ∼100 and 150 ns) to cross the channel from the periplasmic side to the extracellular side.

unit of time. Due to the equilibrium conditions, the number of permeation events in both directions should be nearly similar for bidirectional conduction. Thus, the total number of permeation events is $2q_o$ for bidirectional conduction. In MD simulations, an equal number of bidirectional permeation events may not occur during the simulation time. Thus, an

average of the permeating molecules is considered when calculating q_o.

3.2.6 Residence time (APM 6)

This submodule calculates the residence time (in ps) of one or more molecules inside the channel. It requires a text file with a single column listing the atom number(s) of the permeating molecule(s) as input (Supplementary Table S8). The user can generate this input file based on insights from the "Permeation" submodule outputs. The submodule produces a two-column output text file that lists the atom numbers and their corresponding residence times. The residence time indicates the duration a molecule spends crossing the channel in either direction.

3.2.7 Track molecule (APM 7)

This submodule allows users to fetch the permeation path of molecules along the Z-axis. It requires the same input as the "Residence Time" submodule (Supplementary Table S9) and generates a "Time (ps) versus Z-coordinate (Å)" plot (Fig. 3D, Supplementary Fig. S4D) along with text files for each permeating molecule. To trace the path the permeating molecule takes in the Z-axis, the user can use the atom numbers generated by the "Permeation" submodule (Supplementary Fig. S7).

3.2.8 Diffusion entry/exit (APM 8)

This submodule provides an insight into the number of entries/exits of the molecule(s) of interest through a diffusion point. For this, a diffusion plane is defined, through which the molecule(s) of interest must pass to be considered as a diffusion (Sachdeva *et al.* 2016, Supplementary Fig. S8). A diffusion entry or exit plane is defined by calculating the center of mass (cm) (viz., X_{cm}, Y_{cm}, and Z_{cm} indicated by O in Supplementary Fig. S8) for three or more user-defined amino acids located at the entry/exit point (Supplementary Table S10). A local XY plane (diffusion plane) is subsequently defined with upper and lower bounds of (X_{cm}+5 Å, Y_{cm}+5 Å) and (X_{cm}−5 Å, Y_{cm}−5 Å), respectively. The entry of the molecule of interest into the channel, which is aligned along the Z-axis, is considered as a diffusion only when the molecule has a Z-coordinate greater than Z_{cm} (let us say, Z_2 as indicated in Supplementary Fig. S8) at the nth ps and subsequently falls within the defined XY plane having the Z-coordinate value between Z_{cm} and (Z_{cm}−3.5 Å) (Z_1) and also crosses Z_1. Similarly, the diffusion exit is calculated in such a way that a molecule that has Z-coordinates less than Z_{cm} (let us say, Z_1) at nth ps and subsequently falls within the defined XY plane having the Z-coordinates value between Z_{cm} and (Z_{cm}+3.5 Å) (Z_2) and further crosses Z_2.

3.2.9 Distance calculation (APM 9)

This submodule calculates the distance between the electronegative atom pairs of the molecule(s) of interest and the channel residues (Supplementary Table S11). It then identifies and reports atoms that are within 3.5 Å of each other. This analysis helps determine which channel atoms may be involved in interactions with the molecule(s) of interest. Additionally, it could reveal potential channel residues that might be obstructing the molecule's permeation, thereby improving the understanding of conduction across the channel.

3.3 Test cases

The implementation of the MDTAP methodology is validated by considering two water-conducting channels as the test cases. The first test case is an *E. coli* outer membrane protein Wzi, which conducts water bidirectionally and has a tortuous path (Fig. 3). The second test case is *E. coli* Aquaporin Z (Aqp-Z). In contrast to Wzi, Aqp-Z is a single-file mode unidirectional water transporter with a narrow pore (Supplementary Fig. S4). Wzi represents the channels or transporters with a wide pore, wherein the conduction takes place in bulk, and the molecule(s) freely diffuse across the protein channel. An earlier MD study has shown that Wzi conducts water from the periplasmic side to the extracellular side and vice versa (Sachdeva *et al.* 2016). These calculations were done by manually inspecting the trajectories, which is a laborious task. Here, it is shown from the last 200 ns trajectory of the published Wzi simulation (Sachdeva *et al.* 2016) that MDTAP can automatically fetch the list of permeating molecules (Fig. 3) and the path traced by them in the Z-direction (Supplementary Fig. S7). Thus, the last 200 ns trajectory of the published Wzi simulation is taken to test MDTAP's automation in characterizing the water permeation across Wzi. Similarly, the 50 ns trajectory of Aqp-Z is also used for calculating the permeating events, and the results generated are shown in Supplementary Fig. S4. In summary, the results align with the published values, validating the MDTAP implementation.

4 Conclusion

MD simulation plays a crucial role in understanding the dynamics of various membrane-embedded natural and artificial channels. To reduce the complexity of manual analysis of permeation events from the vast MD trajectories, a user-friendly software package called MDTAP has been developed. MDTAP is flexible enough to take the user-defined inputs specific to a given system. The implementation is validated by considering an *E. coli* outer membrane protein Wzi, which conducts water in bulk through a wider channel bidirectionally, and *E. coli* Aquaporin Z (Aqp-Z), which conducts water in a single file through a narrow channel unidirectionally. MDTAP successfully captures the conduction of molecules across the channel irrespective of the mode of conduction (Supplementary Fig. S5). In summary, MDTAP enables the user to carry out seamless analyses of permeation events across monodirectional/bidirectional channels or heterogeneous membranes by using the MD trajectories (in both DCD and PDB formats), by simply providing the information about the permeating molecules.

Acknowledgements

The authors acknowledge the support provided by the Computer Center, IIT Hyderabad, and NSM, IIT Hyderabad.

Author contributions

Sruthi Sundaresan (Data curation [lead], Formal analysis [lead], Methodology [equal], Software [lead], Validation [lead], Visualization [lead], Writing—original draft [equal], Writing—review & editing [lead]), Raghuvamsi Palur (Data curation [supporting], Methodology [supporting], Software [supporting], Writing—original draft [equal]), and

Thenmalarchelvi Rathinavelan (Conceptualization [lead], Funding acquisition [lead], Investigation [lead], Methodology [equal], Project administration [lead], Resources [lead], Software [supporting], Supervision [lead], Writing—original draft [lead], Writing—review & editing [equal])

Supplementary data

Supplementary data are available at *Bioinformatics Advances* online.

Conflict of interest

The authors declare that they have no known competing interests.

Funding

This work has been supported by the BIRAC-SRISTI GYTI award (PMU_2017_010), BIRAC-SRISTI GYTI award (PMU_2019_010), and SERB (CRG/2022/001825). S.S. and P.V.R. were supported by the Ministry of Education (MoE) fellowships.

Data availability

The source code, installation, and usage instructions for MDTAP are deposited in GitHub and can be accessed at https://github.com/MBL-lab/MDTAP.

References

Acharya A, Kleinekathofer U. Improved free-energy estimates for the permeation of bulky antibiotic molecules through porin channels using temperature-accelerated sliced sampling. *J Chem Theory Comput* 2025;**21**:3246–59.

Niitsu A, Sugita Y. Towards de novo design of transmembrane α-helical assemblies using structural modelling and molecular dynamics simulation. *Phys Chem Chem Phys* 2023;**25**:3595–606.

Beckstein O, Biggin PC, Bond P *et al.* Ion channel gating: insights via molecular simulations. *FEBS Lett* 2003;**555**:85–90.

Bushell SR, Mainprize IL, Wear MA *et al.* Wzi is an outer membrane lectin that underpins group 1 capsule assembly in *Escherichia coli*. *Structure* 2013;**21**:844–53.

Carnevale V, Delemotte L, Howard RJ. Molecular dynamics simulations of ion channels. *Trends Biochem Sci* 2021;**46**:621–2.

Cheatham TE, 3rd, Case DA. Twenty-five years of nucleic acid simulations. *Biopolymers* 2013;**99**:969–77.

de Groot BL, Grubmuller H. Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. *Science* 2001;**294**:2353–7.

Dong C, Beis K, Nesper J *et al.* Wza the translocon for E. coli capsular polysaccharides defines a new class of membrane protein. *Nature* 2006;**444**:226–9.

Dror RO, Dirks RM, Grossman JP *et al.* Biomolecular simulation: a computational microscope for molecular biology. *Annu Rev Biophys* 2012;**41**:429–52.

Feng J, Chen J, Selvam B *et al.* Computational microscopy: revealing molecular mechanisms in plants using molecular dynamics simulations. *Plant Cell* 2019;**31**.

Freites JA, Tobias DJ. Molecular dynamics simulations of the eye lens water channel aquaporin 0 from fish. *J Phys Chem B* 2024;**128**:7577–85.

Gao J, Kuczera K, Tidor B *et al.* Hidden thermodynamics of mutant proteins: a molecular dynamics analysis. *Science* 1989;**244**:1069–72.

Gullingsrud J, Schulten K. Lipid bilayer pressure profiles and mechanosensitive channel gating. *Biophys J* 2004;**86**:3496–509.

Joshi H, Li C-Y, Aksimentiev A. 2023. All-atom molecular dynamics simulations of membrane-spanning DNA origami nanopores. In: Julián V (ed.), *DNA and RNA Origami: Methods and Protocols*. New York, NY: Springer US, 113–128.

Klesse G, Rao S, Sansom MSP *et al.* CHAP: a versatile tool for the structural and functional annotation of ion channel pores. *J Mol Biol* 2019;**431**:3353–65.

Kruse E, Uehlein N, Kaldenhoff R. The aquaporins. *Genome Biol* 2006;**7**:206.

Littlefield P, Liu L, Mysore V *et al.* Structural analysis of the EGFR/HER3 heterodimer reveals the molecular basis for activating HER3 mutations. *Sci Signal* 2014;**7**:ra114.

Lomize AL, Hage JM, Schnitzer K *et al.* PerMM: a web tool and database for analysis of passive membrane permeability and translocation pathways of bioactive molecules. *J Chem Inf Model* 2019;**59**:3094–9.

Lynch CI, Rao S, Sansom MSP. Water in nanopores and biological channels: a molecular simulation perspective. *Chem Rev* 2020;**120**:10298–335.

Marszalek PE, Lu H, Li H *et al.* Mechanical unfolding intermediates in titin modules. *Nature* 1999;**402**:100–3.

Noskov SY, Berneche S, Roux B. Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands. *Nature* 2004;**431**:830–4.

Petrek M, Otyepka M, Banás P *et al.* CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics* 2006;**7**:316.

Pravda L, Sehnal D, Toušek D *et al.* MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update). *Nucleic Acids Res* 2018;**46**:W368–W373.

Rui H, Lee KI, Pastor RW *et al.* Molecular dynamics studies of ion permeation in VDAC. *Biophys J* 2011;**100**:602–10.

Sachdeva S, Kolimi N, Nair SA *et al.* Key diffusion mechanisms involved in regulating bidirectional water permeation across E. coli outer membrane lectin. *Sci Rep* 2016;**6**:28157.

Sachdeva S, Palur RV, Sudhakar KU *et al.* E. coli group 1 capsular polysaccharide exportation nanomachinary as a plausible antivirulence target in the perspective of emerging antimicrobial resistance. *Front Microbiol* 2017;**8**:70.

Schlick T, Portillo-Ledesma S. Biomolecular modeling thrives in the age of technology. *Nat Comput Sci* 2021;**1**:321–31.

Schlick T, Portillo-Ledesma S, Myers CG *et al.* Biomolecular modeling and simulation: a prospering multidisciplinary field. *Annu Rev Biophys* 2021;**50**:267–301.

Smart OS, Neduvelil JG, Wang X *et al.* HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* 1996;**14**:354–60, 376.

Tajkhorshid E, Nollert P, Jensen MØ *et al.* Control of the selectivity of the aquaporin water channel family by global orientational tuning. *Science* 2002;**296**:525–30.

Zhang J, Baker ML, Schröder GF *et al.* Mechanism of folding chamber closure in a group II chaperonin. *Nature* 2010;**463**:379–83.

Zhu F, Tajkhorshid E, Schulten K. Theory and simulation of water permeation in aquaporin-1. *Biophys J* 2004;**86**:50–7.

Bioinformatics Advances, 2025, 00, 1–8
https://doi.org/10.1093/bioadv/vbaf102
Application Note