Uniwersytet im. Adama Mickiewicza w Poznaniu, Polska
Wydział Biologii
Instytut Biologii Molekularnej i Biotechnologii
Zakład Ekspresji Genów
Laboratorium Biomolekularnych Interakcji i Transportu

Rozprawa doktorska

# Ścieżki transportu ligandów w białkach

MSc Carlos Eduardo Sequeiros Borja

Promotor: dr hab. Jan Brezovsky, prof. UAM

ADAM MICKIEWICZ
UNIVERSITY
POZNAŃ

Poznań, 2023

Adam Mickiewicz University in Poznań, Poland
Faculty of Biology
Institute of Molecular Biology and Biotechnology
Department of Gene Expression
Laboratory of Biomolecular Interactions and Transport

Doctoral thesis

# Ligand transport pathways in proteins

MSc Carlos Eduardo Sequeiros Borja

Promotor: dr hab. Jan Brezovsky, prof. UAM

ADAM MICKIEWICZ
UNIVERSITY
POZNAŃ

Poznań, 2023

ACKNOWLEDGEMENTS

**To my supervisor, Prof. UAM dr hab. Jan Brezovsky**

*For the guidance through the PhD studies, support in the scientific research and fruitful discussions not always related to the doctoral studies*

**To dr Karolina Cerbin**

*For all the help in the administrative procedures, not only involving PhD, but also foreign legal affairs*

**To mgr Arleta Kucz and mgr Iwona Kanonik-Jędrzejak**

*For their help to fulfil all bureaucratic requirements related to scientific projects, travels, and conferences*

**To my colleagues from the Laboratory of Biomolecular Interactions and Transport**

*For the fruitful scientific discussions and observations on my research*

**To the members of the Department of Gene Expression at Adam Mickiewicz University in Poznań, Łukasz, Mateusz D, Dawid, Sagar, Mateusz B, Ankur**

*For their friendship and welcoming attitude towards me, and for improving my social life during the PhD studies*

**To my best friend and colleague mgr ing Bartłomiej Surpeta**

*For an awesome and unique friendship, company during the good and bad times of the PhD studies, and mainly for all the scientific and emotional support provided*

**Finally, to my parents**

*Which were my first teachers, and planted the seed of curiosity that guided me to this stage on my professional life*

TABLE OF CONTENTS

## STRESZCZENIE

Analiza tuneli w białkach rozpoczęła się od wizualnej identyfikacji w pojedynczych strukturach statycznych. Jednak wraz ze wzrostem mocy obliczeniowej i lepszymi algorytmami, liczba konformacji, które można uzyskać dla danego białka znacznie wzrosła, głównie dzięki zastosowaniu symulacji dynamiki molekularnej. Ten wzrost ilości danych dostępnych do analizy, spowodował trudności i ograniczenia w badaniu tuneli, przede wszystkim w zasobach i czasie potrzebnym do wykonania takich analiz.

W ramach moich badań doktoranckich przyczyniłem się do rozwoju narzędzi i metodologii analizy sieci tuneli w białkach. Tunele w białkach są istotnymi cechami strukturalnymi, które wpływają nie tylko na procesy katalityczne, ale także na dynamikę i selektywność. Obecnie, wraz ze wzrostem możliwości obliczeniowych, ogromnie wzrosła nie tylko ilość wytwarzanych danych strukturalnych, ale także rodzaj dostępnych informacji, głównie śledzenie ruchu ligandów. Pierwsza z zamieszczonych tu prac pozwala na integracyjne spojrzenie na opis geometrii tuneli i przypisanie do nich transportu ligandów, zapewniając niesubiektywne wyniki i powtarzalność w porównaniu z analizą wizualną. Druga publikacja daje możliwość szybkiej i taniej analizy tuneli z dużych zbiorów danych, eliminując potrzebę posiadania specjalistycznego sprzętu do wykonania tego zadania. Co więcej, metodologia ta pozwala również na identyfikację wcześniej pomijanych wąskich tuneli poprzez zastosowanie mniejszej sondy.

W drugiej części pracy przedstawiam dwa praktyczne zastosowania metodologii opracowanej w pierwszej części. Trzecia publikacja dotyczy selektywnego transportu transportera ABCG46 z Medicago truncatula, gdzie pokazałem jak zmiany w tunelu blokują lub umożliwiają translokację ligandów. Czwarta i ostatnia praca to opracowanie dotyczące transportu wody dla pięciu enzymów hydrolitycznych, gdzie pokazaliśmy jak woda może przemieszczać się przez tunele, których promień najwęższych części jest mniejszy niż pojedyncza cząsteczka wody, podkreślając znaczenie wąskich tuneli.



**Słowa kluczowe**

Bioinformatyka strukturalna, tunele białkowe, transport ligandów, dynamika molekularna

## ABSTRACT

Analysis of tunnels in proteins started with a visual identification in single static structures. However, with the improvement of computational power and better algorithms, the number of conformations that can be obtained for a protein increased considerably, mainly by the usage of molecular dynamics simulations. This increase in data available for analysis, raised difficulties and limitations in the study of tunnels, primarily in the resources and time required to perform such analyses.

In my doctoral research, I have contributed to the development of tools and methodologies for the analysis of tunnel networks in proteins. Tunnels in proteins are relevant structural features that affect not only the catalytic processes but also the dynamics and selectivity. Nowadays, with the increase in computational capabilities, not only the structural data produced has grown immensely, but also the type of information available, mainly the tracking of ligand movement. The first paper included here allows for an integrative view of tunnel geometry description and the assignment of ligand transport to them, providing non-subjective results and reproducibility compared to visual analysis. The second publication grants the possibility of analyzing tunnels from large datasets quickly and cheaply, overcoming the requirements of specialized hardware to perform this task. Furthermore, this methodology also allows for the identification of previously disregarded narrow tunnels by employing a smaller probe.

In the second part of this thesis, I present two practical applications of the methodologies developed in the first part. The third publication focuses on the selective transport of the ABCG46 transporter of *Medicago truncatula*, where I showed how the changes to the tunnel block or allow the translocation of ligands through it. The fourth and final paper is a study about water transport on five hydrolase enzymes, where we showed how water can traverse tunnels of bottleneck radii lower than a single water molecule, highlighting the importance of narrow tunnels.



**Keywords**

Structural bioinformatics, protein tunnels, ligand transport, molecular dynamics

# LIST OF PUBLICATIONS

## PUBLICATIONS INCLUDED IN THE THESIS

1. J. Brezovsky, A.S. Thirunavukarasu, B. Surpeta, **C.E. Sequeiros-Borja**, N. Mandal, D.K. Sarkar, C.J. Dongmo Foumthuim, N. Agrawal, TransportTools: A library for high-throughput analyses of internal voids in biomolecules and ligand transport through them, Bioinformatics. 38 (2022) 1752–1753. https://doi.org/10.1093/bioinformatics/btab872.
   Impact Factor 2021: 6.931
   MNiSW points 2022: 200

2. **C. Sequeiros-Borja**, B. Surpeta, I. Marchlewski, J. Brezovsky, Divide-and-conquer approach to study protein tunnels in long molecular dynamics simulations, MethodsX. 10 (2023) 101968. https://doi.org/10.1016/j.mex.2022.101968.
   Impact Factor 2023: N/A
   MNiSW points 2023: 70

3. K. Pakuła*, **C. Sequeiros-Borja***, W. Biała-Leonhard*, A. Pawela, J. Banasiak, A. Bailly, M. Radom, M. Geisler, J. Brezovsky, M. Jasiński, Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters, Cellular and Molecular Life Sciences. 80 (2023) 105. https://doi.org/10.1007/s00018-023-04751-6.
   Impact Factor 2021: 9.207
   MNiSW points 2023: 140

4. **C. Sequeiros-Borja**, A.S. Thirunavukarasu, C.J. Dongmo Foumthuim, J. Brezovsky, Water will find a way: transport through narrow tunnels in hydrolases, bioRxiv 2023. doi: https://www.biorxiv.org/content/10.1101/2023.05.24.542065v1.
   Impact Factor 2023: N/A
   MNiSW points 2023: N/A

PUBLICATIONS NOT INCLUDED IN THE THESIS

1. **C.E. Sequeiros-Borja**, B. Surpeta, J. Brezovsky, Recent advances in user-friendly computational tools to engineer protein function, Brief. Bioinform. 22 (2021) 1–15. https://doi.org/10.1093/bib/bbaa150.
   Impact Factor 2021: 13.994
   MNiSW points 2021: 140

2. B. Surpeta, **C.E. Sequeiros-Borja**, J. Brezovsky, Dynamics, a Powerful Component of Current and Future in Silico Approaches for Protein Design and Engineering, Int. J. Mol. Sci. 2020, Vol. 21, Page 2713. 21 (2020) 2713. https://doi.org/10.3390/ijms21082713.
   Impact Factor 2020: 5.923
   MNiSW points 2020: 140

3. R.F. Tarabini*, L.F.S.M. Timmers, **C.E. Sequeiros-Borja**, O. Norberto de Souza, The importance of the quaternary structure to represent conformational ensembles of the major Mycobacterium tuberculosis drug target, Sci. Reports 2019 91. 9 (2019) 1–9. https://doi.org/10.1038/s41598-019-50213-0.
   Impact Factor 2019: 3.998
   MNiSW points 2019: 140

4. M. Rotta, L.F.S.M. Timmers, **C. Sequeiros-Borja**, C. V. Bizarro, O.N. De Souza, D.S. Santos, L.A. Basso, Observed crowding effects on Mycobacterium tuberculosis 2-trans-enoyl-ACP (CoA) reductase enzyme activity are not due to excluded volume only, Sci. Rep. 7 (2017) 1–14. https://doi.org/10.1038/s41598-017-07266-w.
   Impact Factor 2017: 4.122
   MNiSW points 2017: 40

# ABBREVIATIONS

MD – molecular dynamics

X-ray – X-ray radiography

NMR – nuclear magnetic resonance

3D structure – three dimensional structure

ATP – adenosine triphosphate

ABC – ATP-binding cassette

ABCG – ABC protein type G

NBD – nucleotide binding domain

TMD – trans-membrane domain

MtABCG46 – *Medicago truncatula* ABCG46 protein

Phe – phenylalanine

Ala – alanine

Leu – leucine

Tyr – tyrosine

μs – microsecond

US – umbrella sampling

$kcal \cdot mol^{-1}$ – kilocalorie per mol

Å – Angstrom

Hal – haloalkane dehalogenase from *Rhodococcus rhodochrous*

Epx – epoxide hydrolase I from *Solanum tuberosum*

Lip - lipase from *Candida rugosa*

H-bond – hydrogen bond

E470G – glutamic acid to glycine mutant of human epoxide hydrolase

# INTRODUCTION

Proteins are molecular machines that allow living organisms to perform the majority of molecular processes required for life. Such processes include but are not limited to the synthesis of biomolecules, modification of chemical compounds, cleavage of chemical bonds, transport of molecules, etc. The biochemical reactions carried by proteins are performed at specific places in their structures, the active site. Is at the active site that the chemical reactions occur, hence, the chemical compounds required for a reaction need to be physically close to the active site. These chemical compounds, usually called ligands, include substrates, co-substrates, cofactors, inhibitors, signaling molecules, and ions among others. Therefore, depending on the localization of the active site, ligands would require to overcome different barriers (or not) before reaching the active site. For proteins with buried active sites, ligands are required to traverse internal protein cavities that connect the active site with the external environment.[1] On the other hand, in proteins that do not perform a catalytic reaction *per se*, but are still responsible for the transport of molecules, these cavities are more like pathways that traverse the whole body of the protein, joining two external regions. The most common example cases of these proteins are ion transporters and aquaporins.[2,3] Hence, protein cavities and pathways are important for the catalysis and transport of ligands in proteins. Given the importance of these structural features of proteins, I decided to study the importance and methods to study protein cavities and pathways.

To obtain a descriptive and detailed picture of these structural features, molecular dynamics (MD) simulations are commonly employed, although some cases where *in vitro* studies were used for this purpose exist, they are scarce and rare.[4] Initially, the study of protein cavities and paths was performed on static structures coming from X-ray or NMR experiments.[5–8] Although this approach was useful and gave new insights into the protein's interior, the information obtained was the result of a static structure, a view that has changed with time. Since proteins are not static immutable objects, we know now that a static view of protein's properties is not accurate enough to produce conclusive results about dynamics, ligand binding affinities, and molecular interactions with other biomolecules. Therefore, MD and other sampling methods have been applied to study the dynamic behavior of proteins.[9–11] This dynamical view brought together, not only new insights and more information but also

new limitations and obstacles to overcome for the study of general properties and tunnel analysis. Therefore, several methods to calculate protein cavities and pathways were developed, each one with its advantages and limitations.[12,13,22–26,14–21]

Nowadays, improvements in algorithms and computing hardware makes it possible to obtain tens or hundreds of thousands of protein states quite easily, leading to a problem in the analysis of results. Here, the complexity is increased from the identification of tunnels in each frame, to the clustering of tunnels across a set of frames, and even further, to the comparison of clusters among different trajectories or related systems. In this sense, the need for a holistic view of the huge amount of structural data available, made me wonder if the solution to this complex problem can be achieved by simplifying it. Hence, I focused the first part of my PhD studies to develop methods to improve the study of cavities and pathways in large datasets. In the second part, I have applied these newly developed methods to two studies: (1) to describe the structural determinants for the differential transport of ligands in a plant ATP-binding cassette transporter wild type and three mutants, and (2) to explore the limits of water transport in three members of the subfamily of hydrolases. Finally, I conclude this dissertation with a personal opinion about the future of the study of tunnels.

# OVERVIEW OF THE DOCTORAL RESEARCH

## CHAPTER 1: CAVITIES IN PROTEINS AND THEIR STUDY

The dictionary defines a cavity as a hollow place, a void or empty space within a solid object. Taking this definition into consideration and simplifying our perspective, if we consider proteins as solid objects, any internal empty space can be defined as a cavity, regardless of its function or geometry. However, a definition as such would become extremely confusing to everyone and also by rejected by the majority (if not all) of researchers in the field. Up today, to the best of my knowledge, there is still no consensus or strict definition of cavities, tunnels, channels, pores, and pockets in proteins, and the available definitions can vary significantly between researchers. Therefore, for this dissertation, I will adopt the definitions of Krone *et al.*[27] for tunnels, channels, pockets, and grooves, with slight and minor personal modifications. An initial classification differentiates between closed and open cavities, with the former being a closed empty space that lacks any contact with the external environment, and the latter includes tunnels, channels, pores, pockets, and grooves (**Figure 1**). Even with this straightforward and simple initial classification, inconsistencies may arise when considering protein dynamics, since an initially closed cavity can open to the exterior environment through gates with variable frequencies, increasing, even more, the complexity of the problem. Although closed cavities are not the major topic of this dissertation, I will refer to these structures as "*enclosed tunnels*", and consider members of this group to be closed cavities that are not naturally open or whose biological function requires them to be cut out from the protein's exterior. The first *enclosed tunnel* described was most likely the tryptophan synthase,[28] and since then several others have been in the spotlight due to their peculiar nature.[29,30] On the other hand, among the open cavities, we have tunnels, channels, pores, pockets, clefts, and grooves. Open cavities with a single entry site can be tunnels, pockets, clefts, and grooves and the differences between them are related to the deepness of each one and can be subtle. Here, I will consider a tunnel as a single entry cavity that connects a buried active site with the exterior (**Figure 1**). Similarly, pockets, grooves, and clefts are single-entry open cavities, however, they connect a more exposed (or less buried if preferred) active site with the exterior, and vary on their deepness, with a decreasing order

respectively. In other words, more than connecting structures, pockets, grooves, and clefts can be seen as "fitting" complementary environments for an exposed active site. In open cavities with multiple entry sites, there are pores (generally straight pathways) and channels, which connect two or more sites of the protein surface (**Figure 1**). As can be seen from these descriptions, the definitions can be subjective to the objectives and background of each researcher, and even in cases where an agreement is achieved, the dynamical behavior of proteins can turn tunnels into channels, pockets into tunnels, and so on. As a deep and thorough definition of these cavities and pathways is out of the scope of this dissertation, the terminology defined previously should be enough to develop a good explanation of the topic.

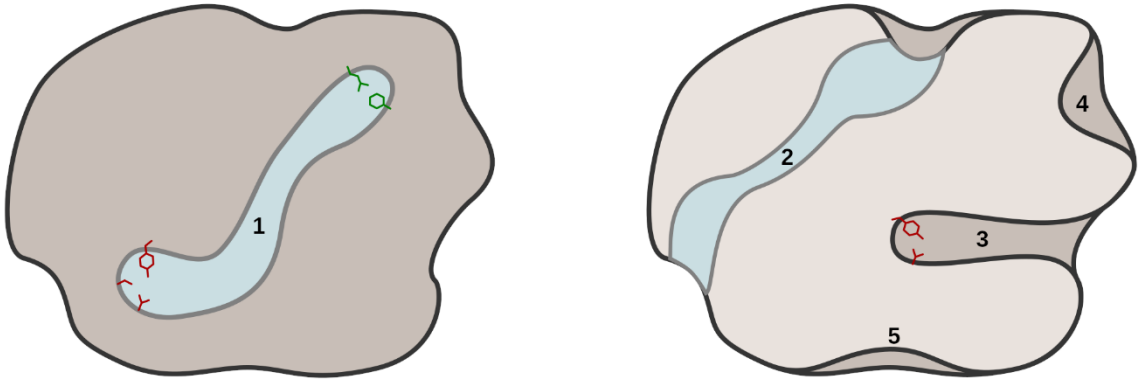



**Figure 1. Cavities in proteins.** Two types of cavities found in proteins are presented. Sliced views of **(1)** an enclosed cavity connecting two active sites (red and green residues) in the same protein, note how the cavity is completely disconnected from the exterior environment and surface. Different types of open cavities: channels connecting two regions of the surface of a protein **(2)**, tunnels **(3)** connecting the exterior with a buried active site (red residues), pockets **(4)** and grooves **(5)** that vary on their deepness.

Historically speaking, the computational study of protein tunnels probably started in the late 70s-80s, when protein cavities became an interesting target of study for protein-ligand interactions.[31–35] Although the study of tunnels did not start itself as such, initially the focus was on visual analysis of surfaces to identify "manually" the existing cavities on single static structures. This visualization was at the beginning made of dots and lines, to be subsequently changed by the more useful Connolly surface, employed even to this day.[31,32,34,35] Probably the first program focused specifically on protein cavities was Cavity Search, a method capable of (given a 3D structure and a starting point) obtaining the shape and electrostatic surface of a cavity.[12] Nevertheless, this method required still an initial

knowledge of the structure to get the starting point, which relied again on an initial surface analysis. Soon after, the POCKET software was developed to overcome this limitation, in which cavities were identified without the need for a starting point by employing a grid approach and representing cavities with a modified marching cubes algorithm.[13] Following efforts did not focus only on the surface representation of the cavities, but on obtaining data from the analysis, like volume,[20] electric potentials,[24] hydrophobicity,[25] likelihood to be targeted as binding site,[36,37] and length and bottleneck radius.[38,39]

Then, molecular dynamics (MD) simulations and other sampling methods shifted the static single view to an ensemble of structures, increasing the amount of information obtained as well as the complexity of the analysis at the same time. Furthermore, with the considerable increase of computation power, especially in MD simulations, the amount of conformations a protein can achieve has grown substantially since its birth, reaching hundreds of thousands of frames in a matter of days depending on the protein size.[40,41] Faced with this new challenge, the study of tunnels has started not only to rely on the protein structure itself but also on the movement of small molecules that form part of the simulation system, these being water molecules, ions, or small chemical compounds. Hence, to date there are mainly two procedures to study protein tunnels: a protein-based approach, which relays on the coordinates of the atoms that form the protein itself,[42] and a small molecule tracking approach, which depends on tracking the movement of a target small molecule through a protein (mainly) across an MD simulation (**Figure 2**). [43,44]

Each of these approaches has its advantages and disadvantages,[46] and the selection would mainly depend on the study to be performed or the availability of resources. The protein-based approach methods are the ones more heavily developed and more diverse considering the algorithms used, data obtained, required inputs, and availability of software.[27,42] Since these methods are extremely diverse depending on the type of cavity to be analyzed (tunnels, pockets, channels, etc.), I will stick mainly to the study of molecular tunnels and leave the other type of cavities aside. The advantage of the protein-based tools for the study of tunnels is that they give a detailed physicochemical description of the tunnels, and although some of the tools can give probabilities or scores for ligand binding or other molecular interactions, they do not provide the actual usage of each tunnel neither direction of movement across.[16,17,47,48,20–24,36–38] On the other side, tracking-based methods are

considerably scarce compared to the previous, and ready-to-use software is even scarcer.[43,44,49,50] Nevertheless, the attractiveness of these methods relies on the results obtained, showing the pathways used by a selected small molecule (most often water). Although the results do not show explicitly a tunnel or channel structure *per se*, like in the protein-based approaches, the information obtained is highly valuable and can be easily linked to functional and biological data.

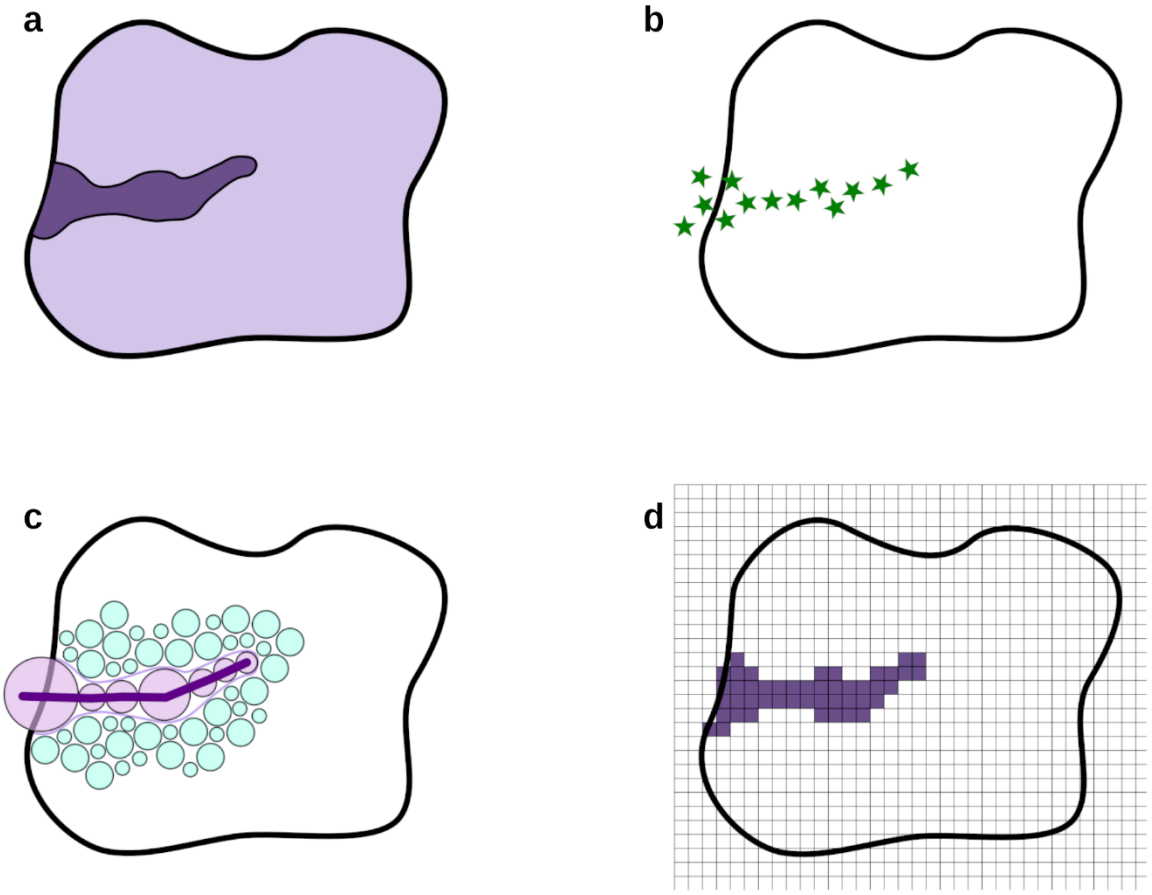



**Figure 2. Methods to identify tunnels in proteins.** Different approaches employed for the identification of protein tunnels are represented in a lateral sliced view. **a)** Surface representation of a protein showing the overall shape of the tunnel (dark purple). **b)** Tracking the movement of small molecules (each green star represent a frame of the simulation, here shown in a single view for simplicity) during MD simulations is used to infer the location of a tunnel. **c)** Voronoi diagrams[45] of proteins are analyzed to find empty spaces (magenta circles) that do not collide with protein atoms (cyan circles), then a tunnel profile can be created around a centerline (thick magenta line). **d)** A grid is constructed to mark all the places where no atoms are present (magenta squares), reconstructing this way a tunnel.

From this short description of the different methodologies to study protein tunnels, it can be inferred that the optimal solution would be a mixture or a combination of both

approaches, yielding detailed information about the tunnel's geometry and physicochemical properties, as well as which tunnels are used for transport of target molecules. Unfortunately (or fortunately depending on the view), to the best of my knowledge, such an approach or methodology did not exist at the moment of starting my PhD studies. Therefore, our group (the Laboratory of Biomolecular Interactions and Transport) focused our efforts on developing a methodology capable of using the outputs of the protein and tracking-based approaches to combine them in an integrative report that would yield information about the transport processes found and the characteristics of the tunnels used by those processes. To achieve this objective, we employed the outputs of the two most popular and easy-to-use software available today: CAVER[47] for the computation of tunnels (**Figure 3a**) and AQUA-DUCT[44] for the tracking of ligand transport (**Figure 3b**). We opted for selecting CAVER from the vast pool of tools available for tunnel identification, since it is the only option available that can handle MD data with considerable simplicity, performs a clustering approach to give a complete view of the tunnels across time, produces detailed geometrical information about individual tunnels in individual frames, is a stand-alone version available in all operating systems, and is well accepted for the community. Nevertheless, even with all the positive features that CAVER possess, it still has some pitfalls and technical limitations that were alleviated in some grade by another research project performed by myself, although I will describe it in the second part of this chapter. For the tracking of ligands, the selection was simpler and straightforward since there is no publicly available implemented software to perform this task other than AQUA-DUCT. The other available options are, either description of methodologies or requests to the implemented code directly to the authors,[43,50] which adds another layer of complication for the future release and usage of the approach developed by our group. The result of this project gave as result the first publication presented in this dissertation: "*TransportTools: a library for high-throughput analyses of internal voids in biomolecules and ligand transport through them*",[51] (**Figure 3**) in **Annex 1**.

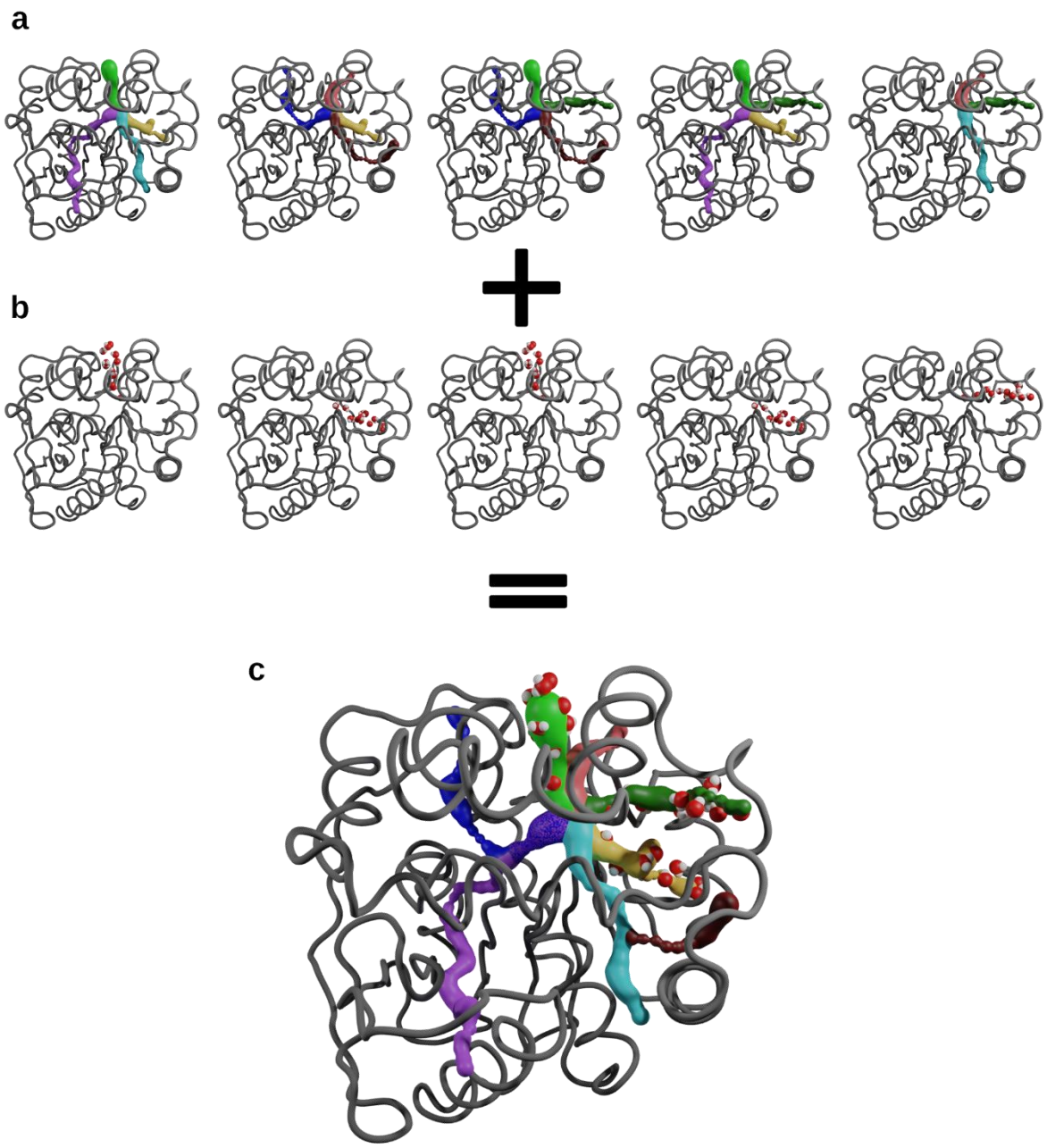



**Figure 3. Overall TransportTools methodology.** The workflow starts by employing a set of molecular dynamics trajectories as inputs, then **a)** the tunnels (colored surfaces) are calculated with CAVER, **b)** and the transport of small molecules (red and white spheres) with AQUA-DUCT. c**)** Finally, TransportTools combines both results in a single view of the tunnel network, the characteristics of each tunnel, and their usage by molecules.

The development of TransportTools took several years to complete the code, test the software and get the paper published in a high-impact journal. Before we published and developed TransportTools, the only approach available to perform a comparative analysis of

tunnels coming from different MD simulations, either repetitions of the same protein, between mutants, or between completely different proteins, was a superposition and alignment of the structures and results, to further analyze visually the localization of the tunnels.[52] Although this visual approach seems to be simple to use, it has considerable limitations and dangers. First of all, since the visual inspection is performed on a case-to-case basis, the results are strongly dependent on the researcher's experience and criteria, leading to subjective results. An improvement of this technique was to employ not only the spatial location of the tunnels but also use the residues or secondary structure motifs in contact with such tunnels to obtain a better description, however, this would be also non-optimal when considering highly-dynamic regions. Furthermore, the application of this technique would be restricted only to repetitions of the same protein or mutants. Second, visual inspection is extremely time-consuming and even when an experienced researcher is performing this task, a high amount of MD simulations make the comparisons extremely confusing and exhausting. Finally, when taking into consideration new datasets available to date that consist of dozens or even hundreds of MD simulations (either short or long) for a system, the visual comparison is already intractable and practically impossible to perform. Therefore, I consider the addition of TransportTools to the analysis of tunnels and transport events to be a major step forward in an integrative view of tunnels' description and their usage in structural bioinformatics, mainly for two reasons. First, since the comparison of tunnel geometries and assignment of transport events are performed by mathematical equations, the subjectivity of the results is considerably reduced and the reproducibility is greatly increased. Second, the automation of the comparative process reduces the burden of the researcher in technical stages and leaves more time for the interpretation of the results and formulation of hypotheses for the phenomena observed. Of course, this does not mean that the method and software are perfect, as stated in the publication, there are still limitations of the method, and room for improvement is acknowledged.

During the development and testing of TransportTools, I realized that another limitation of the current approach to evaluating long MD simulations is the computational resources required to perform this task with CAVER. In the time I spent performing the projects related to this dissertation, I had to analyze hundreds of MD simulations for tunnel analysis. For all those analyses, I was able to count on the infrastructure of our group laboratory, and more

importantly, the Poznan Supercomputing and Networking Center. However, in my personal experience, I know that for a good part of the research community (especially in developing countries), these types of resources are not easily available, or even impossible to acquire, leaving the analysis of tunnels out of their research scopes or performing a limited tunnel analysis. Hence, I envisioned another important application for our TransportTools software, the opportunity to analyze tunnels from long MD simulations with increased speed, considerably low resources, and without losing accuracy. This methodology was successfully published and presented as the second publication in this dissertation: "*Divide-and-conquer approach to study protein tunnels in long molecular dynamics simulations*",[53] in **Annex 2**. In the divide-and-conquer approach, I slice long MD simulations into smaller more tractable pieces and analyze the tunnels in each one individually with CAVER. Following, TransportTools is used to combine the results from all the pieces and yield a unifying single output (**Figure 4**). With this methodology, I showed that the results obtained with the divide-and-conquer approach are comparable to the single CAVER analysis, however, the computational resources are considerably lower, allowing the analysis of tunnels to researchers with a limited budget for computational hardware. Furthermore, this approach allows for a more refined search of tunnels, granting visibility to previously important undetected narrow tunnels. I will show the practical usage and benefits of the divide-and-conquer approach and TransportTools methods in two separate research topics that will cover the second chapter of this dissertation.

## CHAPTER 2: LIGAND MIGRATION THROUGH MOLECULAR PATHWAYS

ATP-binding cassette transporters (ABC) are trans-membrane proteins that make use of ATP as the source of energy, to allow for the transport of molecules across membranes. ABC transporters are extremely diverse and can be found in all living groups. ABC proteins are divided into eight separate groups (from ABCA to ABCH), and this classification is still used today,[54] and although there is a new proposed classification for ABC proteins,[55] I will keep the naming convention for the widely accepted nomenclature in this dissertation. Among all classes, ABCG proteins are well represented in plants, in which these proteins play important roles in response to drought stress, defense against pathogens, transport of metabolites, and

signaling interactions in symbiotic relationships.[56] An important scientific opportunity for me arrived in the form of collaborative research with the group of Prof. dr hab. Michał Jasiński, which focuses on the ABCG46 protein of the plant *Medicago truncatula* (MtABCG46). This protein is involved in the transport of medicarpin precursors in the phenylpropanoid pathway, whose compounds have important roles in plant biology. Structurally speaking, the MtABCG46 protein corresponds to a full-size ABC transporter, which is composed of two "monomers" joined by a linker region. Each of these monomers is formed by two domains: a nucleotide-binding domain (NBD) and a trans-membrane domain (TMD).

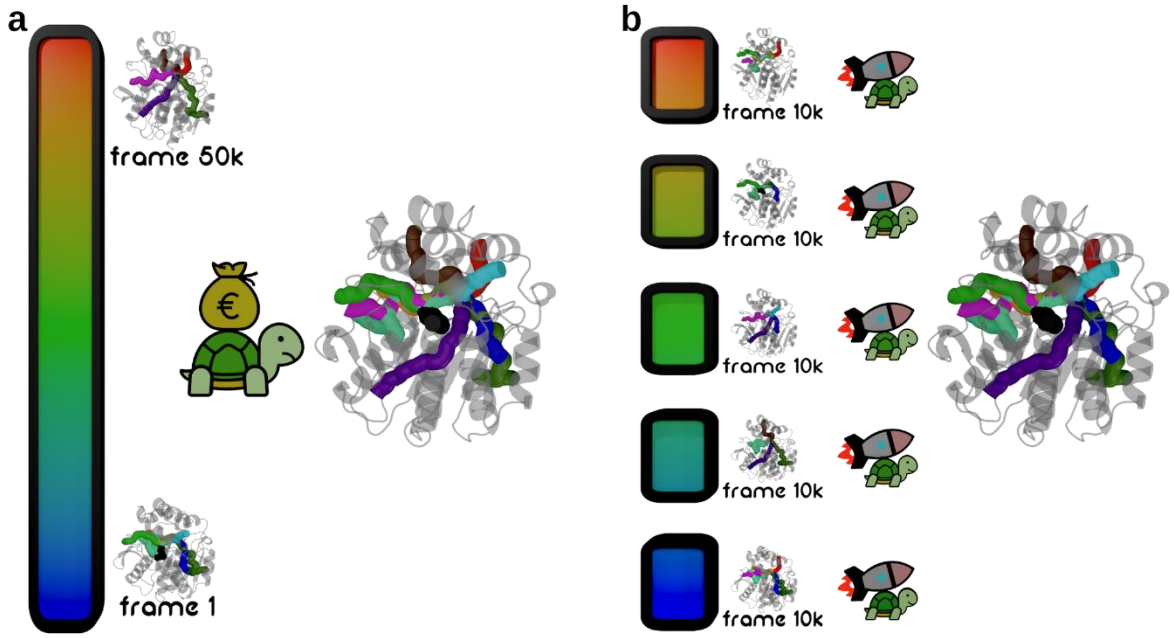



**Figure 4. A comparison of the divide-and-conquer approach to analyze tunnels in long molecular dynamics trajectories to old common approach. a)** The old way requires specialized and costly hardware to run an analysis, which also takes considerable time to yield results. Meanwhile, **b)** in the divide-and-conquer approach, the trajectory is sliced and analyzed individually, then the results are merged with TransportTools to produce results, reducing the time and resources required to yield the same or even better results than the old approach.

Prof. dr hab. Jasiński's group showed that MtABCG46 selectively transports two compounds required for the final production of medicarpin, namely 4-coumarate and liquiritigenin, while at the same time preventing the passage of chemically very similar compounds like isoliquiritigenin and 7,4'-dihydroxyflavone.[57] Although this phenomenon was experimentally observed and verified, there was no structural explanation for this behavior

yet. Furthermore, recent data from the same group showed that from three mutants of MtABCG46 (residue Phe-562 mutated to Ala, Leu, and Tyr), two of them lose the ability to translocate any tested compound (Ala and Tyr mutants), while the remaining (Leu) can transport only liquiritigenin. Hence, my task for the collaboration was first, to obtain the 3D structure of the protein, and then perform mutations and MD simulations to understand the particular behavior of the selective transport. Since there is no experimentally determined structure for MtABCG46, I initially modeled the whole protein with AlphaFold2[58] and later evaluated the resulting models with PROCHECK[59] to select the most promising candidate. Further, I performed the mutations of residue Phe-562 to leucine (Leu), tyrosine (Tyr), and alanine (Ala) and carried out MD simulations for all variants embedded in a membrane for a collective time of 8.8 μs. The tunnel analysis of these simulations using the TransportTools library showed that a tunnel found in the TMD is the most likely to be used for the transport of the active compounds (**Figure 5**). From the results, I showed that mutations to Ala and Tyr alter the geometry and reduce the presence of the tunnel during the simulations and that such alterations lead to higher energy barriers for the passage of the molecules tested, namely liquiritigenin, 4-coumarate, isoliquiritigenin, and 7,4'-dihydroxyflavone. Moreover, I performed umbrella sampling (US) simulations to obtain a more open conformation of the protein to improve the transport calculation and evaluate the energetic cost of such an opening process. The US results showed that Ala and Tyr mutants need to overcome much higher energy barriers to open when compared to the wild-type by about 10 $kcal \cdot mol^{-1}$. Additionally, the tunnel analysis on the open conformations showed that in this conformation, the wild-type protein was able to translocate liquiritigenin and 4-coumarate without major problems, correlating with the experimental observations. Taken together, I showed that small modifications to the TMD tunnel were enough to block any transport at all or to selectively allow the transport of two molecules. A more detailed description can be found in the published paper: "*Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters*", presented in **Annex 3**.

Finally, since water is necessary for life as we know it, I performed research focused on the transport of water across the tunnel network of three enzymes from the hydrolase family. As water is so important for enzymes' functions, its transport across a protein's tunnel network is equally relevant. This transport has proven to be extremely hard to study

experimentally, albeit with exceptions,[4] and the majority of cases use crystallography to obtain static structures and infer tunnel usage.[60–63] Hence, MD simulations are the most attractive methodology to study such phenomena in atomic detail. In the nanomaterial field, the transport of water through artificial pores and membranes has received a spotlight due to its application to water purification.[64–67] Conversely, the transport of water in biological systems has been overlooked and such studies focus mainly on aquaporins.[3,60,68,69] On these proteins, the so-called single-file transport of water has been considered the physical limit for a water molecule to be transported across a protein.[70–72] In this type of transport, water molecules are arranged in a single lane and move across tunnels of a radius of at least 1.4 Å, the accepted radius of a water molecule.[73–75]

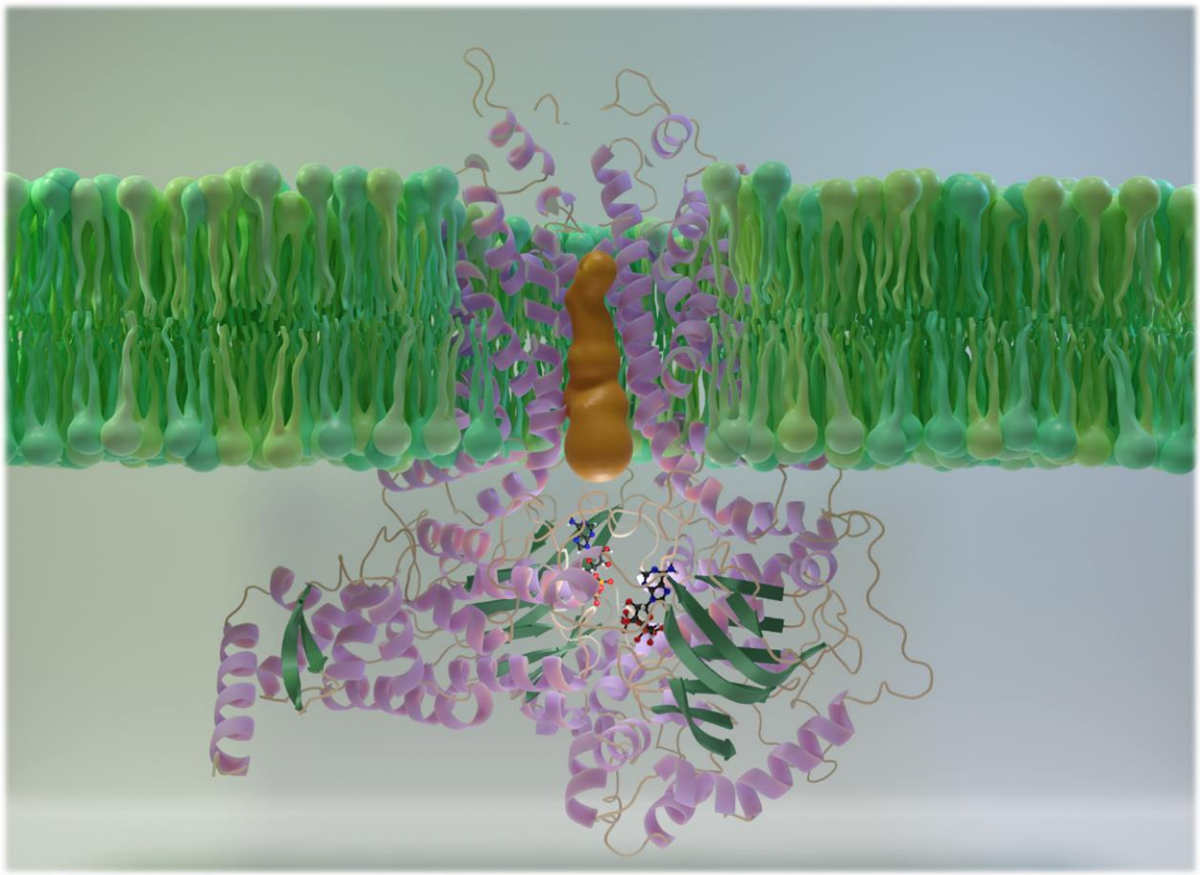


**Figure 5. ABCG transporter of *Medicago truncatula* embedded in a lipidic bilayer displaying the transport tunnel.** A cartoon representation of the ABCG transporter with two bound ATP molecules (the ball and stick representation) is embedded in a lipidic bilayer (green layers of cones with spherical heads), showcasing the most likely tunnel used for the transport (the orange surface in the transmembrane region) of natural substrates.

To improve the knowledge about the transport of water in globular enzymes, we performed extensive MD simulations of a haloalkane dehalogenase from *Rhodococcus rhodochrous* (Hal), epoxide hydrolase I from *Solanum tuberosum* (Epx), and a lipase from *Candida rugosa* (Lip) to evaluate the properties of water transport across their tunnels. Each system consisted of 50 MD simulations, giving as a result a total of 150 MD trajectories collected in 15 μs of simulation time. The analysis of the tunnel network of all these simulations would have been nearly impossible without the development of the TransportTools library, furthermore, the time required for the analysis of potential tunnels using a very small probe with a radius of merely 0.7 Å was reduced considerably by the usage of the divide-and-conquer method. The analysis of the results for the tunnel usage by water in these three proteins yielded interesting and surprising outcomes. First and foremost, I showed that water molecules can traverse tunnels with a bottleneck radius below 1.4 Å, a widely accepted radius for water molecules (**Figure 6**).[73–75] Second, these events account on average for ~20% of all the water transported by the tunnel network of the three enzymes tested. Third, the transport of water through these narrow tunnels is characterized by a high number of H-bonds with the protein's atoms. Fourth, the majority of the H-bonds are formed between water and the backbone atoms of the protein. Lastly, I have performed another set of MD simulations of human soluble epoxide hydrolase and a natural variant (E470G), known to be associated with an increased risk of ischemic stroke in an African-American population.[76] In the 100 MD trajectories (accounting for a total of 10 μs) collected, the same patterns of water transport identified previously were observed, namely, the passage through narrow tunnels, the high presence of H-bonds, and the role of backbone atoms. Additionally, by comparing the results of the wild-type and mutant, I showed that a subangstrom increase in the bottleneck radius of a mutant tunnel by 0.1 Å, results in a dramatic, 20–fold increase in its utilization by water molecules. A more detailed description of the results and methods are available in the preprint: "*Water will find a way:* transport through narrow tunnels in hydrolases", presented in **Annex 4**.

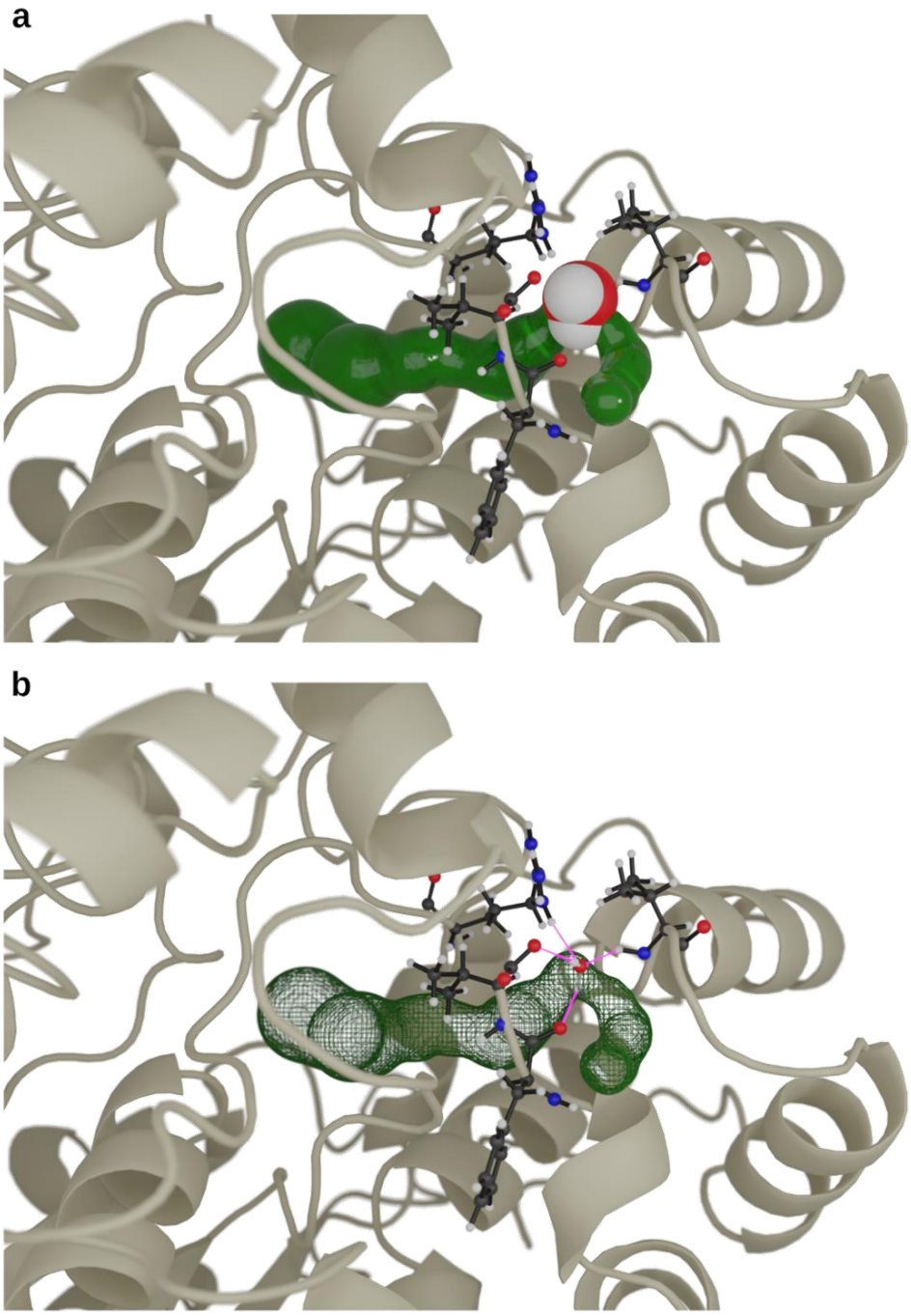



**Figure 6. Transport of water through a narrow tunnel.** A water molecule is transported through a narrow tunnel. **a)** Size comparison of a water molecule (VdW spheres) moving through the bottleneck of a tunnel (the green surface). **b)** H-bond interactions (pink lines) formed between a water molecule and protein residues (the ball and stick representation), within a narrow tunnel (green wires).

# CONCLUSIONS

The static view of a protein and the analysis of tunnels from a single conformation alone is long gone, and the cases that still maintain this analysis today, do it for the sake of description rather than to draw conclusions from them. The dynamical view of proteins and their tunnels is growing constantly nowadays with larger computing power and more efficient algorithms. This advancement leads to an unprecedented amount of data, shifting the bottleneck of the research pipeline to the analysis and understanding of the phenomena in study. Hence, new tools and methods need to be developed for proper and efficient analysis and interpretation of the data produced. We have developed TransportTools to fill this void and improve the analysis of tunnels and transport of small molecules, furthermore, I have prepared a methodology that employs TransportTools to allow for a systematic tunnel analysis of long MD trajectories with limited computational resources or the increased resolution of narrow tunnels.

I have applied the methodologies developed to study the transport of small molecules in a plant ABCG transporter and demonstrated that the identified trans-membrane tunnel is a major contributor to the selective nature observed for this protein. This outcome improves our understanding of the transport process of this protein and represents an interesting target for selective transport modulation. Additionally, I have also studied water transport through the tunnel network of three enzymes and two variants of a human hydrolase. This study shows that water molecules can traverse tunnels narrower than a water molecule, and overcome repulsions by increasing the H-bond interactions with the protein. The insights obtained in the first part of that research were applied to obtain a structural explanation for the differences observed in two variants of the human epoxide hydrolase. These results highlight the importance of previously overlooked narrow tunnels to explain phenotypical differences, changes in catalytic rates, and target regions for mutational studies. Taken together, these two studies showcase the applicability of the developed methodologies, TransportTools and the divide-and-conquer approach, for the study of transport tunnels and the transport of small molecules.

Tunnel analysis started almost half a century ago, from protein surfaces of single static structures to the dynamical view we observe today, where hundreds of thousands of

conformations for a single protein are analyzed. This calls for the requirement of new methodologies capable of handling the current volume of data with acceptable speed. As the study of tunnels changed from surface analysis to grid-based and geometry-based approaches, I believe that the next step of tunnel analysis will incorporate new technologies, like machine learning or artificial intelligence in the coming years. These advances will not only improve the speed and quality of the results but allow for a more complex and maybe interdisciplinary view of tunnels in proteins.

REFERENCES


1. Kingsley, L. J. & Lill, M. A. Substrate tunnels in enzymes: Structure-function relationships and computational methodology. *Proteins Struct. Funct. Bioinforma.* **83**, 599–611 (2015).

2. Gadsby, D. C. Ion channels versus ion pumps: the principal difference, in principle. *Nat. Rev. Mol. Cell Biol.* **10**, 344–352 (2009).

3. de Groot, B. L. & Grubmüller, H. Water Permeation Across Biological Membranes: Mechanism and Dynamics of Aquaporin-1 and GlpF. *Science (80-. ).* **294**, 2353–2357 (2001).

4. Kokkonen, P. *et al.* Molecular Gating of an Engineered Enzyme Captured in Real Time. *J. Am. Chem. Soc.* **140**, 17999–18008 (2018).

5. Taylor, H. C., Komoriya, A. & Chaiken, I. M. Crystallographic structure of an active, sequence-engineered ribonuclease. *Proc. Natl. Acad. Sci.* **82**, 6423–6426 (1985).

6. Getzoff, E. D., Tainer, J. A. & Olson, A. J. Recognition and interactions controlling the assemblies of beta barrel domains. *Biophys. J.* **49**, 191–206 (1986).

7. Moore, J. M. *et al.* Three-dimensional solution structure of plastocyanin from the green alga Scenedesmus obliquus. *Science (80-. ).* **240**, 314–7 (1988).

8. Lear, J. D., Wasserman, Z. R. & DeGrado, W. F. Synthetic amphiphilic peptide models for protein ion channels. *Science (80-. ).* **240**, 1177–81 (1988).

9. Hollingsworth, S. A. & Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **99**, 1129–1143 (2018).

10. Stevens, J. A. *et al.* Molecular dynamics simulation of an entire cell. *Front. Chem.* **11**, 24 (2023).

11. Surpeta, B., Sequeiros-Borja, C. & Brezovsky, J. Dynamics, a Powerful Component of Current and Future in Silico Approaches for Protein Design and Engineering. *Int. J. Mol. Sci.* **21**, 2713 (2020).

12. Ho, C. M. W. & Marshall, G. R. Cavity search: An algorithm for the isolation and

display of cavity-like binding regions. *J. Comput. Aided. Mol. Des.* **4**, 337–354 (1990).

13. Levitt, D. G. & Banaszak, L. J. POCKET: A computer graphies method for identifying and displaying protein cavities and their surrounding amino acids. *J. Mol. Graph.* **10**, 229–234 (1992).

14. Laskowski, R. A. SURFNET: A program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graph.* **13**, 323–330 (1995).

15. Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. P. HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* **14**, 354–360 (1996).

16. Petřek, M., Košinová, P., Koča, J. & Otyepka, M. MOLE: A Voronoi Diagram-Based Explorer of Molecular Channels, Pores, and Tunnels. *Structure* **15**, 1357–1363 (2007).

17. Yaffe, E., Fishelovitch, D., Wolfson, H. J., Halperin, D. & Nussinov, R. MolAxis: Efficient and accurate identification of channels in macromolecules. *Proteins Struct. Funct. Bioinforma.* **73**, 72–86 (2008).

18. Coleman, R. G. & Sharp, K. A. Finding and Characterizing Tunnels in Macromolecules with Application to Ion Channels and Pores. *Biophys. J.* **96**, 632–645 (2009).

19. Pellegrini-Calace, M., Maiwald, T. & Thornton, J. M. PoreWalker: A Novel Tool for the Identification and Characterization of Channels in Transmembrane Proteins from Their Three-Dimensional Structure. *PLoS Comput. Biol.* **5**, e1000440 (2009).

20. Voss, N. R. & Gerstein, M. 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res.* **38**, W555–W562 (2010).

21. Raunest, M. & Kandt, C. dxTuber: Detecting protein cavities, tunnels and clefts based on protein and solvent dynamics. *J. Mol. Graph. Model.* **29**, 895–905 (2011).

22. Paramo, T., East, A., Garzón, D., Ulmschneider, M. B. & Bond, P. J. Efficient Characterization of Protein Cavities within Molecular Simulation Trajectories: trj_cavity. *J. Chem. Theory Comput.* **10**, 2151–2164 (2014).

23. Czirják, G. PrinCCes: Continuity-based geometric decomposition and systematic visualization of the void repertoire of proteins. *J. Mol. Graph. Model.* **62**, 118–127

(2015).

24. Masood, T. Bin, Sandhya, S., Chandra, N. & Natarajan, V. CHEXVIS: a tool for molecular channel extraction and visualization. *BMC Bioinformatics* **16**, 119 (2015).

25. Xu, Y. *et al.* CavityPlus: a web server for protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand binding ability prediction. *Nucleic Acids Res.* **46**, W374–W379 (2018).

26. Garg, P. *et al.* CICLOP: a robust and accurate computational framework for protein inner cavity detection. *Bioinformatics* **38**, 2153–2161 (2022).

27. Krone, M. *et al.* Visual Analysis of Biomolecular Cavities: State of the Art. *Comput. Graph. Forum* **35**, 527–551 (2016).

28. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W. & Davies, D. R. Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from Salmonella typhimurium. *J. Biol. Chem.* **263**, 17857–17871 (1988).

29. Raushel, F. M., Thoden, J. B. & Holden, H. M. Enzymes with Molecular Tunnels. *Acc. Chem. Res.* **36**, 539–548 (2003).

30. Weeks, A., Lund, L. & Raushel, F. M. Tunneling of intermediates in enzyme-catalyzed reactions. *Curr. Opin. Chem. Biol.* **10**, 465–72 (2006).

31. Greer, J. & Bush, B. L. Macromolecular shape and surface maps by solvent exclusion. *Proc. Natl. Acad. Sci.* **75**, 303–307 (1978).

32. Langridge, R., Ferrin, T. E., Kuntz, I. D. & Connolly, M. L. Real-time color graphics in studies of molecular interactions. *Science (80-. ).* **211**, 661–6 (1981).

33. Lerner, R. A. Tapping the immunological repertoire to produce antibodies of predetermined specificity. *Nature* **299**, 592–596 (1982).

34. Connolly, M. L. Analytical molecular surface calculation. *J. Appl. Crystallogr.* **16**, 548–558 (1983).

35. Connolly, M. L. Solvent-accessible surfaces of proteins and nucleic acids. *Science (80-. ).* **221**, 709–13 (1983).

36. Le Guilloux, V., Schmidtke, P. & Tuffery, P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics* **10**, 168 (2009).

37. Schmidtke, P., Bidon-Chanal, A., Luque, F. J. & Barril, X. MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics* **27**, 3276–3285 (2011).

38. Petřek, M. *et al.* CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics* **7**, 316 (2006).

39. Chovancova, E. *et al.* CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* **8**, e1002708 (2012).

40. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J. Chem. Theory Comput.* **9**, 3878–3888 (2013).

41. Le Grand, S., Götz, A. W. & Walker, R. C. SPFP: Speed without compromise - A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput. Phys. Commun.* **184**, 374–380 (2013).

42. Brezovsky, J. *et al.* Software tools for identification, visualization and analysis of protein tunnels and channels. *Biotechnol. Adv.* **31**, 38–49 (2013).

43. Vassiliev, S., Comte, P., Mahboob, A. & Bruce, D. Tracking the Flow of Water through Photosystem II Using Molecular Dynamics and Streamline Tracing. *Biochemistry* **49**, 1873–1881 (2010).

44. Magdziarz, T. *et al.* AQUA-DUCT: a ligands tracking tool. *Bioinformatics* **33**, 2045–2046 (2017).

45. Kim, D.-S. & Sugihara, K. Tunnels and Voids in Molecules via Voronoi Diagram. in *2012 Ninth International Symposium on Voronoi Diagrams in Science and Engineering* 138–143 (IEEE, 2012). doi:10.1109/ISVD.2012.25

46. Mitusińska, K., Bzówka, M., Magdziarz, T. & Góra, A. Geometry-Based versus Small-Molecule Tracking Method for Tunnel Identification: Benefits and Pitfalls. *J. Chem. Inf. Model.* **62**, 6803–6811 (2022).

47. Chovancova, E. *et al.* CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* **8**, e1002708 (2012).

48. Sehnal, D. *et al.* MOLE 2.0: Advanced approach for analysis of biomacromolecular channels. *J. Cheminform.* **5**, 1–13 (2013).

49. Benson, S. P. & Pleiss, J. Solvent Flux Method (SFM): A Case Study of Water Access to Candida antarctica Lipase B. *J. Chem. Theory Comput.* **10**, 5206–5214 (2014).

50. Vad, V. *et al.* Watergate: Visual exploration of water trajectories in protein dynamics. *Eurographics Work. Vis. Comput. Biol. Med.* 33–42 (2017). doi:10.2312/vcbm.20171235

51. Brezovsky, J. *et al.* TransportTools: A library for high-throughput analyses of internal voids in biomolecules and ligand transport through them. *Bioinformatics* **38**, 1752–1753 (2022).

52. Mitusińska, K. *et al.* Structure-function relationship between soluble epoxide hydrolases structure and their tunnel network. *Comput. Struct. Biotechnol. J.* **20**, 193–205 (2022).

53. Sequeiros-Borja, C., Surpeta, B., Marchlewski, I. & Brezovsky, J. Divide-and-conquer approach to study protein tunnels in long molecular dynamics simulations. *MethodsX* **10**, 101968 (2023).

54. Verrier, P. J. *et al.* Plant ABC proteins – a unified nomenclature and updated inventory. *Trends Plant Sci.* **13**, 151–159 (2008).

55. Thomas, C. *et al.* Structural and functional diversity calls for a new classification of ABC transporters. *FEBS Lett.* **594**, 3767–3775 (2020).

56. Hwang, J.-U. *et al.* Plant ABC Transporters Enable Many Unique Aspects of a Terrestrial Plant's Lifestyle. *Mol. Plant* **9**, 338–355 (2016).

57. Biała, W., Banasiak, J., Jarzyniak, K., Pawela, A. & Jasiński, M. Medicago truncatula ABCG10 is a transporter of 4-coumarate and liquiritigenin in the medicarpin biosynthetic pathway. *J. Exp. Bot.* **68**, 3231–3241 (2017).

58. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature*

**596**, 583–589 (2021).

59. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).

60. Murata, K. *et al.* Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599–605 (2000).

61. Brezovsky, J. *et al.* Engineering a de Novo Transport Tunnel. *ACS Catal.* **6**, 7597–7610 (2016).

62. Korasick, D. A., Christgen, S. L., Qureshi, I. A., Becker, D. F. & Tanner, J. J. Probing the function of a ligand-modulated dynamic tunnel in bifunctional proline utilization A (PutA). *Arch. Biochem. Biophys.* **712**, 109025 (2021).

63. Winter, M. B., Herzik, M. A., Kuriyan, J. & Marletta, M. A. Tunnels modulate ligand flux in a heme nitric oxide/oxygen binding (H-NOX) domain. *Proc. Natl. Acad. Sci.* **108**, E881–E889 (2011).

64. Beckstein, O., Biggin, P. C. & Sansom, M. S. P. A Hydrophobic Gating Mechanism for Nanopores. *J. Phys. Chem. B* **105**, 12902–12905 (2001).

65. Secchi, E. *et al.* Massive radius-dependent flow slippage in carbon nanotubes. *Nature* **537**, 210–213 (2016).

66. Xu, J. *et al.* Water transport through subnanopores in the ultimate size limit: Mechanism from molecular dynamics. *Nano Res.* **12**, 587–592 (2019).

67. Coalson, R. D. Driven water/ion transport through narrow nanopores: a molecular dynamics perspective. *Faraday Discuss.* **209**, 249–257 (2018).

68. Saparov, S. M. & Pohl, P. Beyond the diffusion limit: Water flow through the empty bacterial potassium channel. *Proc. Natl. Acad. Sci.* **101**, 4805–4809 (2004).

69. Hadidi, H. & Kamali, R. Molecular dynamics study of water transport through AQP5-R188C mutant causing palmoplantar keratoderma (PPK) using the gating mechanism concept. *Biophys. Chem.* **277**, 106655 (2021).

70. Horner, A. & Pohl, P. Single-file transport of water through membrane channels. *Faraday Discuss.* **209**, 9–33 (2018).

71. Pfeffermann, J., Goessweiner-Mohr, N. & Pohl, P. The energetic barrier to single-file water flow through narrow channels. *Biophysical Reviews* **13**, 913–923 (2021).

72. Jensen, M. Ø. & Mouritsen, O. G. Single-Channel Water Permeabilities of Escherichia coli Aquaporins AqpZ and GlpF. *Biophys. J.* **90**, 2270–2284 (2006).

73. Li, A.-J. & Nussinov, R. A set of van der Waals and coulombic radii of protein atoms for molecular and solvent-accessible surface calculation, packing evaluation, and docking. *Proteins Struct. Funct. Genet.* **32**, 111–127 (1998).

74. Gerstein, M. & Chothia, C. Packing at the protein-water interface. *Proc. Natl. Acad. Sci.* **93**, 10167–10172 (1996).

75. Lynch, C. I., Klesse, G., Rao, S., Tucker, S. J. & Sansom, M. S. P. Water Nanoconfined in a Hydrophobic Pore: Molecular Dynamics Simulations of Transmembrane Protein 175 and the Influence of Water Models. *ACS Nano* **15**, 19098–19108 (2021).

76. Fornage, M. *et al.* The soluble epoxide hydrolase gene harbors sequence variation associated with susceptibility to and protection from incident ischemic stroke. *Hum. Mol. Genet.* **14**, 2829–2837 (2005).

# ANNEXES

# Annex 1

## TransportTools: A library for high-throughput analyses of internal voids in biomolecules and ligand transport through them

Impact Factor 2021: 6.931

MNiSW points 2022: 200

Citations: 2

Main text

https://doi.org/10.1093/bioinformatics/btab872

Supplementary information

https://academic.oup.com/bioinformatics/article/38/6/1752/6491234#supplementary-data

Repository

https://zenodo.org/record/5642954

# Annex 2

## Divide-and-conquer approach to study protein tunnels in long molecular dynamics simulations

Impact Factor 2023: N/A

MNiSW points 2023: 70

Citations: 0

Main text

https://doi.org/10.1016/j.mex.2022.101968

Supplementary information

https://www.sciencedirect.com/science/article/pii/S2215016122003429?via%3Dihub#sec0010

Repository

https://doi.org/10.5281/zenodo.7234699

# Annex 3

## Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters

Impact Factor 2021: 9.207

MNiSW points 2023: 140

Citations: 0

Main text

https://doi.org/10.1007/s00018-023-04751-6

Supplementary information

https://link.springer.com/article/10.1007/s00018-023-04751-6#Sec23

Repository

https://zenodo.org/record/7579354

# Annex 4

## Water will find a way: transport through narrow tunnels in hydrolases

Impact Factor 2023: N/A

MNiSW points 2023: 70

Citations: 0

Main text

https://doi.org/10.1101/2023.05.24.542065

Supplementary information

https://www.biorxiv.org/content/10.1101/2023.05.24.542065v1.supplementary-material

Repository

https://zenodo.org/record/7966082

https://zenodo.org/record/7966059

https://zenodo.org/record/7966092

https://zenodo.org/record/7965661

https://zenodo.org/record/7965881

# CO-AUTHOR STATEMENTS

# TransportTools: A library for high-throughput analyses of internal voids in biomolecules and ligand transport through them

Sequeiros-Borja, C.;

Brezovsky J.;

Thirunavukarasu, A.S.;

Surpeta, B.;

Mandal, N.;

Sarkar, D.K.;

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Carlos Eduardo Sequeiros-Borja, M.Sc.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: carseq@amu.edu.pl

Poznan, March 14, 2023

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication:

Brezovsky J, Thirunavukarasu AS, Surpeta B, Sequeiros-Borja CE, Mandal N, Sarkar DK, Dongmo Foumthuim CJ, Agrawal N, 2021: TransportTools: A Library for High-Throughput Analyses of Internal Voids in Biomolecules and Ligand Transport Through Them. *Bioinformatics 38:* 1752-1753, DOI: 10.1093/bioinformatics/btab872.

I declare the following contribution to this publication:

*contributed to user-testing of the developed tools; implemented algorithms for structural alignment and surface visualization of tunnel clusters; edited the manuscript.*

Carlos Eduardo Sequeiros Borja
.....
Signature

.....
dr hab. Jan Brezovsky, prof. UAM
Promotor

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Jan Brezovsky, Ph.D.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: janbre@amu.edu.pl

Poznan, March 14, 2023

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication:

Brezovsky J, Thirunavukarasu AS, Surpeta B, Sequeiros-Borja CE, Mandal N, Sarkar DK, Dongmo Foumthuim CJ, Agrawal N, 2021: TransportTools: A Library for High-Throughput Analyses of Internal Voids in Biomolecules and Ligand Transport Through Them. *Bioinformatics 38:* 1752-1753, DOI: 10.1093/bioinformatics/btab872.

I declare the following contribution to this publication:

*conceived the research project and the new computational tool; designed and implemented a large majority of algorithms and tests, and prepared the user and technical documentation for the tool; coordinated the work of the project team; analyzed and interpreted the data; and wrote the manuscript.*

.....
Signature

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Aravind Selvaram Thirunavukarasu, M.Tech.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: arathi@amu.edu.pl

Poznan, March 14, 2023

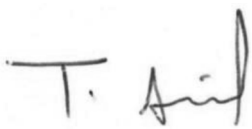
**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication:

Brezovsky J, Thirunavukarasu AS, Surpeta B, Sequeiros-Borja CE, Mandal N, Sarkar DK, Dongmo Foumthuim CJ, Agrawal N, 2021: TransportTools: A Library for High-Throughput Analyses of Internal Voids in Biomolecules and Ligand Transport Through Them. *Bioinformatics 38:* 1752-1753, DOI: 10.1093/bioinformatics/btab872.

I declare the following contribution to this publication:

*contributed to user-testing of the developed tools; generated data for performance evaluation of the TransportTools and Use-case 1 summarized in Supplementary File 5 and 6 and wrote drafts of these two documents.*



…….
Signature

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
Wydział Biologii
Instytutu Biologii Molekularnej i Biotechnologii UAM

Bartłomiej Surpeta, M.Sc.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: bartlomiej.surpeta@amu.edu.pl

Poznan, March 14, 2023

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication:

Brezovsky J, Thirunavukarasu AS, Surpeta B, Sequeiros-Borja CE, Mandal N, Sarkar DK, Dongmo Foumthuim CJ, Agrawal N, 2021: TransportTools: A Library for High-Throughput Analyses of Internal Voids in Biomolecules and Ligand Transport Through Them. *Bioinformatics 38:* 1752-1753, DOI: 10.1093/bioinformatics/btab872.

I declare the following contribution to this publication:

*contributed to user-testing of the developed tools; wrote the draft of state of the art overview in Supplementary File 1; conceived and written the tutorial included in the user guide.*

Bartłomiej Surpeta

.....
Signature

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
Wydział Biologii
Instytutu Biologii Molekularnej i Biotechnologii UAM

Nishita Mandal, M.Sc.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: nishita.mandal@amu.edu.pl

Poznan, March 14, 2023

## CO-AUTHOR CONTRIBUTION STATEMENT

Hereby, I certify that I am the co-author of the following publication:

Brezovsky J, Thirunavukarasu AS, Surpeta B, Sequeiros-Borja CE, Mandal N, Sarkar DK, Dongmo Foumthuim CJ, Agrawal N, 2021: TransportTools: A Library for High-Throughput Analyses of Internal Voids in Biomolecules and Ligand Transport Through Them. *Bioinformatics 38:* 1752-1753, DOI: 10.1093/bioinformatics/btab872.

I declare the following contribution to this publication:

*contributed to user-testing of the developed tools; generated data for Use-case 2 in Supplementary File 7 and wrote the draft of this document.*

Nmandal 14/03/2023
Signature

Dheeraj Kumar Sarkar, M.Sc.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: dhesar@amu.edu.pl

Poznan, March 14, 2023

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication:

Brezovsky J, Thirunavukarasu AS, Surpeta B, Sequeiros-Borja CE, Mandal N, Sarkar DK, Dongmo Foumthuim CJ, Agrawal N, 2021: TransportTools: A Library for High-Throughput Analyses of Internal Voids in Biomolecules and Ligand Transport Through Them. *Bioinformatics 38:* 1752-1753, DOI: 10.1093/bioinformatics/btab872.

I declare the following contribution to this publication:

*contributed to user-testing of the developed tools; generated data for Use-case 3 in Supplementary File 8 and wrote the draft of this document.*

Dheeraj kumar Sarkar

.....
Signature

# Divide-and-conquer approach to study protein tunnels in long molecular dynamics simulations

Sequeiros-Borja, C.;

Surpeta, B.;

Marchlewski, I.;

Brezovsky, J.;

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Carlos Eduardo Sequeiros-Borja, M.Sc.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: carseq@amu.edu.pl

Poznan, May 8, 2023

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the first author of the following publication:

**Sequeiros-Borja CE**, Surpeta B, Marchlewski I, Brezovsky J, 2023: Divide-and-conquer approach to study protein tunnels in long molecular dynamics simulations. *MethodsX 10, 101968*, DOI: 10.1016/j.mex.2022.101968.

I declare the following contribution to this publication:

*contributed to conceptualization and methodology development; software implementation; performed the simulations and analysis; preparation of the figures; writing, reviewing and editing the paper; and making the reviewer's corrections.*

Carlos Eduardo Sequeiros Borja
.....
Signature

.....
dr hab. Jan Brezovsky, prof. UAM
Promotor

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Bartlomiej Surpeta, M.Sc.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: barsur@amu.edu.pl

Poznan, May 8, 2023

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication:

Sequeiros-Borja CE, **Surpeta B**, Marchlewski I, Brezovsky J, 2023: Divide-and-conquer approach to study protein tunnels in long molecular dynamics simulations. *MethodsX 10, 101968*, DOI: 10.1016/j.mex.2022.101968.

I declare the following contribution to this publication:

*contributed to conceptualization; reviewing and editing the paper, prepared guided example of usage.*

Bartłomiej Surpeta
.....
Signature

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Igor Marchlewski, B.S.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: igomar@st.amu.edu.pl

Poznan, May 8, 2023

## CO-AUTHOR CONTRIBUTION STATEMENT

Hereby, I certify that I am the co-author of the following publication:

Sequeiros-Borja CE, Surpeta B, **Marchlewski I**, Brezovsky J, 2023: Divide-and-conquer approach to study protein tunnels in long molecular dynamics simulations. *MethodsX 10, 101968*, DOI: 10.1016/j.mex.2022.101968.

I declare the following contribution to this publication:

*contributed to validation, testing and formal analysis of the method.*

.....
Signature

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Jan Brezovsky, Ph. D.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: janbre@amu.edu.pl, jbrezovsky@iimcb.gov.pl

Poznan, May 8, 2023

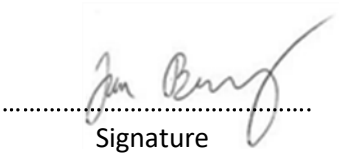
**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication:

Sequeiros-Borja CE, Surpeta B, Marchlewski I, **Brezovsky J**, 2023: Divide-and-conquer approach to study protein tunnels in long molecular dynamics simulations. *MethodsX 10, 101968*, DOI: 10.1016/j.mex.2022.101968.

I declare the following contribution to this publication:

*contributed to conceptualization; writing, revising, and editing the manuscript; supervision of the project, and funding acquisition.*



.....
Signature

# Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters

Sequeiros-Borja, C.;

Pakuła, K.;

Biała-Leonhard, W.;

Pawela, A.;

Banasiak, J.;

Brezovsky, J.;

Jasiński, M.;

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Carlos Eduardo Sequeiros-Borja, M.Sc.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: carseq@amu.edu.pl

Poznan, May 8, 2023

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am a shared first author of the following publication:

Pakula K*, **Sequeiros-Borja CE***, Biala-Leonhard W*, Pawela A, Banasiak J, Bailly A, Radom M, Geisler M, Brezovsky J, Jasinski M, 2023: Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters. *Cellular and Molecular Life Sciences* 80: 105., DOI: 10.1007/s00018-023-04751-6.

I declare the following contribution to this publication:

*contributed to 3D modelling of the protein with all its variants, preparation of the systems for molecular dynamics simulations, and performed all molecular dynamics simulations; analyzed all the structural information from the systems; performed the tunnel network analysis and ligand migration experiments; performed the umbrella sampling simulations and its corresponding analysis; preparation of figures; writing, reviewing and editing the paper; and making the reviewer's corrections.*

Carlos Eduardo Sequeiros Borja
.....
Signature

.....
dr hab. Jan Brezovsky, prof. UAM
Promotor

**INSTITUTE OF BIOORGANIC CHEMISTRY**
Polish Academy of Sciences

Noskowskiego 12/14, 61-704 Poznań
tel.: +48 61 852 85 03, secretariat +48 61 852 89 19
fax: +48 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 VAT no. PL 7770002062
http://www.ibch.poznan.pl

Konrad Pakuła, Ph. D.

Poznań, May 8, 2023

Department of Plant Molecular Physiology, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

Email: kpakula@ibch.poznan.pl

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am a shared first author of the following publication:

**Pakula K***, Sequeiros-Borja CE*, Biala-Leonhard W*, Pawela A, Banasiak J, Bailly A, Radom M, Geisler M, Brezovsky J, Jasinski M, 2023: Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters. *Cellular and Molecular Life Sciences* 80: 105., DOI: 10.1007/s00018-023-04751-6.

I declare the following contribution to this publication:

*contributed to selection of residues for site-directed mutagenesis; designed, performed, and interpreted the results of the transport experiments, BY2 transformation, isolation of microsomes, transport and competition assays; preparation of figures; writing, reviewing and editing the paper; and making the reviewer's corrections.*

Signature

**INSTITUTE OF BIOORGANIC CHEMISTRY**
Polish Academy of Sciences

Noskowskiego 12/14, 61-704 Poznań
tel.: +48 61 852 85 03, secretariat +48 61 852 89 19
fax: +48 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 VAT no. PL 7770002062
**http://www.ibch.poznan.pl**

Wanda Biała-Leonhard, Ph. D.

Poznan, May 8, 2023

Department of Plant Molecular Physiology, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

Email: wbiala@ibch.poznan.pl

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am a shared first author of the following publication:

Pakula K*, Sequeiros-Borja CE*, **Biala-Leonhard W***, Pawela A, Banasiak J, Bailly A, Radom M, Geisler M, Brezovsky J, Jasinski M, 2023: Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters. *Cellular and Molecular Life Sciences* 80: 105., DOI: 10.1007/s00018-023-04751-6.

I declare the following contribution to this publication:

*contributed to selection of residues for site-directed mutagenesis; designed, performed, and interpreted the results of the transport experiments, BY2 transformation, isolation of microsomes, transport and competition assays; performed HPLC/MS experiments and analysis; performed microscopic observations; preparation of figures; writing, reviewing and editing the paper; and making the reviewer's corrections.*

.....
Signature

**INSTITUTE OF BIOORGANIC CHEMISTRY**
Polish Academy of Sciences

Noskowskiego 12/14, 61-704 Poznań
tel.: +48 61 852 85 03, secretariat +48 61 852 89 19
fax: +48 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 VAT no. PL 7770002062
http://www.ibch.poznan.pl

Poznan, May 8, 2023

Aleksandra Pawela, Ph. D.

Department of Plant Molecular Physiology, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland

Email: apawela@ibchp.poznan.pl

## CO-AUTHOR CONTRIBUTION STATEMENT

Hereby, I certify that I am a co-author of the following publication:

Pakula K*, Sequeiros-Borja CE*, Biala-Leonhard W*, **Pawela A**, Banasiak J, Bailly A, Radom M, Geisler M, Brezovsky J, Jasinski M, 2023: Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters. *Cellular and Molecular Life Sciences* 80: 105., DOI: 10.1007/s00018-023-04751-6.

I declare the following contribution to this publication:

*contributed to BY2 transformation, isolation of microsomes, transport and competition assays; reviewing and editing the paper.*

Aleksandra Pawela

.....

Signature

**INSTITUTE OF BIOORGANIC CHEMISTRY**
Polish Academy of Sciences

Noskowskiego 12/14, 61-704 Poznań
tel.: +48 61 852 85 03, secretariat +48 61 852 89 19
fax: +48 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 VAT no. PL 7770002062
http://www.ibch.poznan.pl

Joanna Banasiak, TITLE (Ph. D., MSc, BS)

Department of Plant Molecular Physiology, Institute of Bioorganic Chemistry, Polish Academy of Sciences, 61-704 Poznan, Poland

Email: joaban@ibch.poznan.pl

Poznan, May 8, 2023

## CO-AUTHOR CONTRIBUTION STATEMENT

Hereby, I certify that I am a co-author of the following publication:

Pakula K*, Sequeiros-Borja CE*, Biala-Leonhard W*, Pawela A, **Banasiak J**, Bailly A, Radom M, Geisler M, Brezovsky J, Jasinski M, 2023: Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters. *Cellular and Molecular Life Sciences* 80: 105., DOI: 10.1007/s00018-023-04751-6.

I declare the following contribution to this publication:

*contributed to the selection of sequences to be used for multi-sequence alignments; performed microscopic observations; reviewing and editing the paper.*

J. Banasiak

Signature

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Jan Brezovsky, Ph. D.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: janbre@amu.edu.pl

Poznan, May 8, 2023

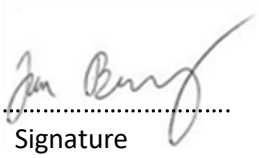
**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am a co-author of the following publication:

Pakula K, Sequeiros-Borja CE, Biala-Leonhard W, Pawela A, Banasiak J, Bailly A, Radom M, Geisler M, **Brezovsky J**, Jasinski M, 2023: Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters. *Cellular and Molecular Life Sciences* 80: 105., DOI: 10.1007/s00018-023-04751-6.

I declare the following contribution to this publication:

*contributed to conceptualization, design, and interpretation of the computational results; writing, revising, and editing the paper; supervision and funding acquisition.*



.....
Signature

**INSTITUTE OF BIOORGANIC CHEMISTRY**
Polish Academy of Sciences

Noskowskiego 12/14, 61-704 Poznań
tel.: +48 61 852 85 03, secretariat +48 61 852 09 19
fax: +48 61 852 05 32, e-mail: ibch@ibch.poznan.pl
REGON 000849327 VAT no. PL 7770002062
**http://www.ibch.poznan.pl**

Michal Jasinski, Ph. D.

Poznan, May 8, 2023

Department of Plant Molecular Physiology, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego12/14, 61-704 Poznan, Poland

Email: jasinski@ibch.poznan.pl

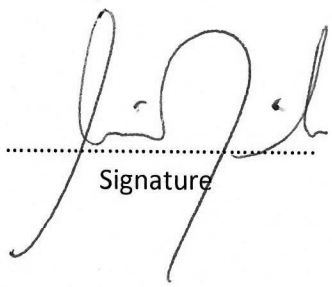
## CO-AUTHOR CONTRIBUTION STATEMENT

Hereby, I certify that I am a co-author of the following publication:

Pakula K*, Sequeiros-Borja CE*, Biala-Leonhard W*, Pawela A, Banasiak J, Bailly A, Radom M, Geisler M, Brezovsky J, **Jasinski M**, 2023: Restriction of access to the central cavity is a major contributor to substrate selectivity in plant ABCG transporters. *Cellular and Molecular Life Sciences* 80: 105., DOI: 10.1007/s00018-023-04751-6.

I declare the following contribution to this publication:

*contributed to conceptualization and interpretation of the results; writing, reviewing and editing the paper; supervision, administration and funding acquisition of the project.*



.....
Signature

# Water will find a way: transport through narrow tunnels in hydrolases

Sequeiros-Borja, C.;

Thirunavukarasu, A.S.;

Dongmo Foumthuim, C.J.;

Brezovsky J.;

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Carlos Eduardo Sequeiros-Borja, M.Sc.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: carseq@amu.edu.pl

Poznan, May 26, 2023

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the first author of the following publication (pre-print):

**Sequeiros-Borja CE**, Thirunavukarasu AS, Dongmo Foumthuim CJ, Brezovsky J, 2023: Water will find a way: transport through narrow tunnels in hydrolases. bioRxiv doi doi.org/10.1101/2023.05.24.542065

I declare the following contribution to this publication:

*contributed to preparation of hEpx and E470G (human epoxide hydrolase and one mutant respectively) systems for molecular dynamics simulations; ran molecular dynamics simulations and perform its initial analysis; performed the tunnel network analysis for all the proteins; performed the overall analysis of transport tunnels and transport of water for all systems; performed a detailed water transport and interactions analyses for all systems; preparation of the figures; writing, revising and editing the paper.*

Carlos Eduardo Sequeiros Borja
.....
Signature

.....
dr hab. Jan Brezovsky, prof. UAM
Promotor

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
Wydział Biologii
Instytutu Biologii Molekularnej i Biotechnologii UAM

Aravind Selvaram Thirunavukarasu, M.Tech.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: arathi@amu.edu.pl

Poznan, May 26, 2023

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication (pre-print):

Sequeiros-Borja CE, **Thirunavukarasu AS**, Dongmo Foumthuim CJ, Brezovsky J, 2023: Water will find a way: transport through narrow tunnels in hydrolases. bioRxiv doi doi.org/10.1101/2023.05.24.542065

I declare the following contribution to this publication:

*contributed to preparation of Hal (haloalkane dehalogenase) system for molecular dynamics simulations; ran molecular dynamics simulations and perform its initial analysis; revising the paper.*

…….
Signature

ul. Uniwersytetu Poznańskiego 6, Collegium Biologicum, 61-614 Poznań
NIP 777 00 06 350, REGON 000001293
tel. +48 61 829 59 50, ibmib@amu.edu.pl

**Cedrix Dongmo Foumthuim, Ph. D.**
*Campus scientifico*
*via Torino 155*
*30172 Venice*
*Italy*
*Email: cedrix.dongmo@roma2.infn.it*

Venice, May 26, 2023

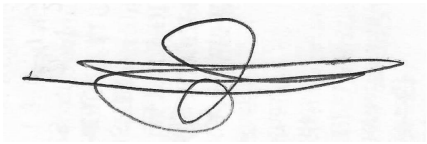
**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication (pre-print):

Sequeiros-Borja CE, Thirunavukarasu AS, **Dongmo Foumthuim CJ**, Brezovsky J, 2023: Water will find a way: transport through narrow tunnels in hydrolases. bioRxiv doi https://doi.org/10.1101/2023.05.24.542065

I declare the following contribution to this publication:

*contributed to the preparation of Lip (lipase) system for molecular dynamics simulations; ran molecular dynamics simulations and performed its initial analysis; revising and editing the paper.*



…….
Signature

UNIWERSYTET IM. ADAMA MICKIEWICZA W POZNANIU
**Wydział Biologii**
**Instytutu Biologii Molekularnej i Biotechnologii UAM**

Poznan, May 25, 2023

Jan Brezovsky, Ph.d.
Faculty of Biology, Adam Mickiewicz University, 61-614 Poznan, Poland
Email: janbre@amu.edu.pl

**CO-AUTHOR CONTRIBUTION STATEMENT**

Hereby, I certify that I am the co-author of the following publication:

Sequeiros-Borja C, Thirunavukarasu AS, Dongmo Foumthuim CJ, Brezovsky J, bioRxiv 2023.05.24.542065; doi: https://doi.org/10.1101/2023.05.24.542065

I declare the following contribution to this publication:

*conceived and coordinated the project; designed the calculations, set up, performed, and analyzed simulations of Epx; interpreted data; co-wrote the manuscript*

.

.....
Signature

ul. Uniwersytetu Poznańskiego 6, Collegium Biologicum, 61-614 Poznań
NIP 777 00 06 350, REGON 000001293
tel. +48 61 829 59 50, ibmib@amu.edu.pl

**www.ibmib.amu.edu.pl**