Eurographics Workshop on Visual Computing for Biology and Medicine (2017)
S. Bruckner, A. Hennemuth, and B. Kainz (Editors)

Protein Tunnel Reprojection for Physico-Chemical Property Analysis

Jan Malzahn[1], Barbora Kozlíková[2] and Timo Ropinski[1]

[1]Visual Computing Group, Ulm University, Germany
[2]Department of Computer Graphics and Design, Masaryk University, Czech Republic

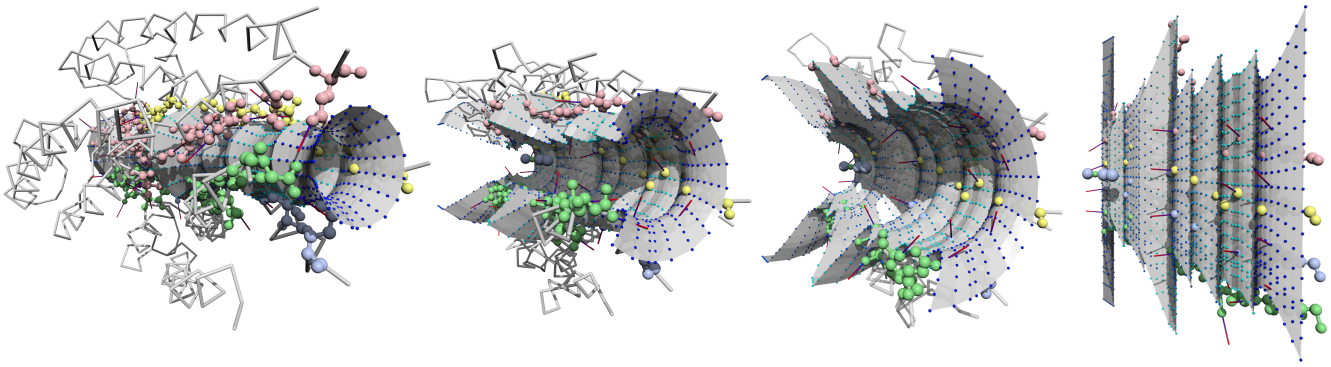


Figure 1: *Demonstration of the proposed tunnel reprojection technique applied to a tunnel of the Potassium channel (PDB ID: 1BL8) protein. The regular outside view on the left is morphed continuously into the fully unfolded view on the right. The morphing animation supports an intuitive association of the reprojected view with the previous locations of the morphed atoms. The coloring emphasizes the four amino acid chains present in this protein with a commonly used chain coloring.*

Abstract

Cavities are crucial for interactions of proteins with other molecules. While a variety of different cavity types exists, tunnels in particular play an important role, as they enable a ligand to deeply enter the active site of a protein where chemical reactions can undergo. Consequently, domain scientists are interested in understanding properties relevant for binding interactions inside molecular tunnels. Unfortunately, when inspecting a 3D representation of the molecule under investigation, tunnels are difficult to analyze due to occlusion issues. Therefore, within this paper we propose a novel reprojection technique that transforms the 3D structure of a molecule to obtain a 2D representation of the tunnel interior. The reprojection has been designed with respect to application-oriented design guidelines, we have identified together with our domain partners. To comply with these guidelines, the transformation preserves individual residues, while the result is capable of showing binding properties inside the tunnel without suffering from occlusions. Thus the reprojected tunnel interior can be used to display physico-chemical properties, e.g., hydrophobicity or amino acid orientation, of residues near a tunnel's surface. As these properties are essential for the interaction between protein and ligand, they can thus hint angles of attack for protein engineers. To demonstrate the benefits of the developed visualization, the obtained results are discussed with respect to domain expert feedback.

Categories and Subject Descriptors (ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Viewing algorithms

1. Introduction

Previous biochemical research has shown that the reactive behavior of proteins depends on a relatively small amount of active sites associated with the macromolecule. These sites can be located on the protein surface or buried deeply in its structure. Many crucial reactions between a protein and a ligand are occurring in buried active sites that are only accessible from the outer environment via tunnels. With respect to a ligand entering and passing through such a tunnel, the physico-chemical properties of amino acids forming the wall of the tunnel have a substantial impact. In other words, when the properties (e.g., atomic charge or hydrophobicity) between the amino acids and the ligand are not compatible, the ligand will never enter the active site via this tunnel or may get stuck along its way.

DOI: 10.2312/vcbm.20171231

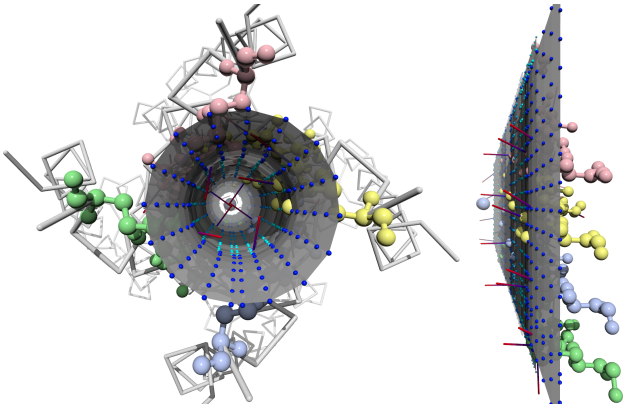


Figure 2: *A look into the tunnel of the Potassium channel (PDB ID: 1BL8) protein. The amino acid side chain orientation markers that point in the direction of the tunnel center can be clearly seen as being parallel to each other in the reprojected view on the right. Using the reprojected visualization of a protein tunnel allows for a quick overview of this important amino acid property.*

Therefore, biochemists are highly interested in studying tunnels and their physico-chemical properties.

While cavities can be sufficiently visualized by regular 3D molecule representations in many cases, this is not the case for tunnels. Due to the tubular shape of tunnels, it is challenging to visualize them in a comprehensible manner, as large parts suffer from occlusions. No matter where a regular camera is positioned in 3D space, some parts of the tunnel will always be invisible. In particular, with standard 3D visualizations it is not possible to depict the entire tunnel wall in a single image, which would facilitate tunnel analysis and comparison.

In other research areas, where tunnel-like structures are under investigation as well, cylindrical projections are often used to enable an occlusion-free view. Many of these approaches have been developed in the context of medical visualization, such as colon flattening [HGQ*06] and curved planar reformations for vessels [KWFG03]. Other domains, as for instance the oil and gas industry [LCMH09], have benefited from similar flattening approaches as well. Unfortunately, when transferring previous techniques and applying them to molecular structures, fundamental problems arise, affecting the intuitive understanding of the obtained reprojections. This is due to the fact that in previous application cases the structures to be reprojected were mostly located on a regular grid. Molecular representations however are different with respect to two criteria. First, atoms forming a complex molecule do not obey a regular structure. Second, the visual representation of a single atom or residue is prone to distortions, avoiding intuitive legibility of the reprojection results. For instance, when applying a ray-casting based flattening, as proposed by Lampe et al. [LCMH09], to a van-der-Waals representation of a molecule, the individual atoms would get distorted into ellipsoids, which results in an unintuitive visualization.

To circumvent these downsides we propose a novel reprojection approach specifically designed for tunnels in molecular structures. Our goal is to preserve detailed molecular representations near the tunnel, while increasing the abstraction level in outer regions. Furthermore, we strive to achieve an occlusion-free overview over all amino acids and their physico-chemical properties near the tunnel to enable protein designers to assess the possibility of a ligand passing through a tunnel. This goal is met by reprojecting the molecular geometry surrounding a tunnel, which is specified by its centerline and hull radius, leading to a visualization that shows the entire tunnel and its immediate surroundings in a single view. Thus we are able to visualize the physico-chemical properties along an entire molecular tunnel. To support intuitive interpretation of the generated visualizations, the presented approach has been developed by keeping domain specific quality criteria in mind. Besides minimizing distortions we also ensure that visibility information and amino acid orientation relative to the tunnel is preserved throughout the reprojection process. Furthermore we have to preserve the proximity of atoms belonging to the same amino acid, as well as amino acids belonging to the same chain, if multiple chains are present in the structure.

The paper is structured as follows. Section 2 gives a brief overview of work related to the proposed technique. Afterwards, Section 3 outlines the tunnel reprojection and introduces the quality criteria informing the presented approach. Section 4 presents the reprojection and morphing algorithms in depth. How we use the reprojection to encode physico-chemical properties on the tunnel hull surface is discussed in Section 5, followed by considerations over visual complexity reduction in the produced visualizations in Section 6. In Section 7 we analyze the obtained results with respect to domain-specific quality criteria and domain expert feedback. The paper concludes in Section 8.

2. Related Work

Analysis and visualization of tunnels in protein molecules has been addressed in many research projects. As our visualization is based on a robust detection of tunnels, we first review literature related to the detection of tunnels, before focusing on tunnel visualization algorithms and the tunnel hull surface.

Tunnel detection. The presence of tunnels in proteins is crucial with respect to their reactivity, therefore many algorithms for their detection were proposed by researchers. Recently Krone et al. [KKL*16] published a survey focusing on analysis and visualization of molecular cavities. It gives a comprehensive overview over all existing types of algorithms for the detection of tunnels, along with different approaches to their visualization as well. Another survey by Brezovský et al. [BCG*13] gives an overview of existing software tools for detection of protein tunnels. A broad overview over the state of the art of biomolecular structure visualization in general is presented by Kozlíková et al. [KKF*16].

The first algorithms for protein tunnel computation were based on grid approaches, such as the first version of CAVER tool published by Petrek et al. [POB*06]. However, these methods suffered from the hardware limitations and inconsistency when using different grid resolution. Therefore, Medek et al. [MBS07] proposed another approach using Voronoi diagrams and Delaunay triangulation. The same approach was also used by Petřek et al. in the MOLE software tool [PKKO07], by Yaffe et al. in their MolAxis tool [YFW*08] and also by Lindow et al. [LBH11]. The Voronoi-based approaches were further extended to be applicable to simulations of molecular dynamics. Among these solutions belongs the CAVER 3.0 tool presented by Chovancová et al. [CPB*12]. The algorithmic details were published by Pavelka et al. [PvK*15].

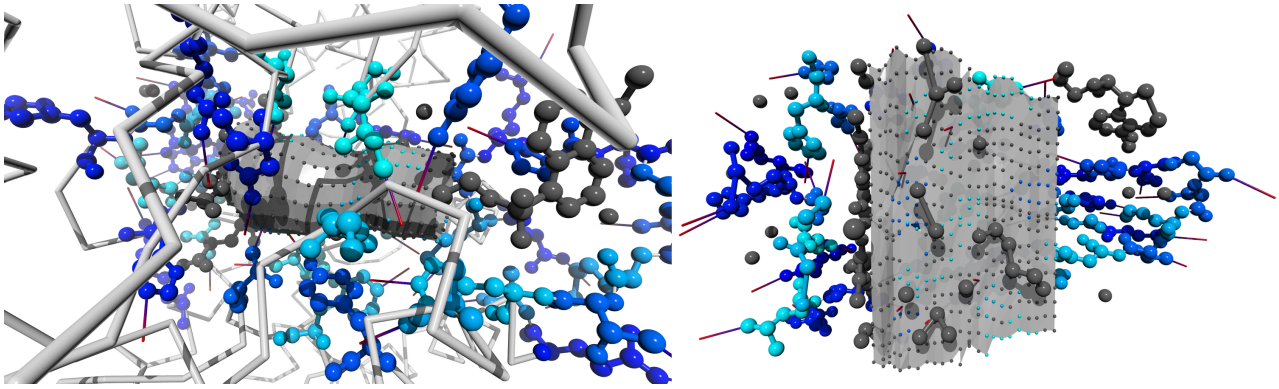


Figure 3: *Example dataset Oxidoreductase (PDB ID: 1RE9). The left view shows the tunnel from an outside point of view, it is deeply buried in the protein. The reprojected tunnel on the right gives an overview over the amino acids near the tunnel in a single image. Additionally, the tunnel hull surface dots visualize the hydrophobicity properties of the nearby amino acids.*

Furthermore, tunnels detected by CAVER 3.0 can be further visualized and explored using the CAVER Analyst tool by Kozlíková et al. [Kvv*14].

Tunnel visualization. Tunnels detected by these methods were initially visualized simply as a set of spheres positioned on the tunnel centerline and touching the surrounding atoms. The first alternative solution aiming to better convey the information about mutual position of similar tunnels was presented by Kozlikova et al. [KAS07]. Yet another representation of the tunnel surface was proposed by Jurcik et al. [JBSK15]. This approach combines the Voronoi-based detection of tunnel centerline and grid-based rasterization of the space surrounding the centerline in order to obtain more precise information about the void space.

Several existing solutions focus on presenting an overview of tunnel behavior in molecular dynamics. Two approaches proposed by Byska et al. [BJG*15, BMG*16] belong in this category. They both present the information about changes of tunnel bottlenecks, their profiles, and tunnel-lining amino acids in a highly abstracted way. In contrast, the focus of our work is to conserve the spatial context. Another approach was presented by Lindow et al. [LBBH12], whereby the authors compute and present the evolution of a cavity over time.

Tunnel hull surface. A very recent publication by Kolesár et al. [KBP*16] presents an approach for comparative visualization of protein tunnel ensembles in a highly abstracted way. The flattened tunnel surface is further equipped with the map of amino acids surrounding the tunnel, along with their extent of influence. While this way the user can see all amino acids and their importance with respect to the tunnel at once, the visualization neglects the spatial context and makes an intuitive understanding of the flattening difficult, as no intermediate representations are provided. This is in contrast to our approach, as the interactive exploration of the tunnel cavity is one of our major goals. We achieve this by introducing an interactively controllable morphing animation that assists the intuitive understanding of the reprojection process. The user can choose to explore a reprojected view that is in between the regular 3D structure visualization and the fully opened up flattened reprojection. This is possible due to seamlessly morphing between both states and enables the user to comprehend how the reprojected view is derived. Another flattening approach, applicable to the overall shape of a molecule, has been proposed by Krone et al. [KFS*17]. As we have developed our reprojection with the goal to visualize physico-chemical parameters, also approaches for visualizing interaction forces between molecules are relevant. Huang et al. [Hua14] review techniques for the analysis of protein-protein interactions, whereby the focus lies on the detection of *interaction hotspots*. They also discuss the vast size differences that are encountered in protein-protein interactions compared to protein-ligand interactions. In contrast to this approach Skanberg et al. have also included the depiction of physico-chemical properties in a 3D visualization [SVGR16]. Specifically, they have exploited diffuse illumination effects to convey interaction forces in real-time. More recently, Hermosilla et al. have presented a visualization approach addressing protein interactions by considering the interaction forces in the visual analysis process [HEG*17]. The presented visualization employs two views, whereby one view conveys the 3D structure of the protein and the ligand, while the other depicts their spatial locations. Thus, they can also preserve spatial structures, but cannot apply their techniques to tunnels, which suffer from occlusion.

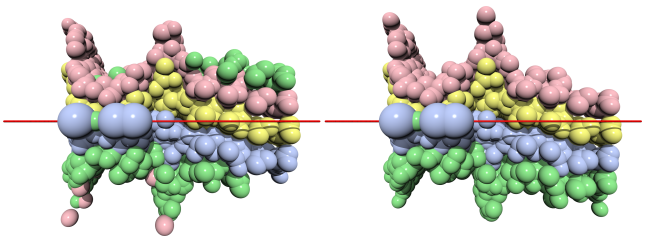


Figure 4: *The Potassium channel (PDB ID: 1BL8) protein contains four amino acid chains. Our technique detects if atoms belong to a macromolecular group (like amino acids or even chains) and respects this affiliation by taking it into account during the cylindrical coordinates computation. The left reprojected view has this feature disabled, the right one has it enabled to depict the difference.*

3. Algorithm Design

Within this section, we motivate our algorithm design. Therefore, we briefly review how tunnels are represented by centerlines within our algorithm, before we discuss quality criteria to which our algorithm sticks in order to generate intuitively interpretable visualizations.

3.1. Tunnel Representation

As stated in the introduction, tunnel visualizations often suffer from occlusions, which make a complete view of the tunnel cavity challenging. No matter whether the virtual camera is placed in or outside the molecule, parts inside of the tunnel are always missing in the resulting image. While a complete overview over one half of the tunnel interior can be obtained by cutting away all residues that are occluding the view to the tunnel's interior, clipping does not allow for visualizing the entire tunnel. While we could place two opposite halves next to each other, such a visualization would suffer from discontinuities at their borders.

Accordingly, alternative visualization techniques are required in order to intuitively visualize an entire tunnel in a single image. Kolesár et al. [KBP*16] have already proposed such a visualization, whereby they employ an abstracted view to show information associated with the tunnel wall. The 2D flattened representation of a tunnel enables to encode the areas of the tunnel surface influenced by surrounding amino acids in a cartogram-like manner. However, the depiction of the individual amino acids, their atoms, and orientation towards the tunnel is completely missing. To get this information users have to go back to the 3D view and explore the tunnel there. In our newly proposed approach we aim to combine the advantages of both 2D and 3D views by animating the tunnel opening to a flattened 2D state while preserving the visualization of the surrounding amino acids. These amino acids can be shown at atomic resolution, using different representations and coloring schemes. Also, mental linking between the flattened tunnel and its original, possibly complex geometry is made possible using this approach. It has been designed to be structure-aware, i.e., we support a continuous reprojection which enables domain scientists to relate features in the reprojection to features in the original molecule structure. To achieve this, we exploit the centerline of the tunnel as a characteristic curve parametrized over $t \in [0,1]$. Since centerlines are sufficient descriptors for protein tunnels, several software systems support the extraction and export of these lines. For the visualizations shown in this paper, we have used the CAVER software package to extract and smooth the tunnels' centerlines. Based on the extracted centerline, we generate a differentiable b-spline curve, of which the first derivative gives us the tangential vector, thus yielding a robust tunnel descriptor that we can employ to obtain the desired reprojection.

3.2. Quality Criteria

To circumvent the aforementioned downsides, and to ensure an algorithm design that is controlled by the needs of domain experts, we have derived the following domain-specific quality criteria which shall improve the readability of the reprojected tunnels.

Structure distortion reduction. Using ray-casting approaches, such as [LCMH09], based on a tunnel's centerline results in perspective distortions on a per pixel basis, due to the centerline's curvature. As the ray origin and direction is calculated for each pixel in the image plane individually, this can result in atom level distortions that change the appearance of a single atom. However, in order to support an intuitive understanding, individual atoms and other structures should not be distorted, and ideally be represented as spheres. Our strategy to fulfill this goal is to exploit an reprojection algorithm that transforms the atom positions directly, in contrast to image-based ray-casting approaches.

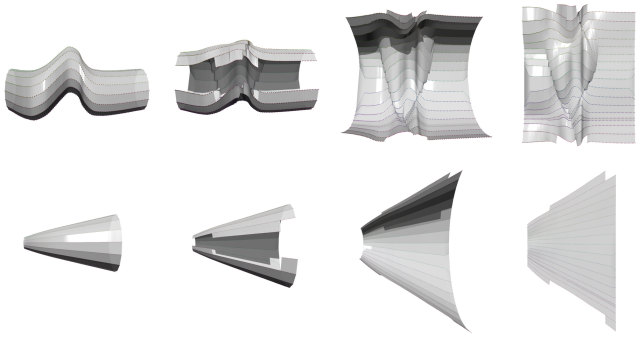


Figure 5: *We used multiple synthetic tunnel datasets to aid in the development of the reprojection technique, two of them are shown here. The top row dataset has very curvy centerline with a constant hull radius to visualize the inevitable hull distortions after the reprojection. The bottom row tunnel uses a straight centerline with an increasing hull radius, demonstrating the correct stretching of the reprojected hull surface.*

Residue preservation. Amino acids form the basic building blocks of protein molecules. Accordingly, amino acids should be inseparable, and even if the molecular structure has to be changed on a global scale, its amino acids shall be preserved. We address this within our algorithm by applying amino acid aware cuts of a tunnel's wall prior to its reprojection.

Comprehensibility. When a visualization has to transform original data into a novel representation, the question arises if the transformed representation is comprehensible to the target audience. Often this is achieved by providing hints on how the new representation relates to the original data. Animations that morph between the original data (and its classical presentation) and the novel visualization can be helpful to improve the comprehensibility. Accordingly, morphed transitions between the original molecule and our reprojection are at the core of the presented algorithm.

In the following section, we will propose our tunnel reprojection algorithm, which has been designed by having these quality criteria in mind.

4. Tunnel Reprojection

To visualize protein tunnels, we present an approach that exploits cylindrical projections to visualize a protein's tunnel interior. The technique is inspired by the following metaphor. Let a virtual camera be located on the centerline of the tunnel to be reprojected. If this camera would rotate one full circle around the centerline for each infinitesimally small step along the centerline, the resulting

image plane would correspond to a virtual cylinder that is bent around the centerline to follow its curvature. While this metaphor helps to understand the overall reprojection goal, a straight forward realization would not take into account the quality criteria discussed above. Instead, we need to compute appropriate transformations for the geometric structures making up the molecule, and transform the structures accordingly. In the following subsection we will first derive a reprojection transformation, which fulfills our quality criteria with respect to structure distortion reduction as well as residue preservation. Afterwards, in the next subsection, we will describe our reformation morphing, which helps us to ensure comprehensibility of the generated visualizations by means of animation.

4.1. Reprojection Transformation

To perform our reprojection transformation, we need to derive some quantities from a molecule's tunnel. As stated above, a tunnel is represented by its centerline. Along this centerline a consecutive set of spheres together with the accompanying sphere radii describes the tunnel hull. As we need a differentiable and continuous centerline curve for our calculations, we smoothly interpolate the centerline sphere centers using b-splines. The same applies to the hull radius for each point on the centerline. This gives us the position on the centerline $\vec{c_p}$ and the hull radius c_r at this position. Additionally, we need the main direction $\vec{m_p}$ of the tunnel. One option for this would be to use the cylindrical coordinate systems local to the centerline position's tangent. However, this approach leads to rifts at focal points of the bent centerline, which is undesirable. Using also the main direction of the tunnel in the computations, as illustrated in Figure 6, helps to solve this problem. Thus we refer to our position on the centerline $\vec{c_p}$, the hull radius c_r and the main direction $\vec{m_p}$ as follows:

$$\vec{c_p}(t_c) \in \mathbb{R}^3,\ t \in [0,1] \tag{1a}$$
$$\vec{m_p}(t_m) \in \mathbb{R}^3,\ t \in [0,1] \tag{1b}$$
$$c_r(t_c) \in \mathbb{R},\ t \in [0,1] \tag{1c}$$

Thus, for any t in the unit interval we get the actual three dimensional position on the centerline, $\vec{c}(0) = \vec{m}(0)$ being the starting point and $\vec{c}(1) = \vec{m}(1)$ being the ending point of the tunnel centerline curve and the tunnel main direction. Based on these values, the proposed reprojection transformation uses cylindrical coordinates, which are a combination of two-dimensional polar coordinates lying in a plane perpendicular to the longitudinal axis, which is the third coordinate.

Cylindrical coordinate system. The underlying coordinate system consists of a plane that is defined by two basis vectors which originate from the same origin: The longitudinal axis vector $\vec{L}$ and a polar axis vector $\vec{P}$. This plane is rotated around the longitudinal axis, referenced against the polar axis, by a rotation angle φ. A cylindrical coordinate then is given by

- The radial distance ρ (travel along the rotated polar axis $\vec{P}$).
- The rotation φ (rotation angle along $\vec{L}$ with respect to $\vec{P}$).
- The height z (amount of travel along the longitudinal axis $\vec{L}$).

Using this definition, the longitudinal axis of the cylinder equals the X axis in Cartesian coordinates, and the angle φ, referenced against an up-vector $\begin{pmatrix}0\\1\\0\end{pmatrix}$, is the rotation around the X axis. Finally, the radial distance ρ is the perpendicular distance to the X axis. This way, the cylindrical coordinate system is defined for intuitive usage with tunnel centerlines, which are predominantly aligned along the X axis in space. If a tunnel dataset isn't aligned along the X axis predominantly, we rotate the complete molecular dataset to achieve this alignment before any further computations.

Mapping world positions to cylindrical coordinates. Based on the underlying coordinate system, we can now transform molecular structures, such as atoms, from their world position into the reprojected position. Therefore, we derive for each entity its cylindrical coordinates with respect to the tunnel centerline and the tunnel's main direction. Hereby, both t and the rotational angle φ are based on the tunnel main direction, only the radial distance ρ and the tunnel hull radius $c_r(t_c)$ are related to the actual centerline curve. As the centerline can be any given curve, we select the nearest point on the centerline as the point of reference for the following calculations. The same is true for the main direction, an illustration of this can be seen in Figure 6.

The appropriate position on the centerline and the main direction, i.e., the nearest point, as well as the corresponding $t_c, t_m \in [0,1]$ are computed for any given atom position. Using the derived cylindrical coordinates for each atom position, we are able to perform a linear mapping to the reprojected world position $\vec{p}$ in Cartesian coordinates as follows:

$$\vec{p}_x = (t_m - 0.5) \cdot c_{length} \tag{2a}$$
$$\vec{p}_y = \varphi_{\vec{m}_p(t_m)} \cdot c_r(t_c) \tag{2b}$$
$$\vec{p}_z = \rho_{\vec{c}_p(t_c)} \tag{2c}$$

whereby c_{length} is the maximum centerline arc length. When transforming all molecular structures to these new coordinates, the scene can be displayed using a regular virtual camera. Our technique also detects if atoms belong to a macromolecular group (like amino acids or even chains) and modifies the angular part of the cylindrical coordinates computation. The effect of this can be seen in Figure 4. Atoms in a group get assigned angles that must have the same sign, this way they get reprojected to the top or bottom side in the reprojected view, but not on both sides at the same time. This makes them stay near to each other in the reprojection as well.

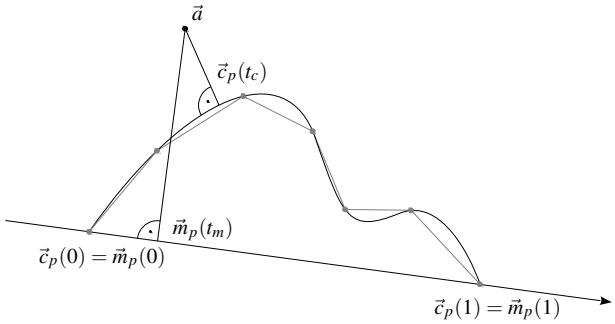


Figure 6: *Illustration of the tunnel's centerline* (curve) *and the tunnel's main direction* (line). *While the radial distance is taken for the nearest centerline point, the position t and rotation angle* ρ *are based on the tunnel's main direction.*

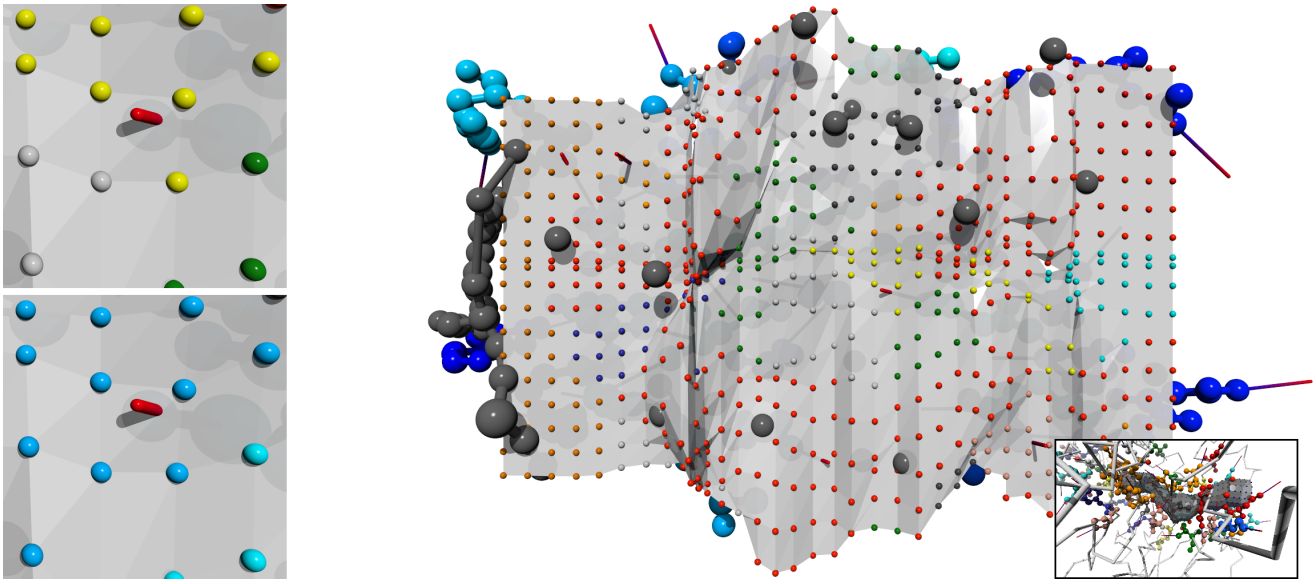


Figure 7: *This is a closeup look on the reprojected tunnel hull of the Oxidoreductase (PDB ID: 1BU7) protein dataset. The overlay on the right shows an outside overview of the molecule. The detail cutouts on the left show two possible color codings of the surface dots we use to visualize amino acid properties near the tunnel hull. The upper cutout indicates nearby amino acids by type, the lower cutout by hydrophobicity. The used shadowing also enhances the perception of the amino acid side chain orientation markers that penetrate the reprojected tunnel hull surface.*

4.2. Reprojection Morphing

Understanding the reprojected visualization can be difficult, as it is hard to visually link atoms and other structures to their previous world positions. Displaying the tunnel centerline and/or tunnel hull helps in that matter. However, we further improve upon this issue by introducing a morphing animation, which smoothly transforms the original world positions into the final reprojected locations. During this morphing the original positions gradually move towards their final positions in the cylindrical view space. Thus, mentally linking them to their original position becomes easier.

Linear interpolation of these positions would be the most straightforward approach for such a morphing. However, due to the nature of the reprojection, this would result in movement paths where atoms would intersect and cross each other, rendering the morphing animation useless. Therefore, a more refined approach is required, which shall resemble a more natural animation. A possible solution would be to linearly interpolate positions in cylindrical coordinates. This yields good results right at the beginning of the animation, but the movement paths also cross later on. To avoid these crossings, and to achieve a more natural looking transition that resembles the unrolling of a sheet of paper that was rolled into a cylindrical form, we exploit a combined interpolation scheme. Based on the standard linear interpolation function given as:

$$\alpha(\vec{x},\vec{y},f) = (1-f)\cdot\vec{x} + f\cdot\vec{y} \quad \text{with } \vec{x}\in\mathbb{R}^3,\ f\in[0,1], \tag{3}$$

we compute a world space position $\vec{w}(f)$ as:

$$\vec{w}(f) = \alpha(\vec{a},\vec{p},f), \tag{4}$$

whereby $\vec{a}$ is the original position in world coordinates and $\vec{p}$ is the reprojected position derived from cylindrical coordinates. The interpolation factor $f\in[0,1]$ corresponds to the position in the blending process.

While this interpolation treats all positions equally, independent of their distance from the centerline, a natural unrolling requires a weighting with respect to this distance. Therefore, we damp positions that are further away from the centerline, to unroll the outer regions later in the animation. We achieve this by treating the z component of the positional interpolation differently:

$$\vec{w}_z(f) = \alpha(\vec{a}_z,\vec{p}_z,f^2) \tag{5}$$

Here, the reprojected component $\vec{p}_z$ equals the distance to the centerline, and it is blended with the parabolic shaped function f^2. To avoid atoms crossing each other, we need to also bring in an interpolation in cylindrical coordinates. We achieve this by using the following equation:

$$\vec{r}(f) = \text{rotate}(-(1-f)\cdot\varphi_{\vec{m}_p(t_m)}, \begin{pmatrix}1\\0\\0\end{pmatrix})\cdot\vec{p}, \tag{6}$$

whereby rotate$(angle, axis)$ yields an appropriate rotation matrix. Thus, the reprojected position $\vec{p}$ is rotated around $\begin{pmatrix}1\\0\\0\end{pmatrix}$, being the longitudinal axis of the cylindrical coordinate system, by the amount of $\varphi_{\vec{m}_p(t_m)}$ which equals the rotation angle around the centerline. Both parts are then blended together by the following interpolation, yielding the final blended position $\vec{b}(f)$:

$$\vec{b}(f) = \alpha(\vec{w}(f),\vec{r}(f),f) \tag{7}$$

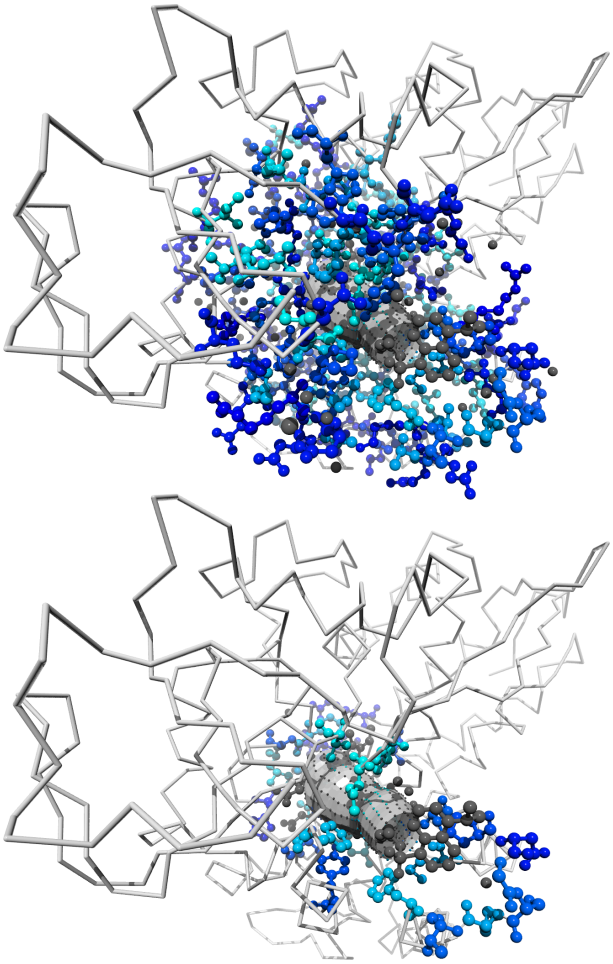


Figure 8: *The Oxidoreductase (PDB ID: 1RE9) protein depicted in this figure is shown with two different abstraction levels: the top one has the near/far threshold for nearby amino acids set very high, while the bottom one shows amino acids in detail only with a low distance threshold. This threshold can be set interactively to the user's preference.*

Thus, we effectively combine an interpolation in cylindrical coordinates with an interpolation in Cartesian coordinates to achieve the desired effect. The resulting movement paths are not intersecting and form a visually appealing unwrapping animation, which can be controlled by the user through modifying f. Figure 1 as well as the accompanying video show the effects of this morphing animation. More results obtained with all technique steps applied are presented in Figure 3 and Figure 7, Figure 5 shows a series of synthetic tunnel hull datasets used during development, while Figure 9 suggests an improved depth perception when using shadowing.

5. Encoding Physico-Chemical Properties

Now that we are able to reproject a molecules tunnel, we can visualize parameters previously not accessible to the domain experts. Domain experts in the field of biochemistry and protein engineering are interested in specific properties of the amino acids connecting to the tunnel hull surface. One of such properties is the orientation of the amino acid side chain relative to its backbone, i.e., the alpha-carbon in the group. The chemically reactive part of amino acids, while being part of a larger peptide bonded amino acid chain, is always located at the outermost end of the side chain. If this part is pointing away from the tunnel cavity, a physico-chemical interaction with a ligand traveling through the tunnel is impossible or at least very unlikely in this configuration. The opposite is true if the side chain points towards the tunnel, this denotes an amino acid that is potentially interesting for further investigation. Thus, conveying this information in the tunnel representation is of great importance. Encoding orientation or angular difference of structures in three-dimensional space as color gradients results in unintuitive and unfamiliar visualizations which do not convey the information sufficiently. Based on this observation and backed by similar challenges in other domains, as for example in vector field visualization, we decided to use orientation markers that point in the side chain's direction relative to its alpha-carbon. This orientation marker can effectively visualize the amino acid orientation if provided with an accompanying tunnel hull surface that serves as a reference plane to the orientation marker. This can be seen in Figure 2, where amino acid side chains that are pointing away from the tunnel can be quickly discarded visually at a glance.

Besides the markers encoding the side chain orientation, the unfolded tunnel wall also enables us to visually encode relevant properties. We have decided to use this opportunity to visualize further physico-chemical properties that are relevant to domain experts when investigating a protein tunnel. These properties include (but are not limited to) hydrophobicity and electronegativity, which can be mapped to the tunnel's surface with adequate color mappings. However, their effect diminishes with distance, it is therefore relevant to visualize their strength relative to their location with respect to the tunnel hull. Initially we mapped these values directly to the reprojected tunnel hull surface, but this overly cluttered the visualization, especially as we intended to display the orientation markers at the same time. To avoid this downside we decided to only depict these properties at discrete locations on the tunnel's surface. We employ small surface mounted dots that carry properties like hydrophobicity of amino acids in proximity in color coded form, while the tunnel hull surface itself is colored neutrally and translucently to aid the orientation markers. The dots are placed on a regular grid that encloses the tunnel hull, which leads to a textured-like appearance of the tunnel hull surface. Research by Interrante et al. [IFP97] suggests that textured surfaces are able to convey their 3D shape better than just smooth semitransparent surfaces, which is what we found as well.

6. Visual Complexity Reduction

One important aspect of our visualization is the varying significance of the various molecular structures in a protein molecule. Based on previous feedback from domain scientists we learned that details down to the atomic level are only of interest near the tunnel, while the outer regions of the protein could potentially be completely left out in such a visualization. However, as we intend to reproject the molecular structure to reveal the tunnel's interior wall in a single view, too much context information would be lost in the process. The structural information of the outer regions is needed to locate the tunnel and its surroundings on a global scale, and as a contextual reference for the morphing animation. However, especially the morphing process is highly unintuitive without surround-

ing context. Therefore, we decided to include the outer regions in the visualization as well, albeit on a higher abstraction level. Van der Zwan et al. [vdZLBI11] describe this approach as continuous abstraction. To achieve this, we represent the amino acid chain only by its backbone, while seamlessly blending into a detailed ball-and-stick atom representation near the tunnel at the same time, whereby the threshold for this transition is user selectable. This can be seen in Figure 8.

In the process of the morphing animation we fade out the outer backbone representation, as it gets squished together by the reprojection and loses its value for the visualization in the completely reprojected state. The backbone fade out, depth cueing and shadowing enhancements are all shown in the morphing sequence in Figure 1.

7. Evaluation

To evaluate the obtained tunnel reprojections, we have analyzed them in two different ways. As the visual quality requires special consideration, we compared the discussed technique with two alternative approaches with respect to the introduced domain-specific quality criteria (see Section 7.1). Furthermore, our proposed technique was evaluated by domain experts (see Section 7.2).

7.1. Visual Quality Comparison

To compare the visual quality of the presented approach, we have implemented two alternative reprojection strategies, which were direct applications of the unfolding techniques proposed in other contexts. We have applied these techniques to vdW representations, as these make the differences most prominent. Before we compare the results of these techniques with ours, we briefly discuss these approaches.

Ray-casting. The regular 3D hardware rendering pipeline allows only for a certain subset of achievable projections (such as orthographic or perspective projection), in particular those expressible by a 4x4 transformation matrix. To realize a cylindrical projection, a rendering system that allows for more arbitrary projection is needed. Several authors present how to exploit ray-casting to achieve cylindrical projections, e.g., Lampe et al. [LCMH09]. To transfer this concept to molecular data, the ray-casting fragment shader is additionally supplied with the tunnel's centerline control points which are smoothed using B-spline interpolation yielding a smooth and continuous curve. Inside the fragment shader, the ray

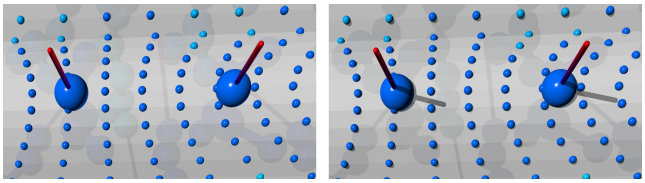
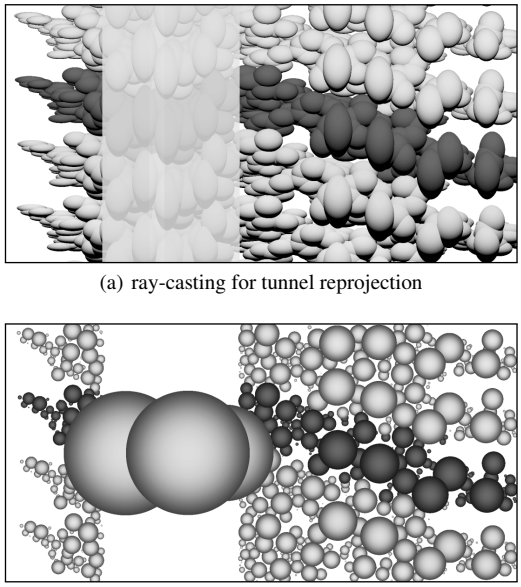


Figure 9: *This a detail view of the reprojected Potassium channel (PDB ID: 1BL8) tunnel hull surface shown in Figure 1 with hydrophobicity color coding. The left detail has shadowing disabled, in the right one it is enabled. Shadowing enhances the depth perception considerably, and is especially useful for the amino acid side chain orientation markers in our case.*



(a) ray-casting for tunnel reprojection

(b) postprocessed ray-casting for tunnel reprojection

Figure 10: *Two alternative approaches have been implemented to facilitate a visual comparison of our presented flattening maps. A standard ray-casting approach as used in other flattening scenarios (*a*), and a post processed ray-casting approach accounting for atom distortions (*b*).*

generation parameters are selected based on the coordinates of the currently processed fragment. The position on the tunnel centerline corresponds to the x axis of the resulting image plane. The leftmost pixel column relates to the beginning of the tunnel, the rightmost pixel column to the closing end of the tunnel. The angle relative to the up-vector, rotating around the tangential vector at this point corresponds to the y axis of the resulting image plane. Thus, all pixels in one row yield the same rotation angle used in the ray generation. With this approach reprojections can be produced, such as the one depicted in Figure 10(a).

Postprocessed ray-casting. As can be seen in Figure 10(a), direct application of ray-casting for tunnel reprojection results in strong atom distortions. To account for this effect, it is possible to perform a postprocessing which modifies the resulting image. The purpose of this postprocessing is to represent the area occupied by one atom in the ray-casted image in a more meaningful way, whereby atoms resemble circles, shaded as spheres with a matching area size. A circle's center represents the weighted average of all pixels belonging to this atom. This can be achieved by rendering a map with a unique color for each individual sphere with no shading applied, such that the color directly corresponds to the index of each individual atom in the molecular structure. During the postprocessing, all pixels corresponding to one unique atom can be summed up, giving the total area occupied in the image. Then, all pixel positions of the respective atom can be averaged, yielding the center of the distorted representation of the atom in the ray-casted image. Using this information, a new image can be generated, where for each atom a circle with matching center and area size is drawn. Figure 10(b) shows this approach applied to the same data set that

was used to generate Figure 10(a). As it can be seen, atom distortion is eliminated, while the molecule distortion still remains.

Technique comparison. To compare the visual quality of the investigated techniques, we have classified them based on the quality criteria postulated in Section 3.2, as well as other desired properties. Table 1 summarizes this classification. As it can be seen in the table, among the tested techniques, ours is the only one scoring on all quality criteria. Especially, it is the only one, which reduces molecule distortion. Furthermore, the animation included in our approach supports comprehension, as it allows for a mental linking of atoms to their original position. This is hard or even impossible to achieve by just taking into account the alternative reprojections shown in Figure 10. Due to proposed morphing which avoids intersections, we can further score with respect to visibility and residue preservation.

7.2. Domain Experts Feedback

To evaluate the usefulness of our proposed technique, we interviewed four protein engineering experts from the Loschmidt Laboratories at the Masaryk University and an expert in physical chemistry from Palacky University Olomouc. All experts have been confronted with static and dynamic reprojection results as generated with the presented approach. After looking at the results, they have been asked the following questions:

- Which type of spatial information, besides structure, regarding the tunnels is relevant for you?
- Would you like to use 2D tunnel projections to show this information?
- Which of the shown visualizations allows you best to relate to the 3D structure of the molecule under investigation? Why?
- Could you imagine to use a 2D tunnel projection as an intermediate view?
- Is the animation in the morphing helpful?
- Is the interactive control of the unfolding important to you?
- Does the unfolded view of a tunnel give you a useful overview over the tunnel's interior structure?

In the following paragraphs we summarize the obtained feedback, collected through these interviews. All experts agreed that the idea of opening the molecule which is controlled by the tunnel is very promising. They admitted that the traditionally used 3D representation serves for better demonstration of the real situation but for data analysis and testing the 2D representation of the reprojections is much better. They also stated, that the reprojections can be very useful for comparative purposes. An example which was brought up was, that when a biochemist wants to compare five proteins with different tunnels, this would be extremely difficult in 3D. A similar situation appears when exploring large simulations of molecular dynamics containing thousands of time steps. Tracking a tunnel in this situation and searching for its extreme anatomies is almost impossible using traditional 3D representation.

When evaluating our technique, all experts agreed that the animation of the opening is crucial for understanding how the reprojection was created. This is also inline with our comprehensibility quality criteria. They also admitted that without this animation it is very hard to understand the flattened structure, and that the interaction with the system is crucial as well. In their daily work, when trying to understand molecular structures, the domain experts have to use different representation with different levels of abstractions. Thus, they also commonly use the cartoon representation, where all atoms and even the amino acids are replaced by helices, arrows, and coils. Thus, they appreciated that the presented technique can be applied to any of these visualization techniques without modification.

The biochemists also see high potential in using varying coloring methods on the tunnel's surface. They liked the possibility to color the tunnel-lining amino acids by different physico-chemical properties, such as hydrophobicity, charge, or hydration. This helps them to see if there are some points which may create a barrier for the ligand to pass (due to the incompatibility of these properties with ligand), and thus such information is crucial in the next step when they want to modify tunnels (i.e., mutate the tunnel-lining amino acids) to change the protein properties or function, such as influence the protein activity, stability, or resistance to organic cosolvents. The experts from Loschmidt Laboratories would further appreciate to combine our technique with the animation of the ligand or flow of water molecules through the tunnel. Visualization of the water stream sliding along the morphed structure could be very interesting, e.g., to see potential barriers.

In summary, all addressed domain experts confirmed that the presented approach is a very promising and yet unexplored direction of tunnel visualization and exploration, which potentially could be highly beneficial for protein engineers and drug designers.

	Technique		
	ours	ray-cast	postproc.
Structure distortion reduction	high	none	high
Residue preservation	high	none	none
Comprehensibility	good	bad	weak
Molecule distortion reduction	high	none	none
Visibility preservation	good	medium	none

Table 1: *We have compared the two implemented alternatives to our proposed approach. The rows show how the individual techniques behave with respect to selected quality criteria.*

8. Conclusions and Future Work

Within this paper we have introduced a novel algorithm for protein tunnel reprojection. The algorithm combines a reprojection transformation together with an interactive morphing animation to reveal structures occluded inside protein tunnels. The proposed algorithm has been developed with respect to domain-specific quality criteria, which ensures that the resulting visualizations are effective. Thus, we were able to introduce a novel protein tunnel reprojection approach, which generates convincing results at interactive frame rates. Once the tunnel's interior is revealed, we are able to visually encode physico-chemical properties which are of relevance to domain experts. Besides taking into account the postulated quality criteria, we have also analyzed the achieved results with respect to expert feedback. The interviewed domain experts confirmed that our reprojections are a very promising and yet unexplored direction of tunnel visualization and exploration, and that they may be highly beneficial for protein engineers and drug designers. We believe this positive feedback results from the fact, that the presented

visualizations on the one hand, do not suffer from misleading distortions, and that the animation does give a better intuition of the transformation process.

In the future we would like to investigate further possible improvements based on the proposed technique. When applying the morphing, for each atom there are various possible paths to follow. To oppose the rift that opens up around sharp bends of the centerline, the nearest point on centerline metric could be enhanced, and all atom positions could be anchored in a neighborhood mesh system that constraints the possible movement of atom positions. While all relative distances to each other would have to stay the same, angular changes between atom links in the mesh would be allowed. This approach would resemble a woven fabric that is unrolled, where all knots stay at the same distance relative to each other. This would alleviate discontinuities and improve the uniformity of the property mapping on the tunnel surface. Based on this other visualizations of mapped properties (fully coloring the tunnel surface instead of or in combination with our surface samples, with and without translucency) could be explored and compared to each other. Furthermore, we plan to investigate how to visually encode additional physico-chemical properties in the context of the unfolded tunnel.

References

[BCG*13] BREZOVSKÝ J., CHOVANCOVÁ E., GORA A., PAVELKA A., BIEDERMANNOVÁ L., DAMBORSKÝ J.: Software tools for identification, visualization and analysis of protein tunnels and channels. *Biotechnol. Adv. 31*, 1 (2013), 38–49. 2

[BJG*15] BYŠKA J., JURČÍK A., GRÖLLER M. E., VIOLA I., KOZLÍKOVÁ B.: MoleCollar and Tunnel Heat Map Visualizations for Conveying Spatio-Temporo-Chemical Properties Across and Along Protein Voids. *Comput. Graph. Forum 34*, 3 (2015), 1–10. 3

[BMG*16] BYŠKA J., MUZIC M. L., GRÖLLER M. E., VIOLA I., KOZLÍKOVÁ B.: AnimoAminoMiner: Exploration of Protein Tunnels and their Properties in Molecular Dynamics. *IEEE Trans. Vis. Comput. Graphics 22*, 1 (2016), 747–756. 3

[CPB*12] CHOVANCOVÁ E., PAVELKA A., BENEŠ P., STRNAD O., BREZOVSKÝ J., KOZLÍKOVÁ B., GORA A., ŠUSTR V., KLVAŇA M., MEDEK P., BIEDERMANNOVÁ L., SOCHOR J., DAMBORSKÝ J.: CAVER 3.0: A tool for the analysis of transport pathways in dynamic protein structures. *Pathways in Dynamic Protein Structures, PLoS Computational Biology 8: e1002708* (2012). 2

[HEG*17] HERMOSILLA P., ESTRADA J., GUALLAR V., ROPINSKI T., VINACUA Á., VÁZQUEZ P. P.: Physics-based visual characterization of molecular interaction forces. *IEEE Trans. Vis. Comput. Graphics 23*, 1 (2017), 731–740. 3

[HGQ*06] HONG W., GU X., QIU F., JIN M., KAUFMAN A.: Conformal virtual colon flattening. In *Proceedings of the 2006 ACM symposium on Solid and physical modeling* (2006), ACM, pp. 85–93. 2

[Hua14] HUANG S.-Y.: Search strategies and evaluation in protein-protein docking: principles, advances and challenges. *Drug Discov. Today 19*, 8 (2014), 1081–1096. 3

[IFP97] INTERRANTE V., FUCHS H., PIZER S. M.: Conveying the 3D shape of smoothly curving transparent surfaces via texture. *IEEE Transactions on Visualization and Computer Graphics 3*, 2 (Apr 1997), 98–117. 7

[JBSK15] JURČÍK A., BYŠKA J., SOCHOR J., KOZLÍKOVÁ B.: Visibility–based approach to surface detection of tunnels in proteins. In *31th Proceedings of Spring Conference on Computer Graphics* (Bratislava, Slovakia, 2015), Jorge J., Santos L. P., Ďurikovič R., (Eds.), Comenius University, pp. 85–92. 3

[KAS07] KOZLÍKOVÁ B., ANDRES F., SOCHOR J.: Visualization of tunnels in protein molecules. In *Proceedings of the 5th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa* (New York, NY, USA, 2007), AFRIGRAPH '07, ACM, pp. 111–118. 3

[KBP*16] KOLESÁR I., BYSKA J., PARULEK J., HAUSER H., KOZLÍKOVÁ B.: Unfolding and Interactive Exploration of Protein Tunnels and their Dynamics. In *Eurographics Workshop on Visual Computing for Biology and Medicine* (2016), Bruckner S., Preim B., Vilanova A., Hauser H., Hennemuth A., Lundervold A., (Eds.), The Eurographics Association. 3, 4

[KFS*17] KRONE M., FRIESS F., SCHARNOWSKI K., REINA G., FADEMRECHT S., KULSCHEWSKI T., PLEISS J., ERTL T.: Molecular surface maps. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 701–710. 3

[KKF*16] KOZLÍKOVÁ B., KRONE M., FALK M., LINDOW N., BAADEN M., BAUM D., VIOLA I., PARULEK J., HEGE H.-C.: Visualization of biomolecular structures: State of the art revisited. *Computer Graphics Forum* (2016). 2

[KKL*16] KRONE M., KOZLÍKOVÁ B., LINDOW N., BAADEN M., BAUM D., PARULEK J., HEGE H.-C., VIOLA I.: Visual Analysis of Biomolecular Cavities: State of the Art. *Computer Graphics Forum* (2016). 2

[Kvv*14] KOZLÍKOVÁ B., ŠEBESTOVÁ E., ŠUSTR V., BREZOVSKÝ J., STRNAD O., DANIEL L., BEDNÁŘ D., PAVELKA A., MAŇÁK M., BEZDĚKA M., BENEŠ P., KOTRY M., GORA A. W., DAMBORSKÝ J., SOCHOR J.: CAVER Analyst 1.0: Graphic tool for interactive visualization and analysis of tunnels and channels in protein structures. *Bioinformatics 30*, 18 (2014). 3

[KWFG03] KANITSAR A., WEGENKITTL R., FLEISCHMANN D., GRÖLLER M. E.: Advanced curved planar reformation: Flattening of vascular structures. In *Proceedings of the 14th IEEE Visualization 2003 (VIS'03)* (2003), IEEE Computer Society, p. 7. 2

[LBBH12] LINDOW N., BAUM D., BONDAR A.-N., HEGE H.-C.: Dynamic Channels in Biomolecular Systems: Path Analysis and Visualization. In *IEEE Symposium on Biological Data Visualization* (2012), pp. 99–106. 3

[LBH11] LINDOW N., BAUM D., HEGE H.-C.: Voronoi-Based Extraction and Visualization of Molecular Paths. *IEEE Trans. Vis. Comput. Graphics 17*, 12 (2011), 2025–2034. 2

[LCMH09] LAMPE O. D., CORREA C., MA K. L., HAUSER H.: Curve-centric volume reformation for comparative visualization. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (Nov 2009), 1235–1242. 2, 4, 8

[MBS07] MEDEK P., BENEŠ P., SOCHOR J.: Computation of tunnels in protein molecules using delaunay triangulation. *Journal of WSCG, University of West Bohemia, Pilsen 15(1-3)* (2007), 107–114. 2

[PKKO07] PETŘEK M., KOŠINOVÁ P., KOČA J., OTYEPKA M.: MOLE: A Voronoi Diagram-Based Explorer of Molecular Channels, Pores, and Tunnels. *Structure 15*, 11 (2007), 1357–1363. 2

[POB*06] PETŘEK M., OTYEPKA M., BANÁŠ P., KOŠINOVÁ P., KOČA J., DAMBORSKÝ J.: CAVER: A New Tool to Explore Routes From Protein Clefts, Pockets and Cavities. *BMC Bioinform. 7*, 1 (2006), 316. 2

[PvK*15] PAVELKA A., ŠEBESTOVÁ E., KOZLÍKOVÁ B., BREZOVSKÝ J., SOCHOR J., DAMBORSKÝ J.: Caver: Algorithms for analyzing dynamics of tunnels in macromolecules. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on PP*, 99 (2015), 1–1. 2

[SVGR16] SKANBERG R., VAZQUEZ P.-P., GUALLAR V., ROPINSKI T.: Real-time molecular visualization supporting diffuse interreflections and ambient occlusion. *IEEE Trans. Vis. Comput. Graphics 22*, 1 (2016), 718–727. 3

[vdZLBI11] VAN DER ZWAN M., LUEKS W., BEKKER H., ISENBERG T.: Illustrative molecular visualization with continuous abstraction. In *Proceedings of the 13th Eurographics / IEEE - VGTC Conference on Visualization* (Chichester, UK, 2011), EuroVis'11, The Eurographs Association & John Wiley & Sons, Ltd., pp. 683–690. 8

[YFW*08] YAFFE E., FISHELOVITCH D., WOLFSON H. J., HALPERIN D., NUSSINOV R.: MolAxis: Efficient and accurate identification of channels in macromolecules. *Proteins: Structure, Function, and Bioinformatics 73*, 1 (2008), 72–86. 2