Documentation for "slabsax" SAXS Fitting Code

Laurence Lurio

July 29, 2021

1 Overview

The slabsaxs code fits small angle x-ray scattering from spherical liposomes to a slab model of the density profile. The code is built from a number of python files which contain command files which are edited by the user to control the fitting and subroutine files, which contain the code used to control the fits and ideally does not need to be edited. The code is presently written to analyse SAXS measurements taken at the 12-IDC/D beamline at the APS on liposomes in water. However much of the code is sufficiently general as to be modified for other data formats and systems.

2 Data Analysis

The data analysis chain consists of the following steps:

1. **Data Preparation:** Summing and averaging the raw SAXS data from the synchrotron, normalizing the data into absolute units and writing the summed data results to a numpy binary file ".npy"
2. **Model Construction:** Building a slab model to simulate the liposome structure
3. **Fitting:** Non-linear least squares fitting of the slab model to experimental data
4. **Documenting and Archiving:** plotting and saving fit results

The files and subroutines used for these steps are described in detail below.

2.1 Data reduction

Routines for averaging and normalization of data are contained within the python script **liposome_sum_code.py**. The routines contained within this file should not need to be modified. Rather, they are employed in a script file used for controlled the data reduction. An example script file is contained

in the file **sum_liposome_files.py**. The subroutines contained within **liposome_sum_code.py** are detailed in appendix A.

This file can be broken down into five parts. In part A the names of the datafiles that need to be reduced, the directory containing the data and the number of measurements per dataset are specified. In part B the data are summed and added to a single dictionary containing all the file data as datasets of which themselves are a dictionary of q, I and dI. The data are also normalized in absolute units using a normalization constant calculated in the file **water_diffuse_calc.py**. The details of the normalization calculation are given in appendix B

2.2 Constructing a model

The model for the liposome structure is based on the scattering from a step in density of the spherical liposome. In the simplest case, consider a uniform density sphere of radius R. This is described by a a single density step from ρ_{sphere} for $r < R$ to ρ_{water} for $r > R$. If we want to describe a hollow lipisome, of uniform density ρ_{lip} and thickness d we could do that with two steps, a step from ρ_{water} to ρ_{lip} at $r = R - d$ and a second step from ρ_{lip} to ρ_{water} at $r = R$. More complicated models can be built up from multiple steps, and the small angle x-ray scattering from these models can then be constructed by the square of the sum of the SAXS from each step in model.

We designate a step as a "slab" and specify it with three parameters, an amplitude change $\Delta\rho$ a radius at which the step occurs, r_i and a transition width σ over which the change in density occurs. The transition width is given by assuming that the density change is given by a gaussian

$$d\rho/dr = \frac{\Delta\rho}{\sqrt{2\pi\sigma^2}} e^{-(r-r_i)/2\sigma^2} \tag{1}$$

The python file **liposome_saxs_funs.py** contains the definition of two objects, a "slab" and a "profile". A slab defining a density step as described above and a profile being a collection of slabs. The profile object has the ability to calculate the x-ray scattering from the collection of slabs specified in the profile. The code in the file liposome_saxs_funs.py should apply to all models of liposomes and should not need to be changed. A specific model and its fitting parameters is specified in the file **liposome_model.py** which can be copied and changed to represent different models. The derivation of the liposome form factor for SAXS is given in appendix C

Within liposome_model.py the specific model is defined within the subroutine saxsfit. The user should infrequently have to redefine models, most of the time the original model should be kept unchanged and just the fitting parameters and other fitting options modified as will be discussed in the next section. In part (A) of the file a model (saxsfit) is defined which maps parameters of the model, onto

specific slab parameters which can then be fed to the profile object to calculate the reflectivity. In addition, the model defines other fitting parameters, such as amplitudes and backgrounds. Part (B) defines a utility function which does the actual translation of parameters and part (C) converts the model into an lmfit model structure, which can be used in the fits. The details of the specific model implemented in liposome_model.py are given in appendix 3,

2.3 Fitting data to the model

The file **run_liposome_fit.py** imports the model defined in liposome_model.py and runs the fit on the defined experimental data. In part (A) a binary .npy file saved by sum_liposome_files.py is read in, and a list of file names to be fit is defined from that set of files. In part (B) names are defined for where to save the data and fit results. In part (C) the parameters defining the fit are specified. This includes defining the subset of the selected files that will actually be fit, the number of times to shorten the data by averaging adjacent q-values using "half", the minimum and maximum q range of the fit, and the two files used to define the background (one being an empty capillary, and the other a capillary filled with water, but no lipid). This section also defines, which fit method is used (linear regression or differential evolution, the population size (psize) for differential evolution and the maximum number of iterations (nfev). In part (D) of the file the parameter file used by lmfit is defined. This gives initial values for all the parameters, sets which parameters are to be varied and gives minimum and maximum values for all the parameters. After these variables are defined the file calls the functions fit_liposome and plot_final_results to perform the fit and plot the results. These functions are defined in the file **liposome_fit_funs.py** and should ideally not need to be changed by the user.

The details of the fitting model and parameters are described in section 3.

3 Model and parameters

We model liposome electron density profile using slabs. Each slab is parameterized by a A_i a σ_i and an r_i as described above. In the computer code the slab is defined as a class which has four parameters, an amplitude a sigma a center and a name representing these quantities and the name to specify the slab. Thus creating a slab could be done via a call such as

```
s1 = slab(A_H,-Win/2,sig,'water to inner head')
```

This would define a slab called "water to inner head" with amplitude A_H, position -Win/2, and width sig. Slabs can be assembled into a profile using the profile class. A profile is an ordered list of slabs, with the order determined by the center positions of each slab. The profile class has a built in function to calculate the predicted SAXS based on eq. C.1 in addition to several utility functions to plot profiles and list the contents.

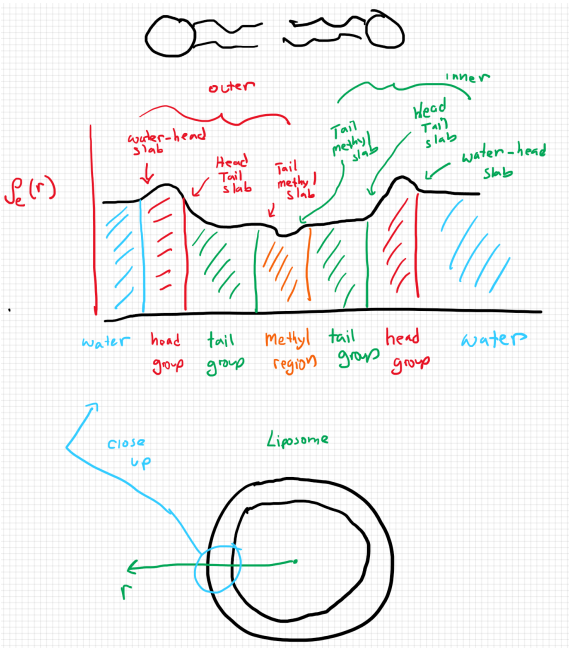

Figure 1: Schematic of bilayer density profile, regions and slabs used to define regions

The bilayer is defined using 7 regions. Each leaflet is defined by a headgroup region and a tail group region (4 regions). There is water on the outside and inside of the bilayer (2 regions) and there is a methyl region between the bilayers (1 region). The slabs define the density changes between regions and the bilayer can be defined with 6 slabs as outlined in fig. 1. Since each slab is defined with 3 parameters it requires 18 parameters to specify a model of the bilayer. In principle the cholesterol would require additional regions but we include the cholesterol density within the regions specifying the rest of the bilayer.

To fully describe the x-ray scattering we need additional parameters. This includes the average radius of the liposomes, R_0 and the second moment of the distribution of radii about their average σ_R. In addition the scattering is proportional to the concentration of liposomes, so there is an additional scale factor for the intensity, or equivalently the concentration, which we denote I.

3.1 Background Scattering

Ideally, subtraction of background would not require the addition of any new parameters. Background subtraction consists of measuring a capillary filled with water measured under identical conditions and then subtracting the water background. The background consists of a few different components:

- Small angle scattering from the surface of the glass capillary.
- Thermal diffuse scattering from the water
- Compton scattering from the water
- The small Q portion of the liquid structure factor peak from the water

3.1.1 Thermal diffuse scattering from water

The scattering from pure water in the capillary at low angles is due to thermal diffuse scattering. There is also background scattering at very small angles, most likely due to the surface of thee capillary glass. The scattering intensity from thermal diffuse scattering is given by:

$$N_w = I_t \Lambda \frac{1}{V} \frac{d\Sigma}{d\Omega} d\Omega = I_t \Lambda \rho_e^2 r_0^2 k_B T \chi_T d\Omega \tag{2}$$

Here N_w is the number of photons detected per second in a detector pixel subtending solid angle $d\Omega$, I_t is photon flux transmitted through the sample of thickness Λ, $\frac{1}{V}\frac{d\Sigma}{d\Omega}$ is the differential scattering cross section per unit volume, χ_T is the isothermal compressibility , T, the temperature, k_B, Boltzmann's constant, ρ_e the electron density and $r_0 = 2.82 \times 10^{-15}$m the classical electron radius.

At $T = 23°$C, $\chi_T = 0.45$ GPa^{-1}, $\rho_e = 335.5e^-/\text{nm}^3$ giving a scattering cross section of $\frac{1}{V}\frac{d\Sigma}{d\Omega} = 1.64 \times 10^{-2}\text{cm}^{-1}$

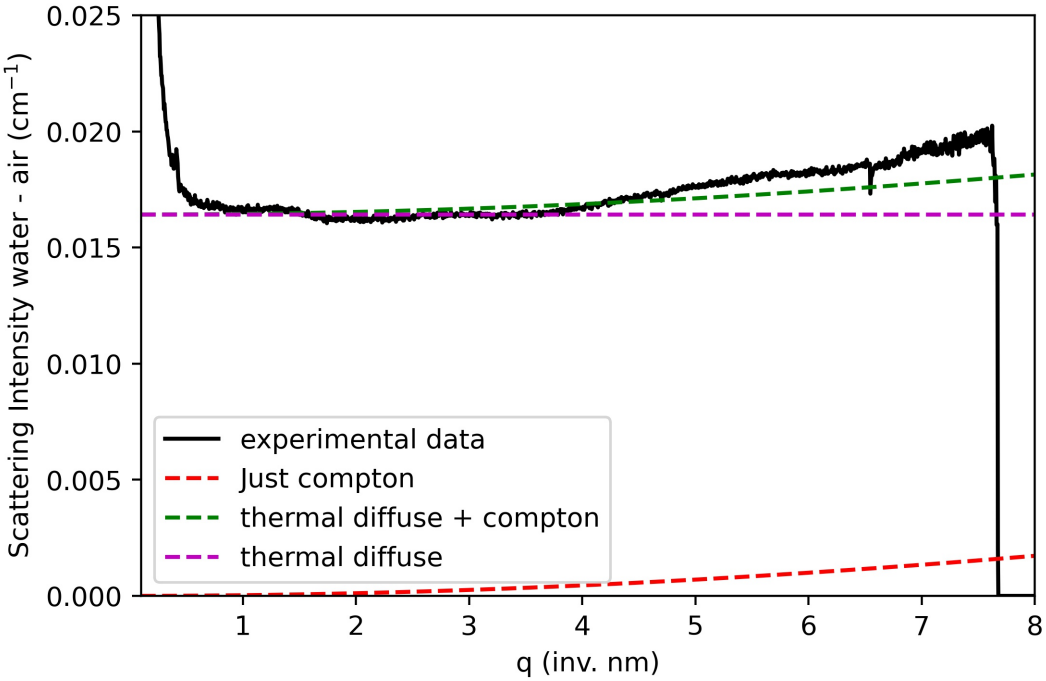

Figure 2: Diffuse scattering from water with empty capillary background subtracted. Data normalized by 0.185 in order to match plateau region of scattering to theoretical cross section for water. Dashed magenta line shows thermal diffuse cross section, while inclusion of the expected compton scattering is shown in the dashed green line.

3.1.2 Using thermal diffuse scattering to place scattering intensity on an absolute scale

The values returned from the experimental SAXS data are in units of $C_j N_w / (C_t I_t)$. Here C_j is a setup dependent proportionality factor between detector counts and photons/s and C_t is a setup dependent proportionality constant between transmitted beam counts and photons/s in the transmitted beam. The value of C_j/C_t can be found by comparing the measured diffuse scatterng from water with the experimentally measured value. This comparison gives $C_t/C_j = 1.148 \times 10^{-10}$. We expect C_t to be of order 1, and C_j to be large, since the transmitted beam flux is measured in a low sensitivity photodiode detector. Details of this conversion are in the python file "water_diffuse_calc.py."

3.1.3 Compton scattering from water

In addition to the thermal diffuse scattering from water, there will be incoherent scattering due to compton scattering. The compton (or modified scattering) cross section is approximately given by

$$\frac{1}{V}\frac{d\Sigma}{d\Omega} = r_e^2 \rho_e \left[\frac{\Sigma_i n_i \left(Z_i - f_i\right)}{\Sigma_i n_i Z_i}\right]$$

Here f_i is the atomic scattering factor for atom i, n_i is the stochiometry coefficient for atom i. This formula ignores polarization and quantum quantum effects since the scattering is at relatively low angles. The dashed green line in fig. 2 shows the expected compton cross section, which agrees reasonably well with the slow rise in the scattering at larger q

The python code for these calculations is in water_diffuse_calc.py in the directory liposomes.

3.2 Fitting the background

Since the sample is approximately 10 mg/ml, approximately 1% of the water is displaced by lipid. This implies that subtracting the water background will remove slightly too much of the last three items, however, as the lipid will have its own contributions to these terms it is only the difference in those terms between water and lipid which would need to be compensated for. Furthermore, the accuracy of the background subtraction depends on a number of factors:

- Hitting the capillary in the same place so the thickness is identical and the surface scattering is identical.
- Not having contamination (due to damaged lipids) stick to the walls of the capillary
- Accurately normalizing the transmitted beam intensity

In order to compensate for these factors we allow several parameters to adjust the background subtraction. These include

- A scale factor to adjust the water background subtraction. This should be 1, but typically takes values within a few percent of 1. (bg1sf)
- A scale factor to adjust the background subtraction from an empty capillary (air). In principle you should not need to subtract an empty capillary background. The argument is that if there is a difference due to small angle scattering from the capillary surface then it may not be possible to simultaniously subtract the water thermal diffuse background and the surface background with a single scale factor. The empty capillary background gives a way to subtract the surface scattering without subtracting additional thermal diffuse. In the end, this option turns out not to be particularly useful, but the option is left in the code. (bg2sf)

- A polynomial background. The idea here is that perhaps errors in background subtraction could be removed by allowing a low order polynomial to be added to the fit. The code allows for a constant (bg), linear (lbg) and quadratic (qbg) term to be added to the fit. In the end, this option turns out not to be particularly useful and is not used.

3.3 Constrained model for bilayer

In order to reduce the total number of adjustable parameters the different parameters in the layer can be linked using a constrained model of the bilayer. This constrained model, together with auxilary parameters such as the background multipliers and lipsome size distributions are defined in a function called saxsfit. The saxsfit function is then turned into a fitting model using the Model call from the lmfit package.

The number of parameters is reduced by making the following assumptions.

- Assume the bilayer is symmetric, so that the inner and outer leaflets are identical.
- Assume that the transitions between all layers are equally rough (e.g. only one value for σ
- Approximate the density of the methyl region as zero. This seems like a rather bold approximation, but since the methyl region is very narrow, the width and the amplitude tend to be indistinguishable and the only net effect on the profile is the product of the two. This prevents having two closely related parameters varied independently in the fitting program which would lead to a singularity in the error matrix.
- Keep the electron density of the water fixed to the value of 334 e^-/nm^3 for bulk water.

This reduces the parameters needed to describe a profile to 5. We designate the parameters as follows:

- position of the water-headgroup transition z_H
- the position of the headgroup-tail group transition z_T
- the position of the tailgroup to methyl group transition z_M
- the amplitude of the head group A_H
- the amplitude of the tail group A_T

We can rewrite these values in terms of thickness of each layer since the headgroup thickness, $d_H = z_H - z_T$, and similarly the tail group thickness $d_T = z_T - z_M$ and the methyl group thickness $d_M = 2z_M$. We can define the total bilayer thickness as $W = 2z_H$ and then we can define the bilayer in terms of the 5 parameters W, d_H, d_M, A_H and A_T.

Of these five parameters, not all need to be freely varied. We fix the headgroup thickness d_H at 0.70 nm, based on x-ray reflectivity studies of supported bilayers. (PCCP Vega, 2020) This is a reasonable assumption since the head group should not change significantly as opposed to the lipid tails which can fold in different configurations and also mix with cholesterol. In addition, we can also fix the electron density of the head group A_H. This is because the scattering only depends on the relative amplitude of all the layers, leaving one of the amplitude parameters arbitrary. We fix $A_H = 441e^{-1}/\text{nm}^3$ based on the electron density of of the DPPC head group region measured in (PCCP Vega). (Note this includes water associated with the headgroup region). Thus we are left with only 3 variable parameters to describe the bilayer W, d_M and A_T.

3.3.1 Asymmetry

It is possible, however, that the bilayer may not be symmetric. The inner leaflet has negative curvature and the outer leaflet positive curvature which could effect packing and distribution of cholesterol. In order to account for this we allow for an asymmetry between the two leaflets. We define an asymmetry function $f_a(p)$ which can go from 0 to 2 and then take the inner leaflet thickness to be $d_{in} = W f_a(p)/2$ and the outer leaflet thickness to be $d_{out} = W(2 - f_a(p))/2$. We then use this to calculate the thickness of the lipid tail region to be $d_{T,in} = d_{in} - d_H - d_M/2$ and similarly with the outer leaflet lipid tail region. For the asymmetry function we take

$$f_a(p) = (1 + 2\arctan(p)/\pi) \tag{3}$$

This takes advantage of the fact that the range of arctan is from $-\pi/2$ to $\pi/2$ to map the full x-axis onto the finite range [0,2].

Making the bilayer asymmetric has the effect that the center of the profile, is not the midpoint between the two ends. In order to redefine the center to be the midpoint, we need to offset the profile by an amount given by $dr = (W - W_{out})/2$

The specific code is given by

```
Wout = W*(1-2*np.arctan(W_asym)/np.pi)
Win = W*(1+2*np.arctan(W_asym)/np.pi)
dr = (W-Wout)/2
```

In a similar manner we allow for an asymmetry in the amplitude of the tail region. The electron densities should also be taken relative to the electron density of water, since the liposome is surrounded by water. Thus the head group electron density in the model (the one fixed parameter) should be taken as $A_H = 441e^{-}1/\text{nm}^3 - 334e^{-}1/\text{nm}^3 = 107e^{-}1/\text{nm}^3$.

Accounting for this offset we have $A_{T,in} = (A_T + 334)f(p') - 334$ and $A_{T,out} = (A_T + 334)(2 - f(p')) - 334$. Here p' is an independent parameter from p.

The specific code is

```
A_T_out = (A_T+334)*(1-2*np.arctan(A_T_asym)/np.pi) -334
A_T_in = (A_T+334)*(1+2*np.arctan(A_T_asym)/np.pi)  -334
```

3.4 Fitting parameters

A list of all the parameters, whether they are varied and their initial vales is given in the table below.

Parameter	Variable in code	default value	varied	comments
$bg1sf$	$bg1sf$	1.0	yes	scale factor for water bg
$bg2sf$	$bg2sf$	0	0	scale factor for air
bg	bg	0	no	constant background
lbg	lbg	0	no	linear background
qbg	qbg	0	no	quandratic background
W	W	4.37	yes	bilayer width
d_H	d_H	0.7	no	headgroup width
d_M	d_M	0.1	yes	methyl width
A_H	A_H	107	no	headgroup electron density (-water)
A_M	A_M	-334	no	methyl electron density (-water)
σ	sig	.3	yes	bilayer roughness
I	I	1	yes	amplitude
R_0	$R0$	fixed from DLS	no	liposome mean radius
σ_R	$Rsig$	fixed from DLS	no	liposome radius distribution
p	W_asym	0	maybe	width asymmetry
p'	A_T_asym	0	maybe	density asymmetry

3.5 Fit range

The range of the data runs from 0.033 nm^{-1} to 7.99 nm^{-1}. However the range of fitting was smaller than this. The data at small q suffered from strong parasitic small angle scattering due to scattering from the capillary tube surface. While this could be approximately subtracted out using the water background, the parasitic scattering was not exactly reproducible. This could be due to the beam moving on the surface of the capillary between measurements, or possibly aggregation of lipids on the capillary walls that was not removed during flushing between samples. Consequently data was only fit from 0.3 nm^{-1}. At large q it was also difficult to fit the data due to the approximately constant diffuse background scattering from water. The background is approximately equal to the signal at $q = 2.5nm^{-}1$ and is is approximately 1000x larger than the signal at the largest q fit (5 nm^{-1}).

3.6 Fit weights

Appendices

A Function descriptions for liposome_sum_code

- **find_file(ddir,searchstring):** This is a utility to locate a group of files with the same name but different file numbers. It takes the name of the data directory to search "ddir" and the string to find "searchstring" and returns all files in that directory which have the searchstring in their name.
- **cmbwe(yin1,ein1,yin2,ein2):** This is a utility to to combine two measurements of the same data set. It assumes that you have a data vector, yin1, and its error ein1 and you want to combine it with yin2 and ein2 in such a way that the result is the optimally weighted combined data. The combined error bar will be lower than ein1 or ein2.
- **sum_files(ddir,fbase,nfiles):** This routine sums "nfiles" files which all have the same base name "fbase" located in directory "ddir". The result is a dictionary of q, I, dI values. Here "q" is the scattering vector, "I" the averaged intensity, and "dI" the error in "I" after averaging.
- **cmb_set(s1,s2):** This is the same as cmbe but takes dictionaries of q, I, and dI values, unpackeges them, combines them with cmbwe and then repackages the result in a dictionary.
- **deglitch(ds1,ds2,cut):** This routine is used to remove glitches in the data resulting (typically) from cosmic ray hits on the detector during the measurement. ds1 and ds2 are two data dictionaries containing q, I and dI. "cut" is a threshold used to compare the data. If the first data set exceeds the second by more than the fractional difference specified by the cut value, then those points in ds1 are replaced by the corresponding points from ds2. This routine is used in the routine **sum_files_glitch** to avoid summing glitchy points when combining two files.
- **sum_files_glitch(ddir,fbase,nfiles):** This routine is identical to sum_files except that it deglitches the files using deglitch before adding them to the sum.
- **half(dset):** This routine takes a dataset (e.g. dictionary of q, I and dI) and merges all pairs of adjacent points in q, so that there are half as many q points with correspondingly smaller error bars.
- **shorten(dset,nhalf)**
- **subtract_sets(set1,set2,SF=1):** Subtracts the intensity in set2 from the intensity in set1 after scaling set1 by 1/SF. The errors are combined

appropriately to yield a (larger) final error.

B Normalization of the Scattering Intensity from Liposomes

B.1 Scattering cross section of suspension

The scattering cross section per unit volume for a suspension of liposomes of number density ρ_n and form factor $F(q)$ is given by

$$\frac{1}{V}\frac{d\Sigma}{d\Omega} = r_e^2 \rho_n \left|F(q)\right|^2$$

Here r_e is the Thompson scattering length. The number density of liposomes can be related to the mass per unit volume of the liposomes in mg/ml as follows. Let W_l be the average molecular weight of the lipid, S_l be the area density of lipids, and R be the average radius of the liposomes. Then the liposome mass is

$$M_l = W_l \times 2 \times \left(4\pi R^2\right) \times S_l \tag{4}$$

If ρ_m is the mass density of liposomes and ρ_n the number density of liposomes, then

$$\rho_n = \frac{\rho_m}{M_l} \tag{5}$$

Thus

$$\frac{1}{V}\frac{d\Sigma}{d\Omega} = r_e^2 \frac{\rho_m}{8W_l \pi R^2 S_l} \left|F(q)\right|^2 \tag{6}$$

If we have a distribution of liposomes centered on radius R_0 with standard deviation σ_R then the liposome mass is averaged over the distribution of radii

$$M_l = \frac{8W_l \pi S_l}{\sqrt{2\pi\sigma_R^2}} \int_{-\infty}^{\infty} R^2 e^{-(R-R_0)^2/2\sigma_R^2} dR = 8W_l \pi S_l \left(R_0^2 + \sigma_R^2\right) \tag{7}$$

For eggPC the molecular weight is $W_l \approx 746$. For the area per lipid we will assume the value for DPPC $\frac{1}{S_l} = 0.48$ nm^2 This gives $8W_l \pi S_l = 6.49 \times 10^{-5}$ kg/m^2. If we put in values for $R = 62$ nm and $\rho_m = 10$ mg/ml we get

$$\frac{1}{V}\frac{d\Sigma}{d\Omega} = 3.2 \times 10^{-12} \left|F(q)\right|^2 \tag{8}$$

See file liposome_normalization.py for details.

C Calculation of form factor of liposome slabs

The liposome form factor is given by:

$$|F(q)|^2 = \left| \frac{4\pi}{q} \int_0^\infty \Delta\rho(r) r \sin(qr)\, dr \right|^2$$

We are approximating $\rho(r)$ by a series of slabs. In particular, the transition in density for each slab is given by

$$\frac{d\Delta\rho_i(r)}{dr} = \frac{A_i}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(r - R - r_i)^2}{2\sigma_i^2} \right]$$

Here r_i is the position relative to the liposome wall center, and R is the radius of the liposome (distance from center of liposome to center of wall). Since its easier to work with the derivative $\frac{d\Delta\rho_i(r)}{dr}$ we can integrate by parts and throw away the boundary terms, since the profile is zero at both $r = 0$ and $r = \infty$:

$$\frac{4\pi}{q} \int_0^\infty \Delta\rho(r) r \sin(qr)\, dr = \frac{4\pi}{q^2} \int_0^\infty \frac{d\Delta\rho_i(r)}{dr} \left[\frac{1}{q} \sin(qr) - r\cos(qr) \right] dr$$

This can be further simplified by recognizing that $\frac{d\Delta\rho_i(r)}{dr}$ is zero sufficiently far from the liposome wall so that the integral can be extended to infinity. We can then change variable to $y_i = q(R + r_i)$, $x + y_i = qr$ and $\epsilon_i = q\sigma_i$, This gives:

$$\frac{4\pi A_i}{q^3\sqrt{2\pi\epsilon_i^2}} \int_{-\infty}^\infty \exp\left(-\frac{x^2}{2\epsilon_i^2} \right) \left[\sin(x + y) - (x + y_i)\cos(x + y_i) \right] dx$$

We can expand $\cos(x + y_i) = \cos(x)\cos(y_i) - \sin(x)\sin(y_i)$ and $\sin(x + y_i) = \sin(x)\cos(y_i) + \cos(x)\sin(y_i)$. Since $\exp(-x^2/2\epsilon_i^2)$ is an even function, the odd terms in x give zero, leaving:

$$\frac{4\pi A_i}{q^3\sqrt{2\pi\epsilon_i^2}} \int_{-\infty}^\infty \exp\left(-x^2/2\epsilon_i^2\right) \left\{ \cos(x) \left[\sin(y_i) - y_i \cos(y_i) \right] + x \sin(x)\sin(y_i) \right\} dx \tag{9}$$

The term in $x\sin(x)$ can be simplified via integration by parts (again throwing away the boundary terms)

$$\int \exp\left(-x^2/2\epsilon_i^2\right) x \sin(x)\sin(y_i) dx = -\int \epsilon_i^2 \exp\left(-x^2/2\epsilon_i^2\right) \cos(x)\sin(y_i) dx \tag{10}$$

So that the integral becomes:

$$\frac{4\pi A_i}{q^3\sqrt{2\pi\epsilon_i^2}}\int_{-\infty}^{\infty}\exp\left(-x^2/2\epsilon_i^2\right)\cos(x)\left[\sin(y_i)\left(1-\epsilon_i^2\right)-y\cos(y_i)\right]dx \tag{11}$$

The cosine transform is given by:

$$\int_{-\infty}^{\infty}\exp\left(-x^2/2\epsilon_i^2\right)\cos(x)dx=\sqrt{2\pi\epsilon_i^2}e^{-\epsilon_i^2/2}$$

This gives the final result

$$f_i(q)=\frac{4\pi A_i}{q^3}\left[\sin(y_i)\left(1-\epsilon_i^2\right)-y_i\cos(y_i)\right]e^{-\epsilon_i^2/2} \tag{12}$$

with $F(q)=\left|\sum_i f_i(q)\right|^2$

C.1 Distribution of Radii

There is typically a distribution of liposome radii. Let us assume the probability density of liposomes with radius R is given by:

$$n(R)=\frac{1}{\sqrt{2\pi\sigma_R^2}}\exp\left[-\left(R-R_0\right)^2/2\sigma_R^2\right] \tag{13}$$

Putting all the terms together we get

$$\begin{aligned} I\propto\frac{16\pi^2}{q^6\sqrt{2\pi\sigma_R^2}}\sum_{i,j}A_ie^{-\epsilon_i^2/2}A_je^{-\epsilon_j^2/2}\\ \int_{-\infty}^{\infty}\exp\left[-\left(R-R_0\right)^2/2\sigma_R^2\right]\left[\sin(y_i)\sin(y_j)\left(1-\epsilon_i^2\right)\left(1-\epsilon_j^2\right)-\right.\\ \left.2\sin(y_i)\cos(y_j)y_j\left(1-\epsilon_i^2\right)+\cos(y_i)y_i\cos(y_j)y_j\right]dR \end{aligned} \tag{14}$$

Change variables to $z=q(R-R_0)$, $\phi=q\sigma_R$, $w_i=y_i-z=q(R_0+r_i)$, and $\delta_i=\left(1-\epsilon_i^2\right)$. This gives:

$$\begin{aligned} I\propto\frac{16\pi^2}{q^6\sqrt{2\pi\phi^2}}\sum_{i,j}A_ie^{-\epsilon_i^2/2}A_je^{-\epsilon_j^2/2}\int_{-\infty}^{\infty}\exp\left[-z^2/2\phi^2\right]\left[\sin(z+w_i)\sin(z+w_j)\delta_i\delta_j\right.\\ \left.-2\sin(z+w_i)\cos(z+w_j)(z+w_i)\delta_j+\cos(z+w_i)\cos(z+w_j)(z+w_i)(z+w_j)\right]dz \end{aligned} \tag{15}$$

Each of these three terms can be simplified by expanding the trig functions and discarding terms odd in z.

- First term

$$\cos(w_i)\cos(w_j)\,\delta_i\delta_j - \cos(w_i+w_j)\cos^2(z)\,\delta_i\delta_j$$

- Second term

$$2\cos(w_i)\sin(w_j)w_j\delta_i - \sin(2z)\cos(w_i+w_j)z\delta_i - 2\cos^2(z)\sin(w_i+w_j)w_j\delta_i \tag{16}$$

- Third term

$$[\cos^2(z)\cos(w_i+w_j)+\sin(w_i)\sin(w_j)](z^2+w_iw_j) - \sin(2z)\sin(w_i+w_j)zw_i \tag{17}$$

Collecting terms in similar z simplifies to:

$$I \propto \frac{16\pi^2}{q^6\sqrt{2\pi\phi^2}}\sum_{i,j} A_i e^{-\epsilon_i^2/2} A_j e^{-\epsilon_j^2/2}\int_{-\infty}^{\infty} e^{-z^2/2\phi^2}\left(C_1 + C_2\cos(z)^2 + C_3 z^2\cos^2(z) + C_4 z\sin(2z) + c_5 z^2\right)dz$$

Here we have

- C_1

$$\cos(w_i)\cos(w_j)\delta_i\delta_j + 2\cos(w_i)\sin(w_j)\delta_i w_j + \sin(w_i)\sin(w_j)w_iw_j$$

- C_2

$$\cos(w_i+w_j)(w_iw_j - \delta_i\delta_j) - 2\sin(w_i+w_j)\delta_i w_j$$

- C_3

$$\cos(w_i+w_j)$$

- C_4

$$-\cos(w_i+w_j)\delta_i - \sin(w_i+w_j)w_i$$

- C_5

$$\sin(w_i)\sin(w_j)$$

Evaluating the integrals gives

$$\frac{16\pi^2}{q^6}\sum_{i,j} A_i e^{-\epsilon_i^2/2} A_j e^{-\epsilon_j^2/2}\{C_1 + \frac{1}{2}(1+e^{-2\phi^2})C_2 + \frac{\phi^2}{2}[1+(1-4\phi^2)e^{-2\phi^2}]C_3 + 2\phi^2 e^{-2\phi^2}C_4 + \phi^2 C_5\}$$