

AreTomo User Manual

Shawn Zheng

Howard Hughes Medical Institute

University of California San Francisco

szheng@msg.ucsf.edu

Revised: September 26, 2022

Version: 1.3.0

AreTomo, the abbreviation for **A**lignment and **R**econstruction for **E**lectron **T**omography, is a GPU accelerated software package that provides a fully automated processing pipeline for the generation of cryoET tomograms from raw tilt series. AreTomo implements our newly developed algorithm for marker-free alignment (Zheng et al., 2022). It also implements weighted back projection (WBP) and simultaneous algebraic reconstruction (SART), resulting in a seamless integration of tomographic alignment and reconstruction in a single software package. Our goal is set to not only generate the high-quality tomograms but also attain high throughput to meet the pressing need of subtomogram averaging. Our alignment scheme is comprehensive that corrects various defects arising from the data collection including translational misalignment, in-plane rotation, tilt-angle offset, and anisotropic local motion due to beam induced motion. The entire workflow can be easily configured via command line parameters to suit various practical needs. It is our best wish that this new release can expedite your quests for uncovering structures at high resolutions and promote the advancement and the interest of cryoET.

AreTomo is compiled on Centos 7 for CUDA 10.1, 10.2, 11.1, 11.2, 11.3, 11.4, 11.5, 11.6, and 11.7. For running on other CUDA platforms, please contact the author for a customized build.

AreTomo is free for academic use only. For commercial use, please contact **David Agard** or **Yifan Cheng** for licensing:

agard@msg.ucsf.edu

Yifan.Cheng@ucsf.edu

Contents

1. System Requirement and Installation.....	4
2. Quick Start.....	4
3. Input of Tilt Angles.....	4
3.1. -TiltRange	5
3.2. -AngFile	5
4. -VolZ.....	5
4.1. Control volume thickness.....	5
4.2. Generate aligned tilt series.....	5
4.3. Reconstruct aligned tilt series.....	6
5. Input tilt axis	6
6. Measure and correct tilt angle offset	6
7. Flip intensity: -FlipInt	7
8. Flip volume: -FlipVol.....	7
9. Reconstruction schemes	7
10. Local motion correction.....	7
10.1. -Patch	8
10.2. Isolated or Sparse Sample: -RoiFile.....	8
11. Dose Weighting.....	8
12. Alignment file.....	9
13. Volume z height for alignment: -AlignZ	9
14. Reconstruction without alignment: -Align 0.....	10
15. Tilt axis determination: -TiltAxis initVal refineFlag	10
16. Output Imod files for subtomogram averaging.....	10
16.1 -OutImod 1 (Relion4)	11
16.2 -OutImod 2 (Warp).....	11
16.3 -OutImod 3 (Global- and local-aligned tilt series).....	11
16.4 Suggestions for subtomogram averaging	11
17. Remove dark images.....	12
Reference	12
1.0.12 Release Note:	13
1.1.0 Release Note:	13
1.2.0 Release Note:	13

1.2.5 Release Note: 13
1.3.0 Release Note: 13
Technical Support 13

1. System Requirement and Installation

AreTomo is a GPU accelerated application compiled on Centos 7. Therefore, a Linux system equipped with an advanced NVidia GPU card is needed. The CPU memory allocated in AreTomo is roughly twice the size of the input tilt series calculated based upon 4 bytes per pixel. For a K3 tilt series containing 120 projection images of 5760x4092 pixels, 21 GB of CPU memory needs to be allocated. Therefore, a Linux system with 32 GB CPU RAM is deemed necessary. Since the current implementation has not optimized yet for GPU, any GPU card that is as good as GTX 1080 card or better is sufficient.

AreTomo is a single-program application. Once unzipped, no further installation is needed if Cuda driver and toolkit have been installed already.

2. Quick Start

AreTomo can be invoked from Linux command line. The version number and built date can be found by running the following command.

AreTomo --version

Users can get a quick help information about the command line parameters by running the following command.

AreTomo --help

When the above two commands can be executed, it is a good sign that the underlying libraries required by AreTomo are properly installed. The following is an example that takes an input tilt series and generated a 3D volume.

AreTomo -InMrc tiltSeries.mrc -OutMrc volume.mrc -VolZ 1200 -OutBin 4

As can be seen, AreTomo requires all the projection images be stored in a single MRC file. In this example the input MRC file is **tiltSeries.mrc**. The output file, **volume.mrc**, stores the generated 3D volume. **-VolZ 1200** sets the volume height or the z dimension to 1200 unbinned voxels. **-OutBin 4** causes the generated volume to be 4x binned before it is saved to disk. Assuming 3838x3710 pixels per tilt image, the unbinned volume has 3838x3710x1200 voxels. **-OutBin 4** reduces the volume to 958x926x300 voxels.

It is important to note that in this example tilt angles are expected being stored as the first float in the extended headers of **tiltSeries.mrc**. If not, users must use other means described in the next section to pass tilt angles into AreTomo. Otherwise, the program will exit immediately.

3. Input of Tilt Angles

Tilt angles are required input parameters by AreTomo. When tilt angles are NOT stored in the extended headers, there are two ways for users to pass them into AreTomo.

3.1. -TiltRange

This option is best suited for a tilt series collected with a fixed angular step and tilt images in the MRC file are sorted according to their tilt angles. **-TiltRange** should be followed by two end angles as shown in the following example.

AreTomo -InMrc tiltSeries.mrc -OutMrc volume.mrc -VolZ 1200 -OutBin 4 -TiltRange -60 60

-TiltRange -60 60 informs AreTomo that **tiltSeries.mrc** is collected in the range of $[-60^\circ, 60^\circ]$ and the tilt images are arranged in ascending order of their tilt angles.

3.2. -AngFile

When a tilt series is not collected with a fixed angular step or its tilt images are not sorted, an angle file is needed that is a single-column text file with all the tilt angles listed in the same order as how the tilt images are stored in the input MRC file. Paired with **-AngFile** as in the following example, this angle file allows AreTomo to link each tilt image with its corresponding tilt angle.

AreTomo -InMrc tiltSeries.mrc -OutMrc volume.mrc -VolZ 1200 -OutBin 4 -AngFile tiltAngles.txt

An angle file can contain an **optional** second column that lists the accumulated dose on sample at each angular step. More details can be found in section 12 Dose Weighting.

4. -VolZ

4.1. Control volume thickness

The z axis refers to the one perpendicular to the projection images. As mentioned earlier, **-VolZ** provides a means for users to specify z height of the reconstructed volume. Ideally, we would like to choose a value such that the reconstructed volume just contains the entire structure with neither clipping nor too many empty voxels. In practice, users can estimate the z dimension based upon sample thickness and the pixel size of projection images and then add 200 voxels for precaution. Another option is to run AreTomo with an educational guess combined with a large binning for **-OutBin**, say 4 or even 8. Upon completion, review the x-z slice of the reconstruction and make corresponding adjustment based upon what you see. Note again the input unit for volume z height is unbinned voxel.

4.2. Generate aligned tilt series

Aligned tilt series will be saved in the output MRC file when **-VolZ 0** is included in the command. The aligned tilt images are sorted in ascending order of tilt angles in the output MRC file with tilt angles saved in MRC extended headers.

When dose weighting is enabled at the command line, the output is the dose-weighted aligned tilt series. If such a tilt series is intended to be reconstructed in the future, make sure to disable dose weighting at the command line. See section 12 for details.

An angle file is also created in the same directory along with the output MRC file. This file has the same name as the output MRC file but ended with “.tlt”. This file can be edited to have a second column for accumulated dose should the aligned tilt series be reconstructed with dose weighting enabled. See section 12 for details.

4.3. Reconstruct aligned tilt series

If users want to reconstruct the aligned tilt series, it can be done by adding **-Align 0** and **-VolZ** followed by a positive integer in the command line. **-Align 0** prevents the alignment process from running on the already aligned tilt series.

5. Input tilt axis

In general, AreTomo does not need to know the orientation of tilt axis. However, if there are several projection images bearing large translational shifts, the tilt axis may not be determined accurately. If you are not satisfied with the quality of the reconstructed volume, you may want to try to bootstrap AreTomo with **-TiltAxis** followed by the user provided value.

```
AreTomo -InMrc tiltSeries.mrc -OutMrc volume.mrc -VolZ 1200 -OutBin 4 -TiltRange -60 60 \  
-TiltAxis 10.5
```

Note that the orientation of tilt axis is relative to the y-axis (vertical axis of tilt image) and rotates counter-clockwise.

6. Measure and correct tilt angle offset

Tilt offset refers to the difference between the actual tilt angles and those read from microscope during data collection. This single-value offset arises due to either sample mounting or milling angle. For a slanted sample, a larger z height than the sample thickness is needed for reconstruction without clipping the reconstructed structure, resulting a larger MRC file to store the bigger volume. See Fig. 1 for the effect of the correction of tilt offset.

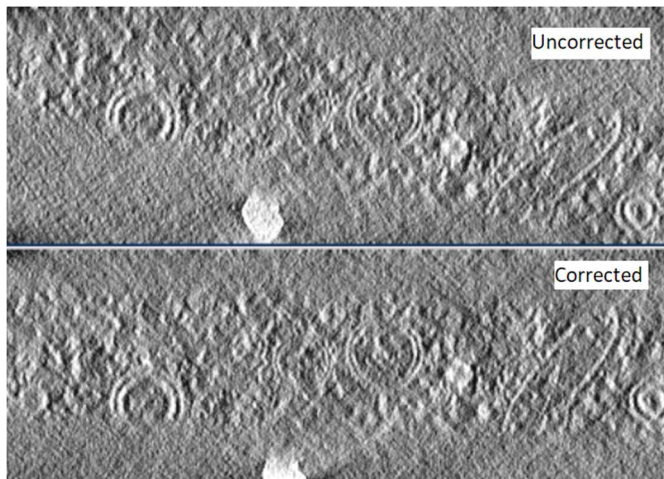


Fig. 1 (Top) Slanted sample reconstructed without the correction of tilt offset. (Bottom) Reconstruction with the correction of tilt offset.

Please note that the orientation of the missing wedge will be changed as a result of the correction of tilt offset. For subtomogram averaging, tomograms reconstructed from tilt series collected within the same tilt range may have different orientations of missing wedges.

-TiltCor can be followed by three values. **-TiltCor 1** results both the measurement and correction of tilt offset. **-TiltCor 0** only measures the tilt offset and is the default setting. The measured offset is corrected only for alignment but not for final reconstruction. **-TiltCor -1** disables the measurement and the correction. Occasionally, the measurement is erroneous and can impair the alignment accuracy.

In the case the tilt offset is not measured properly, an estimated tilt offset can be entered as the second parameter to **-TiltCor**. For example, **-TiltCor 0 10** adds 10° tilt offset to each tilt angle only for the

alignment. **-TiltCor 1 10** adds 10° tilt offset to each tilt angle for both alignment and reconstruction. The tilt offset can be estimated by performing a global only alignment and a quick reconstruction using 4x or even higher binning (**-OutBin 4**) followed by reviewing the xz slice.

7. Flip intensity: **-FlipInt**

By default, the reconstructed volume and the input tilt series use the same grayscale that makes dense structures dark. **-FlipInt 1** makes the dense structures white instead.

8. Flip volume: **-FlipVol**

By default, the x-z slices of the reconstructed volume are saved according to their y coordinates in the output MRC file. **-FlipVol 1** saves x-y slices instead according to their z coordinates.

9. Reconstruction schemes

By default, 3D volumes are computed by SART reconstruction. However, weighted back projection can be enabled by setting **-Wbp 1**.

SART is the default reconstruction with the number of iterations and the number of projections per subset set to 20 and 5, respectively. If users want to change these two settings, **-Sart** can be used for this purpose followed by the number of iterations and number of projections per subset.

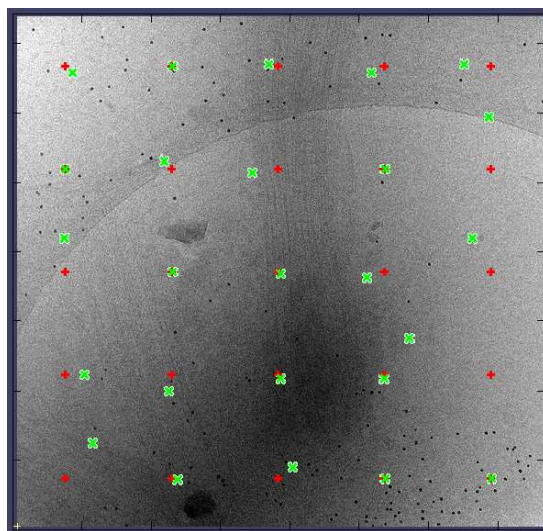


Fig. 2 Uniform and smart grids for local motion measurement and correction. In uniform grid the red-labeled patch centers are evenly distributed as opposed to the green-labeled patch centers in smart grid where patch centers are kept away from empty areas.

10. Local motion correction

Beam induced motion significantly complicates the situation in cryoEM tomography. Specimen movement results from not only stage tilting but also beam induced motion. Tomographic alignment premised only on single-axis rotation, regardless fiducial based or not, are not accurate enough for high-

resolution tomography if beam induced local motion is left uncorrected. AreTomo provides two means to correct the local motion.

10.1. -Patch

-Patch is followed by two positive integers, e.g. **-Patch 5 4**, that specify number of patches in x and y directions, respectively, where local motions are measured and then corrected. Prior to version 1.2.0, the grid of patches is uniform where the patch centers are evenly distributed. AreTomo 1.2.0 implements the smart grid that dodges the empty areas for local motion measurement. The uniform and

10.2. Isolated or Sparse Sample: -RoiFile

-RoiFile requires users to provide a list of x and y coordinates and put them in a two-column text file, one column for x and the other for y. Each pair defines a region of interest (ROI). AreTomo then performs focused alignment as described in section 11 on the structure in each of these ROIs. Local motion is then interpolated based upon a distance-based scheme and corrected at each pixel of each projection image before 3D reconstruction is performed. **Fig. 2** shows an example where nine targets were selected along the elongated structure. As mentioned before, the origin of the coordinate system is at the image's lower left corner.

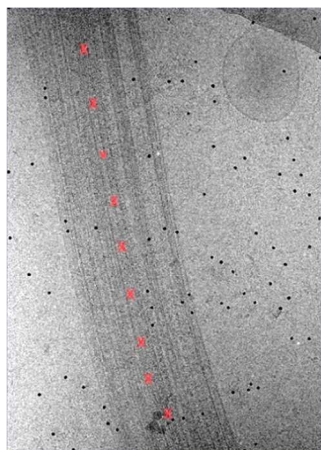


Fig. 3 Illustration of target selection along an elongated structure required to activate local motion correction.

11. Dose Weighting

Once the input tilt series is aligned, dose weighting can be applied to the aligned tilt series before it gets reconstructed. To activate dose weighting, high tension (**-Kv**), pixel size in angstrom (**-PixSize**), and the accumulated dose in e/A^2 at each tilt angle are needed. AreTomo uses the angle file (**-AngFile**) described in section 3.2 to associate accumulative dose with tilt angle. Since the angle file lists all the tilt angles of the images in a single column, we can add a second column for the accumulated dose to associate it with the corresponding tilt angle. The following is an example of dose augmented angle file. This column indicates that the tilt images in MRC file was sorted in ascending order of tilt angles from -

60.0° to 36.0°. The second column shows that the first acquisition took place at -12.0°. An individual tilt image received a dose of $2e/A^2$.

-60.0	66.0
-57.0	64.0
-54.0	58.0
.....
-15.0	8.0
-12.0	2.0
-9.0	4.0
.....
30.0	54.0
33.0	60.0
36.0	52.0

The required parameters for dose weighting should be similar to the following example. Missing any of them will disable dose weighting. The second column is only needed for dose weighting. Otherwise, it is optional.

-Kv 300 -PixSize 3.5 -AngFile PathToAngFile/myAngFile.txt

12. Alignment file

AreTomo automatically generates a text file containing both global and local alignment parameters. Named after the input MRC file, this file uses **aln** as file extension and is saved in the output directory containing the reconstructed volume. Should users decide to reconstruct the volume with a different setting, the corresponding aln file can be loaded by means of **-AlnFile** to bypass the lengthy alignment process, as is illustrated in the following example.

**AreTomo -InMrc tiltSeries.mrc -OutMrc ~/Temp/volume.mrc -AlnFile ~/Temp/tiltSeries.aln **

-VolZ 1200 -OutBin 1

13. Volume z height for alignment: -AlignZ

This function specifies the z height of the temporary volume reconstructed for projection matching as part of the alignment process. This value plays an important role in alignment accuracy. The default value is 600. This z height should be always smaller than that specified by **-VolZ**, as illustrated in Fig. 4. As can be seen, the z height for alignment is preferred to be close to the sample thickness whereas the z height for final reconstruction should be larger.

In practice we can perform global only alignment and 3D reconstruction with a large binning. The sample thickness can then be measured in the xz slice of the volume.

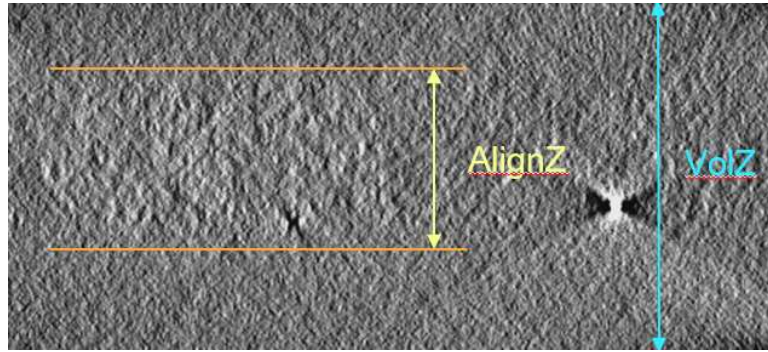


Fig. 4 Illustration of **-AlignZ** and **-VolZ** in a xz slice of a reconstructed volume.

14. Reconstruction without alignment: **-Align 0**

This option lets users reconstruct tomogram without aligning the input tilt series, which must be, therefore, pre-aligned either by AreTomo or some other software. One application of this option is that a raw tilt series is first aligned and then CTF-corrected. The aligned and CTF-corrected tilt series is finally input into AreTomo to reconstruct CTF-corrected tomogram.

15. Tilt axis determination: **-TiltAxis initVal refineFlag**

Tilt axis determination is a two-step processing in AreTomo. A single tilt axis is first calculated followed by the determination of how tilt axis varies over the entire tilt range. The “initVal” lets users enter their estimate and AreTomo refines the estimate in $[-3^\circ, 3^\circ]$. If users do not know, they can enter 0 instead and AreTomo will search an optimal value in full range. The “refineFlag” can be one of -1, 0, and 1. The -1 instructs AreTomo to the user input value without any refinement. The default setting is 0 and 1.

refineFlag	Meaning
-1	Use the input value without any refinement
0	Refine the input value and use the refined value for the entire tilt series
1	Refine the input value and calculate tilt axis at each tilt angle.

16. Output Imod files for subtomogram averaging

Since version 1.1.0 AreTomo can output the files needed for Relion4 and Warp to start subtomogram averaging. The setting **-OutImod** is used to generate the needed files for this purpose. A subfolder named after the output MRC file is created, if not present, under the output directory. For example, “**-OutMrc MyDir/MyTomo.mrc -OutImod 1**” creates the subfolder **MyTomo_Imod** in **MyDir**. The generated files listed in the following table are placed in the subfolder.

xxx.st	MRC file of either aligned or raw tilt series without dark images
xxx.tlt	Single-column text file containing the tilt angles
xxx.xf	Imod transformation file

newst.com	lmod file
tilt.com	lmod file

Note that -Outlmod 0 disables this function and does not create the subfolder and the corresponding files and this is the default setting.

16.1 -Outlmod 1 (Relion4)

This setting is for Relion4 subtomogram averaging and does not generate [xxx.st](#) file since the original tilt series is used as input. The path of the original tilt series is given in the [newst.com](#).

AreTomo may reject the dark images in its alignment process when they are detected, see Section 17 for details. In this case, [tilt.com](#) contains a line (EXCLUDELIST) that lists all the rejected images. To maintain the data consistency across different files, [xxx.xf](#), while containing the transformation matrices for the included images, also have the unit matrices, for the rejected image. [xxx.tlt](#) lists the tilt angles of all images including the rejected ones. The tilt angles are ordered according to the images in the original tilt series.

When -TiltCor 1 is used, AreTomo may change the tilt angle of each image. In this case, the angles in [xxx.tlt](#) are the adjusted values that are different from those in [xxx.rawtlt](#) file if there is one. Use the [xxx.tlt](#) file as input to Relion4 if AreTomo generated tomograms are used to select particles.

16.2 -Outlmod 2 (Warp)

This setting is intended for Warp based subtomogram averaging. It generates a [xxx.st](#) file that stores all the original images except the rejected ones. Correspondingly, [xxx.tlt](#) and [xxx.xf](#) contain only the entries for the included images. [newst.com](#) has a line pointing to the [xxx.st](#) file. [tilt.com](#) does NOT have a line for EXCLUDELIST.

Note that, since the subfolder contains all the files needed for Relion4, this setting is expected to meet the Relion4 requirement, although we have not tested the compatibility.

16.3 -Outlmod 3 (Global- and local-aligned tilt series)

An accurate tomographic alignment is expected to be beneficial for subtomogram averaging, allowing it to start at the vicinity of ground truth. Unfortunately, [xxx.xf](#) file can only have the transformation matrices of the global alignment, leaving no room for the local alignment parameters determined by AreTomo. This setting provides a path for subtomogram averaging to take the advantage of AreTomo's local alignment. To this end, the aligned and dark image-rejected tilt series is generated and saved in the [xxx.st](#) file. As a result, the [xxx.xf](#) file is filled with unit transformation matrices to prevent [xxx.st](#) from being aligned again. Since the interpolation in the generation of the aligned tilt series attenuates the high-frequency information, verison 1.3.0 intentionally enhances the high-frequency components of the raw images beforehand to alleviate the attenuation due to interpolation.

16.4 Suggestions for subtomogram averaging

Since AreTomo may exclude dark images and adjust tilt angles, it is recommended to use AreTomo generated tomograms for particle picking and then feed the needed file in the created subfolder to the subsequent processing pipeline. The mismatch between tilt angles, tilt series, and tomograms, potentially severe for subtomogram average, can be avoided.

17. Remove dark images

Since version 1.1.0, AreTomo allows users to decide how stringent the dark images are detected and removed using “-DarkTol threshold”. The default threshold is 0.7. Lowering the threshold means less images are removed. The removed dark images are listed by their indices in a newly created text file named after the input MRC file name and appended with “_DarkImgs.txt”.

Reference

Zheng S, Wolff G, Greenan G, Chen Z, Faas FGA, Bárcena M, Koster AJ, Cheng Y, Agard DA. [AreTomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction](#). J Struct Biol X. 2022 May 10;6:100068. doi: 10.1016/j.yjsbx.2022.100068. PMID: 35601683; PMCID: PMC9117686.

1.0.12 Release Note:

1. The most important change made in this version is that the handedness of the volume has been changed to the same as that in IMOD. The z axis of volumes is flipped 180° as opposed to those reconstructed by earlier versions.
2. Version 1.0.11 and 1.0.12 fixed a bug that causes random, although rare, crash.
3. Added a new function that allows users to input their own values of tilt axis and decide how they want to refine their values.

1.1.0 Release Note:

1. The most important change made in this version is the new function that outputs the required files needed for Relion4 to perform subtomogram averaging.
2. Added “-DarkTol” to let users decide how dark images are detected and removed.
3. Removed the padding in the x dimension of the reconstructed volume so that the x and y dimensions are the same as the input tilt series.

1.2.0 Release Note:

1. Implemented smart grid for more robust local motion measurement and correction.
2. Improved projection-matching based alignment.
3. Implemented tilt axis refinement based upon local motion measurement.

1.2.5 Release Note:

1. Improved multi-GPU implementation for patch based local motion alignment.
2. Fixed the bug that the flipped volume by “-FlipVol 1” does not match the volume flipped in IMod by rotation around x axis.

1.3.0 Release Note:

1. Reimplemented “-OutImod” to make AreTomo output more easily integrated into subtomogram averaging of Relion4 and Warp.
2. Improved local alignment, less sensitive to “-AlignZ” setting.
3. Improved screening of abnormal local measurements.
4. Applied a weak high-pass filter to enhance the high-frequency components of raw images before they are interpolated to generate the aligned images.
5. Adopted the Imod image coordinate system where the origin is at $(N_x * 0.5, N_y * 0.5)$. The pixel coordinate is at the center (0.5, 0.5). $x=0$, N_x is the left and right edges of the image, respectively.
6. The .aln file contains an extra column to indicate good or bad local measurements.
7. Fixed the bug that misses the mean, max, and min in the MRC header.

Technical Support

Shawn Zheng

S126B, Genentech Hall, 600 16th St, San Francisco, CA 94513

szheng@msg.ucsf.edu