

Vesicle Viewer: Online visualization and analysis of small-angle scattering from lipid vesicles

Aislyn Lewis-Laurent,¹ Milka Doktorova,² Frederick A. Heberle,^{3,*} and Drew Marquardt^{1,4,*}

¹Department of Chemistry & Biochemistry, University of Windsor, Windsor, Ontario, Canada; ²Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, Charlottesville, Virginia; ³Department of Chemistry, University of Tennessee, Knoxville, Tennessee; and ⁴Department of Physics, University of Windsor, Windsor, Ontario, Canada

ABSTRACT Small-angle X-ray and neutron scattering are among the most powerful experimental techniques for investigating the structure of biological membranes. Much of the critical information contained in small-angle scattering (SAS) data is not easily accessible to researchers who have limited time to analyze results by hand or to nonexperts who may lack the necessary scientific background to process such data. Easy-to-use data visualization software can allow them to take full advantage of their SAS data and maximize the use of limited resources. To this end, we developed an internet-based application called Vesicle Viewer to visualize and analyze SAS data from unilamellar lipid bilayer vesicles. Vesicle Viewer utilizes a modified scattering density profile (SDP) analysis called EZ-SDP in which key bilayer structural parameters, such as area per lipid and bilayer thickness, are easily and robustly determined. Notably, we introduce a bilayer model that is able to describe an asymmetric bilayer, whether it be chemically or isotopically asymmetric. The application primarily uses Django, a Python package specialized for the development of robust web applications. In addition, several other libraries are used to support the more technical aspects of the project; notable examples are Matplotlib (for graphs) and NumPy (for calculations). By eliminating the barrier of downloading and installing software, this web-based application will allow scientists to analyze their own vesicle scattering data using their preferred operating system. The web-based application can be found at <https://vesicleviewer.dmarquardt.ca/>.

SIGNIFICANCE Much of the critical information contained in small-angle scattering data is not easily accessible to researchers who have limited time to analyze results by hand or to nonexperts who may lack the necessary scientific background to process such data. Our easy-to-use data visualization software will allow them to take full advantage of their small-angle scattering data and maximize the use of limited resources.

INTRODUCTION

X-ray and neutron scattering data have a wide range of applications in the work of biochemists and biophysicists but remain complex to process and interpret. Whereas mathematical models have been derived for diverse types of samples, applying these models to experimental data is impossible to do by hand and, instead, requires a computer implementation. In the absence of a universal, standardized tool, researchers are often forced to build their own solutions.

There are several reasons why this is a less-than-optimal approach. Researchers are extremely busy and often lack the time to build robust solutions, resulting in computational

tools that are optimized for speed of development rather than ease of use or consistency. This often leads to code that can only be used by the programmer; new team members or other research groups cannot take advantage of the solution, and data are often left to grow stale while waiting for analysis by the developer. Moreover, these solutions may not be prepared with future projects in mind. When newly acquired data require modification of existing models or even novel models, implementing these in the code may not be within reach of the team. This, again, leads to data not being tapped for their full potential with new discoveries and advancements being left on the table.

With the advent of the scattering density profile (SDP) model for the analysis of lipid bilayer structure, the demand for accessible small-angle X-ray scattering (SAXS) and small-angle neutron scattering (SANS) data has increased. The SDP approach was pioneered by Wiener and White (1,2) for analyzing neutron and X-ray scattering data

Submitted June 3, 2021, and accepted for publication September 9, 2021.

*Correspondence: fheberle@utk.edu or drew.marquardt@uwindsor.ca

Editor: Tommy Nylander.

<https://doi.org/10.1016/j.bpj.2021.09.018>

© 2021 Biophysical Society.



obtained from oriented bilayer stacks and subsequently extended to the analysis of solution SANS and SAXS data from lipid vesicles by the efforts of Kučerka and co-workers (3–11). The SDP approach exploits the complementary nature of SAXS and SANS, which are sensitive to different aspects of the bilayer structure, by jointly refining data sets from both techniques to produce volume probability profiles for lipid component groups (e.g., the headgroup and hydrocarbon chains) (12). Because of the archetypal work of Wiener and White, the strategy of modeling data has been adopted by various other scattering techniques, including nondiffraction techniques such as reflectometry (13).

Although the SDP approach, in principle, provides the highest resolution structure of a fluid bilayer, it is technically difficult to implement and prone to overparameterization. For this reason, we previously developed a simpler model referred to as EZ-SDP that has successfully been used to extract bilayer structural information in a wide range of applications (14–22). The EZ-SDP approach requires only two adjustable structural parameters, namely the area per lipid and the headgroup thickness (discussed further below). The strategy of modeling a lipid bilayer as a series of slabs or strips is not a novel approach (23); however, the EZ-SDP method facilitates the determination of volume probabilities for key structural groups through the use of experimentally determined volumes (Eqs. S1–S5). Users also have the ability to render the volume parameters (V_H , V_C , and V_M) adjustable using the advanced options; however, the use of this option is not necessary.

In a bid to make small-angle scattering (SAS) data analysis easier for users and reduce group-to-group variation, we have developed the Vesicle Viewer (VV) online platform, which implements the EZ-SDP model. We have tested this platform with SAS data generated from molecular dynamics (MD) simulations of a single-component lipid bilayer, and multicomponent symmetric and asymmetric lipid bilayers. We further evaluate experimental data from vesicles composed of 1-stearoyl-2-docosahexaenoyl-sn-glycero-3-phosphocholine (SDPC) and directly compare structural parameters obtained with Vesicle Viewer to published results obtained with the SDP analysis.

MATERIALS AND METHODS

Vesicle Viewer models

The EZ-SDP approach determines key structural features that are most sought-after in membrane studies. By constraining lipid component scattering length densities with independent volume measurements, the symmetric bilayer model (Eq. 1) is a function of just two adjustable parameters or four adjustable parameters for the asymmetric bilayer model (Eq. S6) and the scattering vector q . Table 1 summarizes the model parameters, which include area per lipid (A_L), overall bilayer thickness (D_B), bilayer hydrocarbon chain thickness ($2D_C$), and distance between the maxima of the electron density (ED) profile (D_{HH}). Given the experimental considerations when investigating lipid vesicles, we do not account for the q -dependence of the X-ray atomic form factor because this has a negligible

TABLE 1 Relevant parameters in the EZ-SDP model

Parameter ^a	Descriptions
Obtained from independent experiments or literature	
V_C	lipid hydrocarbon chain volume [$\text{\AA}^3$] ^b
V_H	lipid headgroup volume [$\text{\AA}^3$] ^b
V_T	lipid terminal methyl volume [$\text{\AA}^3$] ^b
V_W	water volume [$\text{\AA}^3$]
Calculated from sample composition	
b_C	lipid hydrocarbon chain scattering length ^c
b_H	lipid headgroup scattering length ^c
b_W	water scattering length ^c
Adjustable	
A_L	area per lipid [$\text{\AA}^2$]
D_H	lipid headgroup thickness [$\text{\AA}$]
σ_s	accounts for thermal motion and represents the width of the z -distribution of atoms of an atom and effectively serves as a smoothing parameter for slab model (Fig. S1)

^aCan be varied in advanced options.

^b[fm $\text{\AA}^{-3}$] for neutrons and [$e^- \text{\AA}^{-3}$] for X-rays. Neutron scattering length values are from (25).

^cBilayer parameters double for the asymmetric model. The inner and outer leaflet parameters are signified by “ⁱⁿ” and “^{out},” respectively.

influence (<2%) over the typical experimental q -range in vesicle studies, as demonstrated by Klauda et al. (24). Furthermore, users should be aware that the wavelength dependence of the X-ray scattering length is not accounted for in the software. This omission is consistent with previous works for typical X-ray wavelengths of 1–2 $\text{\AA}$ (3–6,8–12).

$$\begin{aligned}
 F(q) = & \frac{2e^{-(q\sigma_s)^2/2}}{(qD_H A_L V_T V_W (V_C - 2V_T))} \bigg|_{V_T} \\
 & \times \left[b_W (A_L D_H - V_H) (V_C - 2V_T) + V_W b_H (V_C - 2V_T) \right. \\
 & \left. - V_W A_L D_H (b_C - 2b_T) \right] \sin\left(\frac{(qV_C)}{A_L}\right) + V_T (V_C - 2V_T) \\
 & \times (b_W V_H - b_H V_W) \sin\left(\frac{(qD_H + (qV_C))}{A_L}\right) \\
 & + V_W A_L D_H (b_C V_T - b_T V_C) \sin\left(\frac{(2qV_T)}{A_L}\right) \bigg|
 \end{aligned} \quad (1)$$

The models allow for the refinement of A_L and D_H ; the lipid volumes can be optimized in the program's advanced options. Other key structural parameters are subsequently calculated from the optimized A_L and D_H , most notably the total bilayer (Luzzati) thickness (D_B , Eq. 2) and the hydrocarbon thickness ($2D_C$, Eq. 3). Furthermore, the model yields a headgroup-headgroup distance (D_{HH}) calculated from the headgroup probability distribution peak-to-peak distance determined from the optimized fit parameters. We found no measurable discrepancy between the D_{HH} -values calculated by Vesicle Viewer and those obtained from the peaks of the

ED profile traditionally used to locate the electron-rich phosphate groups. In some instances, the number of water molecules per lipid (N_w) is a desired quantity. This can be calculated from the total head volume ($D_H \times A_L$) and the fixed V_H through the relation the $N_w = (D_H A_L - V_H)/V_w$.

Independently derived lipid volumes are required for model fitting. A number of strategies have been used to measure lipid volumes, including densitometry and the neutral flotation method (25), and volumes have now been cataloged for a wide variety of lipids with different headgroup and chain structure (3–11,26). Total lipid volumes (V_L), headgroup volumes (V_H), and hydrocarbon volumes (V_C) are summarized in the work of DiPasquale et al. (26):

$$D_B = \frac{2V_L}{A_L} \quad (2)$$

$$2D_C = \frac{2V_C}{A_L}. \quad (3)$$

The flat bilayer EZ-SDP form factor described above is a suitable approximation under most experimental circumstances. However, Vesicle Viewer does, in fact, provide the option to include vesicle size-curvature effects. Specifically, we implemented the Laplace transform-separated form factor (LP-SFF) equations of Pencic et al. 27 to account for the effects of vesicle size and polydispersity that appear at low q . The LP-SFF approach enables the use of a flat bilayer form factor rather than an exact form factor for concentric spherical shells, for which it is computationally more costly to account for polydispersity (the exact form factor must be numerically integrated over the vesicle size distribution). Pencic et al. convincingly showed that there are negligible differences between the LP-SFF form factor and the exact polydisperse vesicle form factor when the bilayer thickness and mean vesicle diameter are separated by more than an order of magnitude; for a typical 5 nm thick bilayer, the LP-SFF approach is therefore valid for extruded vesicles whose mean size is >50 nm (27).

Vesicle Viewer description

The purpose of the Vesicle Viewer project is to make established mathematical models easily accessible to researchers to increase the efficiency of their workflow and ultimately increase the utility of newly collected and existing SAXS and SANS data. We focused on building both a well-optimized algorithm for data analysis and a carefully constructed and easy-to-understand user interface, as shown in Fig. 1. Important parts of the interface are clearly labeled and well-organized, ensuring that using the tool is intuitive. Source code can be found on GitHub (<https://github.com/AislynLaurent/vesicle-viewer>) along with a listing on Zenodo (10.5281/zenodo.4653420). The webpage can be found at <https://vesicleviewer.dmarquardt.ca/>.

The focus on robustness over flexibility necessitated some restrictions on the user's ability to fine-tune fitting details or modify the scattering models for specialized samples, although some advanced features were included to cater to more experienced users. This compromise ensures that the widest possible audience is able to take advantage of the tool.

Users can input a currently unlimited number of data sets to be individually or jointly optimized using the EZ-SDP model. A standard least-squares algorithm is used to refine the adjustable model parameters to produce optimized scattering curves that best fit the experimental data sets. After refinement, the user is provided with a visual representation of the fit, fit statistics, and a final set of fit parameters. Fitted values are automatically stored, and the user can continue to process the results or export them for subsequent work. Specific inputs and outputs are outlined below.

User input: overview

Users are required to input various types of information during different stages of data processing. Initially, they are required to provide identifying authentication information used on the back end for both security and organization. This information includes a username, email address, and institution name. Additionally, each user must agree to basic terms and conditions. This provides both the user and the developer with clear roles and responsibilities (for example, maintaining user privacy and citing and sourcing results).

Several layers of organization allow users to separate results and prevent repeated processing of the same data. Projects contain high-level information, such as the model used (symmetrical or asymmetrical) and the temperature. Samples contain more detailed information, including isotopic variations in lipids, data sets, and fit parameters. Each of these data elements can be edited at any time. These changes are not retroactively enforced, but are reflected in any new calculations performed.

The fundamental data structure is the fit parameters. These are initially populated with suggested values based on results of previous calculations to provide a reasonable starting point for optimization. Additional parameters are calculated and are therefore only presented to the user after the initial generation of a parameter set. Notably, this includes the lipid volume, which in most cases will not vary and must be calculated based on the temperature of the system. Advanced users may prefer to allow this value to vary for a variety of reasons. An advanced option menu has been provided, in which users can unlock this option.

Each of the data elements is stored in a database until the user chooses to edit or delete it, allowing for quick retrieval and increased efficiency. Vesicle Viewer uses the PostgreSQL database architecture for this purpose.

Calculated output

Volumes are interpolated from temperature-dependent experimental data (4–8,10,11,26) and are only displayed to the user after a parameter set is generated. Although it is possible for the user to input all the fit parameters by hand, this would require significant preprocessing and potentially prevent new or inexperienced users from taking advantage of the tool.

Once a parameter set has been generated, it is fitted simultaneously to all provided data sets. A graphical representation of the fit is displayed for visual analysis. This allows the user to quickly assess the quality of fit. More detailed statistics are also provided, including the initial and final values for each parameter, the reduced χ^2 , and the number of iterations required for convergence of the fit.

The main outputs of interest are the final optimized values for the lipid area (A_L) and the headgroup thickness (D_H). This information, in addition to the calculated scattering intensity versus q for each data set, can be exported as a CSV file for further processing using programs such as Microsoft Excel.

Calculated output values are also stored in the database so that they can be referenced later on or refit when required. It is important to note the Vesicle Viewer is not intended as a permanent repository, and users are encouraged to export any results they intend to keep.

Python libraries

Vesicle Viewer employs several Python libraries to perform data processing and analysis. Many of these are standard Python libraries, such as CSV, and do not provide noteworthy functionality (a full list is available as part of the code posted on [GitHub.com](https://github.com)). Below is a list of the libraries used, the functionality they provide, and the reason(s) for selecting them.

- Django: a web development platform that allows Python code to be run alongside HyperText Markup Language (HTML), Cascading Style Sheets (CSS), and JavaScript and provides necessary core functionality to perform calculations online
- Matplotlib (28): a graphing library with support for a large variety of plotting and visualization options, including adjustment of labeling, line style, and scale

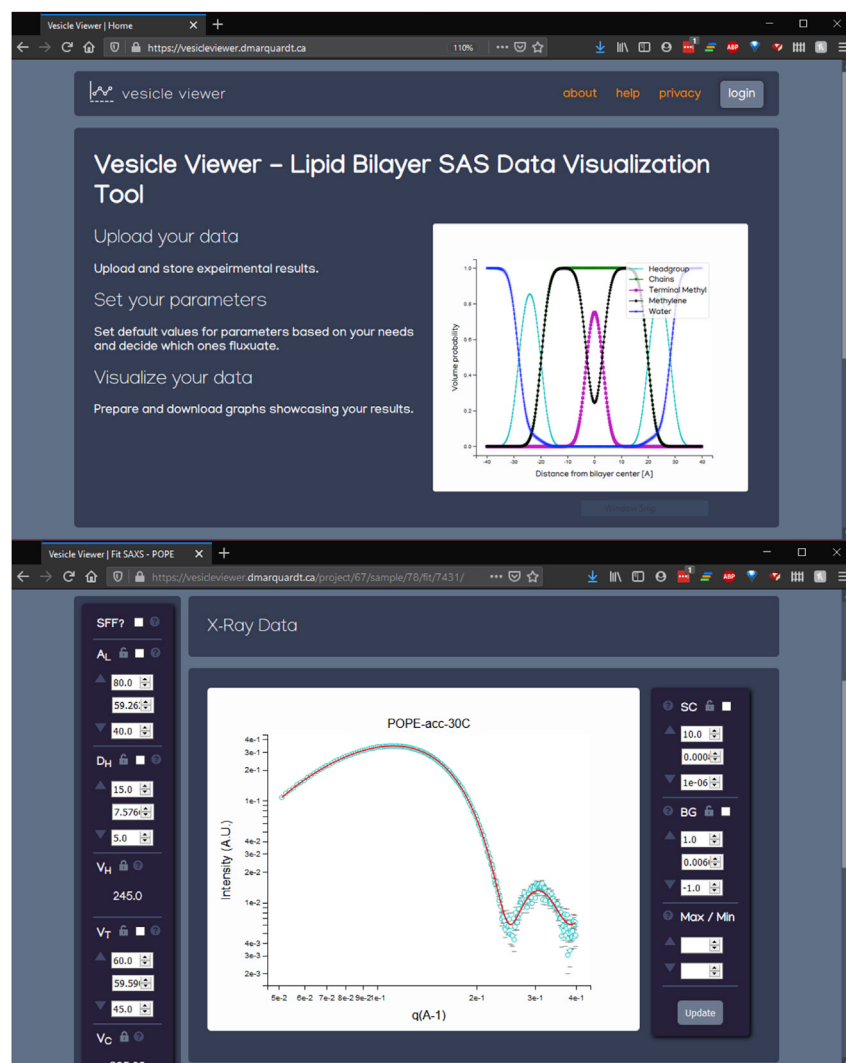


FIGURE 1 Screenshots of (*top*) the main welcome page and (*bottom*) the fit page on the Vesicle Viewer website. The fit page showcases the parameter selection panel (*left*), fit graph, including data points and the fit line (*middle*), and the scale and background adjustment panel (*right*). To see this figure in color, go online.

- **mpld3**: converts graphs generated by matplotlib into HTML, CSS, and JavaScript for display in a web interface; allows for fully featured graphs (including zoom functions) to be displayed in the web-app simply and easily
- **NumPy** (29): mathematical library containing advanced operators and array manipulation modules; seamlessly integrates with other libraries used in the project such as matplotlib and Lmfit
- **Lmfit** (30): a wrapper around the SciPy library module “minimizer” that performs a least-squares minimization using the Levenberg-Marquardt algorithm by default

Workflow

Fig. 2 shows a high-level description of the workflow in Vesicle Viewer. Several layers of processing occur. These stages include user authentication, user input, data preprocessing, fitting, visualization, and output.

User authentication

Data are organized at the highest level by “owner.” This makes data private to each individual user. As such, each page must authenticate the user before rendering its contents. This is done by collecting the user identity

at each page; if the current user is known by the system, their data are displayed.

Users must sign up for an account through a standardized system, made available by Django, which takes their name, email address, institution, and password. Django’s built-in systems automatically encrypt their password so that the information is not directly stored in the database. Once users are ready to begin work, they log in using this same information, allowing each page to check their identity.

User input: detail

Once authenticated, users are provided a list of their current projects and custom-defined lipids. Projects are the second highest level of organization. To create a project, users select the model they would like to use (symmetric or asymmetric), the temperature of their system, and the lipids present in their system. They can also choose to enable advanced options (e.g., varying normally fixed parameters).

After setting up their project, users can create samples. This additional layer is entirely organizational and is designed to prevent the repeated entering of duplicate information. At the sample level, more detailed information is collected about the sample composition, including the mole fraction of each lipid and, in the case of asymmetric systems, their leaflet location (i.e., inner or outer).

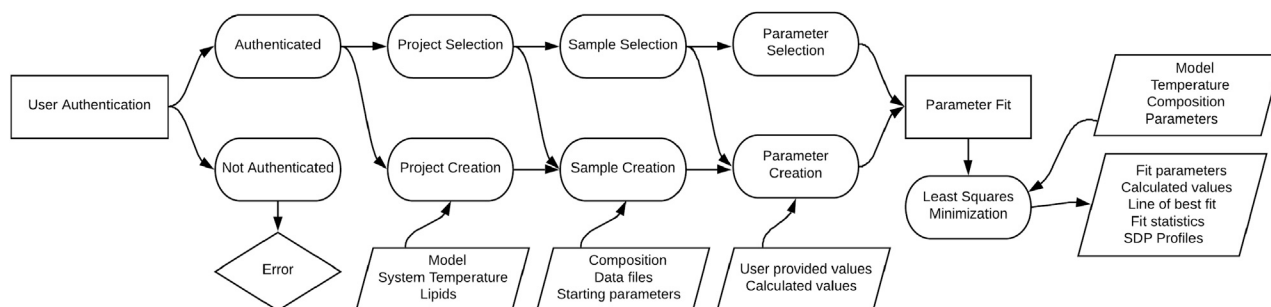


FIGURE 2 A flowchart describing the path taken as the user moves through Vesicle Viewer. The process begins with user authentication and proceeds as various inputs are requested. All data collected from a user are connected to their account and visible only to them. Users may move forward or backward through the program as they like; data are selectively displayed to them to help guide them through the process and keep things organized.

At this point, users can also select from a number of augmentations to the nominal forms of the lipids they have added. A database of common lipids is also available, preventing new users from having to manually search for values. Users can specify lipid modifications, such as deuteration or the addition of heavy metals, without the need to sift through long database tables. More advanced users can enter their own custom augmentations by providing the net scattering length change for each section of the lipid. Although composition information may be more intuitive for chemists to work with, this approach allows for maximal flexibility.

An unlimited number of data sets can be added. More details about how data are stored and used are included in the following sections. Users can specify the experimental data type (neutron or x ray), the proportion of D₂O, and data set name.

Lastly, users generate a parameter set. As with data uploads, some calculations are required at this step. The user is supplied only with parameters that they may wish to specify, such as the area per lipid. Suggested values are supplied so that inexperienced users can still perform analysis without a detailed understanding of what reasonable starting values should look like.

Data preprocessing

To maintain a clean user interface, several operations are hidden in the back end of the application. These range from simple calculations to complex data processing. The first of these is hidden behind the data upload step involved in setting up samples.

When the user chooses to upload a data file, the file is first checked for type and size. Afterward, it is read line-by-line, and three arrays are created based on the extracted values. The header and footer are both discarded because they do not carry usable information to the program. The first column of data is stored as the q value array, the second is stored as experimental intensity values, and the third column is assumed to be error values (if no error column is present, this array is filled with a dummy value of 1). Any additional columns are discarded. These arrays are the only part of the data file kept in the database because they are the only information required later in the process.

Once the lipid composition has been set, users can generate initial parameters for their fit. For unknown values, an estimate is populated as described above. Head- and chain-group volumes are calculated based on a volume equation stored in the database for each lipid. The project temperature and the mole fraction of each lipid is inputted, and a combined volume value for the parameter set is calculated. For custom lipids, the user is not able to enter a volume equation for security reasons (some volume equations will require access to Python functions). As such, they can add the correct lipid volume for the appropriate temperature at the project level.

Fitting

Lmfit is used to perform the fit. The default fitting algorithm is “leastsq,” which is a wrapper around the SciPy function of the same name (30), which is itself a wrapper around the FORTRAN library MINPACK (31).

In addition to what SciPy offers, Lmfit provides a number of convenient features, particularly for parameterization. Our implementation utilizes the default conditions (30). The steps taken to produce a best-fit scattering curve for all models will be described here. More detailed information on the least-squares method used by Lmfit can be found in the documentation on their website (<https://lmfit.github.io/lmfit-py/>) (30).

- 1) The fit function is called when the user presses the “fit” button. This sends the current sample and parameter instances to a separate module for processing. Existing parameters are saved in the database at this stage to allow the user the option to revert to previous fits.
- 2) The database contains the following values:
 - starting values for each parameter.
 - the upper and lower boundaries for each parameter.
 - a Boolean value indicating whether each parameter should vary.

Values may be those suggested by Vesicle Viewer or may be entered by the user before fitting. These values are pulled and converted into a dynamic Lmfit parameter object. The parameter object is used by the minimization algorithm to inject these values into the mathematical model during calculation.

- 3) Scattering lengths (b-values) are calculated by checking for lipid augmentations and the mole fraction of D₂O for the data set. These are added to the parameter object created in the previous step.
- 4) The minimization function is called and is passed with the converted parameters along with an objective function (which is used to calculate the residuals for each data set at each iteration).
- 5) The objective function calls a calculation function that evaluates the appropriate model for each data set with the current parameters (i.e., the current iteration of the minimization). Each residual is scaled by the error for that data set and concatenated to a larger array that will eventually be checked for the quality of fit. The water volume probability (Eq. S5) is also calculated by the objective function, and a penalty is added for any negative (i.e., unphysical) probability.
- 6) Residuals for the current iteration are evaluated for goodness of fit. The minimization function decides whether or not to continue or to stop.
 - Convergence is determined by comparing current values to the supplied data set(s).
 - If the distance between the actual values and the values in the current iteration are within a given tolerance, a series of checks are completed (one against a scaling matrix, a second against normalized values, and a third via a Jacobian matrix) (32).
 - If all checks are passed, the algorithm stops.
 - If one or more checks fail, iterations continue or an error is thrown, depending on the result.
- 7) The final iteration is passed back to the “fit” module, where the newly fitted parameters are stored in the database and displayed for the user. The new parameter set is automatically named using the current system time.

Data visualization

Scatter plots are generated for each data set and displayed for the user before fitting takes place. At this stage, their main utility is for the user to optimize the fitting parameters before fitting. Each graph is generated iteratively by Matplotlib and rendered into HTML and CSS by mpld3.

After fitting, new parameter values are displayed. The best-fit scattering curve is generated by piping these parameter values into the mathematical model used in the fit. The program decides whether or not to display a best-fit curve by checking whether a set of fit statistics exists for a particular parameter set. This allows the best-fit curve to automatically display any changes users make to parameters and also prevents duplicate database entries.

In addition to the data scatter plots and best-fit curves, the program displays fit statistics, volume probabilities, and density profiles. Fit statistics are stored in the database so that they can be displayed and downloaded during the output stage. As with the best-fit curves, probability and density profile graphs are generated based on the current set of parameters.

Output

Users are able to download a variety of data for further processing. CSV files are used because they are easily processed by multiple different software applications, most notably Microsoft Excel. Additionally, plain text files are compact and require relatively little back-end processing to produce. Vesicle Viewer has the ability to automatically generate two different output files:

- 1) a “Fit” file, which contains the current parameter set, fit statistics, such as the χ^2 -value and parameter variance, as well as calculated and experimental $I(q)$ and error values for each data set.
- 2) an “SDP” file containing the values for volume probabilities and separate and combined scattering density profile values.

Between these two files, all useful information displayed to the user as part of the interface is made available for them to download and work with as they please.

MD simulations

We constructed and simulated two symmetric bilayers and one asymmetric bilayer, as summarized in Table 2. The simulation trajectories for the two symmetric bilayers were taken from (33). The asymmetric bilayer was built and simulated in a manner similar to the symmetric bilayers as described in (33). Briefly, the bilayer was constructed with the CHARMM-GUI web server in 2015 (34–38). The system was simulated with NAMD version 2.7 (39,40) and the CHARMM36 force field for lipids and ions (41,42). Because at the time of construction CHARMM-GUI did not provide an established minimization protocol, the bilayer was first energy minimized for 10,000 steps, then simulated for 500 ps with a 1 fs time step. Because no constraints were imposed during these initial equilibration steps, in 2% of all lipids in the bilayer (two 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), four 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE), and two 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoserine (POPS) lipids) the *cis* double bond between carbons 9 and 10 on the sn-2 chain changed its isomerization to *trans* and remained like that for the entire duration of the simulation. The production run was performed at a constant

temperature of 25°C (298 K) and constant pressure of 1 atm under semi-isotropic pressure coupling conditions. The simulation parameters were the same as the ones used for the production runs of the two symmetric bilayers as described in (33).

The number density profile for each bilayer was calculated for every atom in the system with the Density Profile tool in Visual Molecular Dynamics (43) at a resolution of 0.2 Å. In the calculation of the scattering form factors from the simulation, a hydrogen was changed to deuterium by a matter of changing its label and using the appropriate scattering length (44).

Neutron and X-ray scattering intensity from MD simulations. First, an atomic number density profile was generated from the simulation trajectories by binning the *z*-coordinates of the different atoms to create a histogram. This was repeated for all frames in the simulation. The histogram was normalized to the number of MD frames processed. Neutron scattering length density profiles and electron density (ED) profiles were generated from the atomic number density by assigning the appropriate scattering length or number of electrons to the specific atoms and dividing by the bin volume. For example, the hydrogen atoms associated with the palmitoyl chain of POPC would be assigned the scattering length of deuterium for the chain deuterated POPC (dC-POPC). Finally, SANS and SAXS intensity curves are generated from the neutron scattering length density for SANS and ED for SAXS via the transform:

$$I(q) \propto \int (\rho(z) - \rho_s) e^{iqz} dz, \quad (4)$$

where q is the scattering vector, and $\rho(z)$ and ρ_s are the scattering length or ED along the *z*-direction of the bilayer and the solvent, respectively.

Experimental data

The EZ-SDP model has been applied to fairly ubiquitous phospholipid systems in past works (14,18); therefore, we utilize the polyunsaturated fatty acid-containing SDPC that has recently undergone a rigorous analysis using the full SDP model. The SANS and SAXS data were provided from the work of Marquardt et al. (10). The X-ray data were taken at the Cornell High Energy Synchrotron Source G-1 station, and the neutron scattering experiments were performed at the High Flux Isotope Reactor CG-3 Bio-SANS instrument located at Oak Ridge National Laboratory. Detailed sample preparation and instrument conditions are summarized in (10).

RESULTS

Simulated data

The most robust way to assess a new analysis technique or platform is to evaluate data that have an absolute “correct” answer. To this end, scattering data calculated from MD simulations provide “experimental” data, for which the structural parameters of the bilayer are strictly known. In this way we tested whether Vesicle Viewer

TABLE 2 Summary of MD bilayer composition

Bilayer	Lipid composition		Number of lipids		Number of waters per lipid	Simulation time (ns)
	Top system	Bottom	Top	Bottom		
1 ^a	POPC	POPC	208	208	45	520
2 ^a	POPE:POPS 70:30	POPE:POPS 70:30	210	210	82	690
3	POPC	POPE:POPS 70:30	191	220	60	620

^aSimulation trajectories were taken from (33).

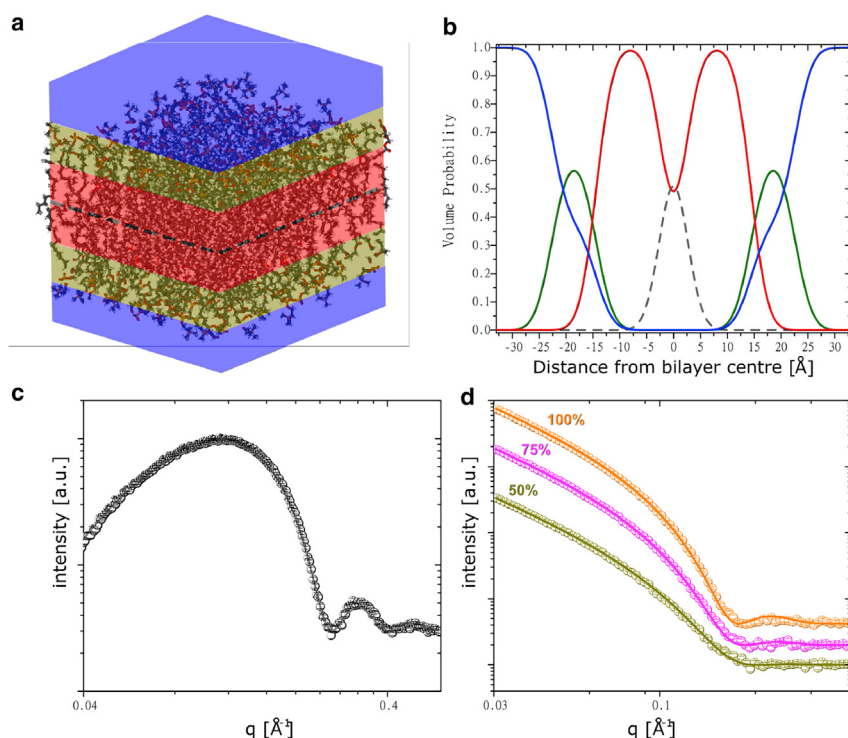


FIGURE 3 (a) Schematic of POPC bilayer indicating the EZ-SDP slabs. Slab colors correspond to the water (blue), headgroups (gold), the methylene (red), and the terminal methyl at the midplane (dashed line). These colors correspond to the volume probability distributions of (b). In (c) and (d), the SAXS and SANS, respectively, are shown. Open symbols indicate data generated from the MD simulations, and the solid lines represent the VV optimized models. The D₂O contrasts are indicated as 100%, 75%, and 50% D₂O. To see this figure in color, go online.

and the EZ-SDP model yield structural parameters of high fidelity.

We first examined a simple case, namely a symmetric POPC bilayer. Fig. 3 *a* illustrates the POPC bilayer broken down into its respective slabs and how these slabs correlate to the functional group volume probabilities and, ultimately, the VV fits of the MD-generated scattering data. Table 3 highlights key structural parameters determined by VV and those extracted from the MD simulations. It is particularly noteworthy that all the values are within 1 Å (or Å²) of each other, and their variance is less than 5%.

We then tested VV with a symmetric lipid mixture of POPE:POPS. It is clear that VV captures the structural parameters with an accuracy of <1.3% (Table 3). Notably, the optimized A_L is accurate to within 0.2%.

Finally, we tested VV on an asymmetric system having dC-POPC on the outer leaflet and a mixture of POPE:POPS

on the inner leaflet. Having one of the leaflets deuterated assists with contrast and is most representative of experimental structural studies of asymmetric membranes (16,18,45,46). Despite the complexity of this system, all optimized values are in good agreement with the actual values extracted from MD, with the most notable deviation occurring in the A_L parameter. We discuss the origin of minor deviations below.

Experimental data

We next used Vesicle Viewer to analyze experimental scattering data from vesicles composed of the polyunsaturated lipid SDPC. These data were previously analyzed using an implementation of the SDP model, which serves as a benchmark for comparison (Table 4) (10). Fig. 4 shows the experimental SAS data for SDPC as well as the optimized scattering density profiles. We note that the 50% of D₂O data appear cut off at $q = 0.15 \text{ Å}^{-1}$ because of the over subtraction of water

TABLE 3 Comparison of EZ-SDP executed through Vesicle Viewer with those directly from MD simulations

Parameter	POPC		POPE:POPS		dC-POPC ^{out} / (POPE:POPS) ⁱⁿ	
	VV ^a	MD	VV ^a	MD	VV ^a	MD
A_L (Å ²)	63.9	64.3	55.6	55.5	66.9/54.1	64.2/55.7
D_H (Å)	8.3	8.8	8.1	8.0	9.5/7.8	8.9/8.0
D_B (Å)	39.1	38.4	41.88	42.2	18.7/21.7	19.3/20.9
D_{HH} (Å)	37.2	38.2	41.2	42.0	18.6/20.8	19.0/20.8
$2D_C$ (Å)	28.79	28.8	32.8	32.9	13.8/17.0	14.5/16.4

^aImplemented using the Vesicle Viewer online platform.

TABLE 4 Comparison of EZ-SDP and full SDP analysis of SDPC at 30°C

Parameter	EZ-SDP ^a	Full SDP
A_L (Å ²)	69.6 (1.3)	70.4 (1.4)
D_B (Å)	39.3 (0.7)	38.8 (0.8)
D_{HH} (Å)	34.5 (0.6)	35.2 (0.7)
$2D_C$ (Å)	30.06 (6)	29.7 (0.6)

The comparison of EZ-SDP and full SDP analysis of SDPC at 30°C is covered in (10).

^aImplemented using the Vesicle Viewer online platform.

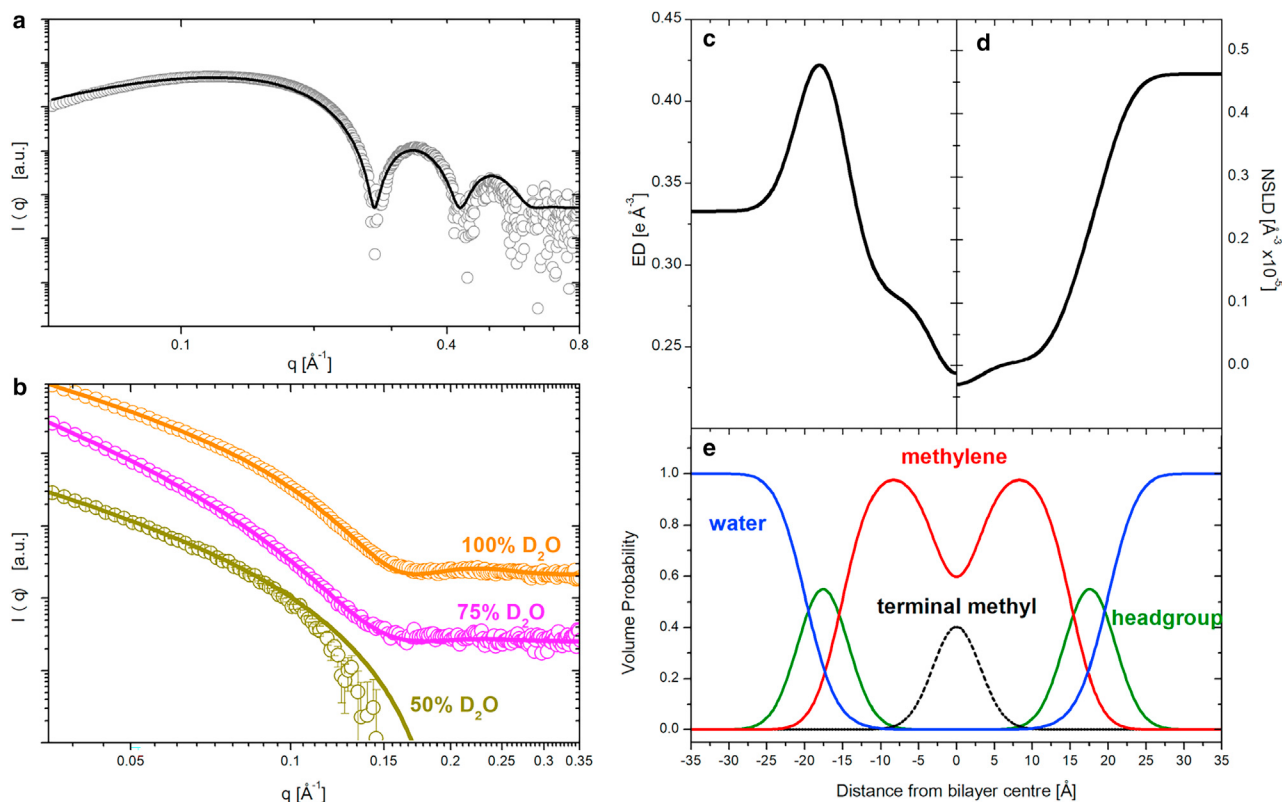


FIGURE 4 In (a) and (b), the SAXS and SANS curves, respectively, are shown. The open symbols are experimental data, and the solid line represent the optimized model. Note, the 50% D_2O SANS curve of (b) appears cut off because values must be >0 to be plotted on a logarithmic scale. In (c) and (d), the electron density (ED) profile and the neutron scattering length density profile from the optimized model, respectively, are shown. To see this figure in color, go online.

background during data reduction and the fact that values on a logarithmic scale must be >0 . Nevertheless, all data points are used in the fit optimization, despite the negative values not shown on the plot with a logarithmic scale.

DISCUSSION

We developed the Vesicle Viewer online platform with novice and experienced scattering users in mind. New users can easily pick up on the streamlined workflow and process data with little oversight, meaning that analysis no longer needs to wait for a more experienced researcher or instrument scientist. Data can be handed to students with little to no experience for quick and easy processing. More experienced researchers can focus their time on their projects and need not be concerned with the nuances of computer programming, data structures, or even the mathematical model itself. This type of task delegation should increase the efficiency of teams, leaving tedious work off the desks of busy, highly skilled individuals and allowing them to focus on higher priority tasks. Moreover, the easy-to-understand interface combined with step-by-step tutorials and guides can also be used as a teaching aid.

We have shown that the EZ-SDP model implemented in the Vesicle Viewer web platform provides accurate bilayer

structural parameters in both simulated and experimental data. Although the full SDP model has proven useful in single-component lipid systems, the large number of parameters creates problems for analyzing multicomponent mixtures and asymmetric bilayers. For example, the asymmetric SDP model employed in Eicher et al. has 12 adjustable parameters, necessitating the joint refinement of multiple differently contrasted data sets (47). The EZ-SDP model eliminates several parameters by combining headgroup components, thus sacrificing fine resolution of headgroup features while maintaining the ability to accurately determine the area per lipid and overall bilayer thickness. For these crucial parameters, the joint refinement of multiple data sets (both neutron and x ray) with the EZ-SDP slab model yields the same result, within error, as the full SDP (47). Although a systematic study of this has not been completed, we are able to comment on the quality of data required to recover values comparable with the full SDP using EZ-SDP. Key lipid parameters were recovered for dipalmitoylphosphocholine using a single SAXS data set from a home source instrument, with usable data to $q = 0.4 \text{ \AA}^{-1}$ (14). Similar results have been obtained for dipalmitoylphosphocholine, with SANS-only data having a usable q -range up to $0.4 \text{ \AA}^{-1}$; we note that, in general, the data range can be limited either by the

sample or the instrument (48). One must bear in mind that mentioned examples were primarily single-lipid systems (phosphatidylglycerol (PG)) was added to ensure unilamellarity (19)), and the complexity of the system being studied will often dictate the quality of the data and the number of data sets required to be jointly refined to achieve meaningful results.

Although VV sufficiently captures the bilayer structures from MD simulations, the very minor deviations between the fitted values and those directly calculated from the simulation can be explained by slight differences in the experimentally determined lipid volumes stored in the VV database and those of the simulated bilayers. For example, the experimentally determined volume of POPC was $1251.5 \text{ \AA}^3$, whereas the MD simulation yielded a volume of $1234.1 \text{ \AA}^3$. Although this is a small difference, it manifests in the value of A_L determined for the asymmetric bilayer. In practice, experimentally determined volumes are necessary for the EZ-SDP model; thus, we do not manually adjust our volumes to improve the results (although the program is capable of doing this) to highlight the robustness of our approach.

A major goal of the VV platform is to make existing scattering models for symmetric and asymmetric bilayers more accessible to the scientific community. Although we believe that VV will prove useful for many applications, some limitations should be kept in mind. First, the EZ-SDP scattering model is a relatively coarse-grained model (i.e., composed of few quasimolecular fragments) and is thus only appropriate for fluid phase bilayers, for which thermal disorder places fundamental limits on the resolution that can be achieved (1). Further, overparameterization is always a concern, especially in the case of asymmetric bilayers for which the number of parameters is effectively doubled. Ultimately, it is the user's responsibility to understand the limitations inherent to their own samples by analyzing appropriate simulated bilayers and checking parameter correlations.

SUPPORTING MATERIAL

Supporting material can be found online at <https://doi.org/10.1016/j.bpj.2021.09.018>.

ACKNOWLEDGMENTS

The authors thank Michael H. L. Nguyen for his critical review and valuable feedback.

The authors acknowledge support from the Natural Sciences and Engineering Research Council (NSERC) of Canada (funding reference number RGPIN-2018-04841) (to D.M.). F.A.H. acknowledges support from National Science Foundation (NSF) grant number MCB-1817929 and National Institutes of Health, National Institute of General Medical Sciences grant R01GM138887. M.D. was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number F32GM134704.

REFERENCES

- Wiener, M. C., and S. H. White. 1991. Fluid bilayer structure determination by the combined use of x-ray and neutron diffraction. I. Fluid bilayer models and the limits of resolution. *Biophys. J.* 59:162–173.
- Wiener, M. C., and S. H. White. 1991. Fluid bilayer structure determination by the combined use of x-ray and neutron diffraction. II. "Composition-space" refinement method. *Biophys. J.* 59:174–185.
- Kučerka, N., J. F. Nagle, ..., J. Katsaras. 2008. Lipid bilayer structure determined by the simultaneous analysis of neutron and X-ray scattering data. *Biophys. J.* 95:2356–2367.
- Kučerka, N., M.-P. Nieh, and J. Katsaras. 2011. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta.* 1808:2761–2771.
- Kučerka, N., B. van Oosten, ..., J. Katsaras. 2015. Molecular structures of fluid phosphatidylethanolamine bilayers obtained from simulation-to-experiment comparisons and experimental scattering density profiles. *J. Phys. Chem. B.* 119:1947–1956.
- Pan, J., F. A. Heberle, ..., N. Kučerka. 2012. Molecular structures of fluid phase phosphatidylglycerol bilayers as determined by small angle neutron and X-ray scattering. *Biochim. Biophys. Acta.* 1818:2135–2148.
- Pan, J., D. Marquardt, ..., J. Katsaras. 2014. Revisiting the bilayer structures of fluid phase phosphatidylglycerol lipids: accounting for exchangeable hydrogens. *Biochim. Biophys. Acta.* 1838:2966–2969.
- Pan, J., X. Cheng, ..., J. Katsaras. 2014. The molecular structure of a phosphatidylserine bilayer determined by scattering and molecular dynamics simulations. *Soft Matter.* 10:3716–3725.
- Pan, J., X. Cheng, ..., J. Katsaras. 2015. Structural and mechanical properties of cardiolipin lipid bilayers determined using neutron spin echo, small angle neutron and X-ray scattering, and molecular dynamics simulations. *Soft Matter.* 11:130–138.
- Marquardt, D., F. A. Heberle, ..., J. Katsaras. 2020. The structures of polyunsaturated lipid bilayers by joint refinement of neutron and X-ray scattering data. *Chem. Phys. Lipids.* 229:104892.
- Doktorova, M., N. Kučerka, ..., F. A. Heberle. 2020. Molecular structure of sphingomyelin in fluid phase bilayers determined by the joint analysis of small-angle neutron and X-ray scattering data. *J. Phys. Chem. B.* 124:5186–5200.
- Heberle, F. A., J. Pan, ..., J. Katsaras. 2012. Model-based approaches for the determination of lipid bilayer structure from small-angle neutron and X-ray scattering data. *Eur. Biophys. J.* 41:875–890.
- Belička, M., Y. Gerelli, ..., G. Fragneto. 2015. The component group structure of DPPC bilayers obtained by specular neutron reflectometry. *Soft Matter.* 11:6275–6283.
- Marquardt, D., F. A. Heberle, ..., G. Pabst. 2017. ^1H NMR shows slow phospholipid flip-flop in gel and fluid bilayers. *Langmuir.* 33:3731–3741.
- Marquardt, D., M. D. Frontzek, ..., J. Katsaras. 2018. Neutron diffraction from aligned stacks of lipid bilayers using the WAND instrument. *J. Appl. Cryst.* 51:235–241.
- Nguyen, M. H. L., M. DiPasquale, ..., D. Marquardt. 2019. Methanol accelerates DMPC flip-flop and transfer: a SANS study on lipid dynamics. *Biophys. J.* 116:755–759.
- Nguyen, M. H. L., M. DiPasquale, ..., D. Marquardt. 2019. Peptide-induced lipid flip-flop in asymmetric liposomes measured by small angle neutron scattering. *Langmuir.* 35:11735–11744.
- Doktorova, M., F. A. Heberle, ..., O. S. Andersen. 2019. Gramicidin increases lipid flip-flop in symmetric and asymmetric lipid vesicles. *Biophys. J.* 116:860–873.
- Scott, H. L., A. Skinkle, ..., F. A. Heberle. 2019. On the mechanism of bilayer separation by extrusion, or why your LUVs are not really unilamellar. *Biophys. J.* 117:1381–1386.
- Rickeard, B. W., M. H. L. Nguyen, ..., D. Marquardt. 2020. Transverse lipid organization dictates bending fluctuations in model plasma membranes. *Nanoscale.* 12:1438–1447.

21. Heberle, F. A., M. Doktorova, ..., I. Levental. 2020. Direct label-free imaging of nanodomains in biomimetic and biological membranes by cryogenic electron microscopy. *Proc. Natl. Acad. Sci. USA*. 117:19943–19952.
22. Chakraborty, S., M. Doktorova, ..., R. Ashkar. 2020. How cholesterol stiffens unsaturated lipid membranes. *Proc. Natl. Acad. Sci. USA*. 117:21896–21905.
23. Kucerka, N., J. F. Nagle, ..., P. Balgavý. 2004. Models to analyze small-angle neutron scattering from unilamellar lipid vesicles. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 69:051903.
24. Klauda, J. B., N. Kucerka, ..., J. F. Nagle. 2006. Simulation-based methods for interpreting x-ray data from lipid bilayers. *Biophys. J.* 90:2796–2807.
25. Greenwood, A. I., S. Tristram-Nagle, and J. F. Nagle. 2006. Partial molecular volumes of lipids and cholesterol. *Chem. Phys. Lipids*. 143:1–10.
26. DiPasquale, M., M. H. L. Nguyen, ..., D. Marquardt. 2021. Chapter TBD. In *Membrane Lipids* C. G. Cranfield, ed.. Springer US, p. TBD.
27. Pencer, J., S. Krueger, ..., J. Katsaras. 2006. Method of separated form factors for polydisperse vesicles. *J. Appl. Cryst.* 39:293–303.
28. Hunter, J. D. 2007. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9:90–95.
29. Harris, C. R., K. J. Millman, ..., T. E. Oliphant. 2020. Array programming with NumPy. *Nature*. 585:357–362.
30. Newville, M., T. Stensitzki, ..., A. Ingargiola. 2014. LMFIT: non-linear LeastSquare minimization and curve-fitting for Python. <https://doi.org/10.5281/zenodo.11813>.
31. Virtanen, P., R. Gommers, ..., P. van Mulbregt; SciPy 1.0 Contributors. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods*. 17:261–272.
32. Moré, J. J., B. S. Garbow, and K. E. Hillstom. 1980. User guide for MINPACK-1 [In FORTRAN]. Argonne National Lab, Lemont, IL..
33. Doktorova, M., D. Harries, and G. Khelashvili. 2017. Determination of bending rigidity and tilt modulus of lipid membranes from real-space fluctuation analysis of molecular dynamics simulations. *Phys. Chem. Chem. Phys.* 19:16806–16818.
34. Jo, S., T. Kim, ..., W. Im. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29:1859–1865.
35. Brooks, B. R., C. L. Brooks, III, ..., M. Karplus. 2009. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30:1545–1614.
36. Wu, E. L., X. Cheng, ..., W. Im. 2014. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* 35:1997–2004.
37. Lee, J., X. Cheng, ..., W. Im. 2016. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* 12:405–413.
38. Lee, J., D. S. Patel, ..., W. Im. 2019. CHARMM-GUI membrane builder for complex biological membrane simulations with glycolipids and lipoglycans. *J. Chem. Theory Comput.* 15:775–786.
39. Phillips, J. C., R. Braun, ..., K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.
40. Phillips, J. C., D. J. Hardy, ..., E. Tajkhorshid. 2020. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* 153:044130.
41. Klauda, J. B., R. M. Venable, ..., R. W. Pastor. 2010. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B*. 114:7830–7843.
42. Klauda, J. B., V. Monje, ..., W. Im. 2012. Improving the CHARMM force field for polyunsaturated fatty acid chains. *J. Phys. Chem. B*. 116:9424–9431.
43. Giorgino, T. 2014. Computing 1-D atomic densities in macromolecular simulations: the density profile tool for VMD. *Comput. Phys. Commun.* 185:317–322.
44. Sears, V. F. 1992. Neutron scattering lengths and cross sections. *Neutron News*. 3:26–37.
45. Heberle, F. A., D. Marquardt, ..., G. Pabst. 2016. Subnanometer structure of an asymmetric model membrane: interleaflet coupling influences domain properties. *Langmuir*. 32:5195–5200.
46. Doktorova, M., F. A. Heberle, ..., D. Marquardt. 2018. Preparation of asymmetric phospholipid vesicles for use as cell membrane models. *Nat. Protoc.* 13:2086–2101.
47. Eicher, B., F. A. Heberle, ..., G. Pabst. 2017. Joint small-angle X-ray and neutron scattering data analysis of asymmetric lipid vesicles. *J. Appl. Cryst.* 50:419–429.
48. DiPasquale, M., M. H. L. Nguyen, ..., D. Marquardt. 2020. The antioxidant vitamin E as a membrane raft modulator: tocopherols do not abolish lipid domains. *Biochim. Biophys. Acta Biomembr.* 1862:183189.