Supplementary data

CAVER Analyst 1.0: Graphic tool for interactive visualization and analysis of tunnels in protein structures

Barbora Kozlikova[1,#], Eva Sebestova[2,#] Vilem Sustr[1], Jan Brezovsky[2], Ondrej Strnad[1], Lukas Daniel[2], David Bednar[2], Antonin Pavelka[1,2], Martin Manak[3], Martin Bezdeka[1], Petr Benes[1], Matus Kotry[1], Artur Gora[2], Jiri Damborsky[2,*], Jiri Sochor[1,*]

[1] Human Computer Interaction Laboratory, Faculty of Informatics, Masaryk University, Botanicka 68a, 602 00 Brno, Czech Republic; [2] Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment RECETOX, Faculty of Science, Masaryk University, Kamenice 5/A13, 625 00 Brno, Czech Republic; [3] Department of Computer Science and Engineering, Faculty of Applied Sciences, University of West Bohemia, Univerzitni 8, 306 14 Plzen, Czech Republic

[#]These authors contributed equally

Contents

Implementation and configuration

CAVER Analyst is the JAVA-based software. Due to its modular architecture, it can easily be customized and distributed as a standalone application. CAVER Analyst runs on a common hardware without the need for special hardware upgrades. This allows for seamless integration into existing IT environments.

The application is supported by the following operating systems: Windows XP, 7 or 8, Mac OS X 10.7.3 or later and major distributions of Linux including Fedora Core, Red Hat and Ubuntu. The application can run on both 32-bit and 64-bit system architectures and requires JAVA version 1.7 or later.

Necessary hardware configuration for processing of small data sets is 32-bit architecture with 2-4 GB RAM. Large datasets require 64-bit architecture with 8 and more GB RAM. AMD Radeon or NVIDIA GeForce dedicated graphics cards are recommended for the utilization of advanced visualization techniques of the CAVER Analyst.

The CAVER Analyst application is distributed as a complete package with all modules required for running the application, including the user guide.

The example data sets are available at http://caver.cz/fil/download/examples/examples.zip.

Case studies

Case study 1 – Engineering enzyme activity

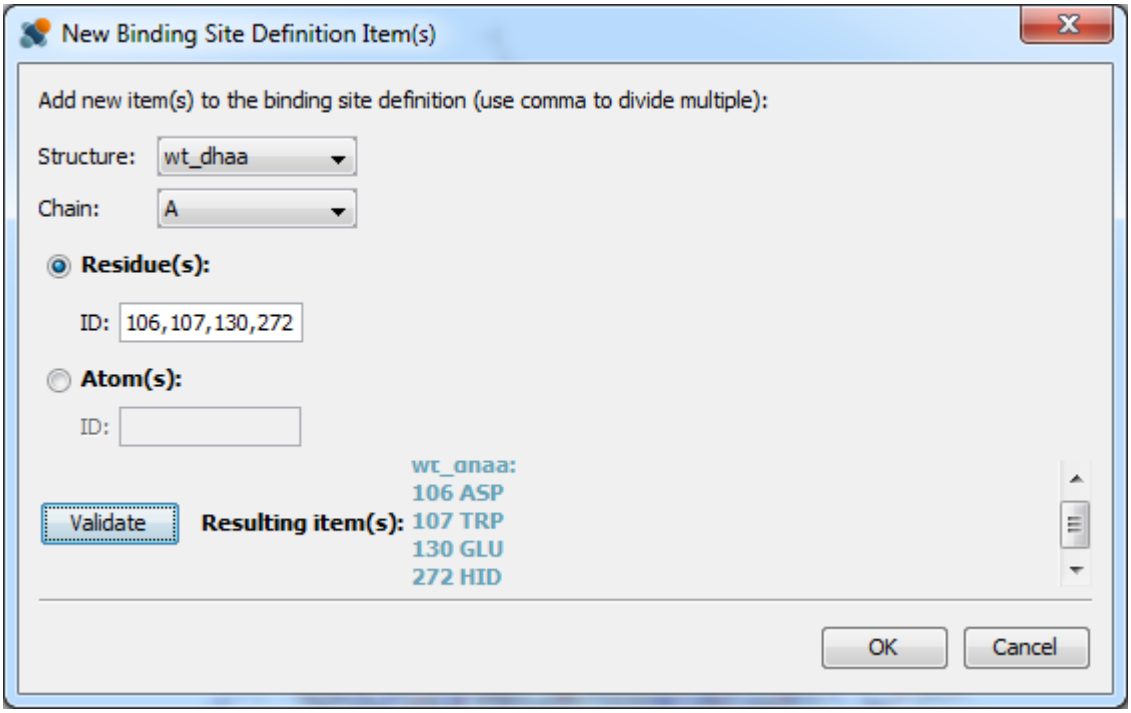
Redesigning access tunnels of the haloalkane dehalogenase DhaA resulted in 32-fold increase in activity with toxic environmental pollutant 1,2,3-trichloropropane (TCP). The study demonstrates the power of combining rational design with directed evolution focused to the access tunnels. Additional information can be found in the article by Pavlova *et al.* published in *Nature Chemical Biology* [1].

1. Loading the structures

The study was performed with wild type and C176Y mutant of the haloalkane dehalogenase DhaA. The users are provided with the structures in PDB format wt_dhaa.pdb and m1_dhaa.pdb. These structures can be obtained from http://caver.cz/fil/download/examples/examples.zip upon unzipping *structures* folder.

2. Setting the binding site for tunnel computation

The CSA database does not contain any records for these structures thus the user must define the starting point manually, e.g., by defining the surrounding residues. In both structures, the active site is defined by the residues 106, 107, 130 and 272. They can be set using the *Add Item…* button in the *Tunnel Computation* window, section *Surrounding items*. It opens the following window:



3. Launching tunnel computation

In this case the parameters present in the *Tunnel Computation* window or *the CAVER 3 Computation Settings* should have the following values:

- *Maximum distance (Å)* = 5
- *Minimal probe radius* = 0.7

The computation is then launched using the "*Compute Tunnels*" button.

This is performed for both structures and then their structure alignment can be achieved using *Structure → Alignment*.

4. Interpretation of results

The active site of wild type DhaA is an occluded cavity with two major access tunnels: (i) the main tunnel and (ii) the slot tunnel (Figure 1). Mutant carrying substitution in the main tunnel (C176Y) limits the access of water molecules to the active site cavity and resulting in higher activity with TCP (see Figure 2).

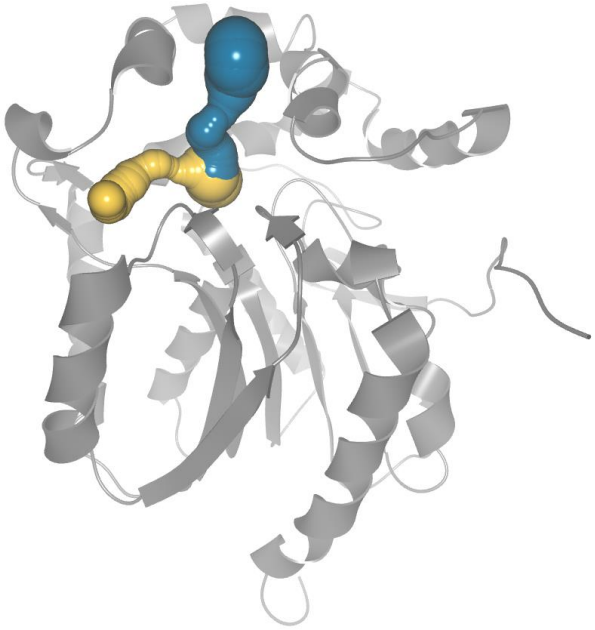


Figure 1. The main (in blue) and the slot (yellow) tunnels of the haloalkane dehalogenase DhaA.

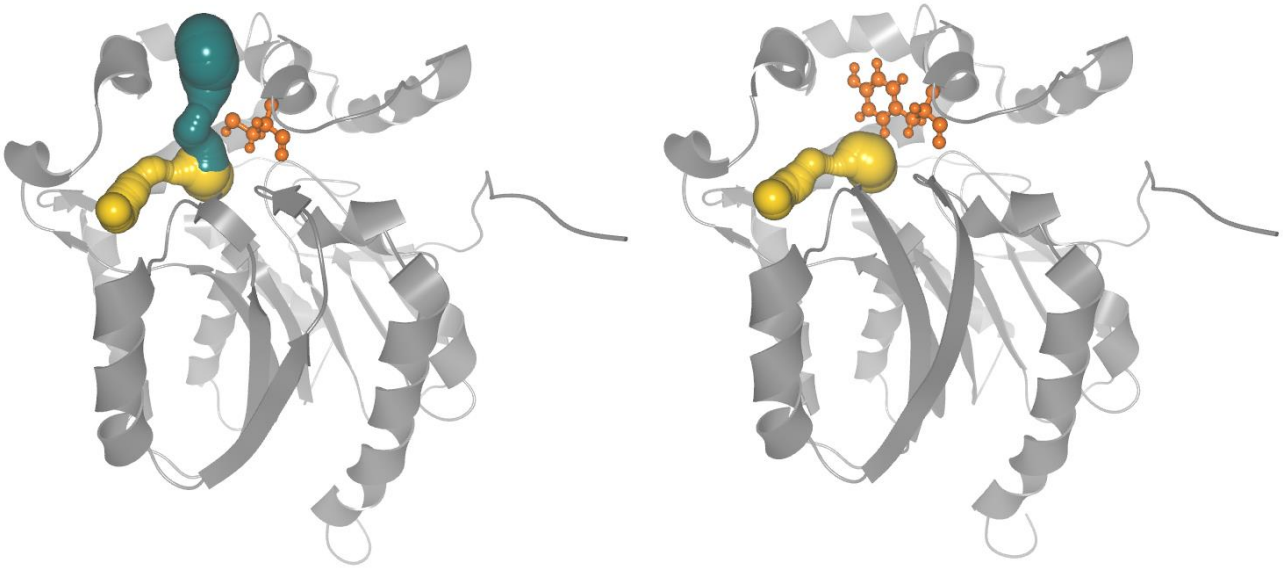


Figure 2. Comparison of the wild type DhaA (left) and the DhaA mutant C176Y (right). The main tunnel is closed by the introduced mutation (in orange).

Case study 2 – Engineering enzyme stability

The modification of residues lining the access tunnel of the haloalkane dehalogenase DhaA increased its melting temperature by 17°C and resistance to co-solvent DMSO (dimethylsulfoxide) 4000-fold. Mutations in the tunnel improved structural and kinetic stability, while the surface mutations did not contribute to protein stabilization. More information can be found in the article by Koudelakova *et al.* published in *Angewandte Chemie* [2] and the US Patent by Damborsky *et al.* [3].

1. Loading the structures

The study was performed with the wild type of haloalkane dehalogenase DhaA and its mutant DhaA80. Protein structures are available in 1cqw.pdb and 4f60.pdb files. Both structures are available at http://caver.cz/fil/download/examples/examples.zip in the *structures* folder created upon unzipping.

Setting the binding site for tunnel computation

The starting point for the 1cqw.pdb structure is automatically loaded from the CSA database. This point can be copied to the other structures using the "*Convert for Another Structure…*" button in the *Tunnel Computation* window.

3. Launching tunnel computation

Following parameters should be used for tunnel computation:

- For both structures the *Maximum distance (Å)* parameter is set to 5
- For the 1cqw the *Min. probe radius* is set to 0.7
- For the 4f60 the *Min. probe radius* is set to 0.5

In the computed tunnel sets, the main tunnel is the first tunnel in the 1cqw and the second tunnel in the 4f60 structure.

4. Interpretation of results

The bottleneck of the main tunnel decreased from 1.4 Å in wild type DhaA (blue) to 0.6 Å in DhaA80 (Figure 3). The introduced mutations: (i) sealed the tunnel and prevented penetration of DMSO into the interior of the active site; and (ii) improved the hydrophobic packing of the tunnel residues improving stability at elevated temperatures.

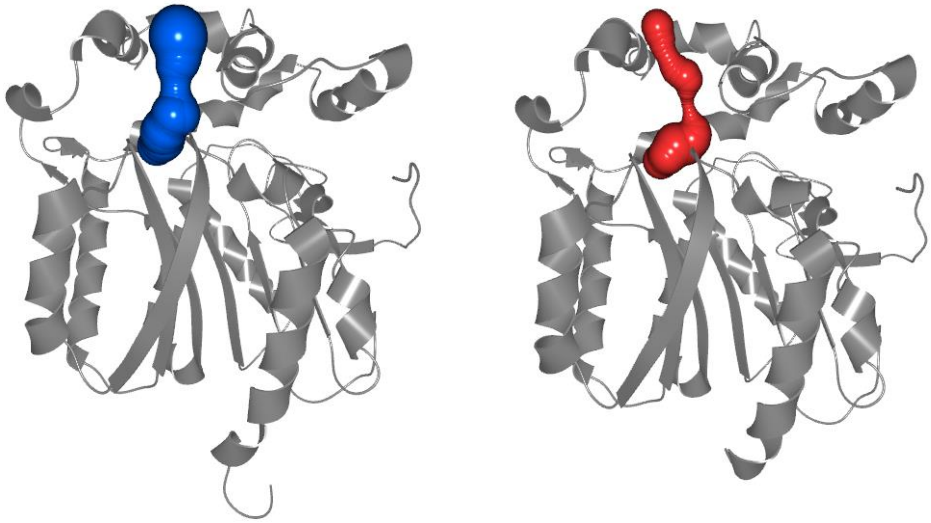


Figure 3. The main tunnel of the wild type DhaA (left) and the mutant DhaA80 (right).

This concept was validated by prediction of stabilizing effects of over 227,000 mutations in 26 proteins from all six enzyme classes. The tunnel mutations provided 2-times higher chance to produce protein variants with significantly improved stability than mutagenesis targeting other protein regions [2].

5. Graphs for tunnels

The graph statistics of computed tunnels can be obtained using the icon for graph visualization (Figure 4).

Figure 4. Blue circle shows the icon for launching the graph visualization of given tunnel set.

This icon opens the *Tunnel Graph* window with the main tunnel from the 1cqw.pdb structure. To add the tunnel from another structure to the same graph, the button "Freeze" must be pressed. Then another tunnel can be added by pressing the same icon at Figure 4 for the second structure, providing the final graph (Figure 5).

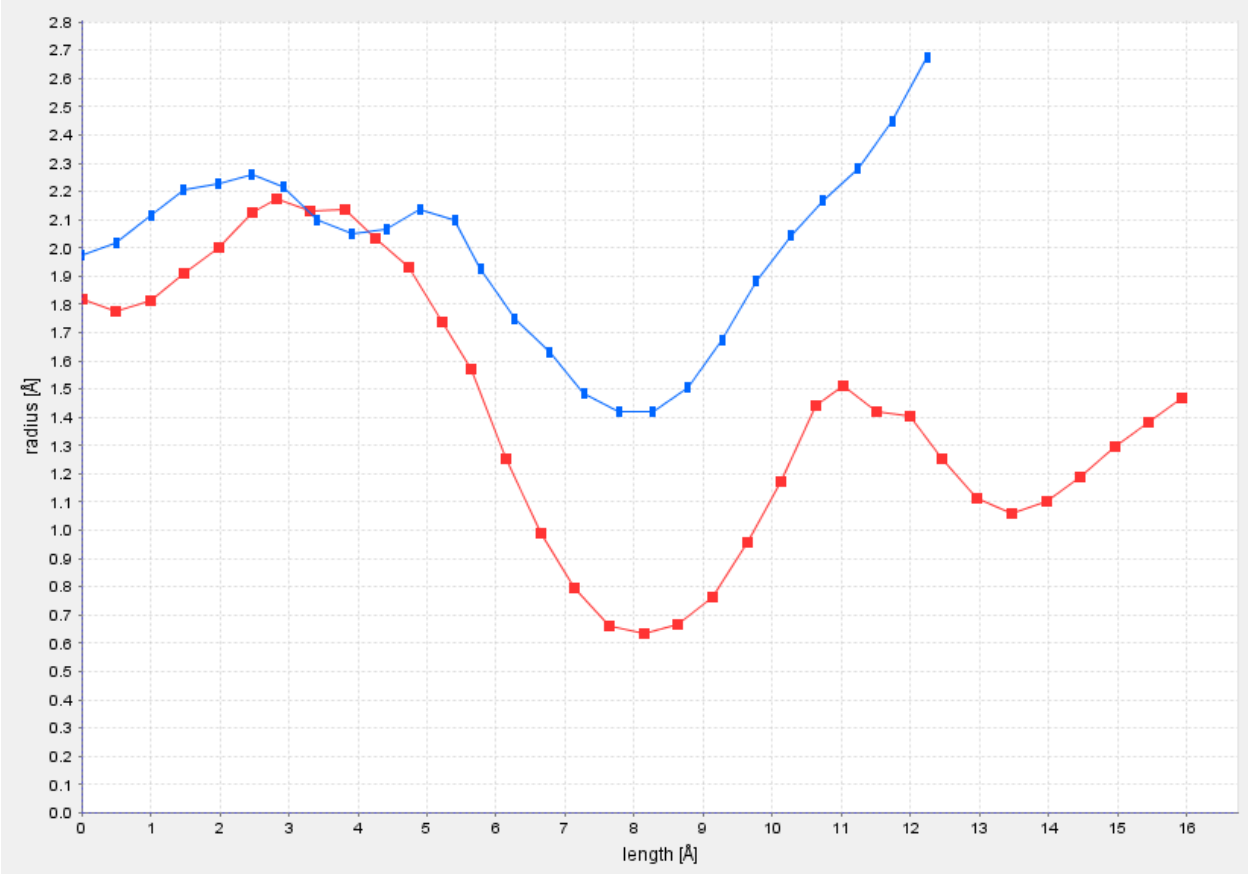
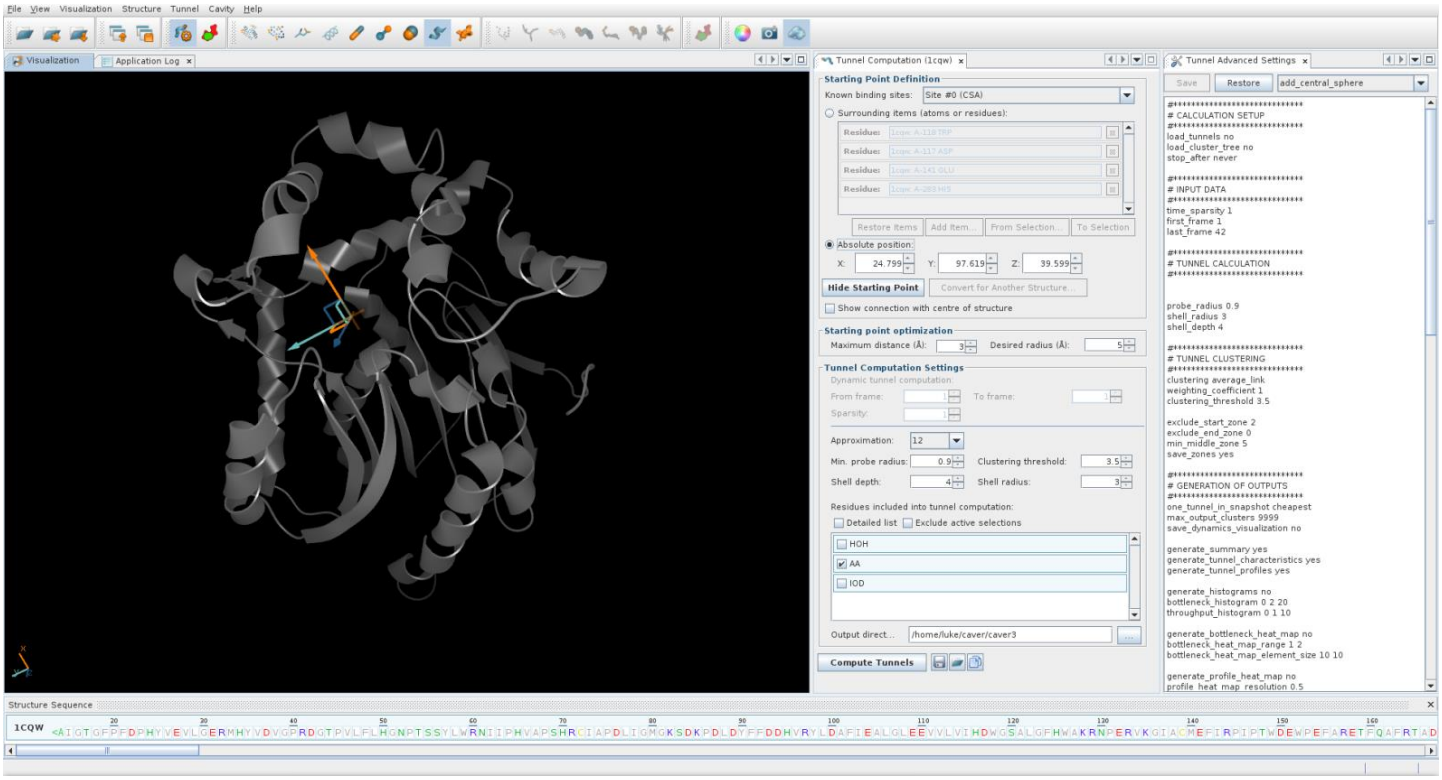
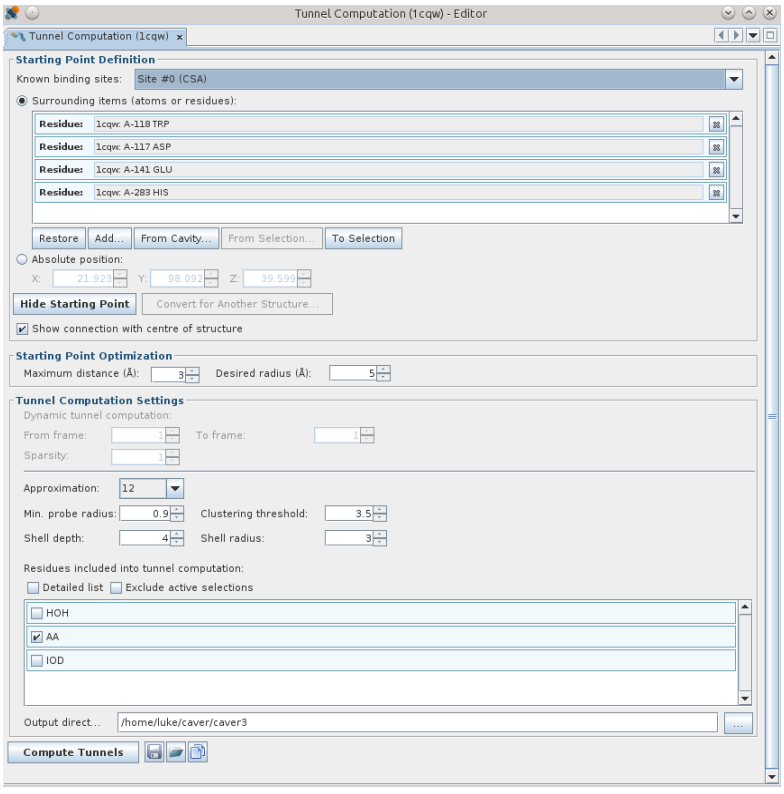


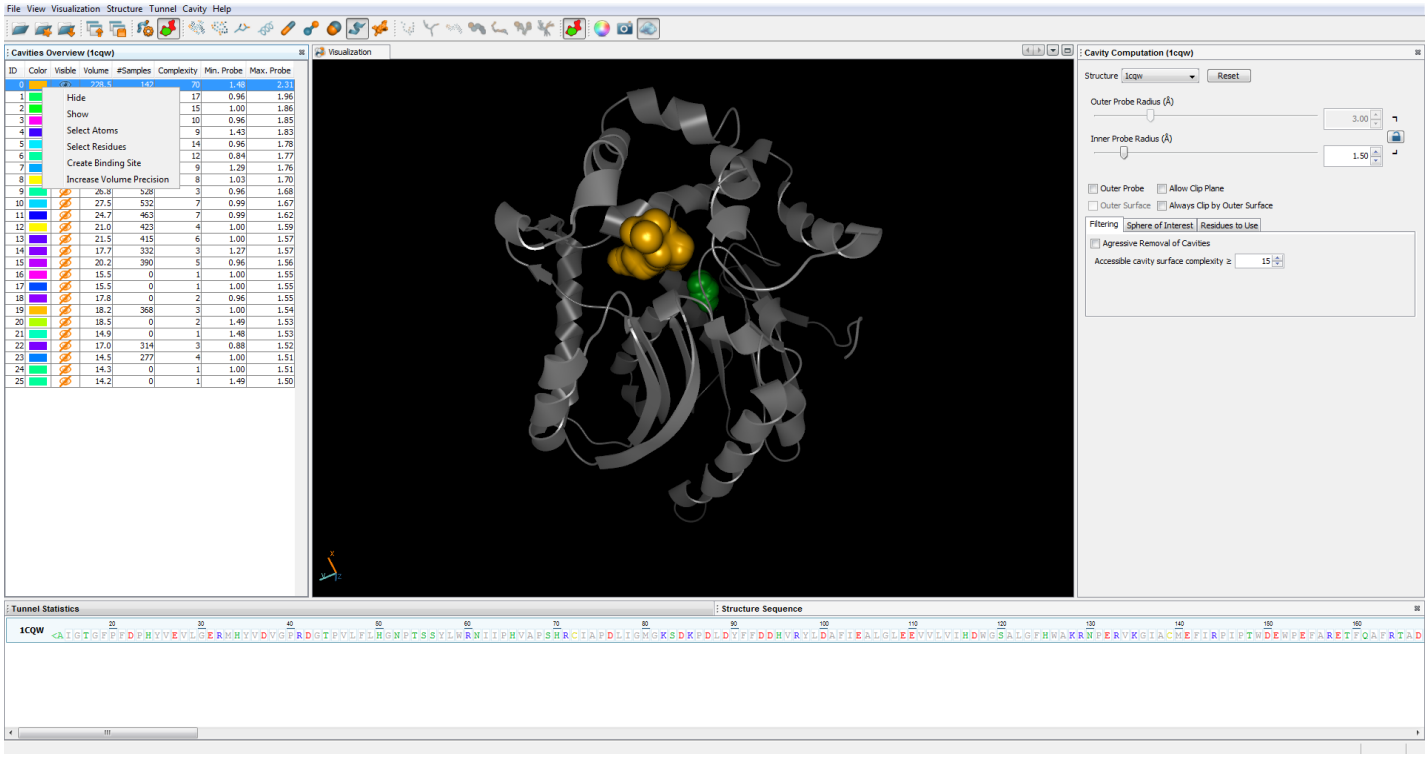
Figure 5. Tunnel graph comparing the length and the radius of two main tunnels for the wild type DhaA (blue) and the mutant DhaA80 (red).

Supplementary figures





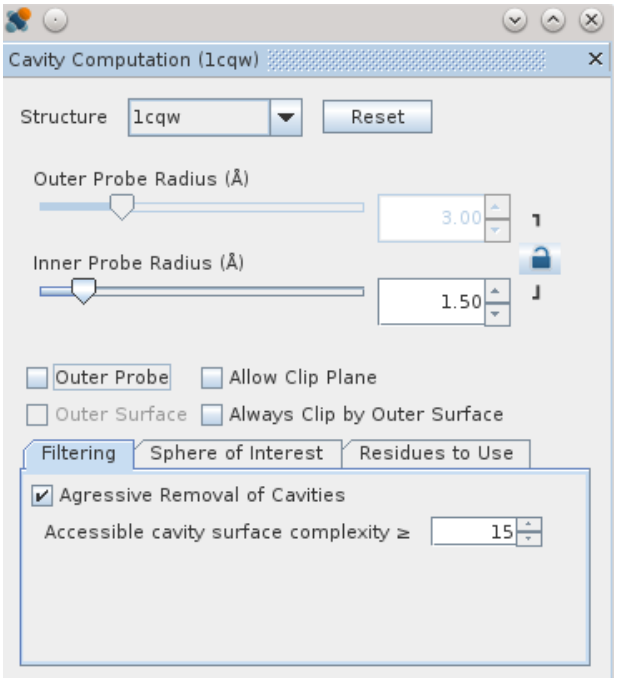
Supplementary figure 1. Tunnel calculation. The starting point for the calculation (origin of arrows) can be specified either automatically (loaded from the Catalytic Site Atlas) or manually (from selections, calculated cavities and specified atoms, residues or coordinates). The starting point can be transferred to any other loaded structure. All calculation settings can be loaded from an external file. Layout of CAVER Analyst (top), detail of tunnel calculation settings (bottom).



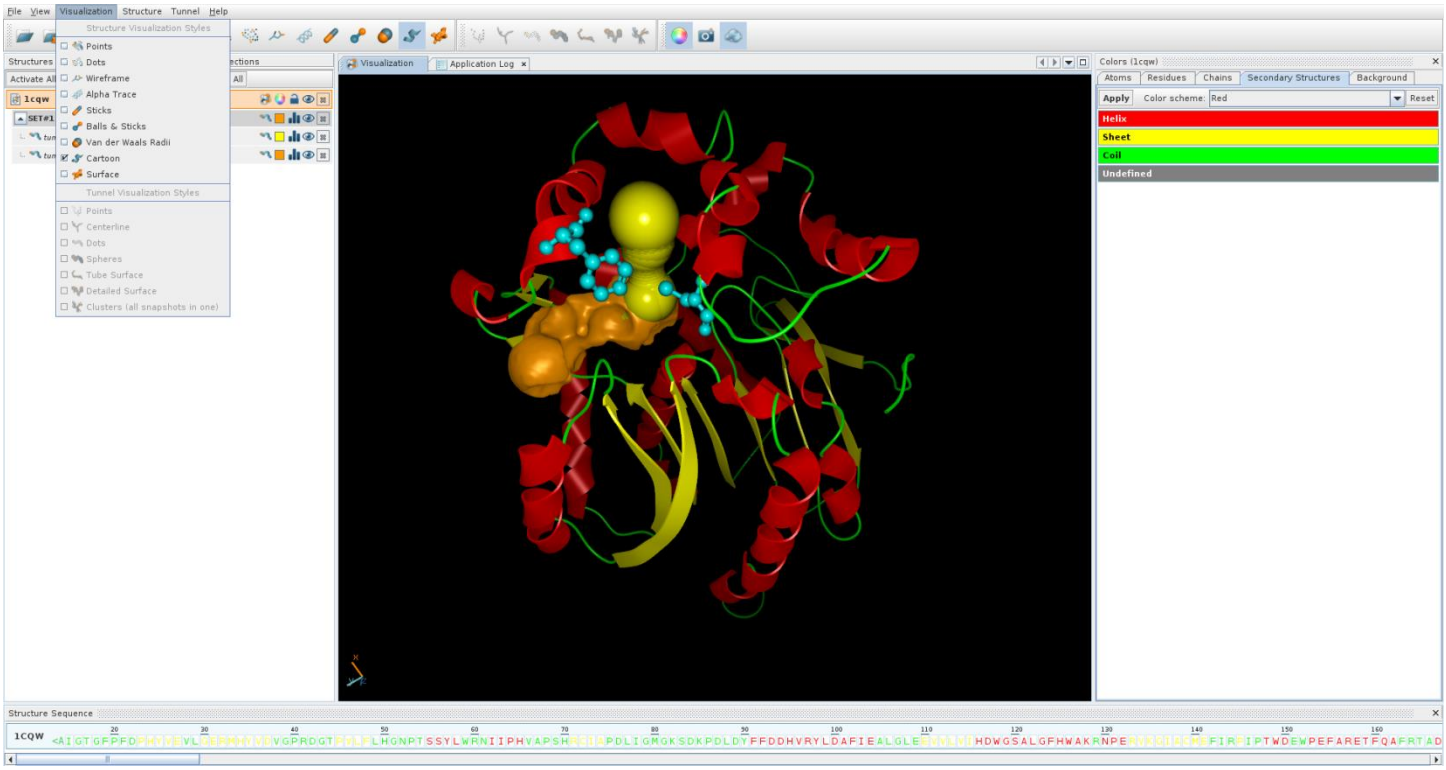
Cavities Overview (1cqw)

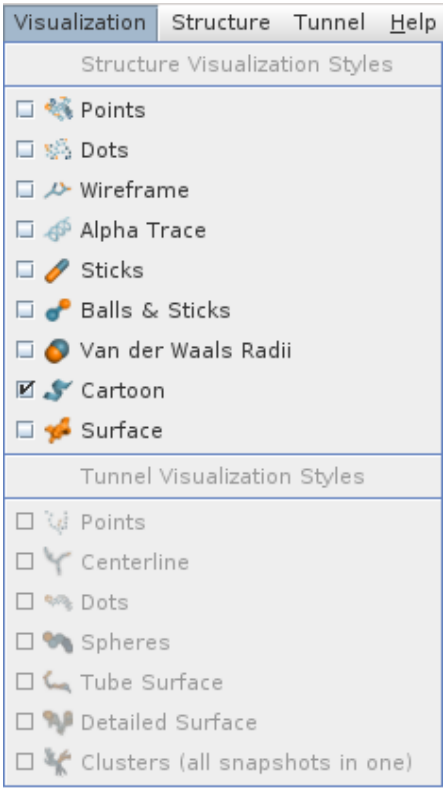
ID	Color	Visible	Volume	#Samples	Complexity	Min. Probe	Max. Probe
0			228.5	142	70	1.48	2.31
1			60.2	588	17	0.96	1.96
2					15	1.00	1.86
3					10	0.96	1.85
4					9	1.43	1.83
5					14	0.96	1.78
6					12	0.84	1.77
7					9	1.29	1.76
8					8	1.03	1.70
9					3	0.96	1.68
10			27.5	532	7	0.99	1.67
11			24.7	463	7	0.99	1.62
12			21.0	423	4	1.00	1.59
13			21.5	415	6	1.00	1.57
14			17.7	332	3	1.27	1.57
15			20.2	390	5	0.96	1.56
16			15.5	0	1	1.00	1.55
17			15.5	0	1	1.00	1.55
18			17.8	0	2	0.96	1.55
19			18.2	368	3	1.00	1.54
20			18.5	0	2	1.49	1.53
21			14.9	0	1	1.48	1.53
22			17.0	314	3	0.88	1.52
23			14.5	277	4	1.00	1.51
24			14.3	0	1	1.00	1.51
25			14.2	0	1	1.49	1.50

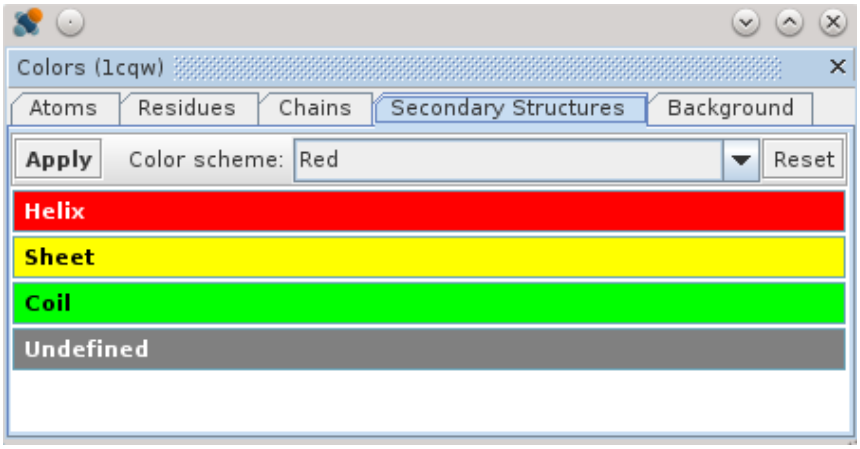
Hide
Show
Select Atoms
Select Residues
Create Binding Site
Increase Volume Precision



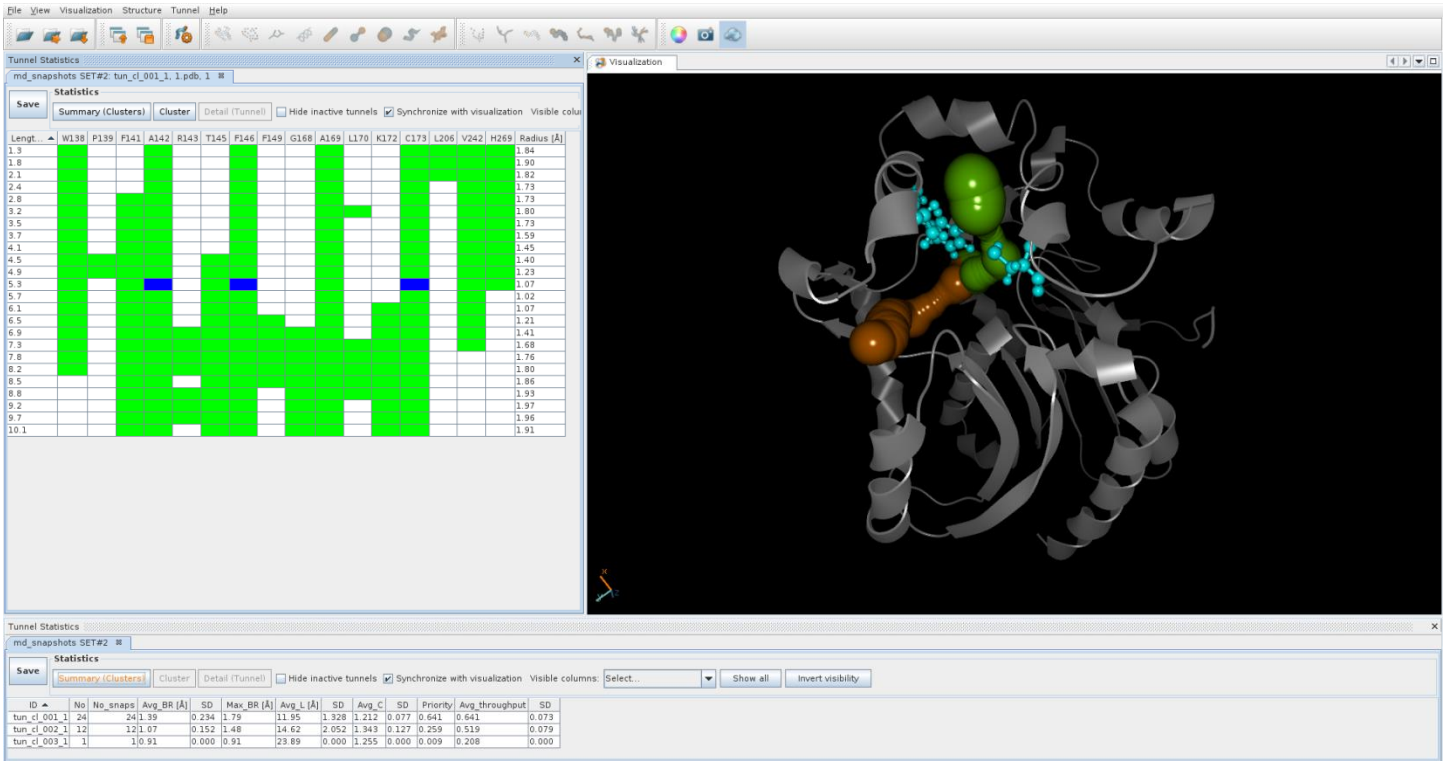
Supplementary figure 2. Cavity calculation. Cavities can be clipped, hidden or selected for estimation of the starting point for the tunnel calculation. The calculated volume of each cavity is provided. Layout of CAVER Analyst (top), detail of characteristics of cavities (bottom left), detail of cavity calculation settings (bottom right).

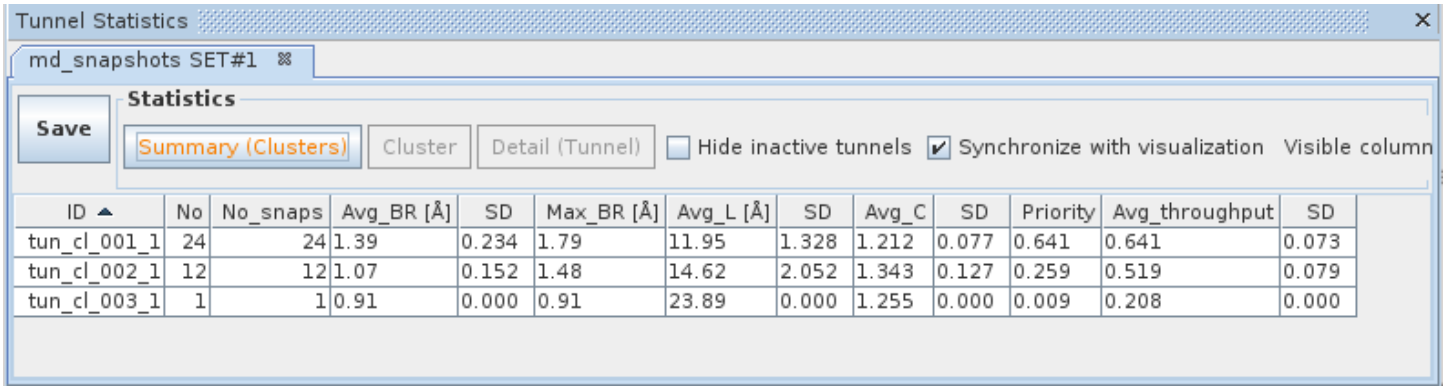






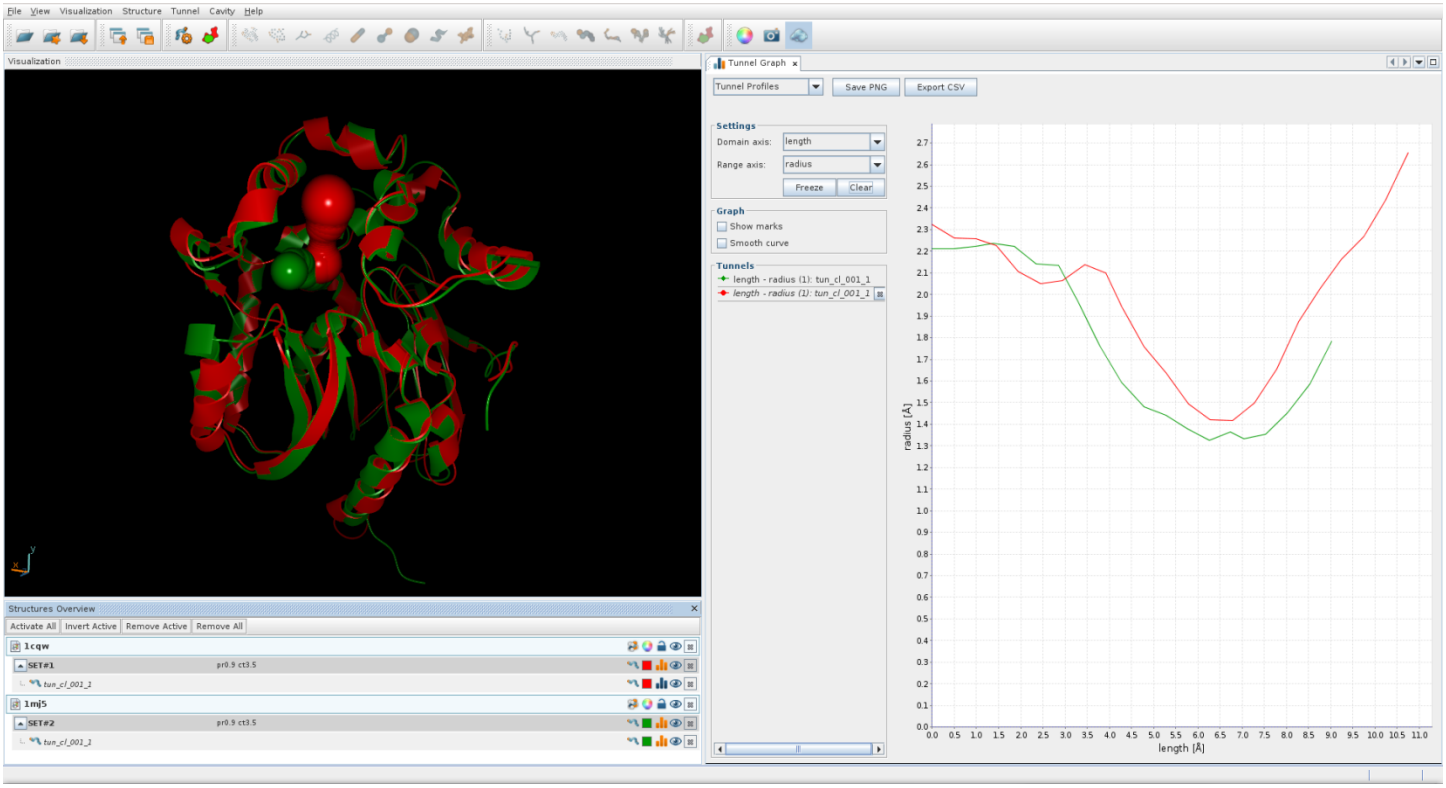
Supplementary figure 3. Visualization styles and coloring schemes. Visualization styles of the protein structure and tunnels can be accessed via: (i) the Visualization button, (ii) the symbolic buttons placed in the top bar and (iii) the visualization button placed in the Structure Overview tab. The coloring scheme is accessible via the top bar and from the Structure Overview tab. All coloring schemes can be adjusted. Layout of CAVER Analyst (top), detail of structure/tunnel visualization techniques (bottom left), detail of coloring schemes (bottom right).

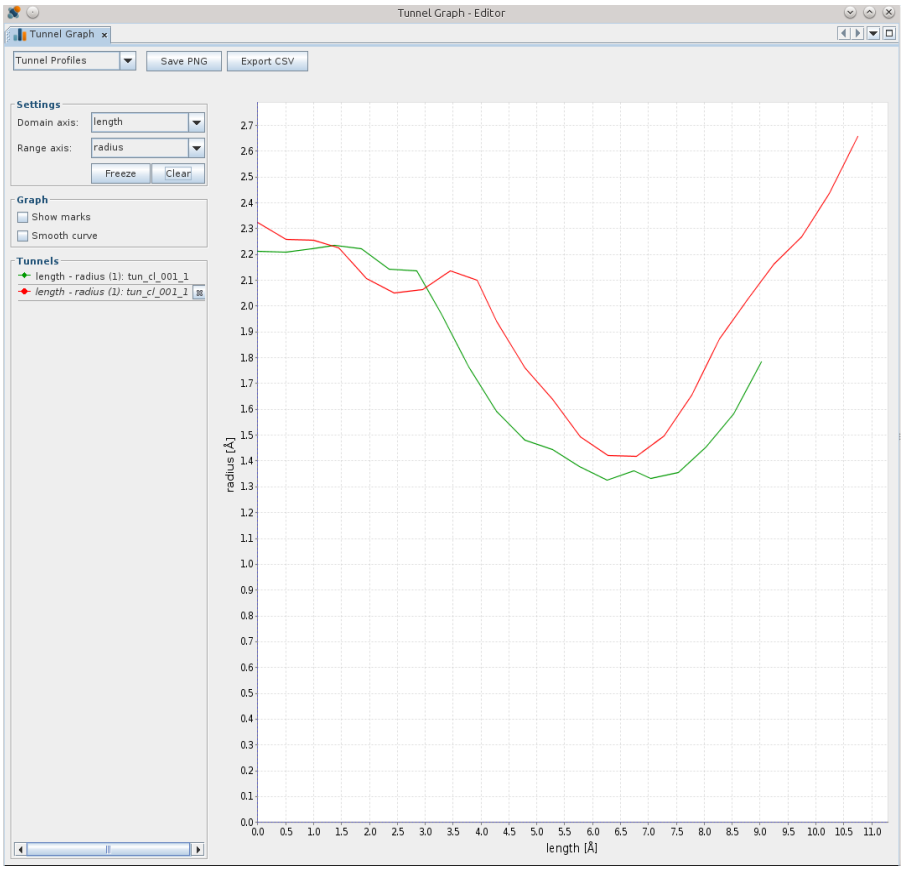




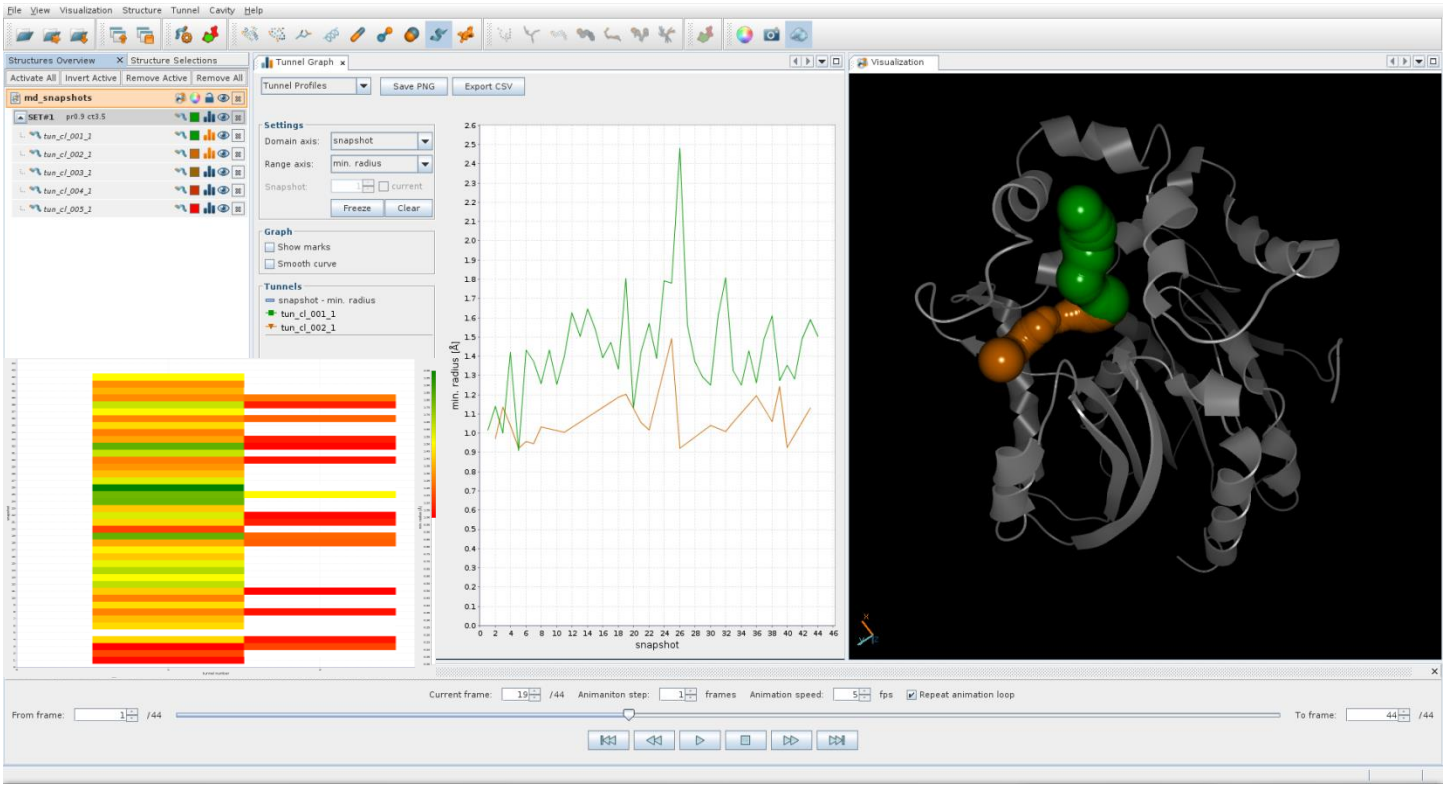
ID	No	No_snaps	Avg_BR [Å]	SD	Max_BR [Å]	Avg_L [Å]	SD	Avg_C	SD	Priority	Avg_throughput	SD
tun_cl_001_1	24	24	1.39	0.234	1.79	11.95	1.328	1.212	0.077	0.641	0.641	0.073
tun_cl_002_1	12	12	1.07	0.152	1.48	14.62	2.052	1.343	0.127	0.259	0.519	0.079
tun_cl_003_1	1	1	0.91	0.000	0.91	23.89	0.000	1.255	0.000	0.009	0.208	0.000

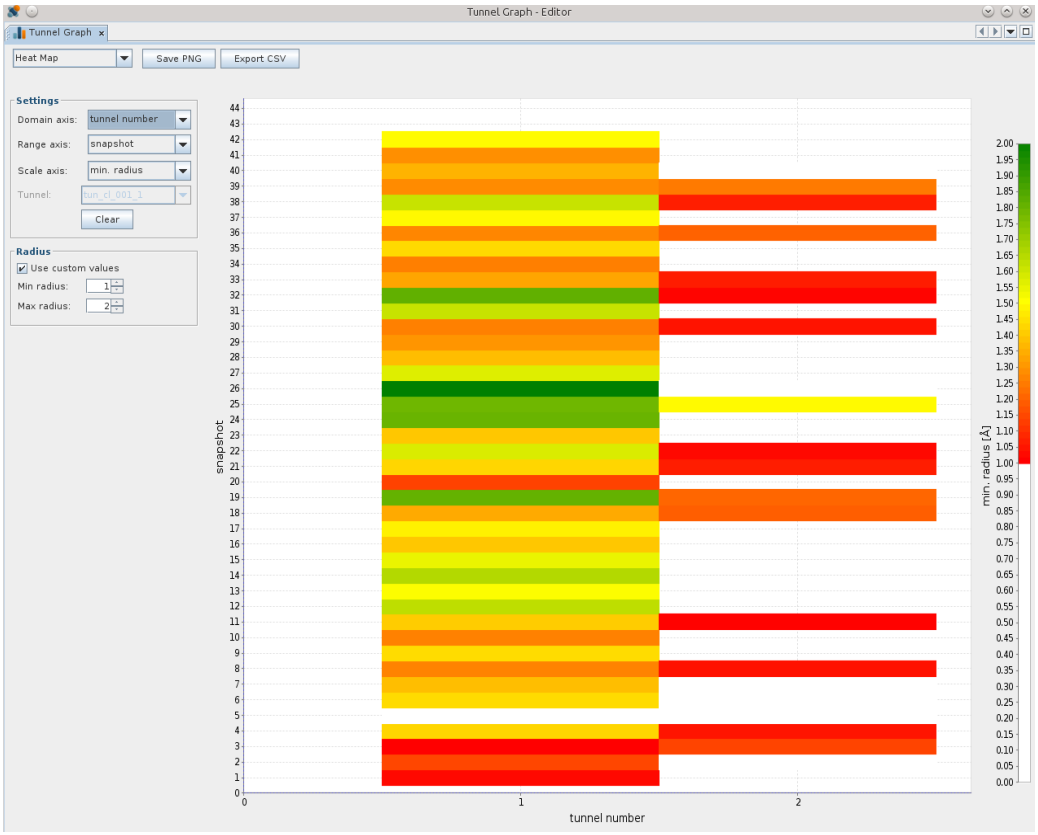
Supplementary figure 4. Statistics of molecular dynamic trajectories. Overview characteristics can be displayed either for all tunnel clusters or for any individual tunnel of a given cluster. Tunnel-lining and bottleneck residues of the main tunnel (green spheres) from a representative snapshot are depicted as green and blue bars, respectively. The tunnels and residues selected in the tables are automatically highlighted in the structure and thus can be simultaneously explored in the tables and visualization window. Layout of CAVER Analyst (top), statistics of tunnel clusters throughout the molecular dynamic trajectory (bottom).





Supplementary figure 5. Comparative analysis of tunnels. Bottleneck radii along the tunnel length (profiles) are plotted as 2D graphs. Tunnel characteristics can be exported either as figures (PNG) or as text (CSV). Layout of CAVER Analyst (top), detail of tunnel graphs (bottom).





Supplementary figure 6. Analysis of molecular dynamic trajectories. Time-dependent evolution of protein structure can be visualized together with dynamic features of tunnels plotted either as 2D graphs or heat plots. All characteristics can be saved as images or exported as a raw data. Layout of CAVER Analyst (top), detail of heat plots (bottom).

References

[1] Pavlova, M., Klvana, M., Prokop, Z., Chaloupkova, R., Banas, P., Otyepka, M., Wade, R.C., Tsuda, M., Nagata, Y., Damborsky, J. (2009) Redesigning Dehalogenase Access Tunnels as a Strategy for Degrading an Anthropogenic Substrate. *Nature Chemical Biology*, 5, 727-733.

[2] Koudelakova, T., Chaloupkova, R., Brezovsky, J., Prokop, Z., Sebestova, E., Hesseler, M., Khabiri, M., Plevaka, M., Kulik, D., Kuta Smatanova, I., Rezacova, P., Ettrich, R., Bornscheuer, U. T., Damborsky, J., 2013: Engineering Enzyme Stability and Resistance to an Organic Cosolvent by Modification of Residues in the Access Tunnel. *Angewandte Chemie International Edition*, 52, 1959-1963.

[3] Damborsky, J., Prokop, Z., Koudelakova, T., Stepankova, V., Chaloupkova, R., Gora, A., Chovancova, E., Brezovsky, J., 2013: Method of Thermostabilization of a Protein and/or Stabilization Towards Organic Solvents. Masaryk University, Brno, Czech Republic. *Patent US 8,580,932*.