

## AUTO3DEM keyword input

AUTO3DEM uses a single keyword-based input file for specifying the AUTO3DEM control parameters, the parameters for the numerically intensive image reconstruction applications that are called by AUTO3DEM (P3DR, PCUT, PO2R, PPFT, and PSF), and the names of the data files. The following rules are applicable to the entire input file.

- **The ordering of lines is unimportant.** However, if a keyword is specified more than once, the value associated with the later occurrence will override that for the earlier occurrence.
- **Extra whitespace is ignored.** Leading/trailing whitespace and blank lines are ignored. Contiguous blocks of whitespace are treated the same as a single space. Embedded whitespace is not allowed in character string input. For example, 'data file 1' would not be a valid string.
- **The hash/number (#) sign indicates start of comment.** Entire lines can be commented out, or comments can be added to the end of a line. Hash signs are not allowed in character strings.
- **Extra fields are ignored.** Input parameters are specified using three fields, while data files only use two fields. Any data appearing after the end of the last required field is ignored. *The one exception to this rule is that the email recipient field may consist of multiple addresses separated by whitespace and/or commas.*
- **Fields are case insensitive, except for character string input.** Identifiers and keywords (described below) are internally converted to lowercase. Character strings specifying directories, file names, and binaries must be typed using the correct case.

### ***AUTO3DEM and program control parameters***

All input parameters, with the exception of the data file names, are specified using a three-field format.

```
identifier key value
```

The first field (identifier) is used to distinguish whether the record contains an AUTO3DEM control parameter or input for one of the image processing programs. The second and third fields form key-value pairs corresponding to the name of the input parameter and its values. The following case-insensitive values are allowed:

- **auto** – AUTO3DEM control parameter
- **all** – input parameter used by all programs
- **p3dr** – P3DR input parameter
- **pcut** – PCUT input parameter
- **po2r** – PO2R input parameter
- **por** – alternative specification for PO2R input parameter to allow back compatibility with earlier versions of AUTO3DEM
- **ppft** – PPFT input parameter
- **psf** – PSF input parameter

A description of the allowed keys for each identifier is given below.

## auto

For the sake of clarity, the AUTO3DEM input parameters are divided into two sets. The first set contains the general parameters, while the second contains the parameters related to particle selection criteria.

### general parameters

- **bin\_reduce** – controls whether or not PPFT bin\_factor should be reduced when resolution of reconstruction fails to improve.
- **boxrad** – radius (pixels) of image box
- **delete\_maps** – non-zero value specifies that the maps generated at intermediate stages of the reconstruction should be deleted.
- **estimate\_res** – non-zero value indicates that resolution estimation is performed.
- **freeze\_annulus** – non-zero value freezes inner and outer radii of the annulus defining the ordered region of the map. This affects the parameters annulus\_low and annulus\_high in PPFT and in\_rad and out\_rad in PCUT.
- **freeze\_res** – non-zero value freezes the resolutions used in PSF, P3DR, PPFT, and PO2R.
- **fsc\_hithresh** – cutoff value for FSC used in estimating map resolution.
- **fsc\_lothresh** – cutoff value for FSC used to set resolution limits in P3DR, PSF, and PO2R.
- **gauss\_adj** – parameter used to set width of Gaussian falloff in P3DR.
- **have\_map** – non-zero value indicates that starting map is available.
- **hollow\_cut\_step** – number of steps used by masking algorithm when generating a hollow map
- **hollow\_cut\_weight** – weight used by masking algorithm when generating a hollow map
- **hollow\_in\_rad** – inner radius of hollowed map
- **hollow\_map** – flag specifying whether or not map should be hollowed
- **hollow\_out\_rad** – outer radius of hollowed map
- **iter\_start** – starting iteration
- **mode** – AUTO3DEM mode of operation. Allowed values = (search, refine).
- **new\_ptles** – flag specifying whether or not new particles should be oriented relative to existing map without updating the map
- **niter** – maximum number of iterations of AUTO3DEM main loop
- **noctf** – if true, disables CTF correction. Overrides CTF mode and sets to zero for programs P3DR, PO2R, PPFT, and PCTFR. Used primarily with image sets which have already been CTF corrected.
- **noise\_suppression** – apply Rosenthal and Henderson JMB 333 721-745 (2003) noise suppression algorithm
- **outfile** – base name used to construct names of log, summary, restart, and continuation files
- **per\_ptle\_ctf** – apply CTF correction on a per-particle basis in P3DR, PO2R, and PPFT. Setting to one overrides per\_ptle\_ctf parameter set for individual programs.
- **quit\_early** – set to non-zero value to have AUTO3DEM quit if the FSC curve never drops below fsc\_hithresh (usual value is fsc\_hithresh=0.5). This option is

normally only used for random model calculations where it is set automatically by setup\_rmc.pl

- **refine\_ctf** – refine CTF parameters at end of each iteration when running in refine mode
- **res\_adj** – parameter that determines the higher resolution to which map will be calculated beyond upper resolution limit used in PO2R.
- **restart** – set to 1 to continue calculation
- **rundir** – directory containing input data (maps, images, particle parameters)
- **start\_map** - name of starting map used by AUTO3DEM.
- **switch\_mode** – if true, allows auto3dem to automatically switch from search to refine mode
- **symm\_code** – symmetry code
- **term\_refine** – allow automatic termination of run when in refine mode (functionality not currently active, added as placeholder)
- **term\_search** – allow automatic termination of run when in search mode (functionality not currently active, added as placeholder)

#### particle selection parameters

- **box\_center\_offset** – maximum allowable distance between the center of the particle and the center of the box; applied separately to each coordinate. Particles with centers too far from center of box are excluded from the model.
- **cmp\_cc\_fraction** – fraction of images to accept on the basis of the CMP correlation coefficient. Makes sense only when parsing particle parameter files generated in search mode.
- **cmp\_cc\_nstd** – number of standard deviations to add to the average CMP correlation coefficient when setting cutoff. Negative values are less restrictive, positive values are more restrictive.
- **global\_select** – if set to true (non-zero) value, then selection criteria are applied globally across particle parameter files. Otherwise, selection criteria applied on a per file basis
- **nselect\_offset** – number of selection criteria to evaluate in each ‘direction’ from the central selection criterion. The total number of selection criteria to be evaluated is  $(2 * nselect\_offset + 1)$ .
- **pft\_cc\_fraction** – fraction of images to accept on the basis of the PFT correlation coefficient. Makes sense only when parsing particle parameter files generated in search mode.
- **pft\_cc\_nstd** – number of standard deviations to add to the average PFT correlation coefficient when setting cutoff. Negative values are less restrictive, positive values are more restrictive.
- **phi\_reject\_lower / phi\_reject\_upper** – range of azimuthal angles ( $\phi\_reject\_lower \leq \phi \leq \phi\_reject\_upper$ ) for which images will be excluded from map construction.
- **prj\_cc\_fraction** – fraction of images to accept on the basis of the PRJ correlation coefficient. Makes sense only when parsing particle parameter files generated in search mode.

- **prj\_cc\_nstd** – number of standard deviations to add to the average PRJ correlation coefficient when setting cutoff. Negative values are less restrictive, positive values are more restrictive.
- **score\_fraction** – fraction of images to accept on the basis of the score generated by program PO2R.
- **score\_nstd** – number of standard deviations to add to the average score when setting cutoff. Negative values are less restrictive, positive values are more restrictive.
- **select\_delta** – the size of the ‘step’ to be used when evaluating multiple selection criteria. For standard deviation-based criteria, adds a fixed number of standard deviations; for fraction-based criteria, adds a fixed fraction.
- **theta\_reject\_lower / theta\_reject\_upper** – range of inclination angles ( $\theta_{\text{reject\_lower}} \leq \theta \leq \theta_{\text{reject\_upper}}$ ) for which images will be excluded from map construction.

### p3dr

- **apo\_border** – width of border region for map apodization
- **bin** – name of P3DR binary
- **ctf\_ff1** – 1<sup>st</sup> CTF filter factor
- **ctf\_ff2** – 2<sup>nd</sup> CTF filter factor
- **ctfmode** – CTF mode (ctf\_mode also accepted)
- **fsc\_file\_name** – name of FSC file to be used when applying noise suppression algorithm. (File format: line 1 = number of FSC records; subsequent lines = spatial frequency ( $\text{\AA}^{-1}$ ) FSC value)
- **filter** – filter mode
- **magfactor** – magnification factor
- **map\_dim** – map dimension
- **max\_cpu** – maximum number of CPUs to be used by P3DR
- **per\_ptle\_ctf** – apply CTF correction on a per-particle basis
- **res\_max** – resolution at end of Gaussian falloff
- **res\_min** – resolution to which map is computed
- **symm\_code** – symmetry code
- **tempfac** – temperature factor
- **zero\_fill** – zero fill for background pixels

### pctfr

- **bin** – name of PCTFR binary
- **ctf\_ff1** – 1<sup>st</sup> CTF filter factor
- **ctf\_ff2** – 2<sup>nd</sup> CTF filter factor
- **ctfmode** – CTF mode (ctf\_mode also accepted)
- **dangle** – CTF parameter step size
- **filter** – filter mode
- **funcmode** – function mode
- **funcweight** – function weight
- **max\_cpu** – maximum number of CPUs to be used by PCTFR
- **nangle** – number of steps in CTF parameters taken in each direction

- **res\_max** – maximum resolution used in image/projection comparison
- **res\_min** - minimum resolution used in image/projection comparison
- **tempfac** – temperature factor
- **zero\_fill** – zero fill for background pixels

## pcut

- **bin** – name of PCUT binary
- **cut\_step** – number of steps used by masking algorithm
- **cut\_weight** – weight used in masking algorithm
- **in\_rad** – inner radius for masking
- **max\_cpu** – maximum number of CPUs to be used by PCUT
- **out\_rad** – outer radius for masking

## po2r

- **bin** – name of PO2R binary
- **ctf\_ff1** – 1<sup>st</sup> CTF filter factor
- **ctf\_ff2** – 2<sup>nd</sup> CTF filter factor
- **ctfmode** – CTF mode (ctf\_mode also accepted)
- **dangle** – angular step size (delta angle)
- **dcenter** – spatial step size (delta xy)
- **filter** – filter mode
- **funcmode** – function mode
- **funcweight** – function weight
- **handtest** – enable/disable handedness tests for images
- **magfstep\*** – size of magnification refinement step
- **max\_cpu** – maximum number of CPUs to be used by PO2R
- **nangle** – number of angular steps taken in each direction
- **ncenter** – number of spatial steps taken in each direction
- **nmagf\*** – number of magnification factor steps
- **per\_ptle\_ctf** – apply CTF correction on a per-particle basis
- **quick\_search** – implement quick approximate search of orientation space
- **res\_max** – maximum resolution used in image/projection comparison
- **res\_min** - minimum resolution used in image/projection comparison
- **symm\_code** – symmetry code
- **tempfac** – temperature factor
- **zero\_fill** – zero fill for background pixels

*\*Not settable through auto3dem. PO2R standalone feature*

## ppft

The PFTsearch/PPFT input parameters **pftrads\_filename**, **pftres1\_filename**, and **pftres2\_filename** are not read from the input parameter file since they are set by AUTO3DEM. They are assigned the values **ppft\_iter\_n.rads**, **ppft\_iter\_n.res1**, and **ppft\_iter\_n.res2**, respectively, where *n* is the iteration number.

- **annulus\_high** – outer radius of annulus for image/projection comparison

- **annulus\_low** – inner radius of annulus for image/projection comparison
- **bin** – name of PPFT binary
- **bin\_factor** – binning factor
- **ctf\_mode** – CTF mode (ctfmode also accepted)
- **delta\_theta** – step size for inclination angle theta
- **filter\_factor\_1** – 1<sup>st</sup> filter factor
- **input\_mode** – input mode
- **jcut** – minimum order Bessel function (J<sub>n</sub>), default strongly recommended!
- **mag\_cen** – midpoint for magnification scale search
- **mag\_norm** – switch used to normalize the MAG scale factors so that the average MAG is 1.0.
- **mag\_num** – extent of magnification search window
- **mag\_step** – grid size of magnification scale search
- **max\_cpu** – maximum number of CPUs to be used by PPFT
- **per\_ptle\_ctf** – apply CTF correction on a per-particle basis
- **pft\_filename** – name of PFT file
- **pftrad\_hi** – outer PFT radius
- **pftrad\_lo** – inner PFT radius
- **pftrad\_step** – PFT radius step size
- **prj\_filename** – file name for prj output
- **quick\_omega** – perform fast approximate search for omega
- **resolution\_high** – upper resolution limit
- **resolution\_low** – lower resolution limit
- **sigcut** – threshold for variance mask when filtering PFT data, default strongly recommended!
- **symmetry** – symmetry code
- **symm\_code** – alternative for specifying symmetry code
- **temperature\_factor** – temperature factor
- **verbose** – verbose factor (controls level of output)

## psf

The PSF input parameters **pixel\_size** is not read from the input parameter file since it is obtained by parsing the first line of the particle parameter files.

- **bin** – name of PSF binary
- **max\_cpu** – maximum number of CPUs to be used by PSF
- **res\_max** – maximum resolution used in FSC calculation
- **res\_min** – minimum resolution used in FSC calculation
- **res\_step** – resolution step size

## Data files

All data files are specified using a two-field format.

```
data filename
```

The first field in the record must be the keyword data. The data lines can appear anywhere in the file and the filenames are case sensitive.

### ***Minimal required input***

Default values can be used for the majority of the auto3dem input parameters, but some values must still be supplied. The listing below shows an example minimal input file.

```
auto      mode          search      # Search mode (using ppft)
auto      niter          5           # 5 iterations
auto      start_map      startmap.pif # Starting map
p3dr      res_min        8.5         # Resolution to which map is computed
data      datafile_1     # At least one data file required
```

### ***Better Performance***

If the dimensions of the virus are known, better performance and reconstruction quality can be achieved by specifying the following parameters

```
auto_freeze_annulus 1 # Don't update annulus_low/high or in_rad/out_rad
ppft annulus_low    n # inner radius of capsid (including protrusions)
ppft annulus_high   n # outer radius of capsid (including protrusions)
pcut in_rad         n # inner radius of capsid (excluding protrusions)
pcut out_rad        n # outer radius of capsid (excluding protrusions)
```

If the images have a small pixel size, can often use the following combination of parameters to start search mode calculations using binned image data

```
auto bin_reduce      1 # Automatically reduce bin_factor
ppft bin_factor      2 # Start with 2x2 binning of images
```

If using all particles for reconstruction in search mode and not concerned about correlation coefficients for particles, can avoid unnecessary calculations using following option

```
ppft verbose         -1 # Skip additional correlation coefficient calculations
```

## ***Complete listing of parameters and default values***

The following table lists the auto3dem input parameters along with their default values. A missing default value means that no default is used. Required input shown in **bold red**.

auto	bin_reduce	0
auto	boxrad	Extracted from PIF <sup>1</sup>
auto	box_center_offset	10000
auto	cmp_cc_fraction	
auto	cmp_cc_nstd	
auto	delete_maps	1
auto	estimate_res	1
auto	freeze_annulus	0
auto	freeze_res	0
auto	fsc_hithresh	0.5
auto	fsc_lothresh	0.3
auto	gauss_adj	0.01
auto	global_select	1
auto	have_map	1
auto	hollow_cut_step	12
auto	hollow_cut_weight	0.001
auto	hollow_in_rad	
auto	hollow_map	0
auto	hollow_out_rad	
auto	iter_start	1
<b>auto</b>	<b>mode</b>	
auto	new_ptles	0
<b>auto</b>	<b>niter</b>	
auto	noctf	0
auto	noise_suppression	0
auto	nselect_offset	0
auto	outfile	Current working directory
auto	per_ptle_ctf	0
auto	pft_cc_fraction	
auto	pft_cc_nstd	
auto	phi_reject_lower	360
auto	phi_reject_upper	-360
auto	prj_cc_fraction	
auto	prj_cc_nstd	
auto	quit_early	0
auto	refine_ctf	0
auto	res_adj	0.01
auto	restart	0
auto	rundir	dat
auto	score_fraction	



auto	select_delta	0
auto	select_nstd	
<b>auto</b>	<b>start_map</b>	See note <sup>2</sup>
auto	switch_mode	1
auto	symm_code	
auto	term_refine	0
auto	term_search	0
auto	theta_reject_lower	360
auto	theta_reject_upper	-360

pcut	bin	PCUT
pcut	cut_step	12
pcut	cut_weight	0.001
pcut	in_rad	boxrad/3
pcut	max_cpu	8
pcut	out_rad	boxrad - 10

p3dr	apo_border	12
p3dr	bin	P3DR
p3dr	ctf_ff1	0.05
p3dr	ctf_ff2	0.1
p3dr	ctfmode	1
p3dr	fsc_file_name	
p3dr	filter	1
p3dr	magfactor	1.0
p3dr	map_dim	0
p3dr	max_cpu	256
p3dr	per_ptle_ctf	0
p3dr	res_max	$(p3dr\{res\_min\}^{-1} + auto\{gauss\_adj\})^{-1}$ (See note <sup>3</sup> )
<b>p3dr</b>	res_min	
p3dr	symm_code	532
p3dr	tempfac	0
p3dr	zero_fill	1

pctfr	bin	PCTFR
pctfr	ctf_ff1	0.5
pctfr	ctf_ff2	0.1
pctfr	ctfmode	1
pctfr	dangle	0.05
pctfr	filter	1
pctfr	funcmode	3
pctfr	funcweight	0
pctfr	max_cpu	256
pctfr	nangle	4

pctfr	res_max	$(p3dr\{res\_min\}^{-1} + auto\{gauss\_adj\})^{-1}$
pctfr	res_min	$(2/5) \cdot boxrad \cdot pixel\_size$
pctfr	tempfac	0
pctfr	zero_fill	1

po2r	bin	POR
po2r	ctf_ff1	0.5
po2r	ctf_ff2	0.1
po2r	ctfmode	1
po2r	dangle	$(1/2) \cdot ppft\{\delta\_theta\}$
po2r	dcenter	0.1
po2r	filter	1
po2r	funcmode	1
po2r	funcweight	0
po2r	handtest	1
po2r	max_cpu	256
po2r	nangle	4
po2r	ncenter	4
po2r	per_ptle_ctf	0
po2r	quick_search	1
po2r	res_max	$(p3dr\{res\_min\}^{-1} + auto\{gauss\_adj\})^{-1}$
po2r	res_min	$(2/5) \cdot boxrad \cdot pixel\_size$
po2r	symm_code	532
po2r	tempfac	0
po2r	zero_fill	1

ppft	annulus_high	boxrad/3
ppft	annulus_low	boxrad - 10
ppft	bin	PPFT
ppft	bin_factor	1
ppft	ctf_mode	3
ppft	delta_theta	0.5
ppft	filter_factor_1	0.1
ppft	input_mode	2
ppft	jcut	1
ppft	mag_cen	1
ppft	mag_norm	1
ppft	mag_num	1
ppft	mag_step	0
ppft	max_cpu	256
ppft	per_ptle_ctf	0
ppft	pft_filename	pft.pfts
ppft	pfrad_hi	auto{boxrad}
ppft	pfrad_lo	1.0
ppft	pfrad_step	1

ppft	prj_filename	pft.prjs
ppft	quick_omega	1
ppft	resolution_high	$(p3dr\{res\_min\}^{-1} + auto\{gauss\_adj\})^{-1}$
ppft	resolution_low	$(2/5) \cdot boxrad \cdot pixel\_size$
ppft	sigcut	0
ppft	symmetry	532
ppft	symm_code	532
ppft	temperature_factor	0
ppft	verbose	2

psf	bin	PSF
psf	max_cpu	8
psf	res_max	$p3dr\{res\_min\}$
psf	res_min	60
psf	res_step	50

<sup>1</sup> Boxrad is normally extracted from the PIF file header field packRadius. The option to specify boxrad in the auto3dem parameter file is provided for use in those cases where this information is missing from the PIF header.

<sup>2</sup> If the have\_map flag is false, then starting map is not required.

<sup>3</sup>Subject to Nyquist condition that resolution is not less than twice the pixel size.