

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

Fully automated virtual screening pipeline of FDA-approved drugs using Caver Web

Milos Musil^{a,b,c}, Andrej Jezik^{a,b,c}, Marie Jankujova^{a,b}, Jan Stourac^{a,b}, Jakub Galgonek^d, Saltuk Mustafa Eyrilmez^a, Jiri Vondrasek^d, Jiri Damborsky^{a,b}, David Bednar^{a,b,*}

^a Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Masaryk University, Brno, Czech Republic

^b International Clinical Research Center, St. Anne's University Hospital Brno, Brno, Czech Republic

^c Faculty of Information Technology, Brno University of Technology, Brno, Czech Republic

^d Institute of Organic Chemistry and Biochemistry of the CAS, Prague, Czech Republic

ARTICLE INFO

Article history:

Received 28 September 2022

Received in revised form 11 November 2022

Accepted 13 November 2022

Available online 17 November 2022

Keywords:

Caver
CaverDock
FDA-approved drug
Channel
Tunnel
Virtual screening
Web

ABSTRACT

Protein tunnels are essential in transporting small molecules into the active sites of enzymes. Tunnels' geometrical and physico-chemical properties influence the transport process. The tunnels are attractive hot spots for protein engineering and drug development. However, studying the ligand binding and unbinding using experimental techniques is challenging, while *in silico* methods come with their limitations, especially in the case of resource-demanding virtual screening pipelines. Caver Web 1.2 is a new version of the web server combining the capabilities for the detection of protein tunnels with the calculation of the ligand trajectories. The new version of the Caver Web server was expanded with the ability to fetch novel ligands from the Integrated Database of Small Molecules and with the fully automated virtual screening pipeline allowing for the fast evaluation of the predefined set of over 4,300 currently approved drugs. The virtual screening pipeline is accompanied by a comprehensive user interface, making it a viable service for the broader spectrum of companies and the academic user community. The web server is freely available for academic use at <https://loschmidt.chemi.muni.cz/caverweb>.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The search for new chemical compounds and their application in therapeutic areas is in the midst of the interest of the pharmaceutical industry. In recent years, a phenotype-based discovery has come at the peak of interest of various research groups [1–2]. However, the continuous efforts in solving protein structures and the wide availability of the modelling tools such as AlphaFold2 [3] made the target-based method a staple procedure in drug discovery. The majority of drugs were obtained using this method [4].

The human genome project has unearthed approximately 20,000 protein-encoding genes [5–6] producing over 10,000 ubiquitously expressed proteins [7–8]. However, less than ten percent

of those proteins are currently targeted by the approved drugs [9] leaving a vast number of potential therapeutic targets that could be potentially exploited. Developing new drugs and their delivery to the market is expensive and laborious. It has been estimated that from the initial inception to the market distribution, the cost of a new drug can go over 2 billion US dollars [10–11]. Therefore, repurposing approved drugs seems a viable strategy to reduce the costs and time demands for novel drug identification.

High-throughput robotics screenings are still time-demanding and require expensive infrastructure. On the other hand, with the recent advances in the fields of information technology and computational biology, *in silico* techniques provide an exciting alternative to experimental screening [12–15]. Those methods can quickly evaluate vast libraries from which only a tiny fraction of the top binders are typically considered for further experimental confirmation [16]. Therefore, virtual screening methods have become a prominent technology in the probing and prioritization of drug-like compounds in the pharmaceutical industry [17]. The vast majority of virtual screening platforms score the binding of inhibitors to the binding or active-site pockets. However, the

Abbreviations: FDA, U.S. Food and Drug Administration; IDS, Integrated Database of Small Molecules; PDB, Protein Data Bank; CIF, Crystallographic Information File; CSA, Catalytic Site Atlas.

* Corresponding author at: Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Masaryk University, Brno, Czech Republic.

E-mail address: davidbednar1208@gmail.com (D. Bednar).

<https://doi.org/10.1016/j.csbj.2022.11.031>

2001-0370/© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

access of ligands to these functionally important sites is rarely considered.

Our *in-house* web tool Caver Web [18] incorporated both Caver [19], software for identification and geometrical analysis of tunnels, and CaverDock [12], software for docking-based analysis of the ligand transport. Therefore, even inexperienced users with limited bioinformatics knowledge can efficiently study the ligand transport process. However, in the current implementation, only a single ligand could be analyzed in each run of Caver Web. This was a limiting step for the users looking to identify the best-fitting inhibitors for their protein of interest.

Here, we present a new version of Caver Web that overcomes those shortcomings, making it a viable service for the scientific community and the broader spectrum of companies. A new version of the Caver Web server provides a fully automated virtual screening pipeline using the predefined set of the medical drugs approved by major world health organizations (FDA+) that contains over 4,300 ligands obtained from the ZINC database [20]. Furthermore, Caver Web was extended with the possibility of fetching the ligands directly from the Integrated Database of Small Molecules (IDSM) [21]. Finally, the graphical user interface of the Caver Web service was expanded with new elements for the visualization of the transition energies and the bottleneck detection based on the transition of the ligands through tunnels. The server is freely available for the scientific community at <https://loschmidt.chemi-muni.cz/caverweb>.

2. Materials and methods

The basic workflow of the implemented virtual screening pipeline is captured in Fig. 1. The virtual screening can be started only after the successful run of the Caver Web calculation, which requires only the protein structure in the PDB or CIF format. The identification of pockets using the Fpocket 2 [22], the search for the essential residues in the mCASA [23] and SwissProt [24], and the calculation of the tunnels employing Caver 3.02 [19] is processed during the standard run of the Caver Web service.

2.1. Virtual screening of the FDA-approved drugs

Several preparation steps must be initiated to proceed with the fully automated virtual screening pipeline. Once the user selects the tunnel in the Caver Web interface and the tunnel data are obtained from the server, the data preparation is initiated. In this step, the `prepare_receptor4.py` script from the MGLTools package [25] is utilized to add Gasteiger charges and AutoDock Vina [15] compatible atom types to every atom in the protein structure. A tunnel discretizer [26] is then employed to cut the tunnel into discrete slices with specified distances. Finally, the predefined set of

the approved drugs (4,381 ligands from the ZINC database) is prepared using the MGLTools `prepare_ligand4.py` script. The discretized tunnel, receptor, and the collection of ligands are then forwarded to be processed by the AutoDock Vina and *in-house* tool CaverDock [12].

Due to the large input library size, it is impossible to evaluate all of the world-wide-approved drugs with the comprehensive and time-demanding CaverDock calculation. Therefore, in the first step, the binding energy of each ligand in the protein's active site is rapidly evaluated by the fast AutoDock Vina algorithm. Ligands in the library are then sorted based on their predicted binding energy, and the top 50 binders are evaluated using the CaverDock algorithm. Users can modify the discretization parameter defining the distance between the centres of two slices of the tunnel and select the residues that should be considered for calculation. Both the bottleneck energy of the lower-bound trajectory and the lowest energy of the bound ligand are provided to the user. However, the binding energy is always set as a default quality metric when sorting the ligands calculated by CaverDock.

2.2. Integrated database of small molecules

The current implementation of the Caver Web service allowed users to start their ligand transport analysis using either PDBQT/MOL, SMILES, or ZINC code identifiers. However, there was no possibility of searching for the ligands of interest in the existing databases. For this reason, we have expanded our server with the IDSM service [21] that combines data from several sources, mainly PubChem [27], ChEMBL [28], and ChEBI [29] small-molecule datasets. It is now possible to fetch ligands directly from the IDSM database using either the ligand's approximate name or its SMILES code. Furthermore, we have implemented a more robust sorting algorithm to obtain the most relevant results on top of the list. As a result, the integration of the IDSM database enables users to start their ligand transport analysis without the need to search for the PDBQT/MOL representation or SMILES codes first.

2.3. Graphical user interface

The interactive user interface for the virtual screening pipeline was implemented as a natural extension of the original Caver Web server. The transport analysis option was expanded with the possibility of fetching ligands directly from the IDSM database (Fig. 2D). The virtual screening of the FDA+ drugs on the target protein can be started by simply clicking the “Start screening of the FDA+ drugs” button (Fig. 2A). Here, selecting the tunnel of interest is possible, together with the setting of the discretization radius and the selection of the residues that should be kept for the calculation (Fig. 2B). Once the virtual screening is finished, it is possible to download raw data as an Excel sheet or generate a full report in PDF. The PDF report contains a plot of lower-bound trajectories for each evaluated ligand and the dot-plot showing the energy minima of all ligands based on their position in the protein tunnel.

By clicking on the book icon, the table with the 50 most promising ligands can also be directly visualized in the web interface (Fig. 2C). The table contains the binding energy of the selected ligands in the protein's active site and in the bottleneck of the analyzed tunnel. Furthermore, by clicking on the book icon, it is possible to open a new tab for any of these ligands to visualize the plot of the calculated binding energies for the whole trajectory. This window allows the user to define points for calculating the activation energy of protein–ligand association and the energy difference between ligand bound on the surface and in the active site (Fig. 2E). The energy profiles of several ligands can be directly compared in a single plot (Fig. 2F). Finally, the ligand transport from the surface to the protein's active site is generated as a PyMOL

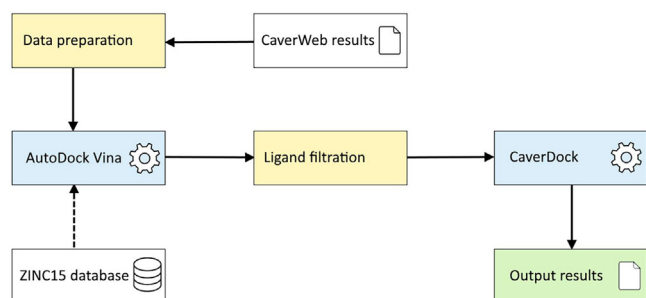


Fig. 1. Simplified workflow of the virtual screening pipeline. The calculation is done in two critical steps: (i) evaluating all FDA+ drugs using the AutoDock Vina and (ii) the assessment of the best hits using the CaverDock calculation.

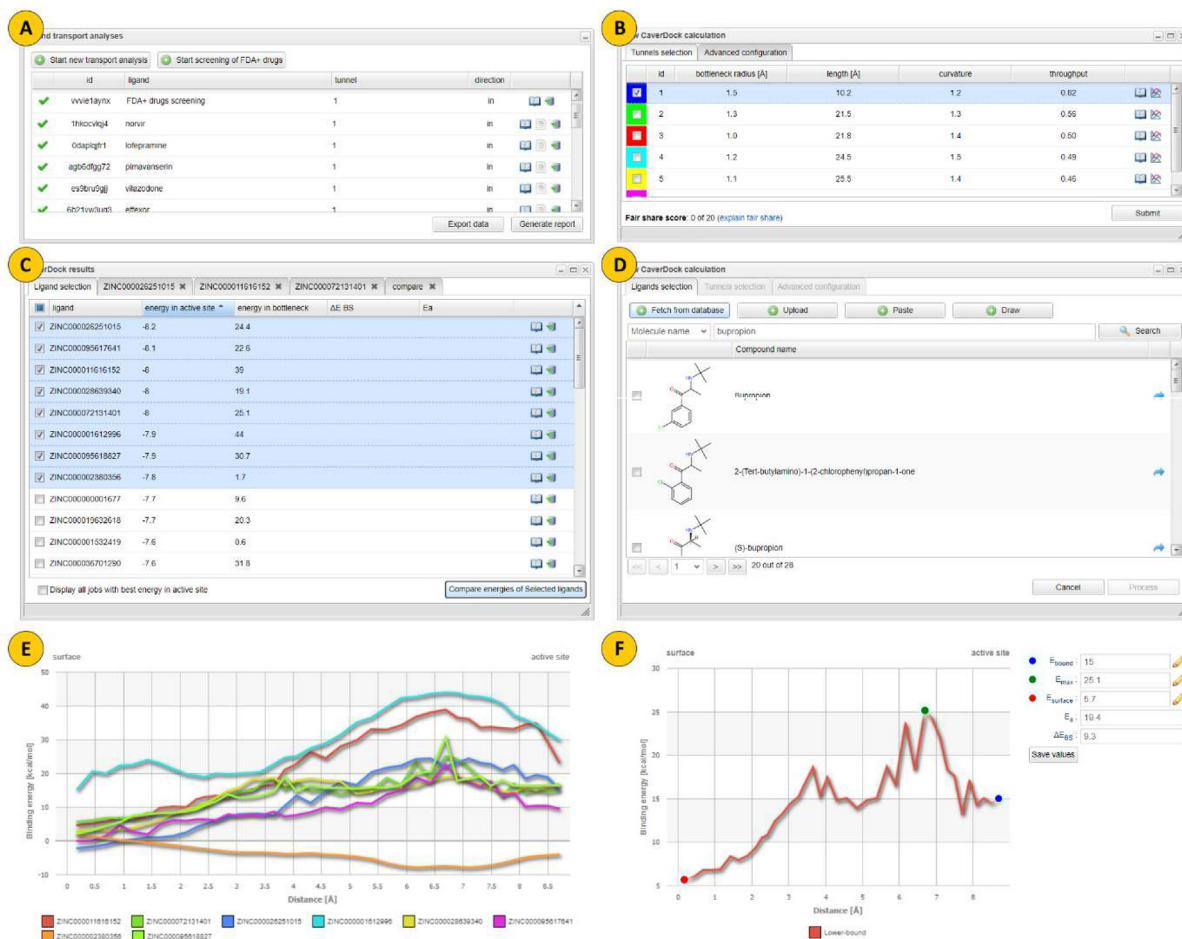


Fig. 2. Graphical user interface of the virtual screening pipeline. A) Ligand transport analysis panel allows users to analyze either a specific ligand or start the virtual screening pipeline using the predefined set of FDA+ drugs. B) Tunnel selection/settings for the virtual screening pipeline. C) The Results table shows the best-fitting ligands. D) Selection of the specific ligand for the CaverDock calculation using the IDSM database. E) Visualization of the calculated binding energies for the whole trajectory. F) Comparison of the energy profiles of the selected ligands.

session [30] for each of the best-performing ligands. It is available via *Download results in a single zip* button.

3. Results

We have evaluated our virtual screening pipeline using the Cytochrome P450 17A1 (CYP17A1). It is one of the most crucial enzymes in the steroidogenic pathway that produces glucocorticoids, androgens, progestins, and estrogens. In humans, CYP17A1 is associated with endocrine effects and steroid hormone metabolism. The mutations in the CYP17A1 gene are therefore connected with the pseudohermaphroditism [31–33], and adrenal hyperplasia [34–35]. Furthermore, the decreased enzyme activity is related to infertility due to hypogonadotropic hypogonadism [36]. CYP17A1 is also an essential target in the treatment of prostate cancer as its inhibition leads to the lower production of androgen required for tumour cell growth [37–40]. The enzyme is expressed by most steroidogenic tissues, including testes, ovaries, and the adrenal cortex. However, it has also been detected in the heart, kidney, and adipose tissue [41].

CYP17A1 crystal structure (PDB ID: 3RUK [42]) was used as an input for the Caver Web calculations. As the enzyme contains an active site that associates with a heme prosthetic group [43], the heme cofactor was selected as a starting point for detecting the active site necessary for tunnel calculation. The integrated Fpock-

et2 software [22] also confirmed the same region. Heme cofactor was set to be preserved in the structure, and the remaining settings were kept in their default values.

Caver Web detected 17 tunnels in total, with only four tunnels being opened more than 1 Å along the whole tunnel. From the list of suggested tunnels, the most biologically relevant tunnel [44] No. 2 was selected for the virtual screening using the predefined set of FDA+ drugs, and the results were compared with the known inhibitors of CYP17A1. We have also evaluated the remaining three well-opened tunnels. However, those tunnels are only modifications of the main tunnel with a smaller opening on the protein's surface or show a significantly higher energetic barrier during the ligand transport.

Abiraterone is the only FDA-approved drug used to treat castration-resistant prostate cancer by inhibiting the CYP17A1, where it binds in its active site and permanently disables the enzyme [42,45]. Abiraterone replaced now rarely applied Ketoconazole which inhibits CYP17A1 competitively, making its effectiveness dependent on the concentration of the administered drug [46]. Two novel drugs, Galeterone and Seviteronel, were investigated as potential inhibitors of CYP17A1. However, those drugs are not included in the set of 4,381 FDA+ drugs as the development of Galeterone was discontinued in 2017.

Firstly, we individually evaluated all four previously described inhibitors by searching them in the IDSM database [21]. The results

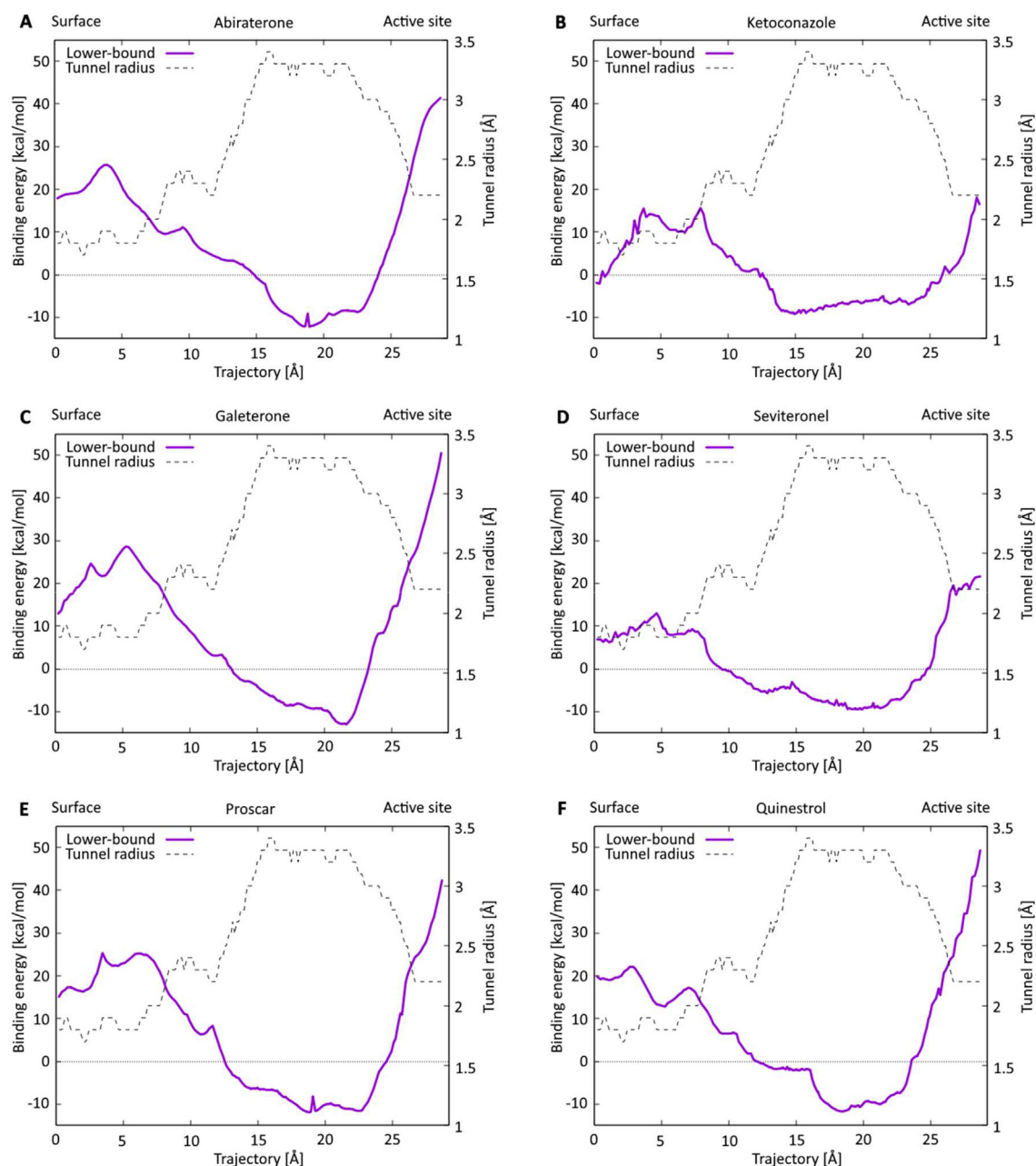


Fig. 3. Lower-bound energies of the ligand transport among the tunnel geometry. The list contains FDA-recognized inhibitors of CYP17A1 Abiraterone (A) and Ketoconazole (B), currently investigated novel inhibitors Galeterone (C) and Seviteronel (D), and the best-scoring ligands from the fully automatized scan of the FDA+ drugs Proscar (E) and Quinestrol (F).

of the ligand transport analysis can be observed in Fig. 3A–D. We can observe that after passing the bottleneck residue Asn202 responsible for the inhibitor's selectivity in the CYP17A1 enzyme [44,47], the Abiraterone binds in the active site of the enzyme with a low binding energy of -11 kcal/mol. A similar result was achieved with Galeterone showing a slightly higher energetic barrier around the gateway residue and lower binding energy of -13 kcal/mol in the active site of the enzyme. The older Ketoconazole shows a lower energetic barrier and a slightly higher binding energy of about -9 kcal/mol. Furthermore, the lowest binding energy of Ketoconazole is reached about 4 and 7 Å closer to the protein surface compared to Abiraterone and Galeterone, respectively. This would suggest that Ketoconazole does not bind directly in the active site and coordinate the heme iron such as Abiraterone through its pyridine nitrogen [42]. Finally, the novel Seviteronel

does not show as promising binding energies as its Galeterone counterpart. However, its energetic barrier is about half of that of Abiraterone. This would make it easier for the ligand to reach the enzyme's active site but also to leave it.

Secondly, we conducted a fully automated scan of the entire FDA+ database to analyze the positions of the known inhibitors in the list of all 4,381 FDA+ drugs. Since the development of Galeterone was discontinued, and Seviteronel is not yet included in the world-wide-approved set of the ZINC database, the analysis was limited to Abiraterone and Ketoconazole. Four variants of Abiraterone are located in the FDA+ list of ligands in the ZINC database. All four were among the top hits of our fully automated scan, with the Abiraterone acetate sold under the brand name Zytiga scoring 15th place from the whole set of 4,381 ligands. Ketoconazole was not placed among the fifty best-scoring ligands with

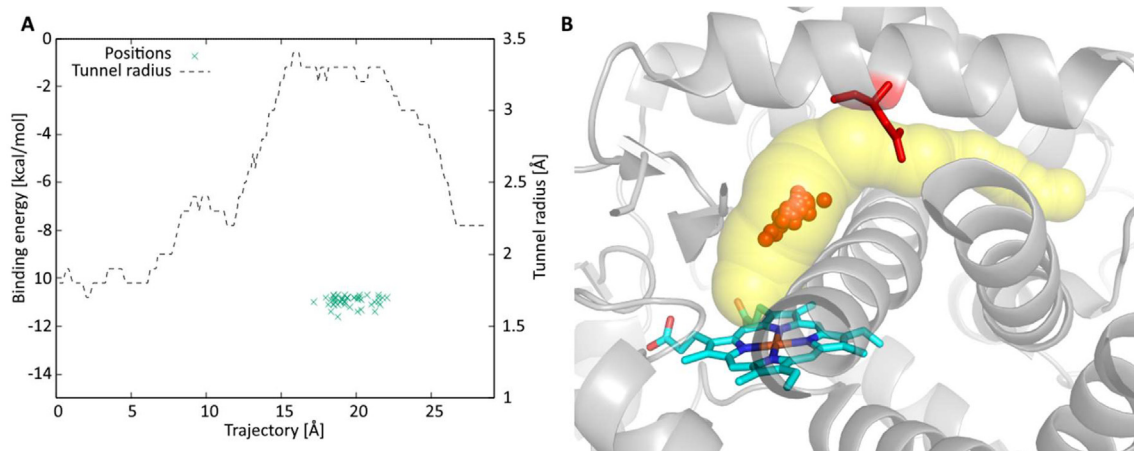


Fig. 4. (A) Position of the lowest binding energies in the proteins tunnel (green crosses) for the set of fifty best-scoring ligands from the fully automatized scan of the FDA+ drugs. (B) The visualization of the enzyme tunnels with the bound best-scoring ligands. The heme cofactor is coloured in cyan; the bottleneck residue Asn202 is highlighted in red. Red spheres denote the positions of bound ligands in the enzyme tunnel. Their locations were calculated as centres of mass of the bound ligands. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a binding energy of 2 kcal/mol lower than Abiraterone. Interestingly, the first two positions on the list were occupied by Proscar (Fig. 3E) and Quinestrol (Fig. 3F). Proscar is a 5 α -Reductase inhibitor that has (similarly to Abiraterone) an antiandrogenic effect and is currently used for the treatment of enlarged prostate. It has been shown to decrease the risk of low-grade prostate cancer. However, the risk of high-grade prostate cancer may increase due to the treatment of some early symptoms [48–49]. Quinestrol is a synthetic estrane steroid already employed as a prostate cancer medication [50]. The top hits list contained several more drugs with a similar purpose (Avodart, Estradiol Benzoate), while the vast majority of the suggested drugs were related to the steroidogenic pathway. Thus, our computational analysis has revealed a large set of drugs that could possibly be repurposed to complement the Abiraterone used as prostate cancer medication. The robustness of our approach could be further substantiated by the fact that Galeterone would be placed on the top of the list if it was already contained in the list of FDA+ drugs.

Finally, Fig. 4 captures the positions of the lowest binding energies of the fifty best-scoring ligands in the enzyme tunnel. From here, it can be observed that all of the suggested ligands are bound directly in the active site of the enzyme, thus interacting with its heme prosthetic group. The ligand transport of the best-scoring drugs can be further analyzed in the [Supplementary material](#) (PyMOL session, MPEGs of the ligands transport, Excel sheet, and raw data in CSV). The data provided in the [supplementary files](#) show almost negligible differences in the top-scoring ligands as the binding energies of the first and the last of the suggested ligands differ in less than 1 kcal/mol. In general, 0.5 kcal/mol deviation can be attributed to the heuristic nature of algorithms, and therefore, the order of ligands can slightly change in each consecutive run. Thus, the user should not consider only one top-scoring molecule as the final inhibitor but rather select several hits for experimental validation.

4. Conclusions

The development of new sequencing strategies and the large genome projects have unearthed the existence of thousands of ubiquitously expressed human proteins from which currently approved drugs target only a tiny fraction. This leaves many proteins to be further exploited by pharmaceutical research. *In silico* techniques provide a fast and cheap alternative to

high-throughput screening. However, they are hard to set up, especially when dealing with the binding sites buried inside the protein structure.

Caver Web is a web tool that incorporates Caver and CaverDock software to provide users with a one-stop-shop solution for the identification and geometrical analysis of tunnels and the calculation of the energy profile of the ligand transport along the selected tunnels. Both aforementioned tools were previously successfully utilized in various studies, including spike glycoprotein of SARS-CoV-2 [51], and leukotriene [44]. However, in the previous version of Caver Web, CaverDock was limited only to a single ligand analysis. A fully automated virtual screening pipeline using the predefined set of 4,381 FDA+ drugs obtained from the ZINC15 database was newly implemented into the Caver Web interface. Users can start the scan of the FDA+ drugs with just a few clicks, and the results are generated automatically. The energy profile of the ligand transport can be visualized for each ligand or compared to other ligands in a single plot. The transport path is generated in the form of a PyMOL session, and all the results can be downloaded as an Excel sheet or PDF report. Furthermore, for the transport analysis of the individual ligands, the IDSM database was integrated to allow the users to fetch ligands outside the ZINC database's scope easily.

Our virtual screening pipeline was evaluated on the cytochrome P450 17A1, and the results were compared with the known CYP17A1 inhibitors. We have analyzed the transport path of Abiraterone, Ketoconazole, Galeterone, and Seviteronel. According to our expectations, Abiraterone has shown superior binding energies compared to the older Ketoconazole. Furthermore, previously investigated Galeterone slightly surpasses Abiraterone binding capabilities (approximately 2 kcal/mol), while Seviteronel offers a significantly lower energetic barrier (approximately 13 kcal/mol) around the enzyme's bottleneck residue. In the second step, we run a fully automated virtual screening pipeline using the entire set of 4,381 FDA+ drugs. All variations of Abiraterone contained in the ZINC database have scored among the best potential inhibitors for the CYP17A1 enzyme with a negligible difference in binding energies of less than 1 kcal/mol compared to the best-scoring ligand. In contrast, many best-scoring drugs show the same or similar function to Abiraterone. This leaves us with a large set of drugs that could be potentially repurposed as Abiraterone alternatives.

To conclude, the new version of Caver Web introduces a fast and effective virtual screening pipeline combining molecular dock-

ing with ligand transport analysis over a wide range of ligands, thus making it a viable service for the broader spectrum of the scientific community. The server is freely available for non-commercial use at <https://loschmidt.chemi.muni.cz/caverweb>.

CRedit authorship contribution statement

Milos Musil: Conceptualization, Supervision, Software, Formal analysis, Writing – original draft, Writing – review & editing. **Andrej Jezik:** Software, Methodology. **Marie Jankujova:** Software, Visualization. **Jan Stourac:** Conceptualization, Software. **Jakub Galgonek:** Software, Writing – review & editing. **Saltuk Mustafa Eyrilmez:** Validation. **Jiri Vondrasek:** Conceptualization, Supervision. **Jiri Damborsky:** Conceptualization, Supervision, Writing – review & editing, Funding acquisition. **David Bednar:** Conceptualization, Supervision, Writing – review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

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

Acknowledgements

The authors would like to express their thanks to the Czech Ministry of Education (TEAMING CZ.02.1.01/0.0/0.0/17_043/0009 632, INBIO CZ.02.1.01/0.0/0.0/16_026/0008451, Exceles Onco LX22NPO5102); the Czech Grant Agency (20-15915Y); the Technology Agency of Czech Republic (FW03010208); Brno University of Technology (FIT-S-20-6293); and the European Commission (TEAMING 857560 and SINFONIA 814418). Computational resources were supplied by the project ‘e-Infrastruktura CZ’ (LM2018140) and ELIXIR CZ (LM2018131).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2022.11.031>.

References

- [1] Wagner BK, Schreiber SL. The Power of Sophisticated Phenotypic Screening and Modern Mechanism-of-Action Methods. *Cell Chem Biol* 2016;23(1):3–9. <https://doi.org/10.1016/j.chembiol.2015.11.008>.
- [2] Zheng W, Thorne N, McKew JC. Phenotypic Screens as a Renewed Approach for Drug Discovery. *Drug Discov Today* 2013;18(21):1067–73. <https://doi.org/10.1016/j.drudis.2013.07.001>.
- [3] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* 2021;596(7873):583–9. <https://doi.org/10.1038/s41586-021-03819-2>.
- [4] Eder J, Sedrani R, Wiesmann C. The Discovery of First-in-Class Drugs: Origins and Evolution. *Nat Rev Drug Discov* 2014;13(8):577–87. <https://doi.org/10.1038/nrd4336>.
- [5] An Integrated Encyclopedia of DNA Elements in the Human Genome. *Nature* 2012;489(7414):57–74. <https://doi.org/10.1038/nature11247>.
- [6] Kim M-S, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A Draft Map of the Human Proteome. *Nature* 2014;509(7502):575–81. <https://doi.org/10.1038/nature13302>.
- [7] Geiger T, Wehner A, Schaab C, Cox J, Mann M. Comparative Proteomic Analysis of Eleven Common Cell Lines Reveals Ubiquitous but Varying Expression of Most Proteins. *Mol Cell Proteomics* 2012;11(3). <https://doi.org/10.1074/mcp.M111.014050>.
- [8] Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, et al. Mass-Spectrometry-Based Draft of the Human Proteome. *Nature* 2014;509(7502):582–7. <https://doi.org/10.1038/nature13319>.
- [9] Uhlen M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-Based Map of the Human Proteome. *Science* 2015. <https://doi.org/10.1126/science.1260419>.
- [10] DiMasi JA, Grabowski HG, Hansen RW. Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs. *J Health Econ* 2016;47:20–33. <https://doi.org/10.1016/j.jhealeco.2016.01.012>.
- [11] Karamehic J, Ridic O, Ridic G, Jukic T, Coric J, Subasic D, et al. Financial Aspects and the Future of the Pharmaceutical Industry in the United States of America. *Mater Sociomed* 2013;25(4):286–90. <https://doi.org/10.5455/msm.2013.25.286-290>.
- [12] Vavra O, Filipovic J, Plhak J, Bednar D, Marques SM, Brezovsky J, et al. CaverDock: A Molecular Docking-Based Tool to Analyse Ligand Transport through Protein Tunnels and Channels. *Bioinformatics* 2019;35(23):4986–93. <https://doi.org/10.1093/bioinformatics/btz386>.
- [13] Lee P-H, Kuo K-L, Chu P-Y, Liu EM, Lin J-H. SLITHER: A Web Server for Generating Contiguous Conformations of Substrate Molecules Entering into Deep Active Sites of Proteins or Migrating through Channels in Membrane Transporters. *Nucleic Acids Res* 2009;37(suppl_2):W559–64. <https://doi.org/10.1093/nar/gkp359>.
- [14] Devaurs D, Bouard L, Vaisset M, Zanon C, Al-Blawi I, Iehl R, et al. MoMA-LigPath: A Web Server to Simulate Protein-Ligand Unbinding. *Nucleic Acids Res* 2013;41(W1):W297–302. <https://doi.org/10.1093/nar/gkt380>.
- [15] Trott O, Olson AJ. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J Comput Chem* 2010;31(2):455–61. <https://doi.org/10.1002/jcc.21334>.
- [16] Slater O, Kontoyianni M. The Compromise of Virtual Screening and Its Impact on Drug Discovery. *Expert Opin Drug Discov* 2019;14(7):619–37. <https://doi.org/10.1080/17460441.2019.1604677>.
- [17] Shoichet BK. Virtual Screening of Chemical Libraries. *Nature* 2004;432(7019):862–5. <https://doi.org/10.1038/nature03197>.
- [18] Stourac J, Vavra O, Kokkonen P, Filipovic J, Pinto G, Brezovsky J, et al. Caver Web 1.0: Identification of Tunnels and Channels in Proteins and Analysis of Ligand Transport. *Nucleic Acids Res* 2019;47(W1):W414–22. <https://doi.org/10.1093/nar/gkz378>.
- [19] Chovanova E, Pavelka A, Benes P, Srnad O, Brezovsky J, Kozlikova B, et al. 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput Biol* 2012;8(10):e1002708.
- [20] Sterling T, Irwin JJ. ZINC 15 – Ligand Discovery for Everyone. *J Chem Inf Model* 2015;55(11):2324–37. <https://doi.org/10.1021/acs.jcim.5b00559>.
- [21] Galgonek J, Vondrášek J. IDSM ChemWebRDF: SPARQLing Small-Molecule Datasets. *J Cheminform* 2021;13(1):38. <https://doi.org/10.1186/s13321-021-00515-1>.
- [22] Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinf* 2009;10(1):168. <https://doi.org/10.1186/1471-2105-10-168>.
- [23] Ribeiro AJM, Holliday GL, Furnham N, Tyzack JD, Ferris K, Thornton JM. Mechanism and Catalytic Site Atlas (M-CSA): A Database of Enzyme Reaction Mechanisms and Active Sites. *Nucleic Acids Res* 2018;46(D1):D618–23. <https://doi.org/10.1093/nar/gkx1012>.
- [24] UniProt Consortium, T. UniProt: The Universal Protein Knowledgebase. *Nucleic Acids Research* 2018;46(5):2699–2699. <https://doi.org/10.1093/nar/gky092>.
- [25] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J Comput Chem* 2009;30(16):2785–91. <https://doi.org/10.1002/jcc.21256>.
- [26] Filipović J, Vávra O, Plhák J, Bednář D, Marques SM, Brezovský J, et al. CaverDock: A Novel Method for the Fast Analysis of Ligand Transport. *IEEE/ACM Trans Comput Biol Bioinf* 2020;17(5):1625–38. <https://doi.org/10.1109/TCBB.2019.2907492>.
- [27] Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem Substance and Compound Databases. *Nucleic Acids Res* 2016;44(D1):D1202–13. <https://doi.org/10.1093/nar/gkv951>.
- [28] Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res* 2012;40(D1):D1100–7. <https://doi.org/10.1093/nar/gkr777>.
- [29] Degtyarenko K, de Matos P, Ennis M, Hastings J, Zbinden M, McNaught A, et al. A Database and Ontology for Chemical Entities of Biological Interest. *Nucleic Acids Res* 2008;36(suppl_1):D344–50. <https://doi.org/10.1093/nar/gkm791>.
- [30] PyMOL | pymol.org. <https://pymol.org/2/> (accessed 2022-06-30).
- [31] Yang J, Cui B, Sun S, Shi T, Zheng S, Bi Y, et al. Phenotype-Genotype Correlation in Eight Chinese 17 α -Hydroxylase/17,20 Lyase-Deficiency Patients with Five Novel Mutations of CYP17A1 Gene. *The Journal of Clinical Endocrinology & Metabolism* 2006;91(9):3619–25. <https://doi.org/10.1210/jc.2005-2283>.
- [32] Gong Y, Qin F, Li W-J, Li L-Y, He P, Zhou X-J. Cytochrome P450 Family 17 Subfamily A Member 1 Mutation Causes Severe Pseudohermaphroditism: A Case Report. *World J Clin Cases* 2022;10(11):3553–60. <https://doi.org/10.12998/wjcc.v10.i11.3553>.
- [33] Wang M, Wang H, Zhao H, Li L, Liu M, Liu F, et al. Prevalence of CYP17A1 Gene Mutations in 17 α -Hydroxylase Deficiency in the Chinese Han Population. *Clin Hypertens* 2019;25(1):23. <https://doi.org/10.1186/s40885-019-0128-6>.
- [34] Mula-Abed W-A-S, Pambinezhuth FB, Al-Kindi MK, Al-Busaidi NB, Al-Muslahi HN, Al-Lamki MA. Congenital Adrenal Hyperplasia Due to 17-Alpha-Hydroxylase/17,20-Lyase Deficiency Presenting with Hypertension and Pseudohermaphroditism: First Case Report from Oman. *Oman Med J* 2014;29(1):55–9. <https://doi.org/10.5001/omj.2014.12>.
- [35] Costa-Santos M, Kater CE, Auchus RJ, Brazilian Congenital Adrenal Hyperplasia Multicenter Study Group. Two Prevalent CYP17 Mutations and Genotype-Phenotype Correlations in 24 Brazilian Patients with 17-Hydroxylase

- Deficiency. *J Clin Endocrinol Metab* 2004;89(1):49–60. <https://doi.org/10.1210/jc.2003-031021>.
- [36] Acién P, Acién M. Disorders of Sex Development: Classification, Review, and Impact on Fertility. *J Clin Med* 2020;9(11):3555. <https://doi.org/10.3390/jcm9113555>.
- [37] Kostin VA, Zolotsev VA, Kuzikov AV, Masamreh RA, Shumyantseva VV, Veselovsky AV, et al. Oxazoliny Derivatives of [17(20)E]-21-Norpregnene Differing in the Structure of A and B Rings. Facile Synthesis and Inhibition of CYP17A1 Catalytic Activity. *Steroids* 2016;115:114–22. <https://doi.org/10.1016/j.steroids.2016.06.002>.
- [38] Bonomo S, Hansen CH, Petrunak EM, Scott EE, Styrisshave B, Jørgensen FS, et al. Promising Tools in Prostate Cancer Research: Selective Non-Steroidal Cytochrome P450 17A1 Inhibitors. *Sci Rep* 2016;6:29468. <https://doi.org/10.1038/srep29468>.
- [39] Mostaghel EA, Marck BT, Plymate SR, Vessella RL, Balk S, Matsumoto AM, et al. Resistance to CYP17A1 Inhibition with Abiraterone in Castration-Resistant Prostate Cancer: Induction of Steroidogenesis and Androgen Receptor Splice Variants. *Clin Cancer Res* 2011;17(18):5913–25. <https://doi.org/10.1158/1078-0432.CCR-11-0728>.
- [40] Cheong EY, Nair PC, Neo RWY, Tu HT, Lin F, Chiong E, et al. Slow-, Tight-Binding Inhibition of CYP17A1 by Abiraterone Redefines Its Kinetic Selectivity and Dosing Regimen. *J Pharmacol Exp Ther* 2020;374(3):438–51. <https://doi.org/10.1124/jpet.120.265868>.
- [41] Storbeck K-H, Swart P, Africander D, Conradie R, Louw R, Swart AC. 16 α -Hydroxyprogesterone: Origin, Biosynthesis and Receptor Interaction. *Mol Cell Endocrinol* 2011;336(1–2):92–101. <https://doi.org/10.1016/j.mce.2010.11.016>.
- [42] DeVore NM, Scott EE. Structures of Cytochrome P450 17A1 with Prostate Cancer Drugs Abiraterone and TOK-001. *Nature* 2012;482(7383):116–9. <https://doi.org/10.1038/nature10743>.
- [43] Vasaitis TS, Bruno RD, Njar VCO. CYP17 Inhibitors for Prostate Cancer Therapy. *J Steroid Biochem Mol Biol* 2011;125(1–2):23–31. <https://doi.org/10.1016/j.jsbmb.2010.11.005>.
- [44] Enzyme Tunnels and Gates As Relevant Targets in Drug Design - Marques - 2017 - Medicinal Research Reviews - Wiley Online Library. <https://onlinelibrary.wiley.com/doi/full/10.1002/med.21430> (accessed 2022-08-04).
- [45] Fernández-Cancio M, Camats N, Flück CE, Zalewski A, Dick B, Frey BM, et al. Mechanism of the Dual Activities of Human CYP17A1 and Binding to Anti-Prostate Cancer Drug Abiraterone Revealed by a Novel V366M Mutation Causing 17,20 Lyase Deficiency. *Pharmaceuticals (Basel)* 2018;11(2):E37. <https://doi.org/10.3390/ph11020037>.
- [46] Eil C. Ketoconazole Binds to the Human Androgen Receptor. *Horm Metab Res* 1992;24(8):367–70. <https://doi.org/10.1055/s-2007-1003337>.
- [47] Liu Y, Denisov I, Gregory M, Sligar SG, Kincaid JR. Importance of Asparagine 202 in Manipulating Active Site Structure and Substrate Preference for Human CYP17A1. *Biochemistry* 2022;61(7):583–94. <https://doi.org/10.1021/acs.biochem.2c00023>.
- [48] Varothai S, Bergfeld WF. Androgenetic Alopecia: An Evidence-Based Treatment Update. *Am J Clin Dermatol* 2014;15(3):217–30. <https://doi.org/10.1007/s40257-014-0077-5>.
- [49] *Finasteride for Prostate Cancer Prevention - NCI*. <https://www.cancer.gov/types/prostate/research/finasteride-reduces-low-grade> (accessed 2022-08-04).
- [50] Zink, C. *Dictionary of Obstetrics and Gynecology*; Walter de Gruyter, 2011.
- [51] *Screening of world approved drugs against highly dynamical spike glycoprotein of SARS-CoV-2 using CaverDock and machine learning - ScienceDirect*. <https://www.sciencedirect.com/science/article/pii/S2001037021002245> (accessed 2022-08-04).