Bioinformatics, 35(7), 2019, 1241–1243
doi: 10.1093/bioinformatics/bty761
Advance Access Publication Date: 31 August 2018
Applications Note

Structural bioinformatics

NanoShaper–VMD interface: computing and visualizing surfaces, pockets and channels in molecular systems

Sergio Decherchi[1,2], Andrea Spitaleri[1], John Stone[3] and Walter Rocchia [1,*]

[1]CONCEPT Lab, Istituto Italiano di Tecnologia, Genoa 16163, Italy, [2]BiKi Technologies s.r.l., Genova 16121, Italy and [3]Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

*To whom correspondence should be addressed.

Associate Editor: Alfonso Valencia

Received on April 18, 2018; revised on July 31, 2018; editorial decision on August 25, 2018; accepted on August 30, 2018

Abstract

Summary: NanoShaper is a program specifically aiming the construction and analysis of the molecular surface of nanoscopic systems. It uses ray-casting for parallelism and it performs analytical computations whenever possible to maximize robustness and accuracy of the approach. Among the other features, NanoShaper provides volume, surface area, including that of internal cavities, for any considered molecular system. It identifies pockets via a very intuitive definition based on the concept of probe radius, intrinsic to the definition of the solvent excluded surface. We show here that, with a suitable choice of the parameters, the same approach can also permit the visualisation of molecular channels. NanoShaper has now been interfaced with the widely used molecular visualization software VMD, further enriching its already well furnished toolset.

Availability and implementation: VMD is available at http://www.ks.uiuc.edu/Research/vmd/. NanoShaper, its documentation, tutorials and supporting programs are available at http://concept.iit.it/downloads.

Contact: walter.rocchia@iit.it

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

The powerful concept of molecular surface (MS) is mostly used for visualization purposes. It has however been successfully exploited for separating high from low dielectric constant regions for Poisson-Boltzmann calculations (Fogolari *et al.*, 2002). NanoShaper (NS) is a standalone software tool for building and analyzing the MS according to several different geometric definitions, including the solvent excluded surface (SES), the Skin and the Blobby Gaussian surfaces (Decherchi and Rocchia, 2013). It has been successfully integrated with the DelPhi Poisson-Boltzmann solver program (Decherchi and Rocchia, 2013; Wang *et al.*, 2013) and used in a MM/PBSA approach (Spiliotopoulos *et al.*, 2016). Among other functionalities, NS allows the identification of pockets and cavities, as well as their volume, surface area and list of constituting atoms. A robust and efficient way to identify and visually represent the MS, pockets, crevices and tunnels is a very desirable feature for structural bioinformatics, enabling a better understanding of the structural and mechanistic details involved in biological processes. The pocket detection ability of NS has recently proven to be key for the development of advanced approaches for monitoring pockets, and their interaction, along molecular dynamics simulations (La Sala *et al.*, 2017). Interestingly, while there are several tools performing pocket identification, the number of those dealing with channels is much smaller (Petrek *et al.*, 2006; Smart *et al.*, 1996).

Here, we present the interfacing of NS with the VMD tool (Humphrey *et al.*, 1996), and show how NS can also identify and visualize channels, further expanding its range of applications. More in detail, we first briefly recall the basis of the NS pocket detection

algorithm, then we show that its integration with VMD addresses a few issues that this latter has been experiencing when dealing with pathological or large-sized systems, and finally we show that with a different choice of the input parameters NS can identify the channel in the human gamma-aminobutyric acid receptor.

2 Methods and implementation

In order to ensure accuracy and manifoldness of the resulting MS also in degenerate cases and for large molecular systems, NS performs analytical computations whenever possible and explicitly considers the potentially critical geometric situations (Decherchi *et al.*, 2013; Decherchi and Rocchia, 2013).

Moreover, NS computes pockets via a very intuitive approach where they are basically defined as the volumetric difference between the space regions enclosed within the SESs of the molecular system obtained with two different probe radii. Although this definition somehow reminds of that in Kawabata and Go (2007), this is the first approach that fully exploits the SES concept. The implementation is grid-based: the raw volumetric map is obtained by flagging grid points that are simultaneously inside the SES calculated with the 'big probe', having radius '*R*', and outside the SES obtained with the 'small probe', of radius '*r*'. Since the raw volumetric map is quite noisy, to achieve a robust identification of the pockets, spurious results are filtered out. To do this, we apply a 'discrete Connolly filter' over all the originally flagged grid points. This filter preserves only points that are either: (i) surrounded by other flagged points for a distance of at least *r*, or (ii) within the *r* distance of points that fulfill condition (i). The default value for *r* is 1.4Å, corresponding to the average radius of a water molecule. In the end, pockets arise as all the unconnected components on the grid that survive the filtering. These components are found via a raster scan of the whole grid. Any time a flagged grid point is found, a flood-fill procedure similar to that adopted in the cavity detection algorithm, as described in Decherchi and Rocchia (2013), is used to label all the points of that pocket. Once a pocket is 'filled up', the pocket counter is incremented and the raster scan continues until the next pocket or the end of the grid are found. The surface area of each pocket is obtained by building the MS of the union of all the spheres centered in grid points of the above-mentioned type (i) and having radius *r*. This auxiliary MS is then triangulated and used for both surface area calculation and for visualization purposes. Further information provided by this procedure includes the list of identifying atoms, the distinction between a closed cavity and a pocket. For the latter, the entrance and its normal vector are also calculated.

From the computational viewpoint, the most demanding memory requirement of this method is the storage of two 3D grids; this scales in memory as $O(n^3)$ where n is proportional to the grid side. The most intense computational tasks are the Connolly filter, the flood fill and the construction of the pocket envelope. To save computing time, the calculation of the surface area can be avoided.

After NS produces surface triangle meshes, VMD reads them and performs post-processing on the meshes to convert them into groups of so-called triangle strips and fans, an efficient mesh representation that reduces the number of vertices processed during visualization to roughly 33% of the original mesh size. The mesh optimization step is particularly beneficial for interactive visualization of large complexes. After the surface mesh has been optimized, VMD builds internal data structures that associate mesh vertices with atoms, enabling VMD to rapidly color (or re-color) the surface mesh according to any atomic property of interest to the user. VMD can also color the surface mesh by properties such as electrostatic potential, or spatial cross-correlation maps using 3D (volumetric) texture mapping (Stone *et al.*, 2014).

3 Results and conclusions

3.1 Robustness of MS construction

The robustness and manifoldness in MS construction is tested in the case of fullerene, where there is a high symmetry in the atomic coordinates, and in the case of the 3.8Å resolution crystal of the adenovirus virion (98 419 atoms), shown in Figure 1. The capability of handling large systems is shown in Figure 2a and b. In Figure 2a NS is used to build the MS of the clathrin protein, a good stress test for MS algorithms due to its very large size and resulting triangle count, and its complex surface. In Figure 2b, the MS of the satellite tobacco mosaic virus (STMV) capsid (147 976 atoms) is depicted (Freddolino *et al.*, 2006). NS correctly computes surfaces for both systems, while in the former case neither Surf nor MSMS are able to produce a surface and in the latter case only MSMS and only if the RNA is present in the interior of the protein.

3.2 Computing and visualizing molecular channels

Analysis of molecular channels is of wide interest, since it allows a better understanding of translocation of ions or small substrates. Interestingly, the same definition used for the pockets was successful, with a different choice of the parameters, namely $R = 10$Å and $r = 3$Å, in detecting, for instance, the channel inside the human gamma-aminobutyric acid receptor.

Access to the usage of NS interface occurs via the new menu entry called NS in the 1.9.4 version of VMD. This provides the scientific community with a robust and efficient tool for visualizing

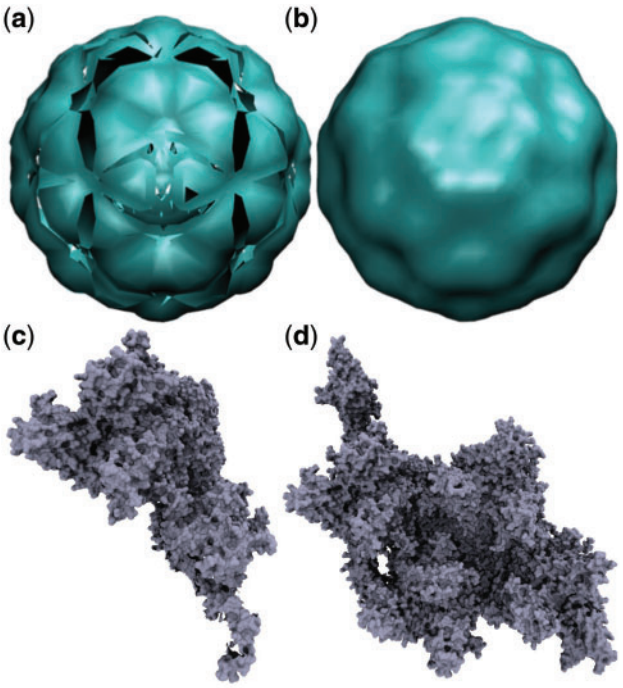

Fig. 1. VMD representation of the MS of fullerene via Surf (**a**) and NS (**b**). In (**a**) the surface is not manifold. In (**c**) and (**d**), the adenovirus virion (PDB code 4CWU) is represented, again with Surf and NS, respectively, from the same viewpoint. In the Surf calculation, part of the MS is missing, while the MSMS option could not handle this system

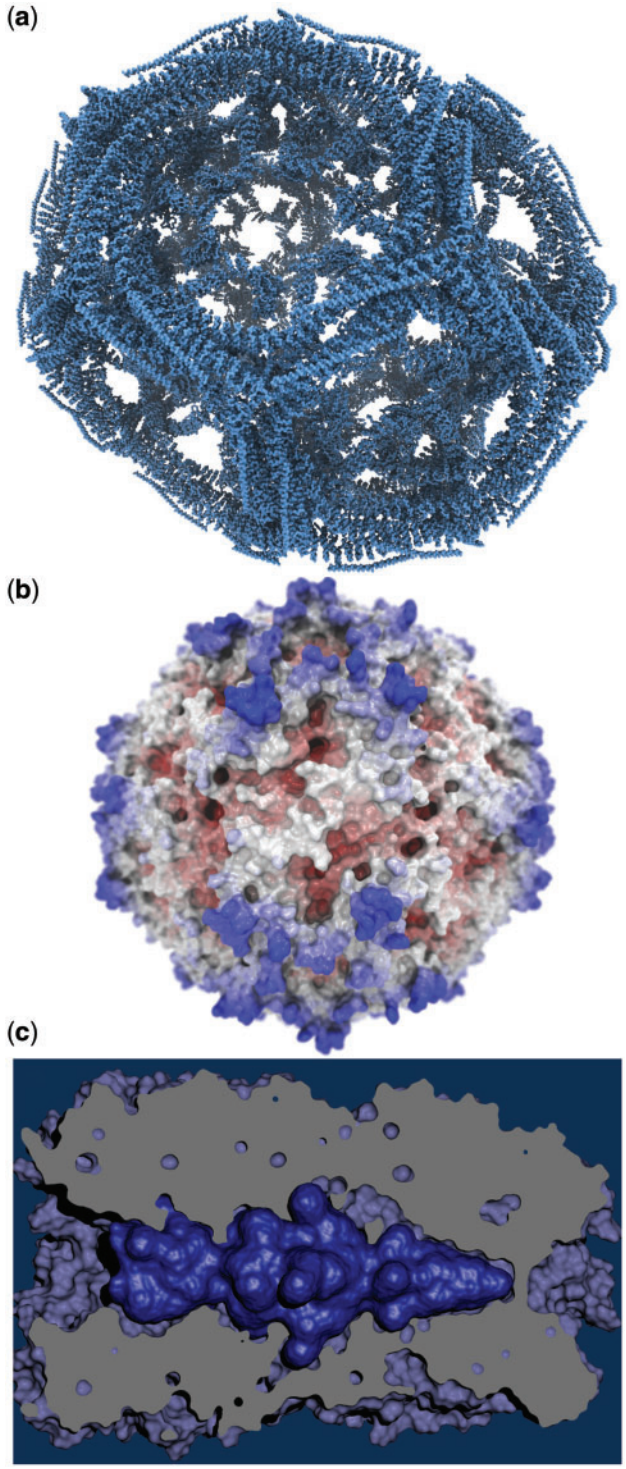

Fig. 2. In (**a**) and (**b**) the MSs of two challenging structures are represented, namely that of the clathrin protein, and that of the STMV. The surface rendering of the STMV capsid surface is colored by position, using a radial distance coloring scheme with a red-white-blue color scale ranging from red in the interior of the capsid to blue at the exterior. In (**c**), the human gamma-aminobutyric acid receptor central channel (PDB code 4cof) was identified via NanoShaper. All structures are visualized via VMD

MSs, internal cavities, superficial pockets and tunnels of various widths such as molecular channels. This in turn makes easier the identification of structural features and functional regions in molecular structures. This approach is heavily based on the concept of the SES, which is calculated with special accuracy, allowing, in addition to the computation of the pocket envelope, a more accurate estimation of volume and surface area, entrance points and normal vector detection and, most importantly for the present application, an explicit representation, via triangulation, which is essential for visualization purposes. It is interesting to note that the manifoldness of the MS built by NS, combined with the exporting capability of VMD in STL format, provides a reliable input for 3D-printing technology. In order to use the integrated software tools, the user needs to download NS form the https://concept.iit.it/ website. Example of configuration files for NS can be found as Supplementary Information.

Funding

For ongoing VMD development: National Institutes of Health [grant P41-GM104601]. NS has been partially supported by NIGMS, National Institutes of Health [grant 1R01GM093937-01].

Conflict of Interest: none declared.

References

Decherchi,S. and Rocchia,W. (2013) A general and robust ray-casting-based algorithm for triangulating surfaces at the nanoscale. *PLoS One*, **8**, e59744.

Decherchi,S. *et al.* (2013) Between algorithm and model: different molecular surface definitions for the Poisson-Boltzmann based electrostatic characterization of biomolecules in solution. *Commun. Comput. Phys*, **13**, 61–89.

Fogolari,F. *et al.* (2002) The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology. *J. Mol. Recognit*, **15**, 377–529.

Freddolino,P.L. *et al.* (2006) Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure*, **14**, 437–449.

Humphrey,W. *et al.* (1996) VMD—visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.

Kawabata,T. and Go,N. (2007) Detection of pockets on protein surfaces using small and large probe spheres to find putative ligand binding sites. *Proteins*, **68**, 516–529.

La Sala,G. *et al.* (2017) Allosteric communication networks in proteins revealed through pocket crosstalk analysis. *ACS Cent. Sci.*, **3**, 949–960.

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

Smart,O.S. *et al.* (1996) HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.*, **14**, 354–360.

Spiliotopoulos,D. *et al.* (2016) dMM-PBSA: a new HADDOCK scoring function for protein-peptide docking. *Front. Mol. Biosci*, **3**, 46.

Stone,J.E. *et al.* (2014) GPU-accelerated analysis and visualization of large structures solved by molecular dynamics flexible fitting. *Faraday Discuss.*, **169**, 265–283.

Wang,L. *et al.* (2013) Using DelPhi capabilities to mimic protein's conformational reorganization with amino acid specific dielectric constants. *Commun. Comput. Phys*, **13**, 13–30.