JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

APL@Voro: A Voronoi-Based Membrane Analysis Tool for GROMACS Trajectories

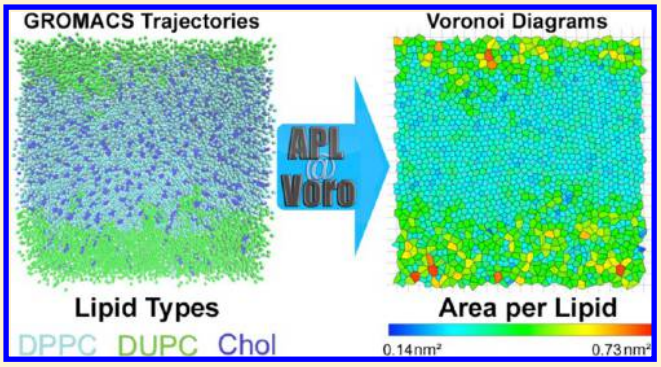
Gunther Lukat,*[,†] Jens Krüger,[‡] and Björn Sommer[†]

[†]Bio-/Medical Informatics Department, University of Bielefeld, Universitätsstr. 25, 33615 Bielefeld, Germany

[‡]Applied Bioinformatics Group, University of Tübingen, Sand 14, 72076 Tübingen, Germany

Supporting Information

ABSTRACT: APL@Voro is a new program developed to aid in the analysis of GROMACS trajectories of lipid bilayer simulations. It can read a GROMACS trajectory file, a PDB coordinate file, and a GROMACS index file to create a two-dimensional geometric representation of a bilayer. Voronoi diagrams and Delaunay triangulations—generated for different selection models of lipids—support the analysis of the bilayer. The values calculated on the geometric structures can be visualized in a user-friendly interactive environment and, then, plotted and exported to different file types. APL@Voro supports complex bilayers with a mix of various lipids and proteins. For the calculation of the projected area per lipid, a modification of the well-known Voronoi approach is presented as well as the presentation of a new approach for including atoms into an existing triangulation. The application of the developed software is discussed for three example systems simulated with GROMACS. The program is written in C++, is open source, and is available free of charge.



1. INTRODUCTION

Molecular dynamics (MD) simulations of lipid bilayers and bilayers holding proteins have proven useful and are widely used.[1−5] The analysis of the resulting trajectories and the conversion from the three-dimensional coordinate files to useful figures can be challenging. The extraction of useful data like the area per lipid and the thickness is a complex task, particularly for mixed bilayers and bilayers holding proteins, and often requires scripting and time-consuming work on snapshots of trajectories. To reduce the laborious work on trajectories, easy-to-use software is needed that produces images and plots and can deal with large trajectories generated by MD software. APL@Voro is a free, open source tool that has been developed to reduce analysis time as well as interpolate the bilayer thickness and calculate the projected area per lipid with the use of planar subdivisions. APL@Voro generates two-dimensional and three-dimensional plots, renders two-dimensional presentations of the bilayer's leaflets, and provides several interactive features to support the analyzing process. The computations carried out by APL@Voro are based on a two-dimensional geometric representation of the leaflet structures known as Voronoi diagrams and Delaunay triangulations.[6,7] These structures have become some of the most useful ones in geometric computations and are well-known in physics and mathematics. In the field of bilayer analysis, Shinoda and Okazaki applied Voronoi diagrams to the calculation of lipid areas by using the lipid's centers of mass to generate the Voronoi diagrams.[3] This approach was refined by Pandit et al. using a special selection of key atoms to derive the membrane thickness.[3,8,9] Nowadays, the Voronoi diagrams are widely used in molecular analysis to examine different structural properties like lipid neighborhoods in mixtures containing cholesterol, phase transitions in suspensions, or lateral organization in lipid bilayers.[10−12] There are even approaches for derived Voronoi structures like the Voronoi S-diagram or combinations of Voronoi diagrams with Monte Carlo integration methods for area per lipid analysis in protein−membrane systems.[13,14] This wide range of applications for Voronoi diagrams shows the great flexibility of this data structure. Although Voronoi diagrams nowadays are often used for the simulation evaluations, the software used to generate the diagrams is seldom mentioned. In some cases, the popular QHull program or the Triangle program are used.[15,16] Both packages can generate Delaunay triangulations or Voronoi diagrams for a set of points given in different file or input formats. In the following sections, a basic description of the algorithms, data structures, and the interactive features of APL@Voro is given. In the section Description of APL@Voro, APL@Voro's internal logic and the used data structures are described. A modified version of the basic Voronoi-based algorithm for bilayer thickness calculation is described in the section Calculating the Membrane Thickness. In the section Triangular Prism Insertion Method (TPIM), a method for including protein

Received: March 22, 2013
Published: October 31, 2013

   dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925

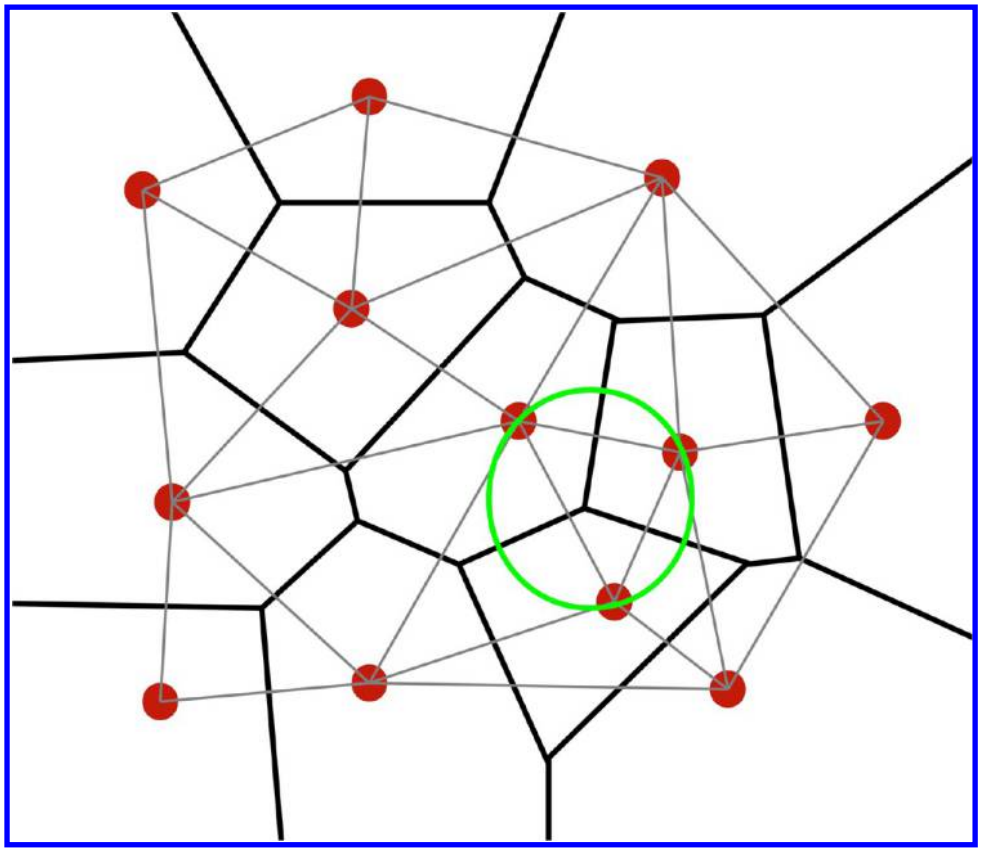


Figure 1. The duality of the Voronoi diagram (black lines) and the Delaunay triangulation (gray lines) for a set of points (red). The green circle marks the circumscribed circle of a triangle in the Delaunay triangulation centered at a vertex of the Voronoi diagram.

atoms in an existing triangulation of a bilayer is described. The interactive features of APL@Voro, like the Voronoi cell selection and plotting, are described in the section User Interfaces and Visualizing Data. The section Validation Simulations provides information on the setup of the three different bilayers, which are processed with APL@Voro in section Results and Discussion.

2. METHODS AND DATA STRUCTURES

2.1. The Voronoi and Delaunay Diagrams. Our tool takes advantage of several geometric structures and algorithms. In general, two main structures are used, the Voronoi diagram and the Delaunay triangulation.[6,7] Voronoi diagrams are widely used in bilayer analysis for calculating the projected area per lipid or the membrane's thickness. Here, we adopt specific algorithms used for Voronoi diagrams and use them for the Delaunay triangulation. A general definition of the Voronoi diagram can be written as follows: Let S denote a set of n points (called *sites*) in $\mathbb{R}^2$. A Voronoi cell $\mathrm{VC}(p, S)$ of a point $p \in S$ is defined as

$$\mathrm{VC}(p, S) = \bigcap_{q \in S - \{p\}} \mathrm{dom}(p, q) \quad (1)$$

Using δ for denoting the Euclidean distance function, $\mathrm{dom}(p; q)$ describes the open half-plane given by

$$\mathrm{dom}(p, q) = \{x \in \mathbb{R}^2 | \delta(x, p) < \delta(x, q)\} \quad (2)$$

Intersecting the $n - 1$ half planes results in the formation of convex polygons, the Voronoi cells. Each point on a boundary of a Voronoi cell called *edge* is equidistant from exactly two *sites*, and each corner of a Voronoi cell called *vertex* is equidistant from at least three *sites*. As a consequence, the Voronoi cells form a polygonal partition of the plane. This partition is called the Voronoi diagram, $V(S)$, of the set of points S. In Figure 1, a Voronoi diagram is shown with *sites* colored in red and *edges* colored in black.

2.1.1. Delaunay Diagrams. The geometric dual graph of the Voronoi diagram, where any two *sites* are connected whose Voronoi cells have an edge in common, was considered by Voronoi himself and is called the Delaunay triangulation.[7] Given a maximal planar subdivision as a subdivision SD such that any edge not in SD intersects one of the existing edges, a triangulation T of a set of points S can be defined as a maximal planar subdivision SD of S whose vertex set is S. For the set of points S, the Delaunay triangulation $\mathrm{DT}(S)$ is a triangulation with the additional criterion that no point $p \in S$ lies inside the circumscribed circle of any triangle in $\mathrm{DT}(S)$.[18] In Figure 1, the Delaunay triangulation (gray lines) and the circumscribed circle of a triangle (green circle) are shown.

The circumscribed circle of a triangle in $\mathrm{DT}(S)$ can be used to derive the Voronoi diagram as the center of that circle marks exactly one *vertex* of the dual Voronoi diagram. This circle can be used to derive the Voronoi diagram from a Delaunay triangulation in a similar way. The circle centered at a *vertex* of the Voronoi diagram passing through the nearest three *sites* is exactly the circumscribed circle of a triangle in $\mathrm{DT}(S)$. In the following sections, this circle will be called the *Delaunay circle*.

2.2. Calculating the Projected Area per Lipid. In single component mixtures such as pure DPPC (1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine) bilayers, the average area per lipid can be computed by simply dividing the simulation box vectors by the number of lipids. Formally, the area per lipid of one layer

dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925

(assuming both layers contain equal numbers of lipids) is then defined as

$$\text{Area} = \frac{\text{Box } X \cdot \text{Box } Y}{N_{\text{lipids}}/2} \tag{3}$$

Of course, this equation is not suitable for bilayers containing different types of lipids with distinct sizes such as CHOL (Cholesterol) and DPPC. One method of calculating the area per lipid for mixed systems is based on Voronoi diagrams. To calculate the projected area per lipid, the coordinates of certain key atoms of the lipids are projected onto a plane, and a Voronoi diagram is calculated using these projected coordinates. Formally, the projected area $A(l)$ of one lipid l present in the leaflet L, represented by one key atom pl, is then defined as the area of the corresponding Voronoi cell (the area of the polygon) $A_{\text{VC}(pl,L)}$ with

$$A_{\text{VC}(pl,L)} = \frac{1}{2}\sum_{i=0}^{N=1}(x_i y_{i+1} - y_i x_{i+1}) \tag{4}$$

where N is the number of vertices of the Voronoi cell and (x_i,y_i) with $i \bmod(N)$ are the coordinates of the vertices of the Voronoi cell. The molecular total area of a lipid $A(l)$ described by N key atoms is then defined as

$$A(l) = \sum_{I=0}^{N} A_{\text{VC}(i,L)} \tag{5}$$

For membranes containing a mixture of lipids and cholesterols, the embedded position of cholesterol molecules inside the layer has to be taken into account for the calculation of the area per lipid. In general, the area per phospholipid, A_P, for membranes containing cholesterol is expected to be lower than for membranes without cholesterol. This is referred to as the condensing effect of cholesterol: the sum of areas of the individual bilayer components is higher than the surface area of a cholesterol-containing lipid bilayer.[19,20] Several methods have been proposed to calculate the area per phospholipid A_P and the area per cholesterol A_{Chol} for membranes containing a mixture of cholesterol and phospholipids.[21−23] Following the *areal-determining plane method* of Alwarawrah et al.[23] for a two component mixture, the total membrane area A in the areal-determining plane is decomposed into the area per phospholipid A_P and the area per cholesterol A_{Chol} with

$$A = A_P + A_{\text{Chol}} \tag{6}$$

The area per cholesterol $A_{\text{Chol}}(x_c)$ for certain cholesterol concentration x_c is calculated on the basis of the average tilt angle of cholesterol $\theta(x_c)$ as a function of x_c and assuming a cross section area of 0.38 nm^2 for cholesterol molecules using

$$A_{\text{Chol}}(x_c) = \frac{A_0}{\cos(\theta(x_c))} \tag{7}$$

The area per phospholipid is then defined as

$$A_P(x_c) = \frac{A(x_c) - N_{\text{Chol}}A_{\text{Chol}}(x_c)}{N_P} \tag{8}$$

where $A(x_c)$ is the total lateral area of the simulation box and N_{Chol} and N_P are the numbers of cholesterols and PCs in each leaflet, respectively.[23]

The calculation of the area per lipid using the proposed Voronoi method in comparison to the *areal-determining plane method* is discussed in the section Mixed Bilayer on the basis of system M2. Using the MARTINI force field for System M2, $\theta(x_c)$ was calculated as the angle between the bilayer normal and the vector connecting R5 and ROH beads.

2.3. Calculating the Membrane Thickness. A common practice used to calculate the average membrane thickness is the peak-to-peak separation of an electron density profile, as the peaks represent the location of electron-rich phosphocholine groups. The findings of Pandit et al. show that this method is unsuitable for the purpose of analyzing domain formation due to possible undulations of the bilayer and the biphasic nature of such systems.[24] Pandit et al.[24] proposed an algorithm based on a Voronoi tessellation to calculate the bilayer thickness and provide thickness data for single lipids. In APL@Voro, Voronoi diagrams are already used for the calculation of the area per lipid and thus can be reused for the calculation of the bilayer thickness. The proposed algorithm is based on finding a vertical neighbor phosphorus atom in one leaflet for a phosphorus atom in the other leaflet (see Figure 2).[8,24] This is done by

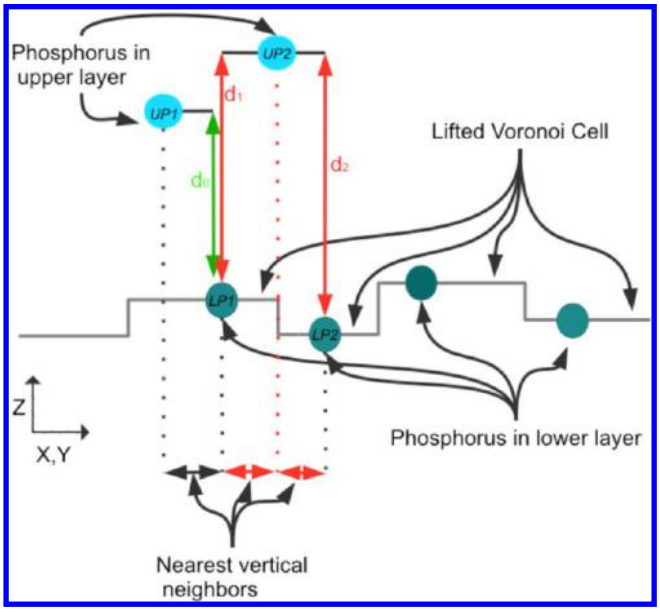


Figure 2. Schematic drawing, describing the method used to calculate the surface-to-point correlation function.[24] The projection of UP1 (phosphorus atom in upper layer) falls inside the Voronoi cell of LP1 (phosphorus atom in lower layer); therefore the identified vertical neighbor is LP1, and the calculated local thickness is d_0. The projection of UP2 falls on an edge of the Voronoi cell of LP1 and LP2. Both LP1 and LP2 can be treated as vertical neighbors of UP2, and therefore d_1 or d_2 can be treated as the local thickness calculated at the coordinate of UP2.

(i) providing a Voronoi tessellation for one leaflet
(ii) projecting the coordinates of the phosphorus atom of the other leaflet onto the Voronoi surface
(iii) identifying the Voronoi cells in which the projected coordinates fall
(iv) defining the z distance between the Voronoi site atom and the projected atom as the local thickness

Given this algorithm, the problem of finding the membrane thickness is reduced to the well-known point-in-polygon problem.[25] This approach may lead to some uncertain results, given the nature of Voronoi diagrams and the implemented

dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925

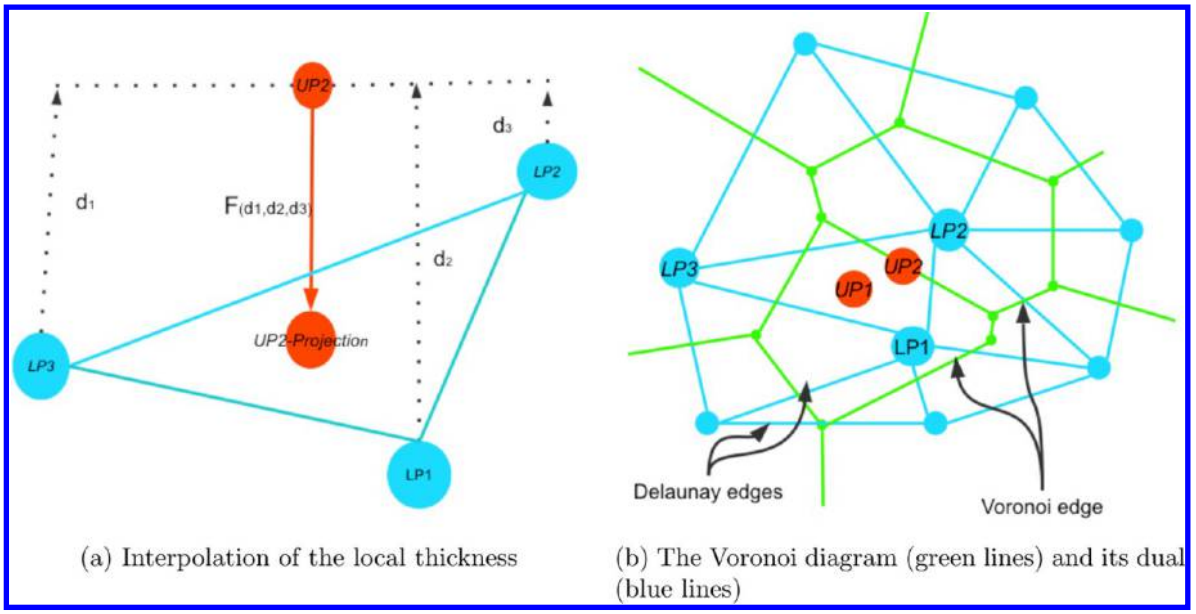


Figure 3. (a) Schematic drawing describing the method used to calculate the interpolated membrane thickness. The projection of UP2 falls inside the triangle defined by LP1, LP2, and LP3. Assuming the surface of the triangle is a plane, the interpolated thickness at the point of UP2 is an interpolation function of d_1, d_2, and d_3. (b) Schematic drawing of the Voronoi diagram and the Delaunay triangulation. Point UP2 shows the pathological case where a point falls on an edge of the Voronoi diagram. The local thickness at point UP2 can be interpolated between LP1, LP3, and LP2.

point-in-polygon test. Two cases can be identified where the given approach can result in uncertain data:

(i) The projected coordinate falls exactly on an edge of a Voronoi cell
(ii) The projected coordinate falls exactly on a vertex of the Voronoi cell

In the first case, the identification of exactly two vertical neighbors is possible, as the common edge of two Voronoi cells marks all points being the same distance from sites to which the cells are related. In Figures 2 and 3, this case is illustrated for UP2 (upper layer phosphorus atom) that is projected onto the Voronoi diagram of the lower layer. Because the projection of UP2 falls directly on an edge of the diagram, LP1 (lower layer phosphorus atom) and LP2 can be treated as vertical neighbors, and thus d1 (distance from LP1 to UP2) and d2 (distance from LP2 to UP2) are possible distance values. In the latter case, one can identify even more than two vertical neighbors, as a vertex of a Voronoi cell marks all points having the same distance from the sites related to the cells sharing the same vertex. Furthermore, the assignment of vertical neighbors depends on the point-in-polygon test algorithm that is used. In case the algorithm treats points on edges or vertices as inside the polygon, it has to be guessed which vertical neighbor to choose. On the other hand, no neighbor can be allocated if the algorithm treats these points as outside the polygon. To deal with this problem, we propose a different approach for calculating the membrane thickness (see Figure 3).

Instead of projecting coordinates onto a Voronoi tessellation, the coordinates are projected onto the Delaunay triangulation. The calculated membrane thickness—the z distance between the projected atom and the vertical neighbor—is then defined as an interpolation function between the three nearest neighbors, thus the calculated values refer to the interpolated thickness or the interpolated z distance. This is achieved by

(i) providing a Delaunay triangulation for one leaflet
(ii) projecting the coordinates of the phosphorus atom of the other leaflet onto the triangulated surface
(iii) identifying the triangle in which the projected coordinates fall
(iv) interpolating the z distance between the corners of the triangle and the projected point

The equation of the plane used to compute the elevation at any point on the triangle is defined by the three vertices of the triangle in which the projected atom $p_a(x_a, y_a, z_a)$ falls

$$z = f(x_a, y_a) = -\frac{A}{C}x_a - \frac{B}{C}y_a - \frac{D}{C} \tag{9}$$

where A, B, C, and D are computed from the triangle coordinates $p_0(x_0, y_0, z_0)$, $p_1(x_1, y_1, z_1)$, and $p_2(x_2, y_2, z_2)$ and the atom coordinates $p_a(x_a, y_a, z_a)$:

$$z_{0i} = |z_0 - z_a|,\ z_{1i} = |z_1 - z_a|,\ z_{2i} = |z_2 - z_a|$$

$$A = y_0(z_{1i} - z_{2i}) + y_1(z_{2i} - z_{0i}) + y_2(z_{0i} - z_{1i})$$

$$B = z_{0i}(x_1 - x_2) + z_{1i}(x_2 - x_0) + z_{2i}(x_0 - x_1)$$

$$C = x_0(y_1 - y_2) + x_1(y_2 - y_0) + x_2(y_0 - y_1)$$

$$D = -Ax_0 - By_0 - Cz_{0i}$$

The ambiguous cases of (i) point on edge or (ii) point on corner are handled separately. In the first case (i), the z distance is interpolated between the starting and end point of the edge on which the projected atom falls. In the latter case (ii), no interpolation is needed, and the z distance is exactly the z distance of the corner atom and the projected atom.

2.4. Triangular Prism Insertion Method (TPIM). Equation 3 relies strongly on averaging, and—in case of protein−membrane systems—it will produce large errors. Several methods have been proposed to access the area per lipid in such systems. One method to calculate the average area per lipid relies on the use of the cross-sectional area of proteins,

the box area, and the number of lipids.[26,27] Another method is to use a grid analysis.[28] The former method lacks accessibility of data for single lipids; the accuracy of the latter strongly depends on the grid size. The simple Voronoi tessellation method cannot be applied directly to protein−membrane systems due to the large size differences between lipid and protein molecules. A more accurate approach is based on a combination of Monte Carlo integration with the Voronoi tessellation method (VTMC method).[13] This method relies on the identification of boundary lipids—lipids neighboring protein atoms. A Voronoi tessellation is performed for the lipids's centers of mass, and protein atoms are projected onto this tessellation. If one of the protein atoms falls inside the *Delaunay circle* (see Figure 1), all lipids corresponding to the Voronoi cells sharing the vertex centered in the *Delaunay circle* are marked as boundary lipids.

In order to calculate the area per lipid for boundary lipids, the Monte Carlo integration method is used, where a Voronoi cell is probed by randomly inserting a pseudo-particle inside a Voronoi cell. The area of the boundary lipid A_{lb} is then defined as a product of the Voronoi cell area A_{VC} and the probability p that a particle insertion missed a projected protein atom.

$$A_{lb} = A_{VC}{}^{*}p$$

$$p = \frac{N_{total} - N_{hit}}{N_{total}}$$

where N_{total} is the total number of inserted particles and N_{hit} is the total number of inserted particles that hit a protein atom.[13] The accuracy of the VTMC method depends on the number of random inserted particles.[13] Although Mori et al. propose to use the lipid's centers of mass to calculate the Voronoi diagram, they mention a (computationally more costly) method to use key atoms. The major reason APL@Voro does not use this method is the missing opportunity to directly derive a new Voronoi cell (for visualization purpose) for a boundary lipid.

We propose another method of (i) calculating the lateral area of a protein and (ii) calculating the area of boundary lipids based on a method we call the *Triangular Prism Insertion Method* (TPIM). Similar to the VTMC method, TPIM relies on the identification of boundary lipids. A lipid is marked as a boundary lipid for a given protein atom, if the atom falls inside a triangular prism constructed from the lipid atom coordinates and a given triangulation. The protein atom is then marked for later insertion into the existing triangulation. Inserting a protein atom into the triangulation will directly result in the reduced size of the derived neighboring Voronoi cells (the lipid Voronoi Cells). To take protein atoms into account that lie slightly above or under the dimension defined by the triangulated lipid key atom coordinates, the height of the prism can be modified. The height of the triangular prism is defined by the coordinates of the lipid atoms and the TPIM offset value. A higher prism will result in more atoms remaining protein atoms and, thus, can reduce the calculated area of the boundary lipids. The offset value can either be a user defined value or be a van der Waals radius corresponding to the respective three lipid atoms. [An offset value of 0.0 will result in a prism height of exactly the maximum z distance of the three contributing lipid atoms.] Formally, the triangular prism P using the two triangular bases $T1$ and $T2$ is constructed given a lifted Delaunay triangle DT(L) for the lipid atom coordinates $L = \{A1(x_1, y_1, z_1), A2(x_2, y_2, z_3), A3(x_3, y_3, z_3)\}$ and the corresponding TPIM offset value $A1_v$, $A2_v$, $A3_v$ with

$$T = A1'(x_1, y_1, z_1 - A1_v),\ A2'(x_2, y_2, z_2 - A2_v),\ A3'(x_3, y_3, z_3 - A3_v)$$

$$T2 = A1''(x_1, y_1, z_1 + A1_v),\ A2''(x_2, y_2, z_2 + A2_v),\ A3''(x_3, y_3, z_3 + A3_v)$$

The basic algorithm for the TPIM test procedure is shown in the flowchart in Figure 4. The TPIM test takes a triangulation

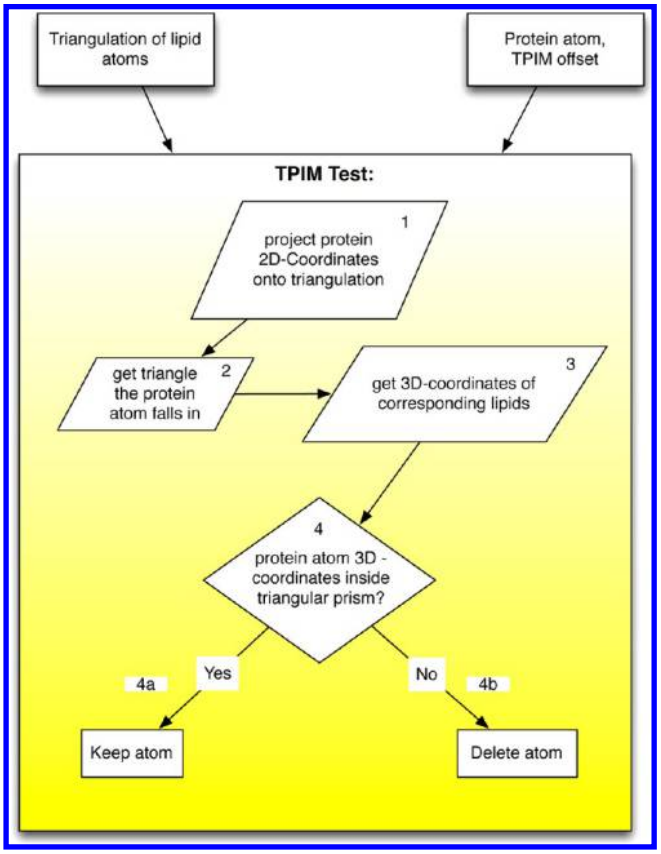


Figure 4. Flowchart of the internal TPIM procedure. The TPIM test takes a protein atom and a Delaunay triangulation of lipids. The test will either keep the atom or delete it (if it does not fall into the triangular prism created with the 3D-coordinates (step 3)).

of lipid atoms for one layer together with a list of the protein atoms and a TPIM offset value. The test starts by projecting the XY coordinates of the protein atom onto the Delaunay triangulation (Figure 4.1). The triangle the protein atom falls in is temporarily stored (Figure 4.2) and the original coordinates in $\mathbb{R}^3$ of the triangle lipids are achieved (Figure 4.3). In the next step, these lipid atom coordinates are used together with an offset value v to construct a triangular prism, and it is tested if the protein atom falls inside the prism (Figure 4.4). For protein atoms falling inside the prism, the test will return true, and the protein atom will remain in the list of protein atoms (Figure 4.4a). Otherwise, the test will return false, and the protein atom is deleted from the list (Figure 4.4b).

Once all protein atoms are processed this way, the remaining protein atoms are inserted in the triangulation, resulting in a reduced size of the derived Voronoi cells for boundary lipids. The use of lifted Delaunay triangles to include protein atoms

ensures that fluctuations in the bilayer thickness or bilayer curvature are taken into account. It can be expected that the lateral area of the embedded protein is reflected more accurately in this way than by using an environment definition based on maximum and minimum coordinates of the whole leaflet.

2.4.1. Selection of Key Atoms. The properties of the Voronoi cells heavily depend on the set of points: thus, the whole Voronoi diagram is a function of a specific set points. Instead of using the centers of mass for the calculation of Voronoi diagrams,[3] the triangulations and Voronoi diagrams used by our tool are calculated for the coordinates of certain key atoms.[3,8,9]

A key atom selection can be understood as the selection of a slice in the z dimension of a layer. Here, different slices are used for the analysis of the three test systems M1, M2, and M3. Independent from the chosen selection, lipids or cholesterols during the flip flop process (a transmembrane lipid translocation) can still lead to erroneous local results.

The Simple Key Atom Model. Using the *simple model* with a lipid headgroup atom, such as the phosphorus atom of PC lipids, a good approximation of the membrane thickness can be achieved.[24] In APL@Voro, the *simple model* can also aid in following a lipid type separation, as Voronoi cells can be colored by the lipid type to which they refer. This technique was used in the section Mixed Bilayer. A respective Voronoi diagram is shown in Figure 13b.

The Boundary Key Atom Model. The boundary key atom selection model or short *boundary model* can be understood as a general recipe for deriving the local membrane thickness and the area per lipid.[24] The selected key atoms are lying approximately at the interface of hydrophobic and hydrophilic portions of each lipid type. This model was tested in the section Mixed Bilayer.

The Chain Key Atom Model. The chain key atom selection model or short *chain model* is a selection model that focuses the observation of phase changes and was proposed by Marrink et al. to distinguish lipids in the fluid from lipids in the gel phase.[29] This model was used in the section Simple Bilayer (Figure 10a,c) to distinguish the different lipid phase states. In general, one can distinguish lipids in the gel phase from lipids in the liquid phase by observing their chain ordering. Using the *chain model*, a part of this information is included in the Voronoi diagram and expressed by the form of single Voronoi cells, as Voronoi cells referring to stretched and straight lipid chains appear in a honeycomb-like form (forming a perfect hexagon).[29]

The Maximum Density Key Atom Model. The maximum density key atom selection model or short *maximum−density model* focuses the calculation of the area per lipid for membranes containing cholesterol. The key atoms for this model are lying approximately at the sterol ring region of embedded cholesterol. The findings of Alwarawrah et al.[23] and Wennberg et al.[30] depict that the area per lipid in PC/cholesterol bilayers is primarily determined by the molecular packing in this region. The calculated area per cholesterol is expected to be higher with this selection model than with the other proposed models. In section Mixed Bilayer, this selection model was tested.

Protein Group Selection. For bilayers with one or more proteins, the previously mentioned selection models can be applied to lipids. The selection of protein atoms in APL@Voro is done using a GROMACS index file. The accuracy of the TPIM procedure strongly depends on the presence of the full set of atoms referring to proteins. In the Voronoi diagrams, every atom that was selected by the TPIM procedure is displayed as one Voronoi cell. The protein group selection enables the user to split a protein into several groups and access protein atoms by their group. Internally, APL@Voro treats every group as a single protein. For single-pass membrane proteins, a selection of all protein atoms comes to mind, while for multipass membrane proteins a selection of several groups can be useful. The selection of protein groups was used in the section Bilayer with Peptide for an example system (M3) holding a bilayer of 122 POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine) lipids with one embedded $Vpu_{1-32}WT$ peptide.[31]

3. DESCRIPTION OF APL@VORO

APL@Voro supports the analysis of molecular dynamics simulations of membranes carried out by GROMACS.[32−34] Voronoi diagrams of the membrane are displayed, and plots of the area per lipid and the membrane thickness can be generated using an interactive user interface described in the section User Interfaces and Visualizing Data. The calculated values for plots can be exported in several file formats for further manipulation with scripts or spreadsheet software. The general workflow of APL@Voro can be described in four steps: a startup step covered in section Starting Procedure, a triangulation step covered in section Generating Triangulations, a step where the calculated data are extracted (see section Extracting Data), and the final visualization step. Binaries of APL@Voro are avialable for OpenSuse Factory, openSuse 12.3, openSuse 12.2, Fedora 18, OSX Lion, and Ubuntu 12.04 (Precise Pangolin). The binary packages vary in size between 10 MB (OSX Lion) and 5 MB (Ubuntu 12.04 (Precise Pangolin)). APL@Voro uses C++11 features and therefore requires the libstdc++ > 4.7 library. The GUI and some other parts of APL@Voro require QT4 ≤ 4.8 (including qt Webkit) libraries as well as qwt ≥ 6.1 libraries for two-dimensional plotting and qwtplot3d ≥ 0.2.7 libraries for three-dimensional plotting. Finally, the Gromacs XTC library (1.1.0) is needed for XTC/TRR parsing. A user's guide, providing extensive descriptions of all options accessible in the tool and on how to run specific analyses, is distributed with the tool [download the manual at www.aplvoro.org].

3.1. Starting Procedure. In general, three different files are used by APL@Voro. A PDB coordinate file that contains a lipid bilayer (oriented normal to the *z* axis), an optional GROMACS-format trajectory file (XTC/TRR) for the bilayer described in the coordinate file, and an optional GROMACS-format index file (NDX), in case the system contains any protein atoms. The coordinate file and the trajectory file must not contain broken molecules. This can be assured by use of the GROMACS tool *trjconv* with option *-pbc whole*. Different options on how to treat periodic boundaries and what algorithm is used to calculate the membrane thickness are set by the user on startup.

Right after the user has finished his entries, the PDB file is parsed by APL@Voro to obtain a list of all residue types and their atoms. This list is presented to the user for the lipid-key-atom selection. Similar to the PDB residue parsing, a given index file is parsed and its contents presented in a separate dialogue for a group selection. Additionally, the TPIM offset method can be selected. Possible values for the TPIM offset are (i) van der Waals radii, (ii) user-defined, and (iii) absolute (see section Triangular Prism Insertion Method (TPIM) for further

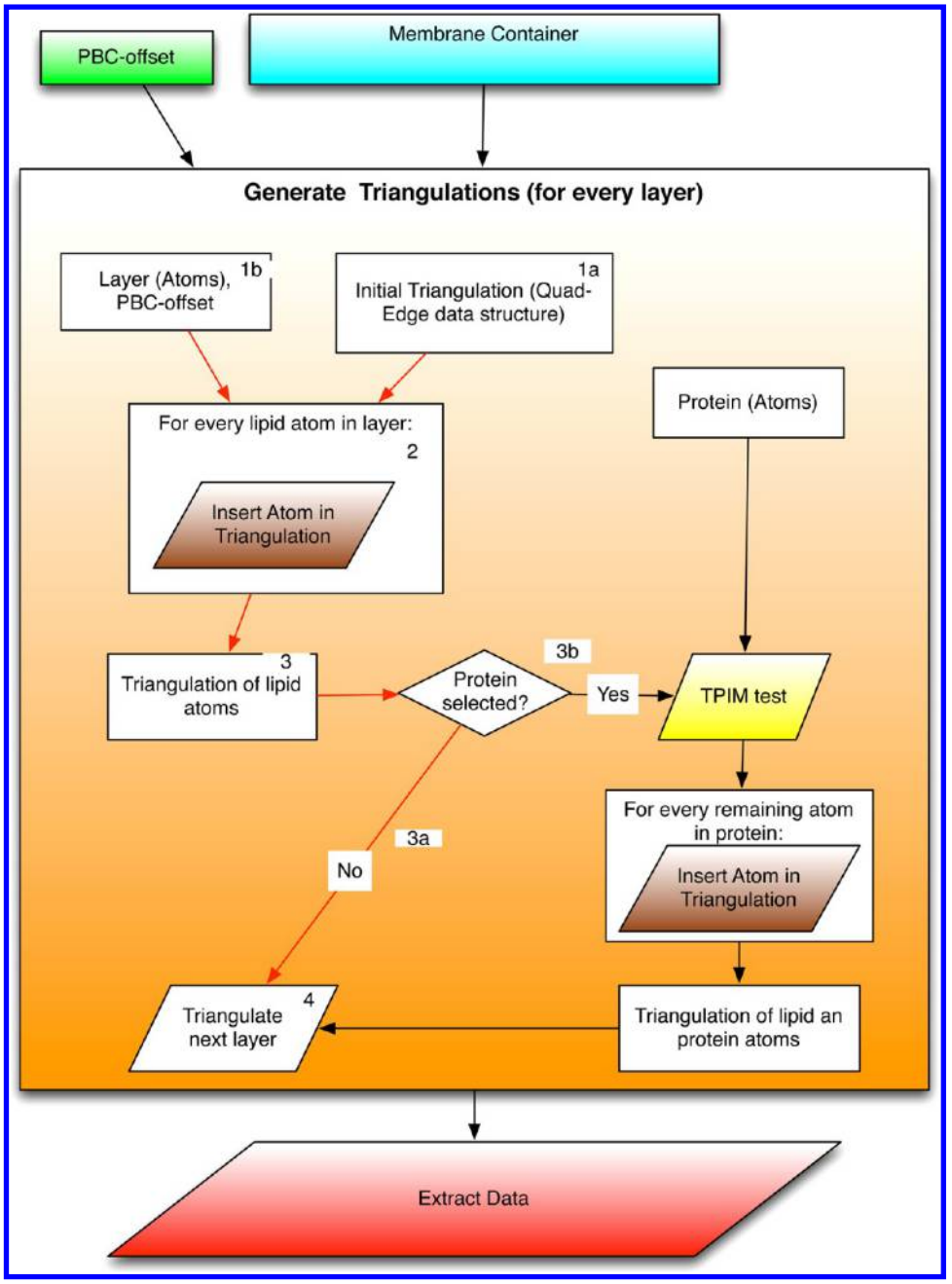


Figure 5. Flowchart of the internal logic of APL@Voro in the program state where triangulations are generated. Rectangles mark internally stored data structures. Parallelograms mark program states, where the calculation is carried out. Rhomboids mark decision states of the program. Red arrows mark the programs flow for a minimal configuration (i.e., without protein). The brown colored function "Insert Atom in Triangulation" is described in the section Insert Atom in Triangulation Procedure; the yellow colored "TPIM test" is described in the section Triangular Prism Insertion Method (TPIM).

details about the TPIM offset). In case a protein is embedded in the membrane and no index file is given, APL@Voro will behave as if there is no protein, and the calculated projected area per lipid for lipids neighboring the protein will be too high. By selecting key-atoms for lipids and groups for a protein, a subset of the contents of the simulation box is defined that will be used for the triangulation procedure and the setup of the internal data structures.

The user-selected lipids and protein groups are read from the PDB file and, if given, the trajectory file. Every lipid is temporarily stored together with all relating atoms. A lipid is assigned either to the upper layer or the lower layer depending on it's orientation (head → tails). Once a lipid is assigned to a layer, all but the previous selected atom types are deleted from the temporary stored lipid. In case a trajectory file is read, the same procedure is applied to every frame that is stored in the trajectory. After all frames are read, the actual triangulation procedure takes place.

3.2. Generating Triangulations. The actual Delaunay triangulation is computed using a Quad-Edge data structure, the preprocessed atom coordinates, and the user-defined PBC offset.[35] [A Quad-Edge data structure is a representation of a two-dimensional map, that is, a graph drawn on a (closed) surface. Here, the structure simultaneously represents both the

dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925

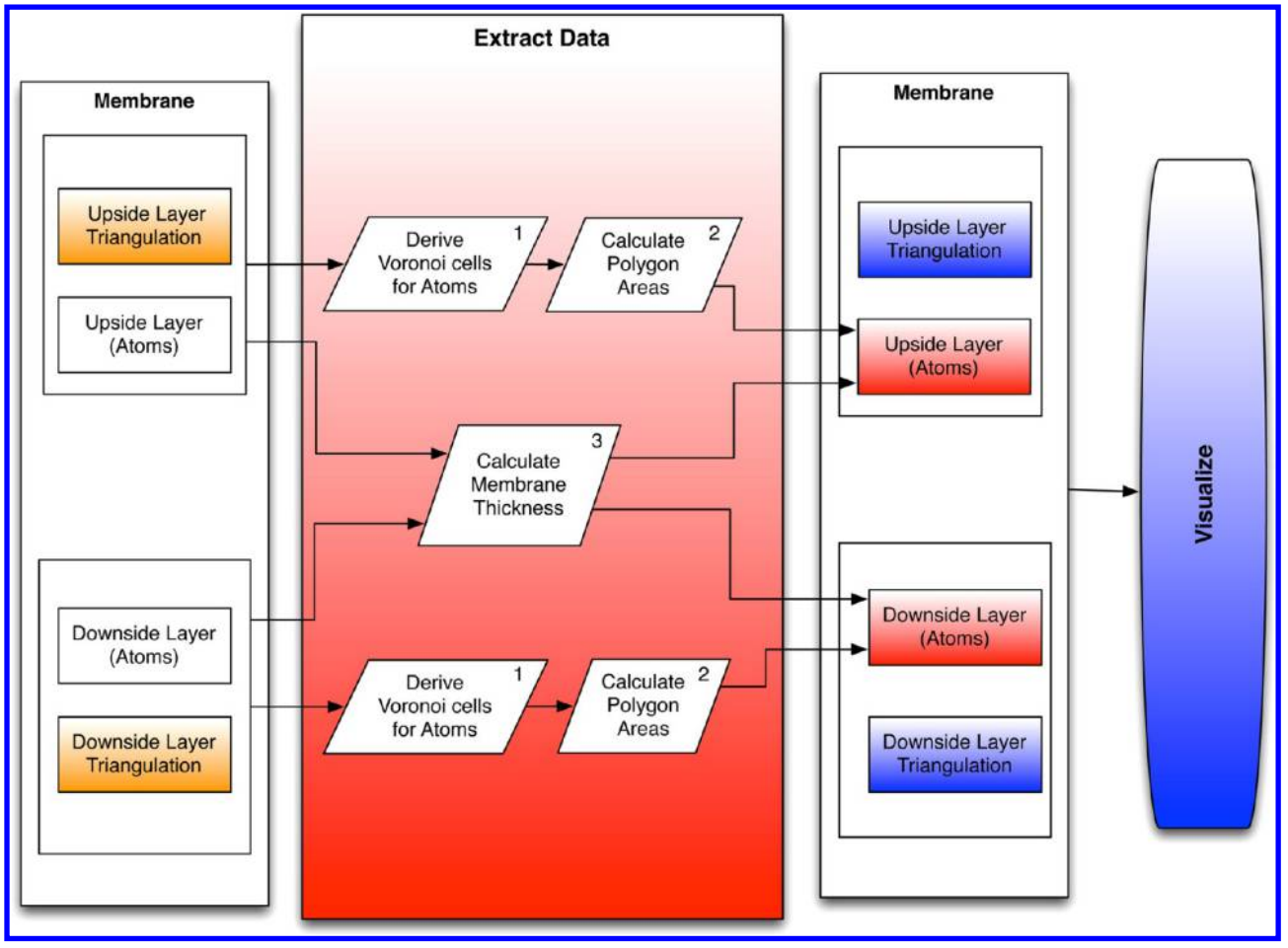


Figure 6. Flowchart of the internal logic of APL@Voro in the program state where the membrane thickness and the area per lipid is calculated. Rectangles mark internally stored data structures. Parallelograms mark program states, where calculation is carried out. The calculation of the area per lipid is performed for every layer separately and follows the methods described in section Calculating the projected area per lipid. The membrane thickness is calculated using both layers at the same time and is carried out following the descriptions in section Calculating the Membrane Thickness.

traingulation and the Voronoi diagram. The Quad-Edge data structure was described by Guibas and Stolfi.[35]] The workflow illustrated in Figure 5 is separately carried out for every layer. The minimum workflow (without protein atoms) is marked with red arrows. The simulation box vectors are used to generate an initial Quad-Edge data structure (Figure 5.1a). This initial triangulation is expanded with the preprocessed lipid atom coordinates (Figure 5.1b) using the "Insert Atom in Triangulation" procedure (described in section Insert Atom in Triangulation Procedure). This procedure is executed for every atom stored in the Layer data structure such that the initial triangulation is iteratively expanded this way.

In the case that no protein atoms have been selected (Figure 5.3a), either the next layer is triangulated using the same procedure or a next membrane (a further frame) is triangulated (Figure 5.4). In the case where a protein was selected, the triangulation of lipid atoms (Figure 5.3) is used together with the protein atoms to perform the TPIM test (Figure 5.3a). The TPIM test reduces the number of stored protein atoms so that only relevant protein atoms remain stored (see section Triangular Prism Insertion Method (TPIM)). These remaining protein atoms are then inserted in the triangulation of lipid atoms using the same "Insert Atom in Triangulation" procedure, resulting in a triangulation of lipid and protein atoms. After the layers in all frames are triangulated, the projected area per lipid and the membrane thickness are calculated.

3.2.1. Insert Atom in Triangulation Procedure. The basic procedure of inserting atoms into the Quad-Edge data structure is used for every lipid atom and every protein atom that passed the TPIM test (Figure 5, brown parallelogram). The procedure takes a triangulation (represented in the Quad-Edge data structure), an atom, and the PBC-offset value in percent entered in step 1 of the general workflow. Periodic boundaries of the simulation box are taken into account for the X/Y dimension. For efficiency, only parts of the system are treated as periodic images. A periodic image is created for every atom lying inside a predefined distance d (defined by the PBC-offset value) to a face of the unit cell. By default, the distance d is set to a length of 20% of the system's X/Y dimension but can be altered by the user on startup.

The procedure starts with the projection of the atom coordinates onto the XY plane ($Z = 0$). If the atom's XY coordinate does not fall into the PBC-offset area, it is directly inserted into the Quad-Edge data structure. Otherwise, an additional periodic image of the atom is inserted together with the original coordinates. Periodic images are taken into account

dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925

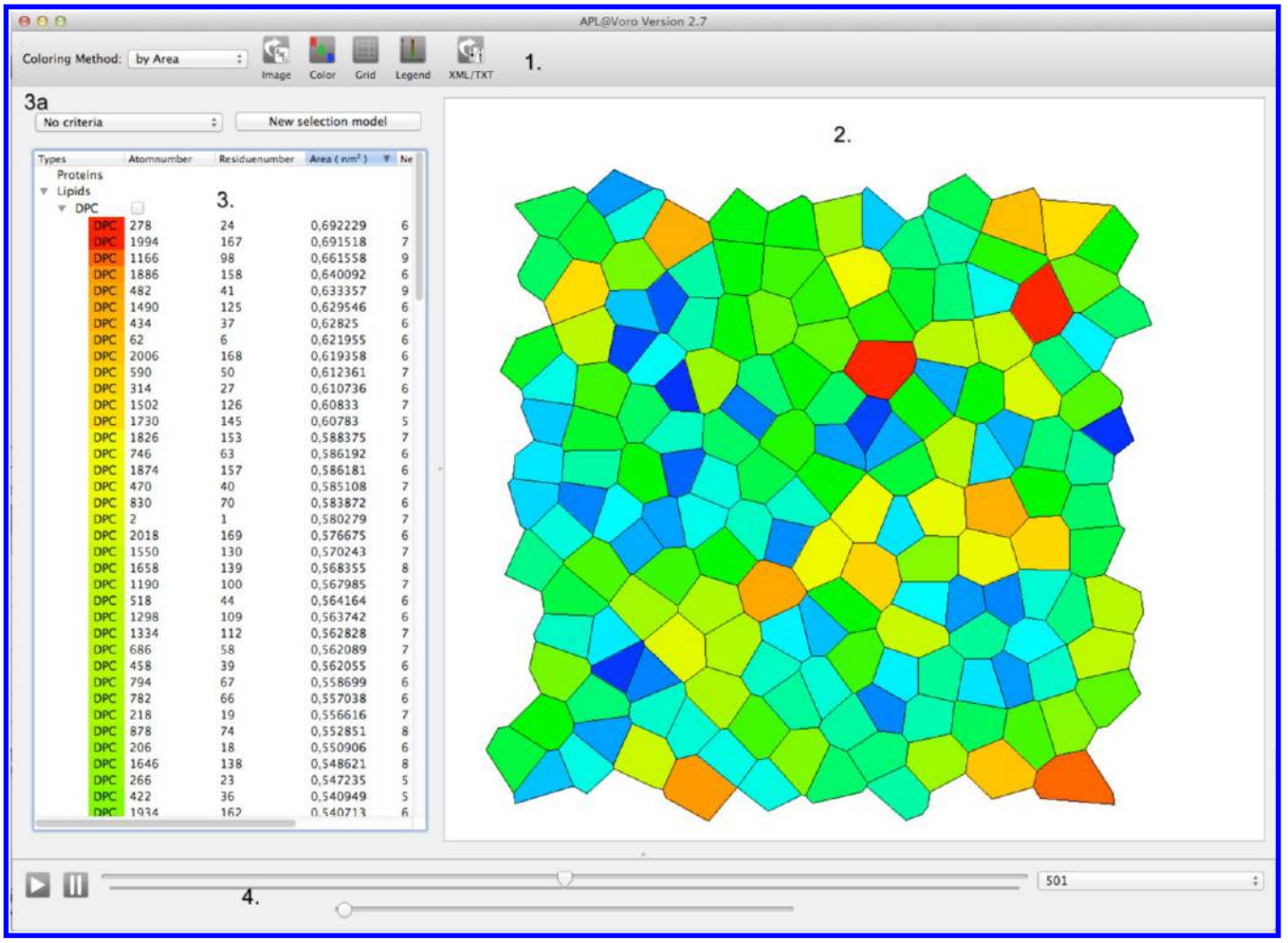


Figure 7. Screenshot of APL@Voro's main window. The Voronoi cells in the Voronoi view (2) are colored using the "by Area" coloring option (see section Colored Voronoi Diagrams). The entries in the tree view (3) are sorted by their calculated area in descending order. Frame number 501 of 1000 frames is selected in this view.

for step four (Extract Data) of the general workflow but are not shown to the user in the final visualization procedures.

3.3. Extracting Data. The "Extract Data" step is the last program step executed before the data are visualized. A flowchart of this step is illustrated in Figure 6. The procedure uses the triangulations of the single layers together with the information stored for every atom to calculate (i) the projected area per lipid and (ii) the membrane thickness. To calculate the projected area per lipid, a corresponding Voronoi cell is derived from the triangulation (Figure 6.1). The projected area per lipid is then calculated as described in the section Calculating the Projected Area Per Lipid (Figure 6.2). Finally, the calculated area is stored.

The calculation of the membrane thickness is done using either the Voronoi or the triangulation method as described in section Calculating the Membrane Thickness. However, two triangulations are needed for this function: one triangulation of the upper layer and one for the lower layer. The thickness values are calculated for every atom stored in the respective layer (Figure 6.3). For the upper layer, the stored atoms are projected onto the triangulation of the lower layer and vice versa. If the method proposed by Pandit et al. (Voronoi method) was selected on startup, the triangulations are converted to Voronoi diagrams before the projection takes place.

3.4. User Interfaces and Visualizing Data. We decided to use a GUI for two main reasons. The first one is the presentation of the colored Voronoi diagrams for every frame in a simple environment. The second reason is to take advantage of the interactivity a GUI-based tool can provide. Two-dimensional and three-dimensional plots are used as interactive elements and are combined with the presented Voronoi diagrams that act as interactive elements themselves. In the following sections, the user interface, the plotting functionality, the interactive Voronoi diagrams, and the export functions of APL@Voro are described.

3.4.1. The Main Window. The main window of APL@Voro is used to display colored Voronoi diagrams for the upper layer, manage selections of Voronoi cells, control the coloring and the displayed elements, and change the displayed frames. Voronoi diagrams for the lower layer can be viewed in a separate window similarly structured to the main window. The main window is divided into four segments: a movable menu bar (Figure 7.1), a Voronoi view (Figure 7.2), a tree view (Figure 7.3) with a menu to define a selection criteria (Figure 7.3a), and a controller area (Figure 7.4). The menu bar provides functions for changing the coloring method, exporting images and ASCII text files (.txt and .xml). The tree view holds contents of the actual frame and is synchronized with the displayed elements in the respective Voronoi view. The

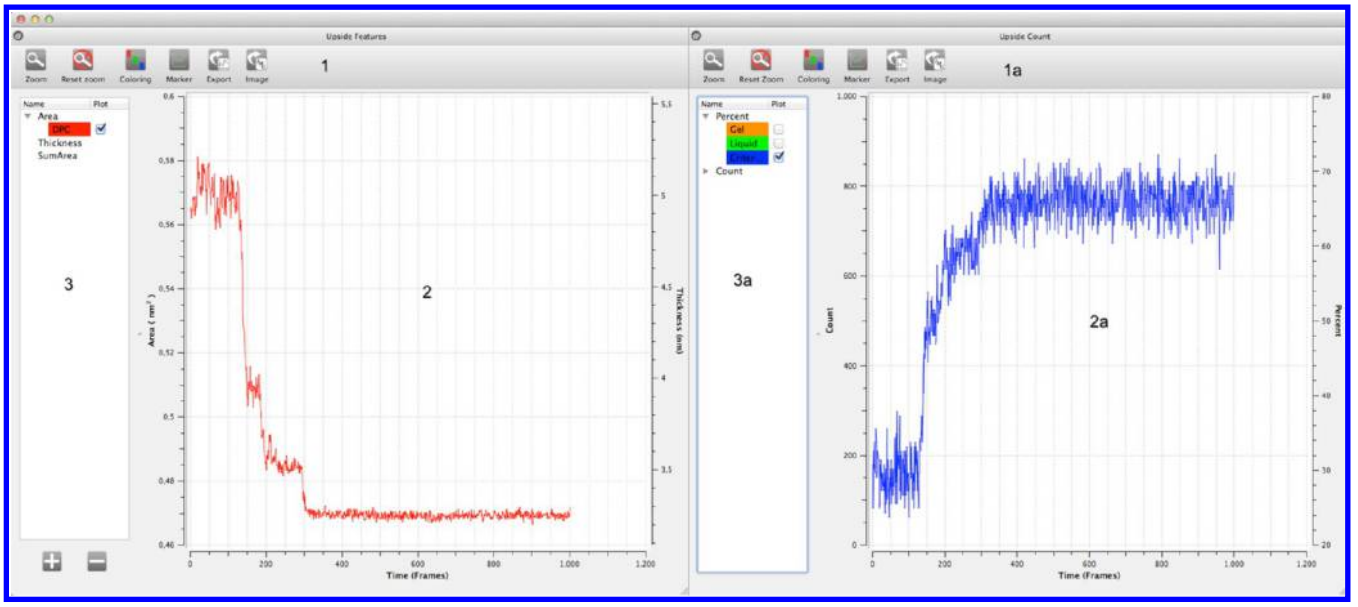


Figure 8. Screenshot of APL@Voro's plotting windows for two-dimensional plots. The left side plotting window shows the "Upside Features" plotting window, and the right side shows the "Upside Count" plotting window. Every plotting window contains a movable menu bar (1, 1a), a plot select area (3, 3a), and the actual plotting area (2, 2a). Using the plotting windows, several plots can be created, colored, and exported as an image or ASCII text file.

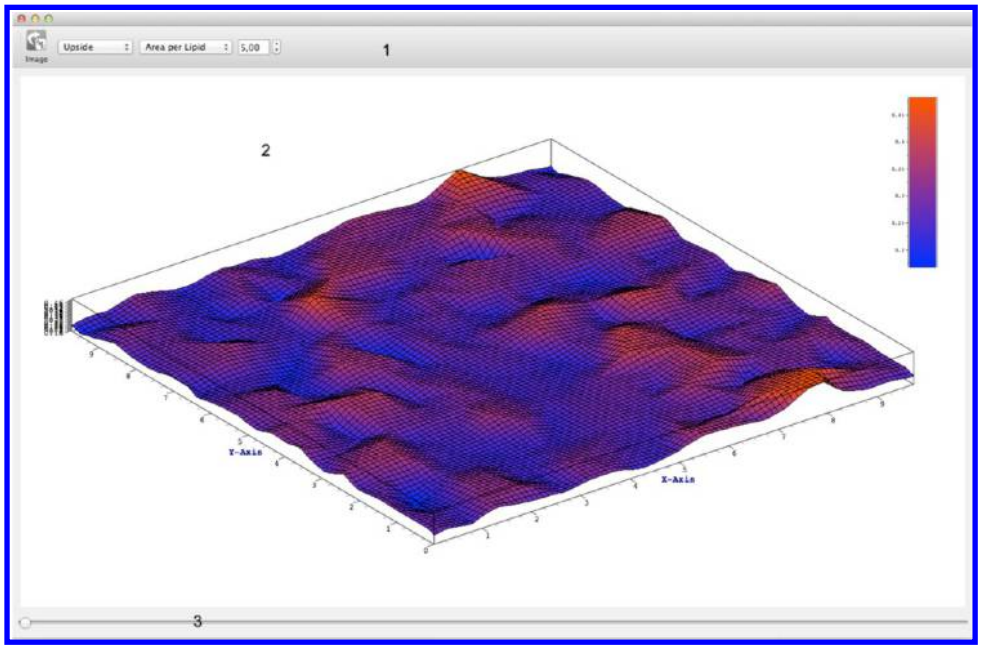


Figure 9. Screenshot of APL@Voro's plotting window for three-dimensional plots. Different plots can be selected and scaled in the movable menu bar (1). Every plot is displayed together with a legend in the plotting area (2). The grid resolution of the surface plot can be changed using the control slider at the bottom of the window (3).

displayed Voronoi diagrams act as interactive elements so that single Voronoi cells or groups of Voronoi cells can be selected in the voronoi view. The displayed frame of the trajectory can be changed in the controller area while keeping a chosen selection.

3.4.2. Colored Voronoi Diagrams. The Voronoi diagrams presented in the Voronoi view can be colored according to their corresponding atoms and features. This is controlled via a selectable coloring method accessible in the drop down menu in the main window's menu bar. Possible selections are "None" (the Voronoi cells will appear in the background color (default white)), "by Residue," "by Area," "by Neighbors," and "by Thickness." Using the coloring method "by Area," "by Thickness," or "by Neighbors" will color the Voronoi cells and the corresponding entries in the tree view with an interpolated RGB color ranging from dark blue for the minimum value (e.g., the minimum area) to red for the maximum value (e.g., the maximum area). The colors in the Voronoi view and the tree view will be updated whenever a new frame is selected and displayed.

3.4.3. Selecting Voronoi Cells. Voronoi cells in the Voronoi view can be selected in different ways and for different reasons.

One reason would be to follow the path a lipid (or Voronoi cell) takes through the entire trajectory or to highlight a specific cell or a group of cells.

In addition to the simple selection via mouse, APL@Voro supports an automatic selection based on user selected criteria. For this, a selection model is created by the user and applied to the layer. A selection model can be created by clicking "New selection model" in the tree view of the main window (Figure 7.3a). A dialogue window is used to specify certain criteria for the automatic selection. Possible criteria are "Residue," "Area," "Thickness," and "Neighbors." Every selected criterion is combined with either the term "is" or the term "is not." The automatic selection will select every Voronoi cell that matches the created criterion. [For example, the criterion "is" and "Residue" and "DPC" will select every Voronoi cell that corresponds to a DPC labeled atom, the criterion "is not," "Area < (smaller than)," and "0.5" will select every Voronoi cell with a calculated area ≤0.5.]

3.4.4. 2D and 3D Plotting. Upon startup, the two-dimensional plotting interface displays two plotting windows for each layer side; a plotting window showing plots as a function of time for calculated values such as the area per lipid ("Upside Features") and another plotting window for the count of Voronoi cells selected by a selection model for the every layer ("Upside Count"). Figure 8 shows the two plotting windows "Upside Features" on the left and "Upside Count" on the right for the upper layer. The plotting windows for the upper and lower layer provide the same functionality, and every plotting window can be used independently from the others. The plotting areas of the "Upside Features" window and the "Upside Count" window differ by the labeling of their *Y* axes. The *Y* axes in the plotting area of the "Upside Features" window (Figure 8.2) refer to the calculated average area in nm^2 on the left and the calculated thickness in nm on the right. The *Y* axes in the plotting area of the "criteria-plot window" (Figure 8.2a) refer to the count in numbers of the selected Voronoi cells on the left and percentage of selected Voronoi cells on the right. While the plots for the criteria-plot window are generated automatically, the plots for the "feature-plot" window must be created using a dialogue. Three different types of plots can be created. A plot of the average area per lipid, a plot of the sum of the area per lipid, and a plot of the membrane thickness. It is possible to extract and display the values of present Voronoi cell selection.

The interface for three-dimensional graphs (Figure 9) supports two different types of graphs for every layer of the membrane, one graph for the area per lipid and one for the thickness. The three-dimensional graphs are updated whenever a new frame of the trajectory is chosen in the main window. The *X* and *Y* axes of the 3D plot refer to the dimension of the layer in nanometers. The displayed graph can be rotated with the mouse.

3.4.5. Exporting Images and ASCII Text Files. Images can be saved for two-dimensional plots, three-dimensional plots, and the Voronoi view in the main window. The file save dialogue is accessible via the menu bar of the respective windows. The supported image formats vary depending on the operating system and the installed libraries. In general, the .png, .jpg, and .tiff formats are supported to export high resolution images.

Two different ASCII text file formats can be saved for displayed Voronoi diagrams: a .txt file and an .xml file. The .txt file holds one line for every atom present in the actual displayed layer. Every line holds the following space-separated values: atom number (as printed in the PDB file), atom name (as printed in the PDB file), residue name (as printed in the PDB file), residue number (as printed in the PDB file), *x* coordinate (in nm), *y* coordinate (in nm), *z* coordinate (in nm), time (in ps), calculated area (in nm^2), and the calculated thickness at this point (in nm). The .xml file holds the same content as the .txt file. Further information regarding the structure of the exported .xml files is available in the Supporting Information.

Two-dimensional plots can be exported as an ASCII text file. The exported file uses the .xvg extension (supported by Grace/XmGrace, http://plasma-gate.weizmann.ac.il/Grace/) and contains a line and space-separated data. Every line contains two space-separated entries. The first entry is the frame number (*x* value), and the second entry is the respective data (area, thickness). ASCII text file export for three-dimensional graphs is not supported in the current version of APL@Voro.

4. VALIDATION SIMULATIONS

The membrane models for systems M1 and M2 were generated using the CELLmicrocosmos MembraneEditor 2.2.1.[36] Using its Molecule Editor, the MARTINI-based 4-to-1 mapping was realized. For this purpose, all-atom PDB models were simplified according to the scheme discussed by Marrink et al.[37] For the membrane packing, the Distributor algorithm was chosen (Random Seed: 220). All coarse-grained simulations were carried out by GROMACS 4.5.5 using the MARTINI force field for coarse-grained lipid models.[32,37]

4.1. Simple Bilayer (System M1). The modeled DPPC bilayer system M1 is a single-component bilayer based on the membrane presented in Marrink et al.[29] It contains 338 DPPC lipids with a layer size of 10 nm × 10 nm. The bilayer was solvated in 4559 MARTINI waters, resulting in a size of 10 nm × 10 nm × 10 nm. It showed domain formation as a result of phase separation of DPPC. During the simulation, the bilayer was cooled down from 338 K to 285 K (≈ 10 K below the MARTINI DPPC transition temperature)[29] within 3 μs using the Berendsen thermostat.[38] The simulation time was set to 5 μs; the first and last microseconds were used as equilibration time. After reaching the target temperature of 285 K, nearly the whole bilayer transformed in the timespan of ∼170 ns from the L_α phase to the L_β phase. During the 170 ns of transformation starting at t = 4130 ns, the area per lipid dropped from 0.56 nm^2 ± 0.08 nm^2 to 0.47 nm^2 ± 0.02 nm^2 at t = 4305 ns, which is close to the experimental measured value for L_β phase DPPC (4.6 nm^2).[39] While the average area per lipid dropped, the membrane thickness raised from ∼4.3 nm to ∼4.8 nm.

4.2. Mixed Bilayer (System M2). System M2 contains 576 cholesterol, 828 DPPC (diC_{16}−PC), and 540 DUPC ($diC_{18:2}$−PC) molecules with a layer size of 21 nm × 21 nm. It is based on the work of Risselada and Marrink.[40] The bilayer was solvated with 9450 MARTINI waters, resulting in a 21 nm × 21 nm × 10 nm simulation box. During the 5 μs simulation, the bilayer was held at 295 K using the Berendsen thermostat.[38] Already well before 100 ns, a separation of DUPC from DPPC and cholesterol as a formation of small clusters of DPPC and cholesterol was observed. During the following simulation time, the small clusters merged to form larger clusters resulting in a single-stripe formed DPPC and cholesterol cluster at t = 4000 ns. The large cluster remained stable until the end of the simulation at t = 5000 ns.

4.3. Bilayer with Peptide (System M3). System M3 was originally created to evaluate the stability of pentameric models of the Vpu pore.[31] It consists of five peptides, 112 POPC, and

dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925

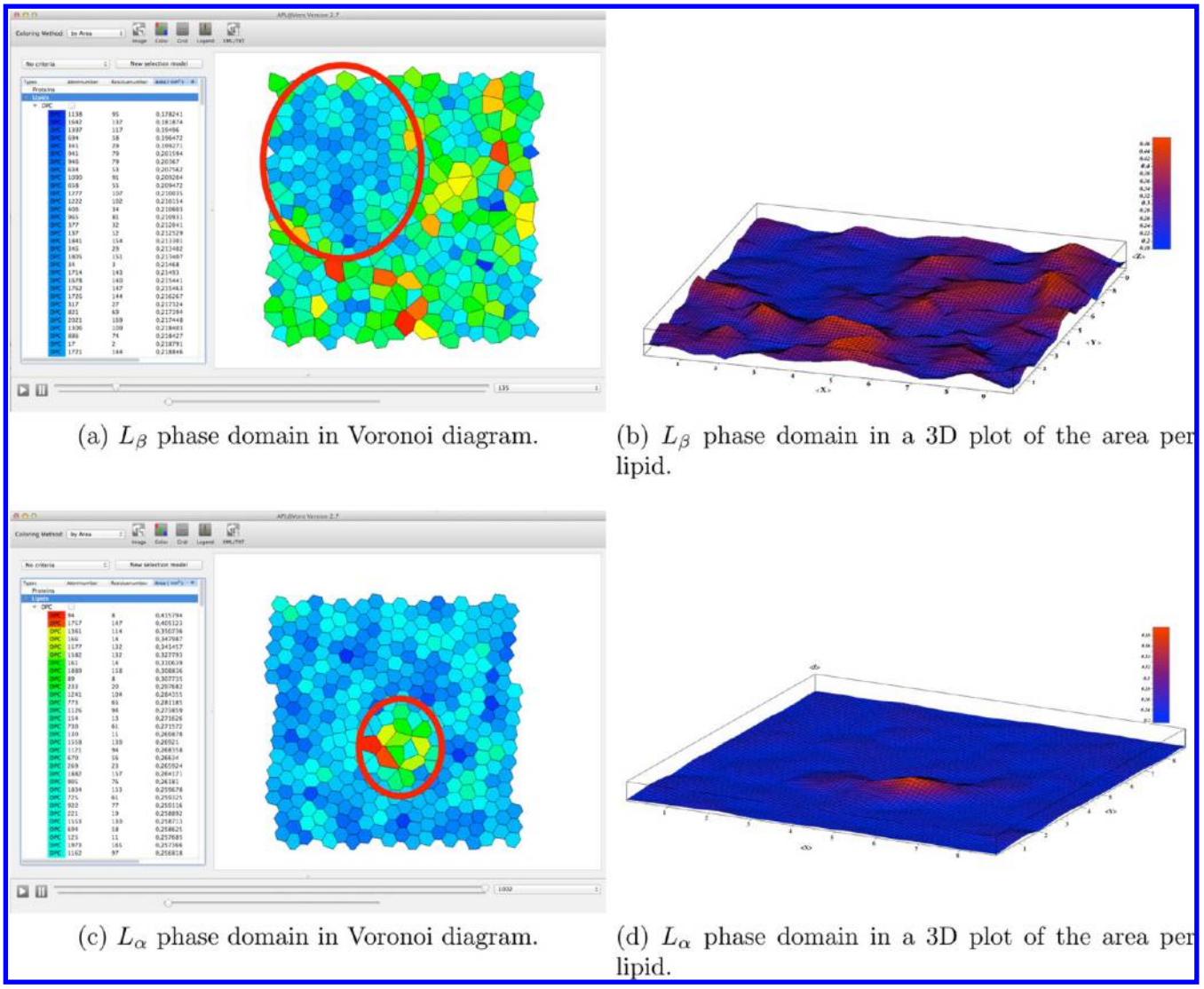


Figure 10. Screenshots of APL@Voro showing colored Voronoi diagrams (a, c) and 3D plots (b, d) of the upside layer of the system M1. Parts a and b refer to the simulation step at t = 4134 ns at a temperature of 285 K. Parts c and d refer to the simulation step at t = 5000 ns at a temperature of 285 K. The formed lipid domains are marked with a red circle in a and c. The domain of L_β phase lipids in a appears as a cluster of nearly perfectly formed hexagons colored in light and dark blue, while the domain of L_α phase lipids in c appears as a cluster of randomly formed Voronoi cells with a higher area per lipid than the surrounding L_β phase lipids. In parts b and d, the X and Y axes refer to the dimension of the layer in nm. The Z axis refers to the calculated area per Voronoi cell in nm^2 (note that area values need to be doubled to take the multiatom selection into account). The higher area appears as a higher z value and the lower area as a lower z value. In part b, the domain of the L_β phase lipids appears as a valley on the upper left corner of the plot, while the domain of the L_α phase lipids in part c appears as a hill near the center of the layer. (b, d) The 3D grid was interpolated from a Delaunay triangulation for selected atoms of type C_2. (a, c) The Voronoi diagrams were constructed for the C_2 sites of the MARTINI DPPC model. Colors are interpolated using linear interpolation between the actual maximum cell area colored red and the actual minimum cell area colored blue. For further information about system M1, see the section Validation Simulations.

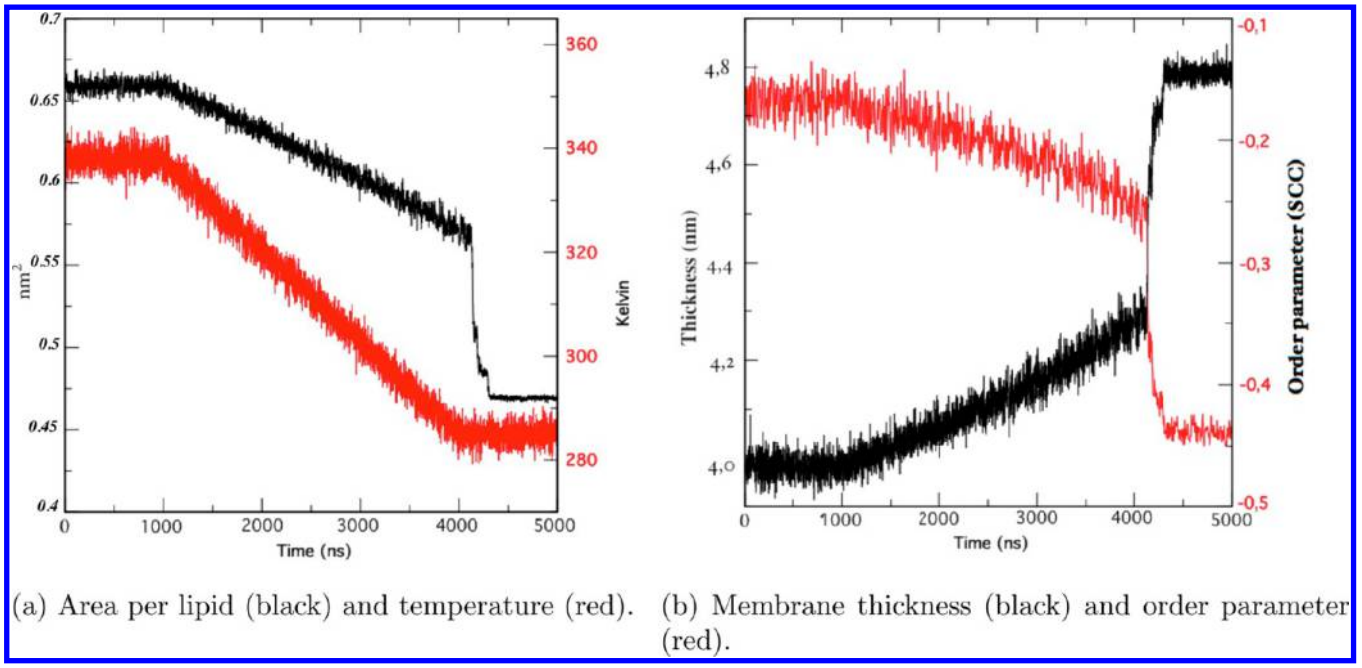
6453 waters using an "all-atom" force field. The topology for the lipids [POPC (16:0−18:1 diester PC, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine)] was created on the basis of the parameters of Chandrasekhar et al.[41] In order to reduce stress induced by lateral pressure fluctuations, the system was simulated with constraints for 1 ns using surface tension pressure coupling with a tension of 37.5 mN/m. The production run of the unconstrained system was carried out for 50 ns with semi-isotropic pressure coupling applying 1 bar in the z direction while keeping the xy plane fixed. The temperature of the peptide, lipid, and the water molecules were separately coupled to a Berendsen thermostat at 310 K with a coupling time of 0.1 ps. The pore remained stable throughout the simulation showing RMSD values between 0.20 and 0.35 nm. The system M3 is already published, and all details are available through Krüger and Fischer.[31]

5. RESULTS AND DISCUSSION

5.1. Simple Bilayer. To observe the structural changes of the bilayer in system M1 during the process of phase transformation from the liquid to the gel phase, we used the *chain model* with selected C2 tail beads. Using this selection model, the colored Voronoi diagrams presented by our tool can be used to identify the formation of a critical cluster of gel lipids (the starting point of the phase transformation)[29] by the color and structure of the single Voronoi cells (see Figure 10a).

The corresponding 3D area per lipid plot (Figure 10b) displays the cluster of gel lipids as a large valley colored in deep

dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925



(a) Area per lipid (black) and temperature (red). (b) Membrane thickness (black) and order parameter (red).

Figure 11. The process of phase transformation expressed (a) in a drop of area per lipid accompanied by (b) increasing chain ordering and membrane thickness. (a) Red graph: Temperature of the system over time in Kelvin. Black graph: The area per lipid measured by APL@Voro as the area of the Voronoi cells for selected sites of type PO_4. (b) Red graph: Second rank order parameter of the SN-1 Chain (C1A-C2A-C3A).[42] Black graph: Membrane thickness measured by APL@Voro with selected sites of type PO_4. The massive drop of the area per lipid (a) followed by nearly constant area suggests a complete transformation from a L_α phase bilayer to a L_β phase bilayer. The increasing chain ordering and membrane thickness (b) reinforce the idea of a complete transformation. Cluster formation during the transformation process can be assumed but neither proven nor located by this kind of plot.

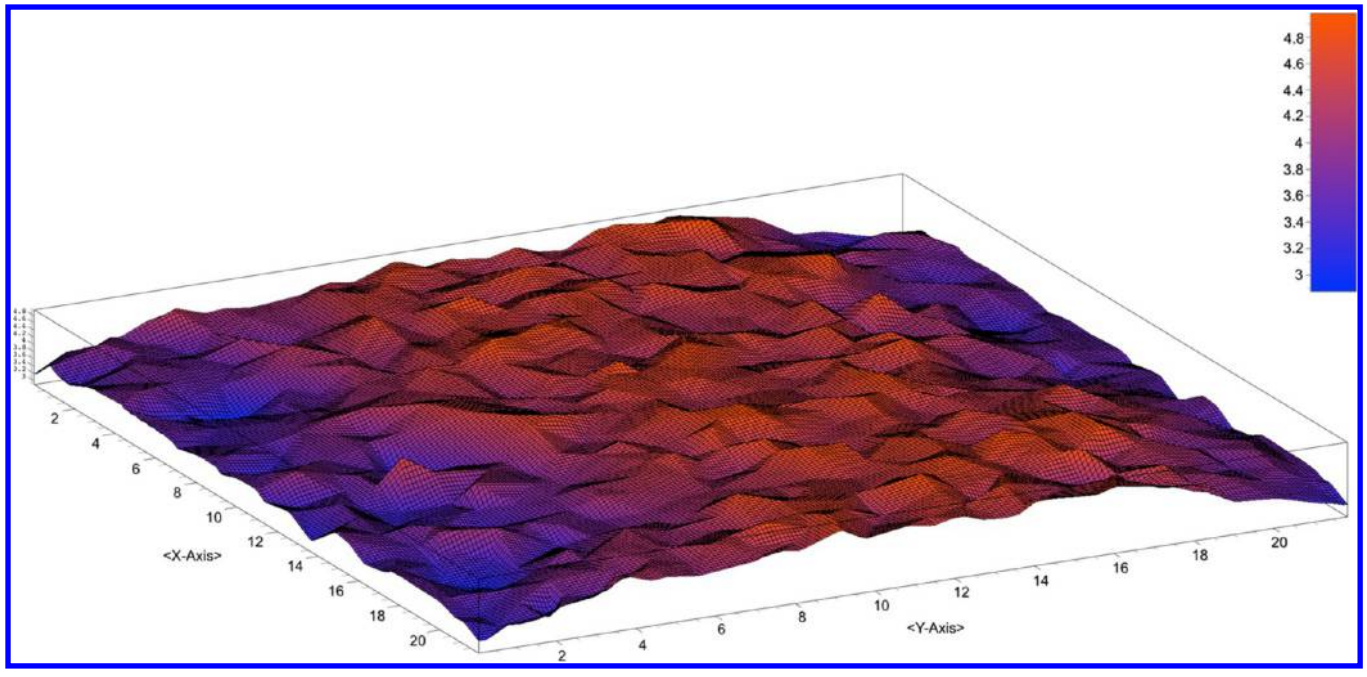


Figure 12. Surface plot of the bilayer thickness of the 0.42:0.28:0.3 DPPC/DUPC/cholesterol mixture at t = 5000 ns. X/Y axes mark the layer dimension in nanometers. The Z axes refer to the calculated bilayer thickness. The increased bilayer thickness from ~3.8 nm at the DUPC rich areas (blue/purple) until ~4.5 nm at the DPPC/cholesterol rich centered stripe is displayed as a hill (orange/red).

blue, while the L_α phase area appears undulating. Both Figure 10a and b state the same facts, a lipid domain with low area per lipid. The data are merely displayed in different ways and based on different subdivisions.

In order to identify the domain formation, it is possible to plot the change in area per lipid or membrane thickness over the simulation time. These plots can give just a hint of domain existence, however, on neither the domain's location nor the domain's dimensions. The change in area per lipid is shown in Figure 11a and displays a massive drop from 0.56 nm^2 ± 0.08 nm^2 at t = 4134 ns to 0.47 nm^2 ± 0.02 nm^2 at t = 4305 ns. The nearly constant area per lipid in the remaining time span of

dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925

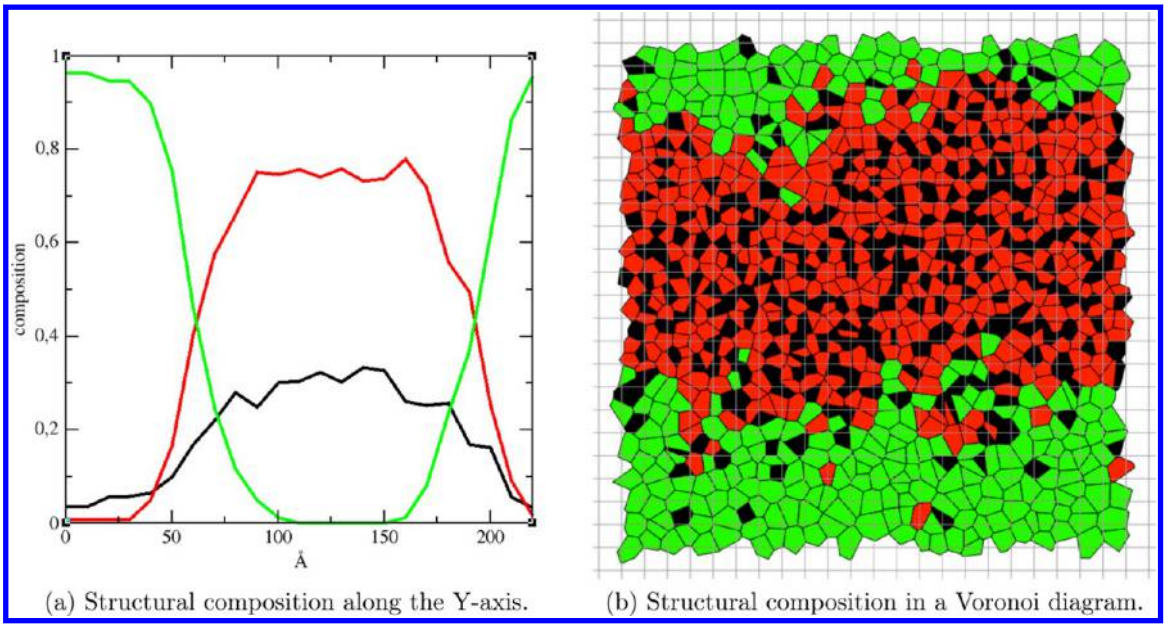


Figure 13. Structural composition of the 0.42:0.28:0.3 DPPC/DUPC/cholesterol mixture at t = 5000 ns at a temperature of 295 K. Both figures use the same color code with red = DPPC, green = DUPC, and black = cholesterol. (a) To measure the change in composition, the bilayer was divided into 1 nm wide stripes along the X axis. For every stripe, the proportion of the components was measured. (b) A residue-type-colored Voronoi diagram exported from APL@Voro with selected sites of type PO_4 for PC lipids and ROH for cholesterol. Both figures display the same circumstances. The formerly randomly distributed lipids appear more organized and divided into three stripes. A central area rich in cholesterol and DPPC and two flanking areas rich in DUPC. While part a gives direct access to the percentual values, part b displays the details of diffuse domain borders in a more intuitive way.

~700 ns suggests a full transformation of the former L_α phase bilayer to a L_β phase bilayer. The same applies to the measured order parameter and the membrane thickness (Figure 11b). While Figure 10a and b show a domain formation during the transformation process, the Figure 10c and d reveal a cluster of remaining L_α phase lipids in a sea of L_β phase lipids. At t = 5000 ns, the small cluster with a size of ~3.0 nm^2 of seven L_α lipids remain in a sea of L_β phase lipids. This L_α phase domain is distinctly marked with a higher average area per lipid of ~0.43 nm^2 and a lower average thickness of ~4.8 nm than the surrounding L_β lipids. The colored Voronoi diagram and the three-dimensional plot of the area per lipid at t = 5000 ns provide direct access to the state of the layer. The formerly suggested conclusion of a complete phase transformation can be revised with this enhanced view of the local properties.

5.2. Mixed Bilayer. The selection of certain atoms from the lipid chains showed good results for the structural analysis of the simple system M1. For more complex systems like system M2, we propose to use different selection models depending on the data one wishes to derive.

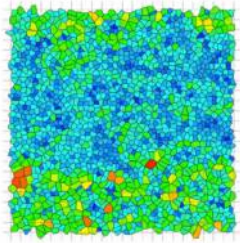
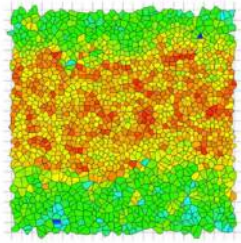
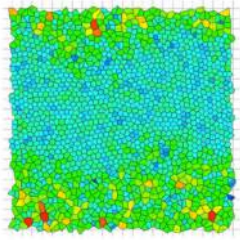
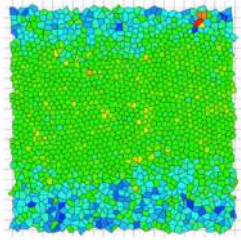
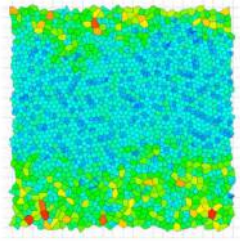
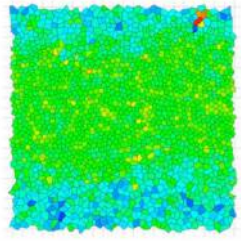
A construction of Voronoi diagrams—for the purpose of following the lipid type separation—can be done with the *simple model* for key atom selection (one atom per lipid). At the end of the simulation of system M2, the membrane appeared as a striped pattern with a high percentage of cholesterol of ~39% and DPPC located in a centered ~6-nm-wide stripe. Such structural changes can easily be analyzed with the *simple model* and one key atom per lipid (Figure 13).

To cover the embedded position of cholesterol for the calculation of the area per lipid, several key atom selections were tested. Table 1 shows the colored Voronoi diagrams generated for system M2 at t = 5000 ns using different selection models. The Voronoi cells are colored using linear interpolation between the minimum observed value (area, thickness) colored in deep blue and the respective maximum observed value colored in red (see section Colored Voronoi Diagrams for more details about the coloring methods). A good visualization of the area per lipid distribution and the membrane thickness is achieved with the *boundary model*, the *chain model*, and the *maximum-density model*. The visualizations using the *chain model* and the *maximum-density model* state the correlation of the lipids chain ordering to the features (area, thickness) colored in the diagrams (Table 1c). Voronoi cells referring to stretched and straight lipid chains appear in a honeycomb-like form (forming a perfect hexagon). On one hand, higher chain order gives rise to smaller areas per molecule (colored in light blue), and on the other hand, a higher chain order leads to an increase in membrane thickness (colored in green). Given the visualized features, the user may deduct the area per lipid, the chain order, and the membrane thickness from one Voronoi diagram displayed in our tool, simply by changing the coloring style.

To compare the results of the Voronoi method with the areal-determining plane method,[23] a ~21 nm × 6 nm wide stripe of the domain completely depleted of DUPC was selected. The calculated areas for this region using different selection models are presented in Table 2. The calculated area for DPPC using the area-determining method with a cholesterol tilt angle of 14.7° for the given cholesterol percentage of ~39% and a cholesterol cross-sectional area of 0.38 nm^2 did not reach the value below 0.4 nm^2 reported by Wennberg et al.[30] Although the calculated areas using the *maximum-density model* do not exactly match the values calculated with the area-determining method, the expected high area for cholesterol and low area for DPPC is reflected. Independent from the method used, the condensing effect of cholesterol is mirrored by the low area per DPPC inside the

dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925

Table 1. Colored Voronoi Diagrams of the 0.42:0.28:0.3 diC_{16}−PC/$diC_{18:2}$−PC/cholesterol Mixture at t = 5000 ns (System M2) at a Temperature of 295 K with Different Key Atom Selection Models That Can Be Used to Display the Area Per Lipid and Membrane Thickness[a]

Selection	Colored by area	Colored by thickness
(a) *boundary model* PC lipids: *GL1 GL2* Cholesterol: ROH		
(b) *chain model* PC lipids: *C2 D2* Cholesterol: R_5		
(c) *maximum − density model* PC lipids: *C2 D2* Cholesterol: *R4 R5*		

[a](a) The *boundary model* shows good visualization results for the area per lipid and the membrane thickness. (b) Additional information about the phase state is provided by the shape of the Voronoi cells. (c) The *maximum-density model* visualizes the expected lower area per cholesterol.

domain and the high area per DPPC outside the domain of ∼0.6 nm^2.

The calculation of the area per phospholipid using the method proposed by Alwarawrah et al. assumes that the area not assigned to cholesterol can be assigned to the phospholipids.[23] For three component bilayers, it is unclear how to distribute the area to the single phospholipid types present in the membrane. Using the Voronoi method, one has access to every single lipid, and therefore individual areas can be calculated. Our tool can show the change in area per lipid over time directly as a graph in a separate window. In Figure 14, the calculated area per lipid using the *maximum-density model* is shown. The calculated average areas per lipid for DPPC of ∼0.48 nm^2 and ∼0.66 nm^2 for DUPC follow the assumption of an existing L_o domain. Because the selection model included two sites for each cholesterol and PC lipid, the graphs show half the area of these.

Accompanying the change in the area per lipid, the raft-like domain discriminates in the measured local thickness. The *simple model* selection can be used to create an intuitive three-dimensional presentation of the local membrane thickness (Figure 12). The calculated local thickness shows a clear accretion in thickness toward the domain from ∼3.8 nm at the DUPC rich areas to ∼4.5 nm in the DPPC/cholesterol rich domain. This follows the proposed sign of L_o domains which appear with a higher local thickness than the surrounding area of lipids.[43]

5.3. Bilayer with Peptide. APL@Voro and the TPIM procedure were tested for membranes holding proteins using Vpu_{1-32}WT embedded in a POPC bilayer (system M3). During the simulation time of 50 ns, the temperature of system M3 was held at 310 K. A multigroup selection was done using one group for every monomer. For the POPC lipids, the phosphorus atom was chosen as reference atom for the tool (simple model). The TPIM procedure was used together with a van der Waals radius of 0.18 nm for the phosphorus atom. In Figure 15, a Voronoi diagram of the first frame of the simulation is shown together with a traditional rendering for comparison reasons. The calculated average area per lipid of this frame was 0.65 nm^2 for the top leaflet and 0.63 nm^2 for the bottom leaflet, with a bilayer thickness of ∼3.15 nm. Over the 50 ns trajectory, the average area per lipid for the top and

Table 2. Area per DPPC Lipid A_{DPPC} and Area Per Cholesterol A_{Chol} Calculated with Different Approaches for a ~21 nm × 6 nm Wide Stripe of DPPC/Cholesterol in the 0.42:0.28:0.3 DPPC/DUPC/cholesterol Mixture at t = 5000 ns[a]

method	A_{Chol}	A_{DPPC}
boundary model	~0.273 nm^2	~0.494 nm^2
key atoms:		
DPPC: GL1, GL2		
Chol: ROH		
chain model	~0.268 nm^2	~0.492 nm^2
key atoms:		
DPPC: C2A, C2B		
Chol: R5		
maximum−density model	~0.324 nm^2	~0.456 nm^2
Key atoms:		
DPPC: C2A, C2B		
Chol: R4, R5		
area-determining plane[23] (see section 2.2)	~0.393 nm^2	~0.421 nm^2

[a]The area-determining plane method (row 4) was used with a cholesterol tilt angle of 14.7° and assuming a cholesterol cross sectional area of 0.38 nm^2.

bottom leaflet was 0.628 nm^2 and 0.625 nm^2, respectively. Together with the calculated bilayer thickness of ~3.25 nm, the L_α phase is proposed. The reported area per lipid for POPC in the L_α phase for experiments at 310 K varied between 0.63 nm^2 and 0.68 nm^2.[2,44−47] The calculated area per lipid for M3 is slightly below the reported reference minimum, which is ascribed to the overlap of protein atoms and lipids as a result of projecting atoms onto an *XY* plane.[3,28] Although APL@Voro is not designed for protein analysis, it supports plotting of the area or thickness of selected protein groups, because the calculation of the Voronoi cell area or the bilayer thickness does not differentiate between protein atoms and lipid atoms. Each group and every inserted protein atom can be accessed in the same way as lipid atoms. This allows not only the calculation of the lateral area of the inserted Vpu_{1-32}WT peptide but also allows calculating the individual lateral area of the selected monomers. The average lateral area over the full trajectory for Vpu_{1-32}WT was ~10.6 nm^2 for the top and ~10.1 nm^2 for the bottom leaflet. The multigroup selection allows calculating the individual area for the monomers A, B, C, D, and E in the top and bottom layers to give a more detailed view of the layers structure: The calculated monomer areas were 2.29 nm^2 for A, 2.40 nm^2 for B, 1.67 nm^2 for C, 2.36 nm^2 for D, and 1.88 nm^2 for monomer E in the top layer and 1.85 nm^2 for A, 1.96 nm^2 for B, 2.22 nm^2 for C, 1.62 nm^2 for D, and 2.45 nm^2 for monomer E in the bottom layer (see Figure 15).

6. SUMMARY

In this work, APL@Voro was presented, a tool to support the process of analyzing membrane simulations. In summary, APL@Voro is capable of fast calculations of the area per lipid, the bilayer thickness, and in visualizing these data. Although this work was focused on biphasic systems, the methods presented in this manuscript are applicable to nearly all kinds of setups, no matter what thermodynamical states are covered in the simulations. APL@Voro is written in C++ and accessible to most of the Unix-like operating systems. It is available for free, is open source, and is published under the GPL3 license. To evaluate the software package, three membrane systems based on published data were used. For all three systems, the analysis with APL@Voro delivered useful additional information. The findings for system M1 and for system M2 have been reproduced in a suitable manner with the data derived from APL@Voro.[29,40] It was shown that our extension to the algorithm presented by Pandit et al. can be used to adequately derive the membrane thickness by avoiding ambiguous constellations. The problematic cases of the original algorithm described in the section Calculating the Membrane Thickness were solved to improve the calculation of the membrane thickness. Moreover, system M2 was used to evaluate three different selection models: the *boundary model*, the *chain model*, as well as the *maximum-density model*. It turned out that the feature of APL@Voro to be able to freely choose a key atom selection is of great use as different selections showed strengths

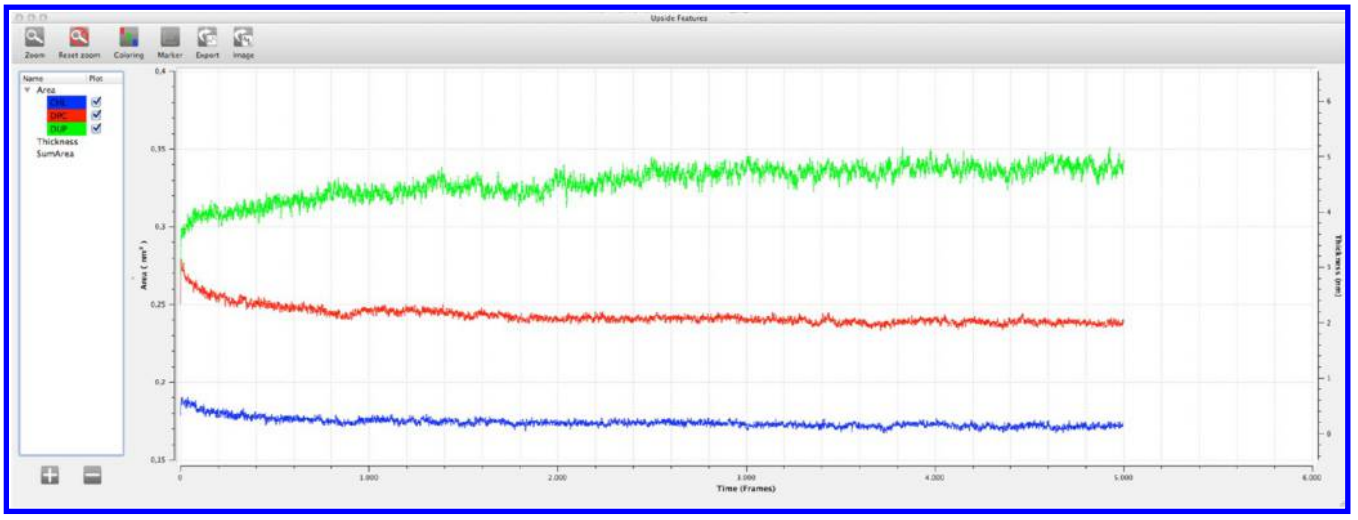


Figure 14. Screenshot of the change of the area per lipid in nm^2 over time as presented by APL@Voro for the *maximum-density model* with selected sites of types *C2A* and *C2B* for DPPC, *D2A* and *D2B* for DUPC, and *R4* and *R5* for cholesterol. The data are directly derived from the Voronoi diagrams created for the selection model. The values need to be doubled as the selection model used two sites for each lipid and cholesterol. Red graph: DPPC. Green graph: DUPC. Blue graph: cholesterol.

dx.doi.org/10.1021/ci400172g | J. Chem. Inf. Model. 2013, 53, 2908−2925

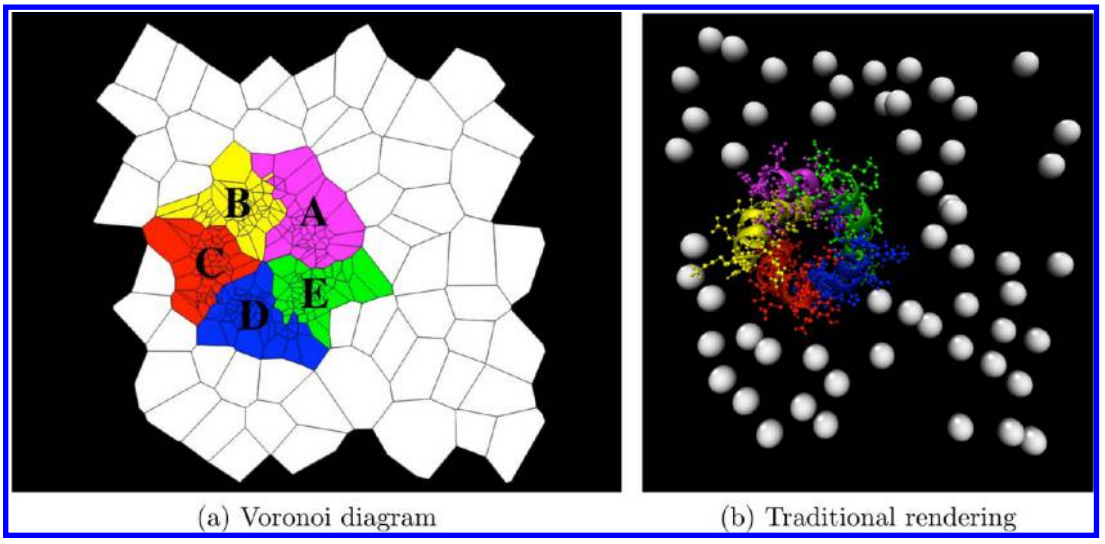


Figure 15. (a) A Voronoi diagram of the top leaflet of system M3 at t = 1 ns. The letters A, B, C, D, and E and the colors magenta, yellow, red, blue, and green indicate the monomers of Vpu_{1-32}WT. The white Voronoi cells refer to the lipids. The Voronoi diagram was constructed and rendered by APL@Voro using multigroup protein selection with one selection group for every monomer of Vpu_{1-32}WT and the phosphorus atom of the POPC lipid as lipid atom selection. The TPIM procedure with van der Waals radii was used to include protein atoms. (b) A traditional rendering of system M3 done with VMD. Every monomer of Vpu_{1-32}WT is colored as referred to in part a. The monomers of Vpu_{1-32}WT are rendered using the *NewCartoon* method for *α helices* and the *CPK method* for the remaining atoms. The POPC phosphorus atoms of the top leaflet are displayed as white spheres. Lipid tails, water molecules, and phosphorus atoms in the bottom leaflet are not displayed.

and weaknesses in various fields. For system M3, APL@Voro's *Triangular Prism Insertion Method* (TPIM) together with the ability to select certain groups of the embedded peptide provided a concise representation of the molecules in the system. The three-dimensional representation and colored Voronoi diagrams presented by APL@Voro can complement the existing toolsets given by the GROMACS package to support the analyzing process. Support for other MD simulation packages is currently not implemented, and coordinate files need to be converted to GROMACS format files. For further information about APL@Voro and download opportunities, please see www.aplvoro.org.

■ ASSOCIATED CONTENT

Supporting Information

Figures for section 3, Description of APL@Voro: the general workflow of APL@Voro, the internal logic during the project creation, internal logic during coordinate file processing, logic of the "Insert Atom in Triangulation" procedure. Figures for section 3.4, User Interfaces and Visualizing Data: startup dialogue, atom selection dialogues, criteria creation dialogue, plotting window for three-dimensional plots, XML Schema for the exported .xml files. This material is available free of charge via the Internet at http://pubs.acs.org

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: gunther@CELLmicrocosmos.org.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors wish to thank the RWTH Aachen for easy access to the RWTH HPC-Cluster and the Paderborn Center for Parallel Computing PC^2 (http://www.cs.uni-paderborn.de/pc2/) for providing computer time.

■ REFERENCES

(1) Klauda, J. B.; Kučerka, N.; Brooks, B. R.; Pastor, R. W.; Nagle, J. F. Simulationbased methods for interpreting x-ray data from lipid bilayers. *Biophys. J.* **2006**, *90*, 2796−2807.

(2) Murtola, T.; Kupiainen, M.; Falck, E.; Vattulainen, I. Conformational analysis of lipid molecules by self-organizing maps. *J. Chem. Phys.* **2007**, *126*.

(3) Shinoda, W.; Okazaki, S. A Voronoi analysis of lipid area fluctuation in a bilayer. *J. Chem. Phys.* **1998**, *109*, 1517.

(4) Sung, B. J.; Yethiraj, A. Lateral diffusion and percolation in membranes. *Phys. Rev. Lett.* **2006**, *96*, 228103−228106.

(5) Esteban-Martín, S.; Salgado, J. The dynamic orientation of membrane-bound peptides: bridging simulations and experiments. *Biophys. J.* **2007**, *93*, 4278−4288.

(6) Delaunay, B. N. Sur la sphère vide. *Bull. Acad. Sci. USSR* **1934**, 793−800.

(7) Voronoi, G. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les parallélloèdres primitifs. *J. Reine Angew. Math.* **2009**, *1908*, 198−287.

(8) Pandit, S. A.; Bostick, D.; Berkowitz, M. An algorithm to describe molecular scale rugged surfaces and its application to the study of a water/lipid bilayer interface. *J. Chem. Phys.* **2003**, *119*, 2199−2205.

(9) Pandit, S. A.; Jakobsson, E.; Scott, H. L. Simulation of the Early Stages of Nano-Domain Formation in Mixed Bilayers of Sphingomyelin, Cholesterol, and Dioleylphosphatidylcholine. *Biophys. J.* **2004**, *87*, 3312−3322.

(10) Suits, F.; Pitman, M. C.; Feller, S. E. Molecular dynamics investigation of the structural properties of phosphatidylethanolamine lipid bilayers. *J. Chem. Phys.* **2005**, *122*, 244714.

(11) Alinchenko, M. G.; Voloshin, V. P.; Medvedev, N. N.; Mezei, M.; Pártay, L.; Jedlovszky, P. Effect of cholesterol on the properties of phospholipid membranes. 4. Interatomic voids. *J. Phys. Chem. B* **2005**, *109*, 16490−16502.

(12) Weiss, J. A.; Oxtoby, D. W.; Grier, D. G.; Murray, C. A. Martensitic Transition in a Confined Colloidal Suspension. *J. Chem. Phys.* **1995**, *103*, 1180−1190.

(13) Mori, T.; Ogushi, F.; Sugita, Y. Analysis of lipid surface area in protein-membrane systems combining Voronoi tessellation and Monte Carlo integration methods. *J. Comput. Chem.* **2012**, *33*, 286−293.

(14) Alinchenko, M. G.; Anikeenko, A.; Medvedev, N. N.; Voloshin, V. P.; Mezei, M.; Jedlovszky, P. Morphology of voids in molecular systems. A Voronoi-Delaunay analysis of a simulated DMPC membrane. *J. Phys. Chem. B* **2004**, *108*, 19056−19067.
(15) Barber, C. B.; Dobkin, D. P.; Huhdanpaa, H. The Quickhull Algorithm for Convex Hulls. ACM Trans. *Math. Softw.* **1996**, *22*, 469−483.
(16) Shewchuk, J. R. Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator. In *Lecture Notes in Computer Science*; Lin, M. C., Manocha, D., Eds.; Springer-Verlag: New York, 1996; pp 203−222.
(17) Weisstein, E. W. Geometric Dual Graph. http://mathworld.wolfram.com/GeometricDualGraph.html.
(18) Okabe, A.; Boots, B.; Sugihara, K.; Chiu, S. Definitions and Basic Properties of Voronoi Diagrams. In *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*; John Wiley & Sons, Inc.: Hoboken, NJ, 2008; pp 43−112.
(19) Falck, E.; Patra, M.; Karttunen, M.; Hyvönen, M. T.; Vattulainen, I. Response to comment by Almeida et al.: free area theories for lipid bilayers−predictive or not? *Biophys. J.* **2005**, *89*, 745−752.
(20) Róg, T.; Pasenkiewicz-Gierula, M.; Vattulainen, I.; Karttunen, M. Ordering effects of cholesterol and its analogues. *Biochim. Biophys. Acta, Rev. Biomembr.* **2009**, *1788*, 97−121.
(21) Edholm, O.; Nagle, J. F. Areas of molecules in membranes consisting of mixtures. *Biophys. J.* **2005**, *89*, 1827−1832.
(22) Hofsäss, C.; Lindahl, E.; Edholm, O. Molecular dynamics simulations of phospholipid bilayers with cholesterol. *Biophys. J.* **2003**, *84*, 2192−2206.
(23) Alwarawrah, M.; Dai, J.; Huang, J. A molecular view of the cholesterol condensing effect in DOPC lipid bilayers. *J. Phys. Chem. B* **2010**, *114*, 7516−7523.
(24) Pandit, S. A.; Vasudevan, S.; Chiu, S. W.; Mashl, R. J.; Jakobsson, E.; Scott, H. L. Sphingomyelin-Cholesterol Domains in Phospholipid Membranes: Atomistic Simulation. *Biophys. J.* **2004**, *87*, 1092−1100.
(25) Eric, H. Polygons and Polyhedra. In *Graphics Gems IV*; Heckbert, P. S., Ed.; Morgan Kaufmann: Burlington, MA, 1994; pp 24−47.
(26) Jo, S.; Kim, T.; Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS One* **2007**, *2*, 1−9.
(27) Kandt, C.; Ash, W. L.; Tieleman, D. P. Setting up and running molecular dynamics simulations of membrane proteins. *Methods* **2007**, *41*, 475−488.
(28) Allen, W. J.; Lemkul, J. A.; Bevan, D. R. GridMAT-MD: a grid-based membrane analysis tool for use with molecular dynamics. *J. Comput. Chem.* **2009**, *30*, 1952−1958.
(29) Marrink, S.-J.; Risselada, H. J.; Mark, A. E. Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chem. Phys. Lipids* **2005**, *135*, 223−244.
(30) Wennberg, C. L.; van der Spoel, D.; Hub, J. S. Large influence of cholesterol on solute partitioning into lipid membranes. *J. Am. Chem. Soc.* **2012**, *134*, 5351−5361.
(31) Krüger, J.; Fischer, W. B. Structural implications of mutations assessed by molecular dynamics: Vpu1-32 from HIV-1. *Eur. Biophys. J.* **2009**, *39*, 1069−1077.
(32) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(33) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(34) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306−317.
(35) Guibas, L.; Stolfi, J. Primitives for the manipulation of general subdivisions and the computation of Voronoi. *ACM Trans. Graphics* **1985**, *4*.
(36) Sommer, B.; Dingersen, T.; Gamroth, C.; Schneider, S. E.; Rubert, S.; Krüger, J.; Dietz, K.-J. CELLmicrocosmos 2.2 MembraneEditor: A Modular Interactive Shape- Based Software Approach To Solve Heterogeneous Membrane Packing Problems. *J. Chem. Inf. Model.* **2011**, *51*, 1165−1182.
(37) Marrink, S.-J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.
(38) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1983**, *81*, 3684−3684.
(39) Tristram-Nagle, S.; Nagle, J. F.; Tristram-Nagle, S. Structure of lipid bilayers. *Biochim. Biophys. Acta, Rev. Biomembr.* **2000**, *1469*, 159−37.
(40) Risselada, H. J.; Marrink, S.-J. The molecular face of lipid rafts in model membranes. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17367−17372.
(41) Chandrasekhar, I.; Kastenholz, M.; Lins, R. D.; Oostenbrink, C.; Schuler, L. D.; Tieleman, D. P.; van Gunsteren, W. F. A consistent potential energy parameter set for lipids: dipalmitoylphosphatidylcholine as a benchmark of the GROMOS96 45A3 force field. *Eur. Biophys. J.* **2003**, *32*, 67−77.
(42) Vermeer, L. S.; de Groot, B. L.; Réat, V.; Milon, A.; Czaplicki, J. Acyl chain order parameter profiles in phospholipid bilayers: computation from molecular dynamics simulations and comparison with 2H NMR experiments. *Eur. Biophys. J.* **2007**, *36*, 919−931.
(43) Rinia, H. A.; Snel, M. M.; van der Eerden, J. P.; de Kruijff, B. Visualizing detergent resistant domains in model membranes with atomic force microscopy. *FEBS Lett.* **2001**, *501*, 92−96.
(44) Kučerka, N.; Tristram-Nagle, S.; Nagle, J. F. Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains. *J. Membr. Biol.* **2005**, *208*, 193−202.
(45) Janosi, L.; Gorfe, A. A. Simulating POPC and POPC/POPG Bilayers: Conserved Packing and Altered Surface Reactivity. *J. Chem. Theory Comput.* **2010**, *6*, 3267−3273.
(46) Hyslop, P. A.; Morel, B.; Sauerheber, R. D. Organization and interaction of cholesterol and phosphatidylcholine in model bilayer membranes. *Biochemistry* **1990**, *29*, 1025−1038.
(47) Smaby, J. M.; Momsen, M. M.; Brockman, H. L.; Brown, R. E. Phosphatidylcholine acyl unsaturation modulates the decrease in interfacial elasticity induced by cholesterol. *Biophys. J.* **1997**, *73*, 1492−1505.