Documentation

MDTAP: Molecular Dynamics Trajectory Analysis of Permeation

A tool to analyze permeation events across membrane proteins

Palur Venkata Raghuvamsi[#], Sruthi Sundaresan[#] and Thenmalarchelvi Rathinavelan*

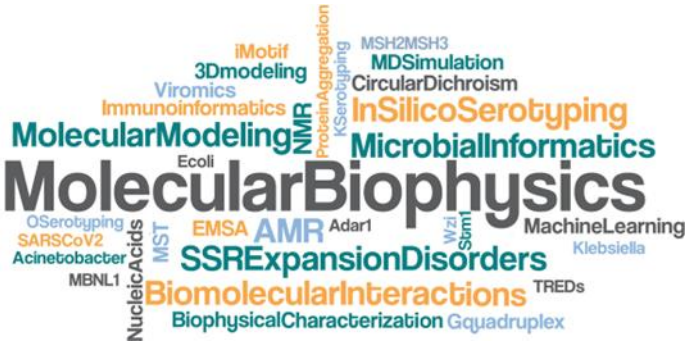
Department of Biotechnology

Indian Institute of Technology Hyderabad

Kandi, Telangana, India - 502284.

[#] Equal contribution

* For correspondence: tr@bt.iith.ac.in



Welcome to MDTAP !!

MDTAP is an MD (molecular dynamics) analysis software that captures and quantifies permeation events across proteins and nucleic acid channels. It allows the user to define a molecule of interest and track its permeation across the channel using the PDB structures generated from MD trajectories. Using the MD trajectories, a methodology is developed here to detect the permeation events across the channel irrespective of their shape, size, and the type of solute molecules permeating. This tool is beneficial in analyzing and calculating the solute/solvent permeations in an automated fashion.

Prerequisites:

1. Linux/Mac-based system
2. Gnuplot

Getting started:

Download and unpack the MDTAP scripts to any folder.

```
unzip MDTAP-main.zip
```

To run/access MDTAP from any folder, follow the instructions given below:

1. Go to the MDTAP scripts folder after unpacking/unzipping

   cd MDTAP-main/MDTAP-scripts/

2. Change the permission of all the scripts in the MDTAP folder

   chmod 777 *.sh

3. Open the ~/.bashrc file to export the path to the scripts and create an alias by pasting the following lines

   export MDTAPpath=/path/to/folder/MDTAP-main/MDTAP-scripts

   alias mdtap='/path/to/folder/MDTAP-main/MDTAP-scripts/mdtap.sh'

4. Save the bashrc file and run it by executing the following command

   source ~/.bashrc

After installing MDTAP, it can be run using the following command to call MDTAP and follow the on-screen instructions to analyze the MD trajectories.

mdtap

How to use MDTAP for the analysis of MD trajectories?

The following section describes the usage of the MDTAP tool. It demonstrates the interface of the software, the inputs required by each module, and the corresponding outputs generated by the various modules.

Demonstration of the usage of MDTAP and the outputs generated:

```
[testuser ~]$ mdtap
================================================================================
            Molecular Dynamics Trajectory Analysis of Permeation (MDTAP)
================================================================================
Modules:
A. Rename PDBs
B. Analyze Permeation
C. Quit
Enter an option (A, B, or C) and press [ENTER]: B
Submodules under analyze permeation:
1. Z-density profile
2. XY-area profile
3. Rate of change of molecules
4. Permeation
5. Net flux and Permeability coefficient (Pd)
6. Ion conductance
7. Residence time
8. Track molecule
9. Diffusion entry/exit
Enter an option (1, 2, or .... 9) and press [ENTER]:
```

Table 1. MDTAP interface with the various modules available. After installation of the MDTAP software, it can be invoked by typing "mdtap". Note that the Analyze Permeation module has nine submodules under it. The user can access the submodules of Analyze Permeation by entering options 1-9.

The next page briefly describes each submodule's functionalities under the “Analyze Permeation” module (APM), as well as the inputs required by each module.

Functionalities of the modules

APM Number	Module 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	Ion conductance	Calculates the conductance of an ion channel by capturing the positive and negative ion flow across the channel
7	Residence time	Calculates the time (in ps) that the permeating molecule resides in the conduction path
8	Track molecule	Tracks the path followed by the permeating molecule along the pore axis (Z-axis) as it passes through the channel
9	Diffusion entry/exit	Calculates the number of molecules that enter or exit through a diffusion plane defined by three amino acids

Inputs required by each module

Required inputs	Sub-modules under APM								
	1	2	3	4	5	6	7	8	9
Molecule of interest	✓	✓	✓	✓					✓
Address/path of the PDBS	✓	✓	✓	✓			✓	✓	✓
Start PDB	✓	✓	✓	✓			✓	✓	✓
End PDB	✓	✓	✓	✓			✓	✓	✓
PDBs to skip	✓	✓	✓	✓			✓	✓	✓
Time difference between PDBs	✓	✓	✓	✓			✓	✓	
Chain ID	✓	✓	✓	✓					✓
Seg ID	✓	✓	✓	✓					✓
Output folder name	✓	✓	✓	✓			✓	✓	✓
Channel limits	✓	✓	✓	✓			✓		
Molar volume of molecule			✓						
Minimum frequency/occurrence				✓					
Number of influx molecules					✓	✓			
Number of efflux molecules					✓	✓			
Time frame					✓	✓			
Valency of permeating molecule						✓			
Applied voltage						✓			
Input file with atom IDs							✓	✓	
3 amino acid residues									✓

<u>Note</u>: Avoid the use of ":" or "step_" or ".pdb" or "." in the name of the input PDB folder (*viz.*, the folder that contains the input PDBs), and the names of the output folders given in each module during the analysis.

```
================================================================================
  A. Rename PDBs
 --------------------------------------------------------------------------------
  This module renames the PDB files to step_i.pdb (1<=i<=n), which is compatible
  with other modules. It takes two inputs: the path (address) where all PDBs are
  stored and the output folder name where all the renamed PDBs will be stored.
  All the renamed PDBs will be stored in the path (address) defined by the user.
================================================================================

Enter the address/path of the PDBs and press [ENTER] (Note: if the PDB files are present in the working directory, just press [ENTER]):
../MDTAP/testPDBs/
Enter the output path along with the folder name and press [ENTER]:
PDBs
The output files are stored in PDBs
1. Continue    2. Back    3. Quit
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

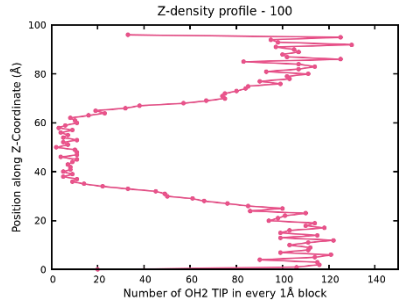
Table 2. The "Rename PDBs" module is used to rename the user's PDB trajectories to make them compatible with other modules. The renamed PDBs are stored in the path defined by the user.

```
==================================================================================
  B. 1. Z-density profile
 ----------------------------------------------------------------------------------
  This module calculates the number of molecules present along the protein axis
  (Z-axis). The user-defined space is divided into slices of 1 Å thickness, and the
  number of molecules in each slice is calculated.
==================================================================================

Enter the molecule of interest (e.g. OH2 TIP for water oxygens) and press [ENTER]:
OH2 TIP
Enter the address/path of the PDBs and press [ENTER] (Note: if the PDB files are present in the working
directory, just press [ENTER]):
PDBs
Enter the start PDB (e.g. if step_10.pdb is the starting PDB, then enter '10') and press [ENTER]:
1
Enter the end PDB (e.g. if step_500.pdb is the final PDB, then enter '500') and press [ENTER]:
200
Enter the PDBs to skip and press [ENTER]:
1
Enter the difference in time (in picoseconds) between each PDB and [ENTER]:
(e.g. If the PDBs are generated at every 10ps interval, then enter '10')
1000
Enter chain ID and press [ENTER] (Note: if chain ID is absent in the PDB files, just press [ENTER]):

Enter seg ID and press [ENTER] (Note: if seg ID is absent in the PDB files, just press [ENTER]):

Enter the output folder name and press [ENTER] (Note: if the folder already exists, it will be
overwritten):
1-Z-density_profile
Recommended dimensions (Xmax Xmin Ymax Ymin Zmax Zmin) for the channel are: 69.070 17.260 66.210 4.400
96.220 -0.930
Enter the channel limits (Xmax Xmin Ymax Ymin Zmax Zmin) and press [ENTER]:
69.070 17.260 66.210 4.400 96.220 -0.930

The output files are stored in 1-Z-density_profile at PDBs/1-Z-density_profile
1. Continue    2. Back    3. Quit
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

Output file name: Z-densityprofile-##.dat

```
This file contains the variation in the number of
OH2 TIP within the channel along the Z-axis.
Column 1: Sections along Z-axis (Å)
Column 2: Number of OH2 TIP
96-95   33
95-94   125
 :       :
2-1     116
1-0     106
```

Output file name: Z-densityprofile-##.png



Output file name: Z-densityprofile-mean_sd.png

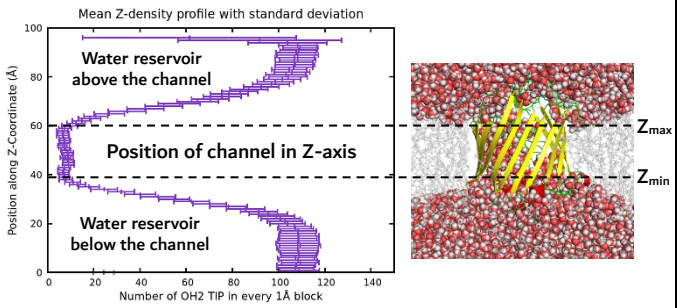


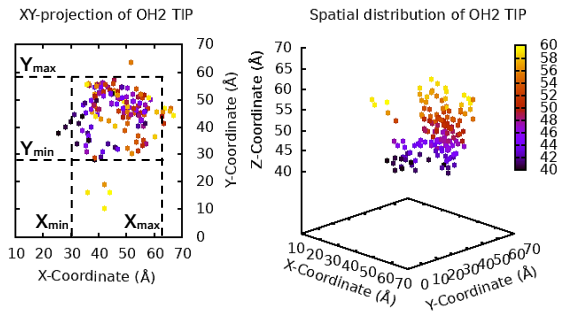
Table 3. The outputs generated by the "Z-density profile" module give an idea of the channel's position in the Z-axis. The example output shown (Z-densityprofile-mean_sd.png) is the front view of the channel. It spans between ~40-60 Å along the Z-axis. These Z-axis values can be used in the subsequent modules to analyze the molecules of interest within the channel. ## indicates that it is a representative example, and the number of files created is equal to the number of input PDBs analyzed. Note that the chain and seg IDs correspond only to the channel (protein) residues.

```
==========================================================================================
  B. 2. XY-area profile
 ------------------------------------------------------------------------------------
  This module identifies the distribution of molecules in the user-defined space.
  The extremities in the X and Y directions are used to calculate the accessible
  area for the molecule of interest.
==========================================================================================

Enter the molecule of interest (e.g. OH2 TIP for water oxygens) and press [ENTER]:
OH2 TIP
Enter the address/path of the PDBs and press [ENTER] (Note: if the PDB files are present in the working
directory, just press [ENTER]):
PDBs
Enter the start PDB (e.g. if step_10.pdb is the starting PDB, then enter '10') and press [ENTER]:
1
Enter the end PDB (e.g. if step_500.pdb is the final PDB, then enter '500') and press [ENTER]:
200
Enter the PDBs to skip and press [ENTER]:
1
Enter the difference in time (in picoseconds) between each PDB and [ENTER]:
(e.g. If the PDBs are generated at every 10ps interval, then enter '10')
1000
Enter chain ID and press [ENTER] (Note: if chain ID is absent in the PDB files, just press [ENTER]):

Enter seg ID and press [ENTER] (Note: if seg ID is absent in the PDB files, just press [ENTER]):

Enter the output folder name and press [ENTER] (Note: if the folder already exists, it will be
overwritten):
2-XY-area_profile
Recommended dimensions (Xmax Xmin Ymax Ymin Zmax Zmin) for the channel are: 69.070 17.260 66.210 4.400
89.110 18.460
Enter the channel limits (Xmax Xmin Ymax Ymin Zmax Zmin) and press [ENTER]:
69.070 17.260 66.210 4.400 60.000 40.000

The output files are stored in 2-XY-area_profile at PDBs/2-XY-area_profile
1. Continue    2. Back    3. Quit
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

Output file name: XY-areaprofile.dat

```
This file contains the area accessible by OH2 TIP
with respect to time.
Column 1: Time (*1000ps)
Column 2: Area (A²)
1    2319.653
2    2533.405
:     :
200   1714.586

The mean area accessible by OH2 TIP is 1968.26 A²,
and the standard deviation is 405.045
```

Output file name: XY-plot-##.png



Output file name: XY-density-Z-##.png

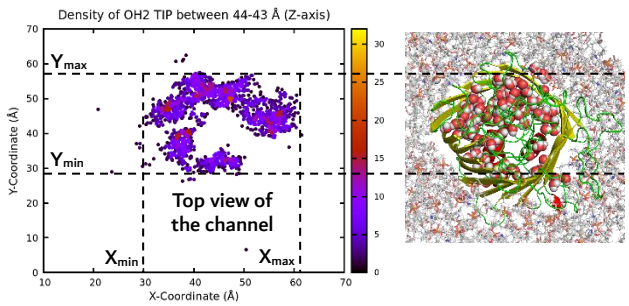


Table 4. The outputs generated by the "XY-area profile" module give an idea of the XY distribution of the molecule of interest with respect to time within the channel. Note that the Z-axis limits (Z_{max} = 60.000 and Z_{min} = 40.000) are entered according to the outputs that were generated in the "Z-density profile" module. The XY values seen from the outputs of this module can be used in the subsequent modules to perform the calculations exactly for the channel dimensions. ## indicates that it is a representative example, and the number of files created is equal to the number of input PDBs analyzed. The chain and seg IDs correspond only to the channel (protein) residues.

```
==================================================================================
  B. 3. Rate of change of molecules
 ----------------------------------------------------------------------------------
  This module calculates the number of water molecules present within the protein
  or in any user-defined area with respect to time.
==================================================================================

Enter the molecule of interest (e.g. OH2 TIP for water oxygens) and press [ENTER]:
OH2 TIP
Enter the address/path of the PDBs and press [ENTER] (Note: if the PDB files are present in the working
directory, just press [ENTER]):
PDBs
Enter the start PDB (e.g. if step_10.pdb is the starting PDB, then enter '10') and press [ENTER]:
1
Enter the end PDB (e.g. if step_500.pdb is the final PDB, then enter '500') and press [ENTER]:
200
Enter the PDBs to skip and press [ENTER]:
1
Enter the difference in time (in picoseconds) between each PDB and [ENTER]:
(e.g. If the PDBs are generated at every 10ps interval, then enter '10')
1000
Enter chain ID and press [ENTER] (Note: if chain ID is absent in the PDB files, just press [ENTER]):

Enter seg ID and press [ENTER] (Note: if seg ID is absent in the PDB files, just press [ENTER]):

Enter the molar volume of the molecule and press [ENTER]:
(e.g the molar volume of Water = 18.07cm^3, Sodium = 23.78cm^3, and Chlorine = 22.40cm^3. Note: The
user is free to enter any value specific to the molecule of interest apart from the examples given.)
18.07
Enter the output folder name and press [ENTER] (Note: if the folder already exists, it will be
overwritten):
3-Rate_of_change_of_molecules
Recommended dimensions (Xmax Xmin Ymax Ymin Zmax Zmin) for the channel are: 69.070 17.260 66.210 4.400
89.110 18.460
Enter the channel limits (Xmax Xmin Ymax Ymin Zmax Zmin) and press [ENTER]:
69.070 17.260 66.210 4.400 60.000 40.000

The output files are stored in 3-Rate_of_change_of_molecules at PDBs/3-Rate_of_change_of_molecules
1. Continue    2. Back    3. Quit
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

Output file name: Rate_output.dat

```
This file contains the number of OH2 TIP present
within the user-defined space with respect to
time.
Column 1: Time (*1000ps)
Column 2: Number of OH2 TIP
Column 3: Volume occupied by OH2 TIP (*10^-23 cm^3)
The mean number of OH2 TIP is 149.731
1       153     458.95
2       148     443.95
3       146     437.95
:        :        :
198     144     431.95
199     152     455.95
200     151     452.95
```

Output file name: Rate_number.png

The mean number of water molecules is ~150

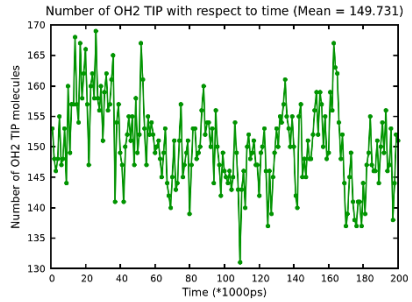


Table 5. The outputs generated by the "Rate of change of molecules" module show the rate of change in the number and volume of the molecule of interest within the channel with respect to time. Note that the Z-axis limits (Z_{max} = 60.000 and Z_{min} = 40.000) are entered according to the outputs that were generated in the "Z-density profile" module. The chain and seg IDs correspond only to the channel (protein) residues.

```
==================================================================================
  B. 4. Permeation
 ----------------------------------------------------------------------------------
  This module captures the molecules that permeate through the channel across
  either direction and lists the atom numbers of the permeating and
  non-permeating molecules.
==================================================================================

Enter the molecule of interest (e.g. OH2 TIP for water oxygens) and press [ENTER]:
OH2 TIP
Enter the address/path of the PDBs and press [ENTER] (Note: if the PDB files are present in the working
directory, just press [ENTER]):
PDBs
Enter the start PDB (e.g. if step_10.pdb is the starting PDB, then enter '10') and press [ENTER]:
1
Enter the end PDB (e.g. if step_500.pdb is the final PDB, then enter '500') and press [ENTER]:
20
Enter the PDBs to skip and press [ENTER]:
1
Enter the difference in time (in picoseconds) between each PDB and [ENTER]:
(e.g. If the PDBs are generated at every 10ps interval, then enter ‘10’)
1000
Enter chain ID and press [ENTER] (Note: if chain ID is absent in the PDB files, just press [ENTER]):

Enter seg ID and press [ENTER] (Note: if seg ID is absent in the PDB files, just press [ENTER]):

Enter the output folder name and press [ENTER] (Note: if the folder already exists, it will be
overwritten):
4temp
Recommended dimensions (Xmax Xmin Ymax Ymin Zmax Zmin) for the channel are: 69.070 17.260 66.210 4.400
89.110 18.460
Enter the channel limits (Xmax Xmin Ymax Ymin Zmax Zmin) and press [ENTER]:
70.000 20.000 60.000 15.000 65.000 35.000
Enter the minimum frequency that should be considered for the calculation and press [ENTER]:
20
The output files are stored in 4temp at PDBs/4-Permeation
The user can verify the PDBs generated by this module and further track the path traced by the atom
numbers listed using the 'Track molecule' (APM #8)
1. Continue    2. Back    3. Quit
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

Output file name: Permeation-dir1.dat

```
This file contains the list of atom numbers of the
molecules that are permeating across the channel
in direction 1.
33263
35978
36941
37400
 : :
```

Output file name: Permeation-dir2.dat

```
This file contains the list of atom numbers of the
molecules that are permeating across the channel
in direction 2.
32954
33320
33416
33704
 : :
```

Output file name: True_permeation.dat

```
This file contains the list of atom numbers of the
molecules that do not permeate across the channel.
32057
32057
32270
32270
 : :
```

Table 6. The outputs generated by the "Permeation" module list the atom numbers of the permeating molecules in both directions (direction 1 and direction 2) and the atom numbers of the molecules that do not permeate across the channel. In addition to this, .pdb files for every permeating and non-permeating module are generated for the user to visualize the localization of each atom within the channel during the simulation time. Note that the X, Y, and Z dimensions are selected after generating outputs from modules 1 and 2 and are appropriately relaxed to capture all the permeating events. The chain and seg IDs correspond only to the channel (protein) residues.

```
================================================================================
  B. 5. Net flux and Permeability coefficient (Pd)
 --------------------------------------------------------------------------------
  This module calculates the net flux and permeability coefficient Pd of the
  permeating molecules.
  Note: Outputs from the permeation module provide atom numbers of molecules that
  completely permeate (influx/efflux). This gives an idea of the number of
  molecules permeating in each direction and can be used as inputs for this
  module.
================================================================================

Enter the number of influx molecules and press [ENTER]:
17
Enter the number of efflux molecules and press [ENTER]:
21
Enter the time frame (ns) over which the above influx/efflux is to be calculated and press [ENTER]:
200

Since the given membrane channel is bidirectional (i.e. both influx and efflux are greater than zero), the diffusion permeability coefficient will be
calculated for both directions.

--------------------------------------------------------------------------------
 The diffusion permeability coefficient is 0.283 x 10^-14 cm^3/s
 The net flux is 0.020 molecules/ns
--------------------------------------------------------------------------------
1. Continue    2. Back    3. Quit
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

Table 7. The "Net flux and Permeability coefficient (P_d)" module calculates the diffusion permeability coefficient (cm^3/s) and the net flux (molecules/ns) of the molecule of interest within the channel in the given time frame. The user can get an idea about the number of influx and efflux molecules across the channel based on the outputs of the "Permeation" module that lists the atom numbers of the permeating molecules in each direction.

```
================================================================================
  B. 6. Ion conductance
 --------------------------------------------------------------------------------
  This module calculates the ion conductance by capturing the positive and
  negative ions that flow across the channel.
  Note: Outputs from the permeation module provide atom numbers of molecules that
  completely permeate (influx/efflux). This gives an idea of the number of
  molecules permeating in each direction and can be used as inputs for this
  module.
================================================================================

Enter the number of influx/efflux molecules, Q1 (e.g. K) and press [ENTER]:

Enter the number of influx/efflux molecules, Q2 (e.g. Cl) and press [ENTER]:

Enter the valency of the permeating ion (e.g. '1' for monovalent, '2' for divalent ions, etc.) and press [ENTER]:

Enter the time frame (ns) over which the above influx/efflux is to be calculated and press [ENTER]:

Enter the voltage applied across the membrane (Vmp) and press [ENTER]:

--------------------------------------------------------------------------------
 The conductance is 0.000 nS
--------------------------------------------------------------------------------
1. Continue    2. Back    3. Quit
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

Table 8. The "Ion Conductance" module calculates the ion conductance of the permeating ion (nS) within the channel in the given time frame. Since there are no permeating ions in this system, the ion conductance is zero. The user can get an idea about the number of influx and efflux molecules across the channel based on the outputs of the "Permeation" module that lists the atom numbers of the permeating molecules in each direction.

```
==================================================================================
  B. 7. Residence time
 ---------------------------------------------------------------------------------
  This module takes atom numbers (as defined in the PDB files) in the form of a
  text file as an input and captures its initial and final PDB files in the
  specific region and calculates time spent in that region.
==================================================================================

Enter the address/path of the PDBs and press [ENTER] (Note: if the PDB files are present in the working
directory, just press [ENTER]):
PDBs
Enter the start PDB (e.g. if step_10.pdb is the starting PDB, then enter '10') and press [ENTER]:
1
Enter the end PDB (e.g. if step_500.pdb is the final PDB, then enter '500') and press [ENTER]:
200
Enter the PDBs to skip and press [ENTER]:
1
Enter the difference in time (in picoseconds) between each PDB and [ENTER]:
(e.g. If the PDBs are generated at every 10ps interval, then enter ‘10’)
1000
Enter chain ID and press [ENTER] (Note: if chain ID is absent in the PDB files, just press [ENTER]):

Enter seg ID and press [ENTER] (Note: if seg ID is absent in the PDB files, just press [ENTER]):

Enter the name of the input file having the atom IDs (Note: the input file should be stored in the
folder where the PDBs are present) and press [ENTER]:
input.txt
Enter the output folder name and press [ENTER] (Note: if the folder already exists, it will be
overwritten):
7-Residence_time
Recommended dimensions (Xmax Xmin Ymax Ymin Zmax Zmin) for the channel are: 69.070 17.260 66.210 4.400
89.110 18.460
Enter the channel limits (Xmax Xmin Ymax Ymin Zmax Zmin) and press [ENTER]:
69.070 17.260 66.210 4.400 60.000 40.000

The output files are stored in 7-Residence_time at PDBs/7-Residence_time
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

Output file name: Residence_time.dat

```
This file contains the residence time of the
provided atom numbers.
Column 1: Atom number
Column 2: Residence time (*1000ps)

33263   74
33923   36
34055   94
  :      :
39938   35
40220   64
```

Table 9. The outputs generated by the "Residence Time" module provide the time that the molecule of interest resides within the channel in picoseconds. In the example output shown, the atom numbers of the molecule of interest given in column 1 were given as an input (input.txt) to calculate how long they reside within the channel. Note that the Z-axis limits (Z_{max} = 60.000 and Z_{min} = 40.000) are entered according to the outputs that were generated in the "Z-density profile" module. The chain and seg IDs correspond only to the channel (protein) residues.

```
==================================================================================
  B. 8. Track molecule
 ----------------------------------------------------------------------------------
  This module takes atom numbers (as defined in the PDB files) in the form of a
  text file as an input and captures the molecule's Z-coordinate position (Å)
  with respect to time and generates a plot. This gives an idea of the migration
  of the molecule with respect to time.
==================================================================================

Enter the address/path of the PDBs and press [ENTER] (Note: if the PDB files are present in the working
directory, just press [ENTER]):
PDBs
Enter the start PDB (e.g. if step_10.pdb is the starting PDB, then enter '10') and press [ENTER]:
1
Enter the end PDB (e.g. if step_500.pdb is the final PDB, then enter '500') and press [ENTER]:
200
Enter the PDBs to skip and press [ENTER]:
1
Enter the difference in time (in picoseconds) between each PDB and [ENTER]:
(e.g. If the PDBs are generated at every 10ps interval, then enter '10')
1000
Enter the name of the input file having the atom IDs (Note: the input file should be stored in the
folder where the PDBs are present) and press [ENTER]:
input.txt
Enter the output folder name and press [ENTER] (Note: if the folder already exists, it will be
overwritten):
8-Track_molecule

The output files are stored in 8-Track_molecule at PDBs/8-Track_molecule
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

Output file name: Track-##.dat

```
This file contains the Z-coordinates (Å) with
respect to time.
Column 1: Time (*1000ps)
Column 2: Z-coordinate (Å)
1    93.6800
2    78.0400
3    5.9400
:     :
198   4.3200
199   22.6500
200   94.9100
```

Output file name: Track-##.png

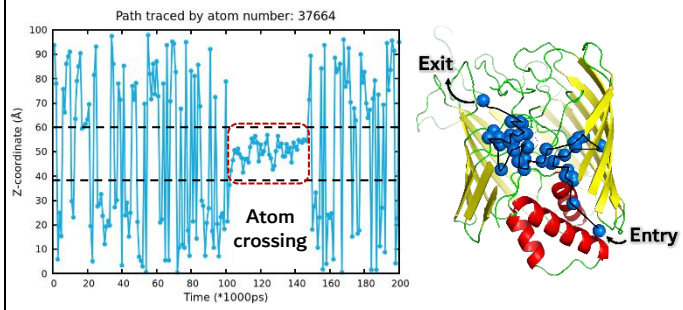
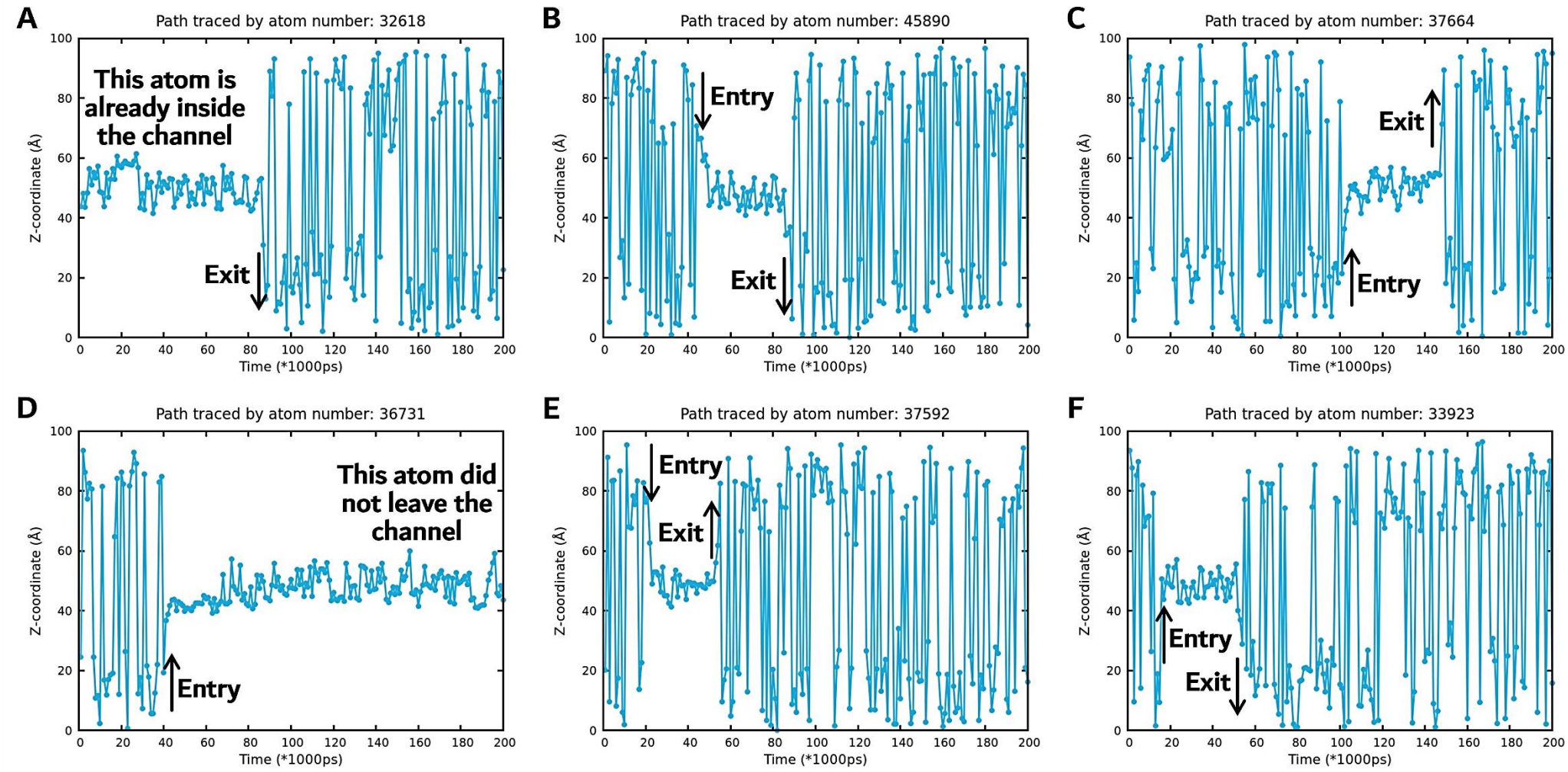


Table 10. The outputs generated by the "Track molecule" module show the path traced by the molecule of interest within the channel over a given time interval. A file with the atom numbers that need to be tracked is given as input to this module. In the example output shown (Track-##.png), it can be seen that the traced molecule enters the channel from the bottom at around 100 ns and completely permeates by around 150 ns (as the position of the channel is known to be between ~40-60 Å along the Z-axis from the results of the "Z-density profile" module). The fluctuation in the graph results from the periodic boundary condition during MD simulations, which makes it appear as though the molecule has crossed the channel instantly when it really hasn't. ## indicates that it is a representative example, and the number of files created is equal to the number of atom numbers provided in the input file.

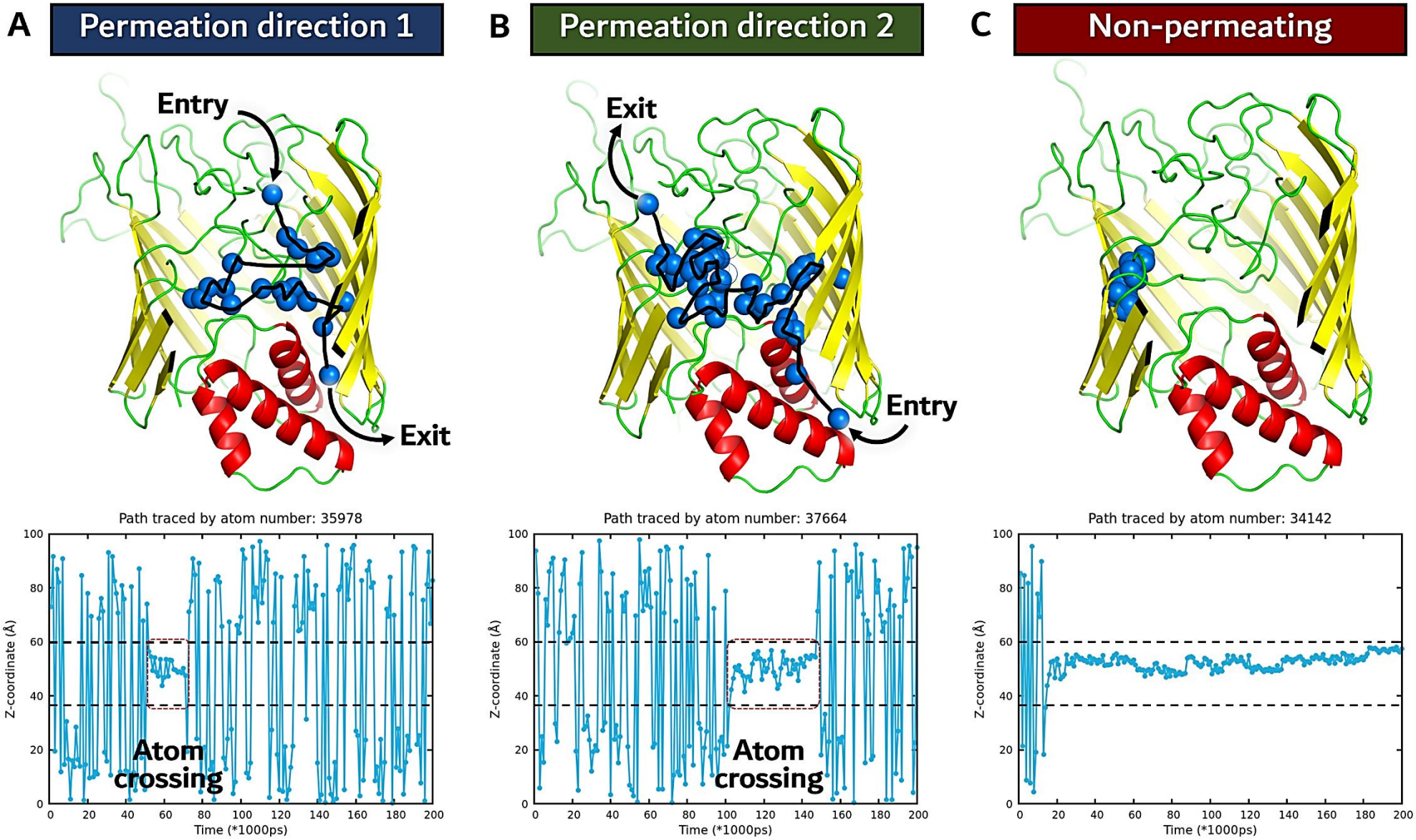
A few examples of plots generated by this module are explained on the next page.

As a permeating molecule enters the channel, it can trace various paths, as shown below:

Among the following plots, B and C are permeation in directions 1 and 2, respectively. A, D, E, and F are examples of non-permeating molecules in the channel.



Here are a few examples of how the "Permeation" module outputs can be correlated to the plots generated by the "Track molecule" module.



```
==================================================================================
  B. 9. Diffusion entry/exit
 ----------------------------------------------------------------------------------
  This module calculates the number of solvent molecules that enter or exit
  through an entry point defined by three amino acid residues.
==================================================================================

Enter the molecule of interest (e.g. OH2 TIP for water oxygens) and press [ENTER]:
OH2 TIP
Enter the address/path of the PDBs and press [ENTER] (Note: if the PDB files are present in the working
directory, just press [ENTER]):
PDBs
Enter the start PDB (e.g. if step_10.pdb is the starting PDB, then enter '10') and press [ENTER]:
1
Enter the end PDB (e.g. if step_500.pdb is the final PDB, then enter '500') and press [ENTER]:
200
Enter the PDBs to skip and press [ENTER]:
1
Enter chain ID and press [ENTER] (Note: if chain ID is absent in the PDB files, just press [ENTER]):

Enter seg ID and press [ENTER] (Note: if seg ID is absent in the PDB files, just press [ENTER]):

Enter 3 amino acid residue numbers (e.g. GLN 102 TYR 380 PHE 427) and press [ENTER]:
GLN 102 TYR 380 PHE 427
Enter the output folder name and press [ENTER] (Note: if the folder already exists, it will be
overwritten):
9-Diffusion

The output files are stored in  9-Diffusion at PDBs/9-Diffusion
Enter an option (1, 2, or 3) and press [ENTER]: 3
```

Output file name: Diffusion_direction1.dat

```
This file contains the atom number(s) of OH2 TIP
that diffuse through the user-defined amino acid
entry/exit points (GLN 102 TYR 380 PHE 427) in
direction 1

35900
37400
44228
55982
```

Output file name: Diffusion_direction2.dat

```
This file contains the atom number(s) of OH2 TIP
that diffuse through the user-defined amino acid
entry/exit points (GLN 102 TYR 380 PHE 427) in
direction 2

53408
60509
```

Table 11. The outputs generated by the "Diffusion entry/exit" module provide the list of atom numbers of the molecule of interest that diffuse through a plane defined by three amino acid residues. Two output files are created, one for each direction of diffusion of the molecule of interest. Note that the chain and seg IDs correspond only to the channel (protein) residues.

Analyzing the MD trajectories using input files

To use the various MDTAP modules to analyze the MD trajectories, the user can either manually enter the details as prompted by the software (refer to pages 8-15 and 18) or provide an input file directly. After downloading and unpacking the files from GitHub, the user can find a folder named “inputs”. The “inputs” folder contains a sample input for each module in MDTAP that uses the trajectories given in a folder named “testPDBs”. The inputs provided are named according to their module numbers. For example, a sample input for the Permeation module is named B4.in, which indicates that it is the 4th sub-module under the Analyze Permeation module (B).

To use these input files, first ensure that the present working directory is inside the “inputs” folder and execute the following command:

```
mdtap < B4.in
```

If the user wishes to run the calculations in the background, an ampersand (&) at the end.

```
mdtap < B4.in &
```

To save the values displayed on the screen during the calculation, it can be written in a file using the following command. This is especially useful for the “Permeation” module.

```
mdtap < B4.in > B4.out &
```

Description of the input file:

The following example is the file (B4.in) that can be given as input for the "Permeation" module. Note that the description of each line is mentioned on the right side after the hash (#).

```
B                                                   # Selecting Analyze Permeation module
4                                                   # Selecting Permeation sub-module
OH2 TIP                                             # Molecule of interest
../testPDBs                                         # Address/path to PDB files
1                                                   # Start PDB
25                                                  # End PDB
1                                                   # PDBs to skip
1000                                                # Time difference between each PDB (ps)
                                                    # Chain ID (Absent in the test case)
                                                    # Seg ID (Absent in the test case)
4-Permeation                                        # Output folder name
69.070 17.260 66.210 4.400 60.000 40.000            # Channel limits
5                                                   # Minimum frequency of occurrence
3                                                   # Exitting the module
```