viewSq Tutorial: Water

The initial viewSq calculations performed for the tutorial should take 1-2 hours to complete on most desktop machines. The tutorial requires <2 GB RAM and <8 GB disk space (additional RAM may be required if running tutorial in a virtual machine). VMD <u>downloads</u> and <u>installation instructions</u> can be found at those links. viewSq installation instructions are in the viewSq README.

The simulation trajectory pdb and ndx files can be found within the /tutorials/spce_water folder of the viewSq github repository.

- 1. Place spce_final_frame.pdb and elements.ndx into a folder. spce_final_frame.pdb contains atomic positions from a molecular dynamics simulation of 1,000 SPC/E water model molecules (2,000 Hydrogen and 1,000 Oxygen atoms). elements.ndx is a Gromacs style .ndx file which maps each atom to its element type. The pdb file contains two force field atom types: type 1 (Hydrogen) and type 2 (Oxygen).
- 2. Open VMD. In the VMD Main window select File, New Molecule..., which will open the Molecule File Browser (Figure 1). In the Molecule File Browser click Browse... and select spce_final_frame.pdb. VMD should automatically determine the file type, however if it does not, select PDB from the drop-down menu. Set Frames: to "First: 0", "Last: -1", and "Step: 1". Select the "Load all at once" radio button, and then press Load. The VMD Display window should now display the water box. Close Molecule File Browser.

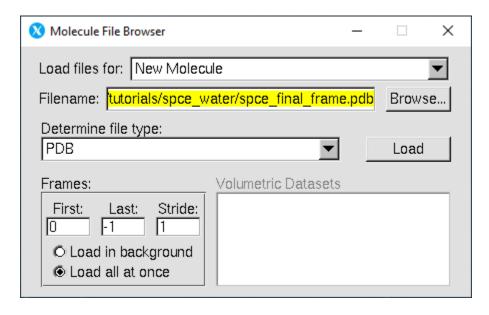


Figure 1. VMD Molecule File Browser.

3. In VMD Main select Extensions, Analysis, viewSq. This opens the viewSq GUI (Figure 2).

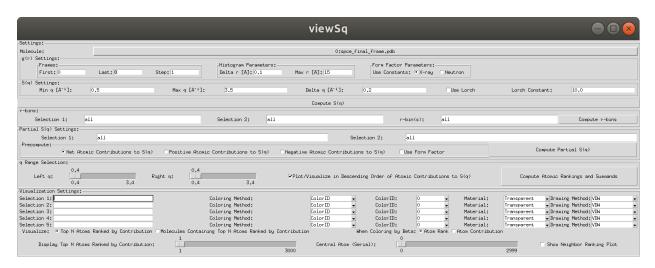


Figure 2. viewSq GUI.

- **4.** In the viewSq GUI select "(none)" next to Molecule and choose "0:spce_final_frame.pdb". If the number zero does not precede the pdb file name, then VMD must be closed and the tutorial restarted from Step #2. viewSq is only capable of performing analysis on VMD molecule index zero.
- **5.** For "g(r) Settings" set Frames settings to "First: 0", "Last: -1", and "Step: 1" (the defaults). This selects the first (and only in this case) frame in a simulation file. The count is zero-indexed, so the first frame is 0, second frame 1, etc. Because there is only one frame, a Step of any positive integer will work. For Histogram Parameters set "Delta r [Å]: 0.1" and "Max r [Å]: 15.0". For "S(q) settings" set "Min q [Å-¹]: 0.5", "Max q [Å-¹]: 3.5", and "Delta q [Å-¹]: 0.2". Increasing Max r (also referred to as r_{max} in this tutorial) increases RAM usage and calculation time, while increasing the number of q adds a small amount of extra RAM usage and additional calculation time and disk space usage. In this case we are setting Max r (r_{max}) to 15 Å, which corresponds to just under half the simulation box length and therefore the longest which can be used for this simulation). For "Form Factor Parameters" make sure "X-ray" is checked. This setting allows choosing between X-ray and neutron parameters for form factor calculations. We will not use a Lorch function for this tutorial, though it can be enabled and it's constant set using "Use Lorch" and "Lorch Constant".
- **6.** Press "Compute S(q)". The calculation will take approximately 1-2 hours. After a few minutes viewSq will begin printing progress updates in the VMD terminal every time 100 of the 3,000 atoms has been processed.

7. When the calculation from Step #6 completes, six new windows containing plots (e.g. Figure 3). The plots include g(r), S(q), the positive and negative components of S(q) with and without form factors, and the sums of the absolute values of the positive and negative components with and without form factors. The "Partial S(q) Settings" portion of the GUI is now accessible. Let us save the data from the S(q) plot. Select the plot, press "File", "Export to ASCII MATRIX" (a text file with columns), and then choose a file name which tells what the calculation is. Using "p" instead of decimal points, the file name could be: multiplot_sq_spce_final_frame_maxr15_deltar0p1_minq0p5_maxq3p5_deltaq0p2.dat. The output file contains three columns in the following order: q, S(q) without form factors, and S(q) with form factors. Plotting them in an external program will reproduce Figure 3.

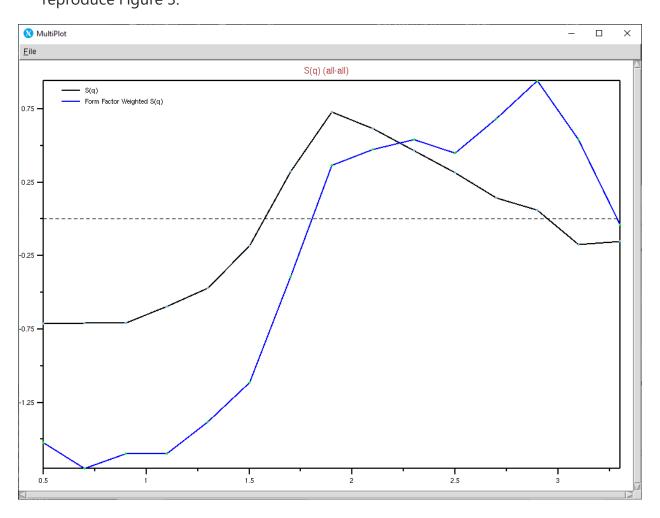


Figure 3. One of the six viewSq windows from Step #7. The plot shows S(q) with and without form factor weights.

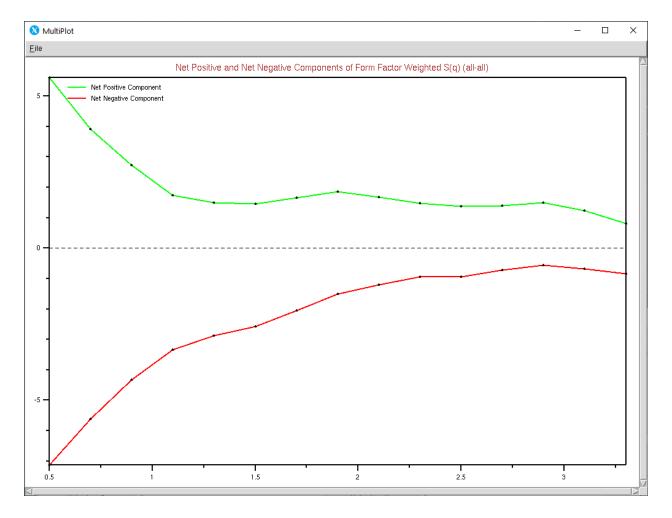


Figure 4. One of the six viewSq windows from Step #7. The plot shows the positive and negative components of S(q) accounting form factor weights. The x-axis is q, and the sum of the positive and negative value at a given q exactly reproduces S(q) at that q.

8. For "Partial S(q) Settings" set "Selection 1: type 1" and "Selection 2: type 2". The "Precompute" settings are irrelevant for this step and can be left as defaults. Press "Compute Partial S(q)". After some time (progress can again be tracked in the VMD terminal) the Hydrogen-Oxygen partial g(r), partial S(q) (Figure 5), and four other plot windows will display. All six plot windows are partial g(r) or partial S(q) analogs of those from Step #7.

In addition to <u>VMD selection methods</u>, viewSq offers an alternative selection method. The format is: A,B,C -- where A is an initial atom serial (i.e. VMD serial), B is an integer skip parameter, and C is the number of skips to use. So, for a 30-atom molecule of which there are 250 such molecules, and atom serial 4 of each molecule is the only to be selected, the selection would be: 4,30,250. Colons can be inserted between entries (e.g. 4,30,250:8,30,250). The selection in Step #8 could be replaced

as follows: "Selection 1: 2,3,1000:3,3,1000" and "Selection 2: 1,3,1000". Selection 1 again corresponds to "type 1" or Hydrogen and Selection 2 to "type 2" or Oxygen.

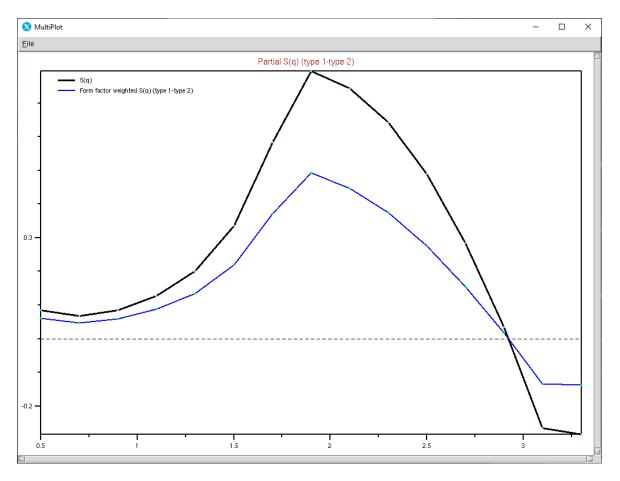


Figure 5. One of the six viewSq windows from Step #8. The plot shows the Hydrogen-Oxygen partial S(q) with and without form factor weights.

9. The "r-bins" portion of the GUI becomes accessible once the initial calculations complete and is independent of any settings located lower in the GUI (e.g. "Partial S(q) Settings"). The r-bins module prints the count and percent of atomic distances for any two atom selections which are associated with a user-defined range (or ranges) of distances (i.e. set of r-bins). This allows g(r) and Fourier transform summands to be interpreted in terms of the percent of different atomic distances of type Selection 1 with Selection 2 found in the r-bin selection. The output is printed to the VMD terminal and consists of three quantities (indicated with ***, as well as various intermediate quantities, as seen in Figure 6): (I) the percent of distances of the chosen type found within the entire box, (II) the percent of distances of the chosen type found within a user-selected set of r-bins.

The default settings in r-bins is "Selection 1: all", "Selection 2: all", and "r-bin(s): all". This means all distances within r_{max} will be used (i.e. spanning all Hydrogen-Hydrogen, Oxygen-Oxygen, and Hydrogen-Oxygen distances of 15 Å or less). Press "Compute r-bins". After a few minutes the VMD terminal will look like Figure 6.

```
1600 out of 3000 atoms processed.
1700 out of 3000 atoms processed.
1800 out of 3000 atoms processed.
1900 out of 3000 atoms processed.
2000 out of 3000 atoms processed.
2100 out of 3000 atoms processed.
2200 out of 3000 atoms processed.
2300 out of 3000 atoms processed.
2400 out of 3000 atoms processed.
2500 out of 3000 atoms processed.
2600 out of 3000 atoms processed.
2700 out of 3000 atoms processed.
2800 out of 3000 atoms processed.
2900 out of 3000 atoms processed.
3000 out of 3000 atoms processed.
Completed!
Total number distances in box: 4498500
Total number atoms in r-bins module selection 1: 3000
Total number atoms in r-bins module selection 2: 3000
Total number of atoms in intersection of selections 1 & 2: 3000
***Percent selected distance types out of all distances in box***: 100.0
Total distances within rmax: 4220820
Count of selected distances within rmax: 4220820
***Percent selected distance types represent out of all distances within rmax***: 100.0
Count of selected distances within selected r-bins: 4220820
count_all_selected_rbins: 4220820
***Percent of selected distances within selected r-bins***: 100.0
```

Figure 6. VMD terminal showing output from the viewSq r-bins module.

Total number distances in box: 4498500

The total number of distances in the simulation box, accounting for the number of frames chosen in Step #2 (for the tutorial that corresponds to one frame).

Total number atoms in r-bins module selection 1: 3000

Selection 1 is "all", representing all 3,000 atoms.

Total number atoms in r-bins module selection 2: 3000

Selection 1 is "all", representing all 3,000 atoms.

Total number of atoms in intersection of selections 1 & 2: 3000

Both Selection 1 and Selection 2 contain all atoms, so their intersection is all atoms.

Percent selected distance types out of all distances in box: 100.0

Both Selection 1 and Selection 2 contain Hydrogen and Oxygen, and all distances within the box are either Hydrogen-Hydrogen, Oxygen-Oxygen, or Hydrogen-Oxygen. Therefore the above value is 100%.

Total distances within rmax: 4220820

The above count represents the number of distances within the box which are of length 15 Å or less.

Count of selected distances within rmax: 4220820

Both Selection 1 and Selection 2 contain all atoms, so the above count represents the fraction of distances within the box which are of length 15 Å or less.

Percent selected distance types represented out of all distances within rmax: 100.0

Both Selection 1 and Selection 2 contain all atoms, and therefore they represent all distances of length 15 Å or less.

Count of selected distances within selected r-bins: 4220820

The r-bin(s) selection is "all", and Selection 1 and Selection 2 are both "all", so the above count represents all distances of length 15 Å or less (4220820 distances, matching the same count a few lines above).

count all selected rbins: 4220820

There are 4220820 distances of length 15 Å or less.

Percent of selected distances within selected r-bins: 100.0

The r-bin(s) selection is "all", and Selection 1 and Selection 2 are both "all", so the above count represents all distances of length 15 Å or less (100%).

10.r-bins are selected by entering comma separated values of r or comma separated ranges of r. The width of r-bins was set in Step #2 using Delta r. Therefore bin 6.5 includes distances 6.5 Å <= X < 6.6 Å, where 6.5 Å is included in bin 6.5 and 6.6 Å is associated with bin 6.6. The r_{max} of 15.0 Å selected in Step #2 corresponds to an r-bin selection of 0.0 - 14.9. Set "r-bin(s): 5.0-6.0, 7.5, 8.8-9.3" and press "Compute r-bins". The first eight lines of output will match that from Step #9 (because Selections 1 and

2 are still both "all", and r_{max} is still 15 Å), however the remaining lines will reflect values associated with the selected distances:

Count of selected distances within selected r-bins: 333684

There are 333684 distances of selection (type all-all) within the r-bins 5.0-6.0, 7.5, and 8.8-9.3.

count_all_selected_rbins: 333684

There are 333684 distances in the r-bins 5.0-6.0, 7.5, and 8.8-9.3.

Percent of selected distances within selected r-bins: 100.0

There are 333684 distances of lengths within r-bins 5.0-6.0, 7.5, and 8.8-9.3.

11.Covalent bonds in the simulation are in r-bins 0.9 and 1.0, which corresponds to distances 0.9 Å <= X < 1.1 Å (see radial distribution function from Step #7). Let us use the r-bins module to confirm something known -- that all atomic distances within these bins are Hydrogen-Oxygen (i.e. force field type1-type2). Set "Selection 1: type 1", "Selection 2: type 2", and "r-bin(s): 0.9-1.0".

Total number distances in box: 4498500

Unchanged from Step #9

Total number atoms in r-bins module selection 1: 2000

Type 1 represents Hydrogen and there are 2,000 Hydrogen atoms in the simulation box.

Total number atoms in r-bins module selection 2: 1000

Type 2 represents Hydrogen and there are 1,000 Oxygen atoms in the simulation box.

Total number of atoms in intersection of selections 1 & 2: 0

Selections 1 and 2 are disjoint, having no atoms in common. Therefore, the above count is zero.

Percent selected distance types out of all distances in box: 44.459264199177504

Hydrogen-Oxygen distances represent about 44.46% of all distances in the simulation box.

Total distances within rmax: 4220820

Unchanged from Step #9.

Count of selected distances within rmax: 1877358

There are 1877358 Hydrogen-Oxygen distances within length 15 \AA (r_{max}).

Percent selected distance types represent out of all distances within rmax: 44.478513653745004

Hydrogen-Oxygen distances represent about 44.48% of all distances within length 15 Å (r_{max}). This percent can be compared with the percent from a few lines above (44.46%) to gauge whether r_{max} is sufficiently long to capture the character of the simulation box.

Count of selected distances within selected r-bins: 4000

The 4,000 covalent bonds between Hydrogen and Oxygen in the 1,000 water molecules.

Percent of selected distances within selected r-bins: 0.09476831516150891

- **12.**For "Partial S(q) Settings" set "Selection 1: all" and "Selection 2: all". Set the "Precompute" radio button to "Net Atomic Contributions to S(q)" and check the box next to "Use Form Factor". Press "Compute Partial S(q)". Because the selections are all-all, the six plots which appear when the calculations complete are identical to the six plots which appeared after the initial calculations (Step #7). Whenever Selections 1 and 2 are both set to "all", viewSq is able to reuse portions of the initial calculation (Step #6) in manners which allow the new calculation to finish more quickly than when using any other selections (e.g. Step #8). Another optimization speeds up calculations when Selection 1 equals Selection 2 (e.g. type1-type1).
- **13.**For "q Range Selection" set the "Left q" slider to 0.5 and the "Right q" slider to 0.5 (or 0.4 and 0.6, respectively, if 0.5 is not accessible -- either set of settings are identical for the tutorial, as 0.5 is the only q we selected in Step #4 within the range of 0.4 and 0.6, inclusive). This allows examination of S(q) at $q = 0.5 \text{ Å}^{-1}$. Ensure the radio button is set to "Net Atomic Contributions to S(q)". Press "Compute Atomic Rankings and Summands".
- **14.** When the calculation from Step #13 completes, three windows containing two graphs and one table will appear (Figures 7, 8, and 9). Exporting the atom ranking plot (Figure 8) also exports VMD information regarding each ranked atom (type, name, residue, resname, resid, and chain), which can be useful for analysis of the different atoms and motifs which are contributors to S(q) at a given q or range of q.

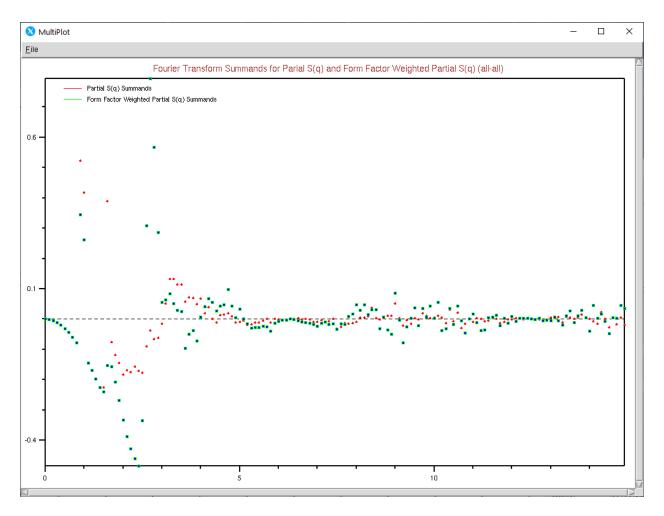


Figure 7. viewSq plot showing Fourier transform summands for $q = 0.5 \text{ Å}^{-1}$. The sum of all shown summands exactly produces S(q) at $q = 0.5 \text{ Å}^{-1}$. The sum of the positive summands (or negative summands) yields the positive (or negative) component of S(q) at $q = 0.5 \text{ Å}^{-1}$.

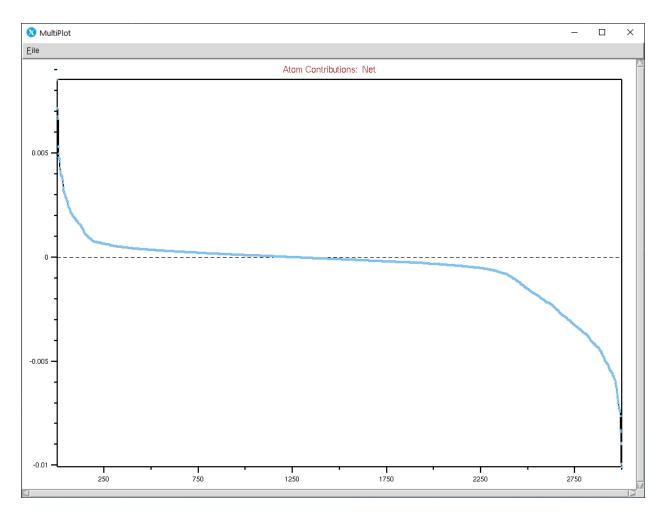


Figure 8. viewSq plot showing atoms ranked by their contributions to S(q) for $q = 0.5 \text{ Å}^{-1}$. The x-axis shows the 3,000 atom rankings, while the y-axis is the corresponding atom's contribution (negative or positive) to S(q) at $q = 0.5 \text{ Å}^{-1}$. The sum of all shown values is exactly equal to S(q) at $q = 0.5 \text{ Å}^{-1}$.

	_	- 🗆 X
Selections:-		
	No Form Factor	Form Factor
S(q):	-0.713	-1.522
Net Positive Component:	2.666	5.604
Net Negative Component:	-3.379	-7.126
Sum of Net Positive and Magnitude of Net Negative Components:	6.045	12.73
Sum of Atom Contributions = S(q):		-1.522
ок		

Figure 9. viewSq table which appears when "Compute Atomic Rankings and Summands" is pressed. Because $q = 0.5 \text{ Å}^{-1}$ was used, the first row corresponds to S(q) at $q = 0.5 \text{ Å}^{-1}$. The second and third rows are the positive and negative components of

S(q) at $q = 0.5 \text{ Å}^{-1}$ (the form factor entries exactly match the two $q = 0.5 \text{ Å}^{-1}$ values in Figure 4). Summing the net positive and net negative components in the first column reproduces S(q) without form factors, while summing those in the second columns reproduces S(q) with form factors. The fourth row is the sum of the magnitudes of the net positive and net negative components (the two preceding values in each column). These values are indications of total ordering at a given q, because the positive and negative components are not allowed to cancel (as when they are directly added to reproduce S(q)). The final row is the sum of all values from the plot, and exactly reproduces S(q) with form factor weights, because "Use Form Factor" was selected in the GUI.

- **15.** The "Visualization Settings" menu allows up to five VMD-style visualization commands (inspired by the visualization settings offered by VMD in the VMD Main windows's Graphics, "Representations..." menu). For the first row of "Visualization Settings" set the following: "Selection1: type 1", "Coloring Method: ColorID", "Color ID: 0", "Material: Transparent", "Drawing Method: VDW". Set the "Visualize" radio button to "Top N Atoms Ranked by Contribution". "When Coloring by Beta" can be left at the default setting. For the second row use all the settings for the first row except set "Selection 2: type 2" and "Color ID: 1".
- **16.** The bottom left slider named "Display Top N Atoms Ranked by Contribution" can now be moved to update VMD's visualization window. Move the slider a bit to the right, and then all the way to the left so the slider shows 1. VMD's window now shows a single red atom, corresponding to the atom at the x=1 position in Figure 8. This atom makes the most positive net contribution to S(q) at $q=0.5 \ \text{Å}^{-1}$. In "VMD Main" open the "Mouse" menu and select "Query". Now hover the mouse over the single red atom and press the left mouse button. The VMD terminal shows that atom is index 198 and is type 2 (Oxygen). Return to the "Mouse" menu and select "Center".
 - Moving the slider to the right visualizes additional atoms, in order of their net contribution as in Figure 8. The "Use Form Factor" checkbox in "Precompute" decided whether the visualization would be form factor weighted.
- 17. In Visualization Settings change "Coloring Method" to "Beta" for the first two rows. Moving the bottom left slider now shows atoms colored according to their "Atom Rank", with dark blue being the top ranked atom (index 198) and the atoms becoming lighter blue, then white, light red, and finally dark red to indicate the ranking of atom contributions. The atom to the furthest left in Figure 7 (and analogs of Figure 7) will be dark blue, and the atom to the furthest right will be dark red. Changing "When Coloring by Beta" to "Atom Contribution" colors atoms by their contributions to S(q) in Figure 7, rather than by their ranking in Figure 7. Toggling

the radio button to "Molecules Containing Top N Atoms Ranked by Contribution" displays entire water molecules in order of contributions rather than individual atoms.

- **18.** Return the "Visualization Settings" as they were in Step #15 (i.e. change both "Coloring Method" menus to "ColorID"). The "When Coloring by Beta" setting becomes irrelevant again.
- **19.** When the data from the plot corresponding to Figure 7 is exported it shows that the Oxygen atom with VMD serial 397 is the six greatest net contributor to S(q) at $q = 0.5 \text{ Å}^{-1}$. We can confirm this by moving the "Display Top N Atoms Ranked by Contribution" to 6 and using VMD to query the sixth atom which appears. The query confirms in the VMD terminal that the sixth ranked atom has index 396, which is equivalent to serial 397. We choose atom with serial 397 because it is a highly ranked net contributor to S(q) and because it is conveniently located in the box for the visualization in Step #20.
- **20.** Move the "Display Top N Atoms Ranked by Contribution" to 1. Move the "Central Atom (Serial)" slider to 397 (hint: clicking to the left or right of the slider moves the slider to 1). This selects the atom discussed in Step #19 and renders it as white. There should be one red (Oxygen) atom adjacent to the white atom. The red atom is the atom within r_{max} of the white atom which makes the greatest net contribution to S(q)at $g = 0.5 \text{ Å}^{-1}$ with the white atom. Moving the "Display Top N Atoms Ranked by Contribution" slider to 2 displays a second red atom, which is the second greatest net contributor to S(q) at $q = 0.5 \text{ Å}^{-1}$ with the white atom. Check the "Show Neighbor Ranking Plot" checkbox and move the "Display Top N Atoms Ranked by Contribution" to 3. A plot will appear which ranks atoms by their net contributions to S(q) at $q = 0.5 \text{ Å}^{-1}$ with the white atom (Figure 10). Some atoms make net positive contributions, while others make net negative contributions. At the "Precompute" step had we selected "Positive Atomic Contributions to S(q)" and generated the plot as in Figure 10, all the same atoms (the atoms within rmax of the white atom) would have had different ranking and the y-axis would have had all positive values. Similarly, if we had selected "Negative Atomic Contributions to S(g)" then the same set of atoms would have been ranked according to their negative contributions on the y-axis (we might have wanted to uncheck "Plot/Visualize in Descending Order of Atomic Contributions to S(q)" for the negative case, as it would make the most negative atoms appear on the left of the ranking plot, and the visualization using the "Display Top N Atoms Ranked by Contribution" would have had 1 be the most negatively contributing atom with the white atom, 2 the second most negatively

contributing atom with the white atom, etc.). What is shown in Figure 10 is the sum of those two plots.

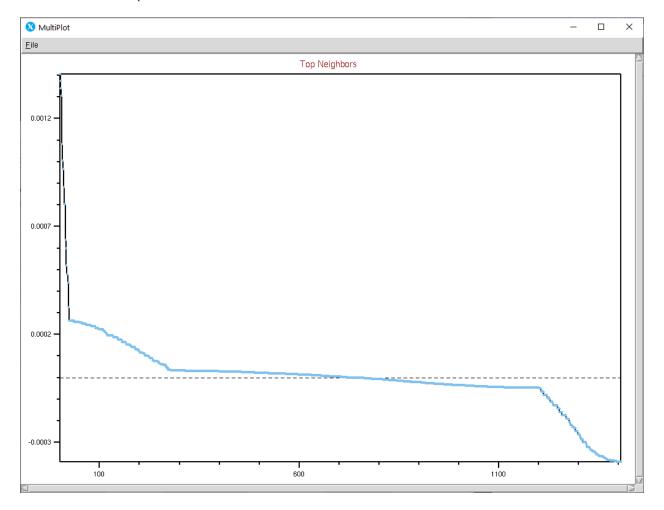


Figure 10. viewSq plot showing the atoms within r_{max} of the "central" Oxygen atom with serial 397 ranked by their net contributions to S(q) with the central Oxygen atom. The x-axis is the atom ranking (showing about 1,400 atoms are within r_{max} of the central Oxygen atom) and the sum of the y-values is equal to the net contribution the central Oxygen atom makes to S(q) at $q = 0.5 \text{ Å}^{-1}$. Some atoms make net positive contributions, while others make net negative contributions.

21. In VMD Main click "Display" and ensure "Depth Cueing" is selected, which will give some depth perception to the visualization scene. Under "Axes" in the same menu select "Off", which will turn off visualization of the axes. Next, in VMD Main click "Graphics", "Colors" and under "Categories" Click Display. In the "Names" column select "Background" and under the "Colors" column select "8 white". This makes the visualization scene have a white background. In VMD Main select "Mouse" and ensure "Rotate Mode" and "Center" are selected.

- 22. In the viewSq window, under "Visualization Settings", for the first row set "Selection 1: type 1", "Coloring Method: ColorID", "ColorID: 0", "Material: Opaque", and "Drawing Method: VDW". For the second row set "Selection 1: type 2", "Coloring Method: ColorID", "ColorID: 1", "Material: Opaque", and "Drawing Method: VDW". Set the "Visualize" radio button to "Top N Atoms Ranked by Contribution" and uncheck "Show Neighbor Raking Plot".
- **23.** Start the "Display Top N Atoms Ranked by Contribution" slider at 1 and move it to the right while watching the VMD visualization window. The mouse can be used to drag and rotate the visualization scene. Figure 11 shows the scene when the slider is set to 100. Note that the central atom has a cluster of red (Oxygen) atoms around it and a red atom shell beginning to form further out. The emerging pattern is representative of the structural ordering underlying (form factor weighted) S(q) at $q = 0.5 \text{ Å}^{-1}$.

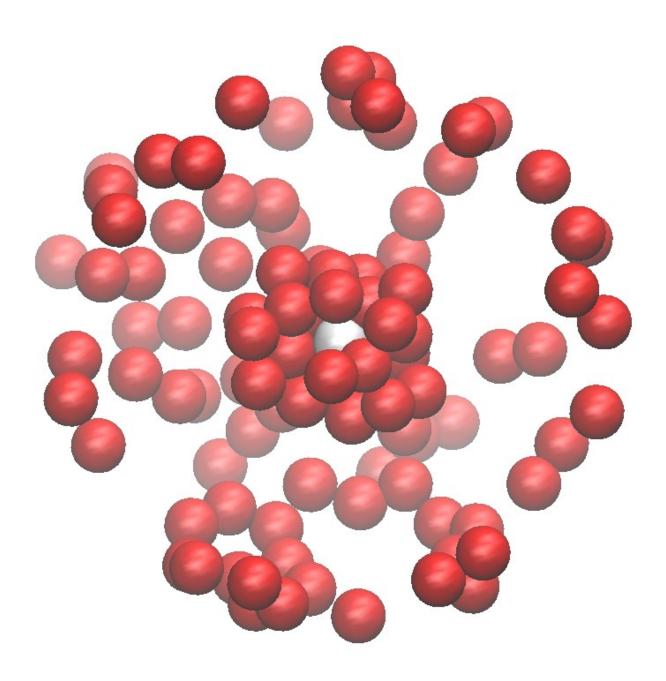


Figure 11. VMD visualization showing the top 100 atoms within r_{max} of the "central" Oxygen atom with serial 397 ranked by their net contributions to S(q) at q = 0.5 Å⁻¹ with the central Oxygen atom. Only red (Oxygen) atoms are visible, meaning no Hydrogen atoms are within the top 100 net contributors, likely because form factors are being accounted for.

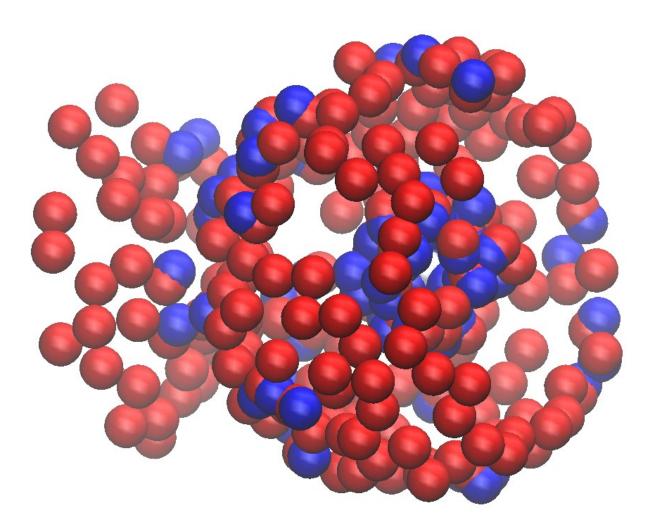


Figure 12. VMD visualization showing the top 325 atoms within r_{max} of the "central" Oxygen atom with serial 397 ranked by their net contributions to S(q) at $q = 0.5 \text{ Å}^{-1}$ with the central Oxygen atom. A mixture of red (Oxygen) and blue (Hydrogen) atoms can be seen. The atoms on the left form are part of the outer "shell" layer but are seen from behind because of period boundary conditions. Effectively those atoms are the "cap" from the shell missing from the right side, which conveniently offers a glimpse inside the outer shell.

24. Under "Visualization Settings" change both of the "Drawing Methods" from "VDW" to "Surf". If your hardware configuration does not show any neighboring atoms for "Surf" you can instead try "QuickSurf", which will be similar. Setting "Display Top N Atoms Ranked by Contribution" to 250 (and with proper rotation) will recreate Figure 13, while setting the slider to 720 will recreate Figure 14.

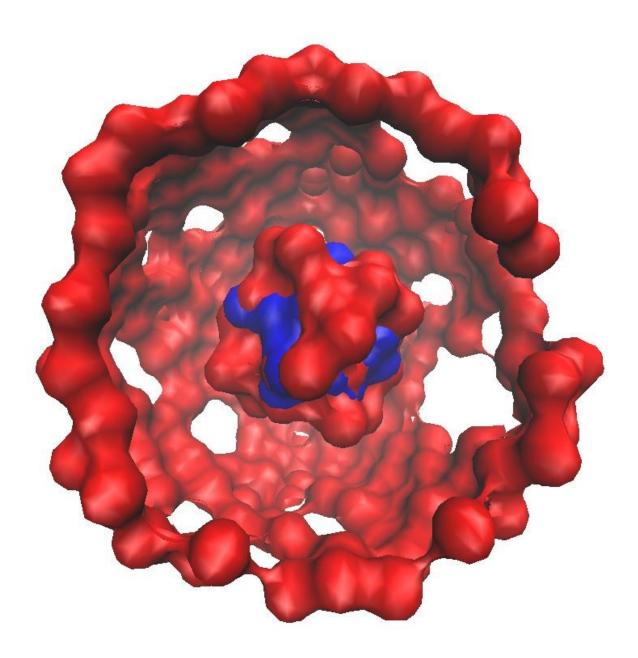


Figure 13. VMD visualization showing the top 250 atoms within r_{max} of the "central" Oxygen atom with serial 397 ranked by their net contributions to S(q) at q = 0.5 Å⁻¹ with the central Oxygen atom. VMD's Surf surface with a probe radius of 1.4 Å was used in order to clarify the inner and outer shells surrounding the central Oxygen atom.

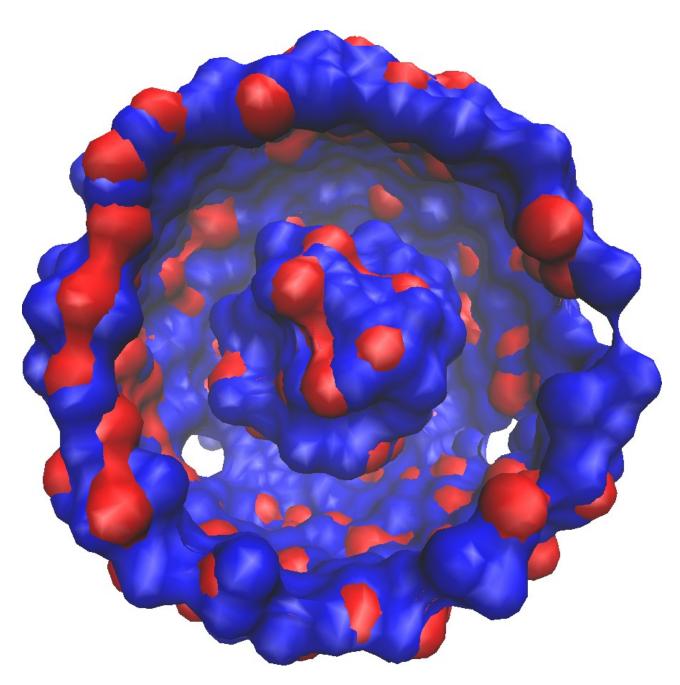


Figure 14. VMD visualization showing the top 720 atoms within r_{max} of the "central" Oxygen atom with serial 397 ranked by their net contributions to S(q) at q = 0.5 Å⁻¹ with the central Oxygen atom. VMD's Surf surface with a probe radius of 1.4 Å was used in order to clarify the inner and outer shells surrounding the central Oxygen atom.