

---

# CTFMIX

---

Notes by James Conway, 26-Mar-2001. (Not complete)

## Contact me:

Laboratoire de Microscopie Electronique Structurale  
 Institut de Biologie Structurale  
 41 rue Jules Horowitz  
 38027 Grenoble Cedex 1  
 France  
 Telephone: +33-(0)476-884568  
 Fax: +33-(0)476-885494  
 Email: James.Conway@ibs.fr

## Contents:

1. Overview .....	1
2. Sample Data .....	2
3. CTFMIX & single datasets – phase correction .....	2
4. CTFMIX & defocus pairs – phase correction .....	5
5. CTFMIX & defocus pairs – deconvolution.....	6
6. References .....	9

---

## 1. Overview

The purpose of CTFMIX is to apply a correction for the contrast transfer function of the electron microscope to the scanned 2D images, so as to allow for data beyond the first phase reversal to be used in orientation refinement and 3D reconstruction. A write-up is available in reference [1]. This current document explains use of the CTFMIX program. Some knowledge of the reconstruction methodology is assumed. I refer to other programs that I wrote myself, as well as the ‘core’ software for orientation refinement (PFT) and 3D reconstruction (EM3DR) by Tim Baker (Purdue University, Indiana) and colleagues. A list is given here in the order the programs would be used:

- x3d – extract images of particles from a digitized micrograph
- SumPS – calculate average power spectrum from the extracted particle images
- ctfzeros – fit defocus to the positions of phase reversal assuming phase contrast
- ctfmix – apply corrections for CTF and a Gaussian decay envelope
- PFT – polar Fourier transform-based orientation determination
- EM3DR – Fourier-Bessel reconstruction

Separate documents outline use of these programs – contact me for copies. PFT and EM3DR software and documents are available from Tim Baker (see references [2] & [3], and <http://bilbo.bio.purdue.edu/~baker/>).

CTFMIX produces a series of output images that are to be used by PFT and EM3DR in place of the ‘raw’ images. Several variants are described, according to my own use of these programs. In particular, the correction for phase reversals is a relatively benign operation which allows spatial frequencies beyond the first reversal to be used in orientation determination. Initially I operate on a single dataset at a time, so as to find a good estimate for the phase origin of each particle image using PFT. At this stage I repeat the phase-reversal correction, but typically use defocus pairs of images which can now be combined into a single corrected image. The benefits of this approach are firstly to fill in missing or weak spatial frequencies of one image with corresponding information from

the pair, and secondly to increase the signal-to-noise ratio by averaging the two images. The phase-corrected-and-mixed images are further refined by PFT using data to 17Å and sometimes beyond, although the signal becomes rather dominated by noise. In parallel, a second version of the combined-and-mixed images is made in which a full deconvolution of the CTF is done. The results are very noisy but should have a more accurate representation of the projected density, and these images are destined for the 3D reconstruction. The orientations determined for the phase-corrected-and-mixed images are directly applicable to these deconvolved images, and reconstructions can be calculated to considerably higher resolution than the frequency limit used in the refinement.

The methodology has been successful in studying several virus capsids to resolutions of 9-15Å, including hepatitis B virus,<sup>4-7</sup> bacteriophages<sup>8</sup> and HK97,<sup>10</sup> and rhinovirus HRV2.<sup>9</sup>

---

## 2. Sample Data

The data are the same as described in the “SumPS” document. The sample is HK97 Prohead I imaged on a Phillips CM20 FEG at 38,000, scanned at 7µm (1.842Å/pixel). Micrographs 5881 and 5882 are a focal pair (closer-to- & further-from-focus, respectively). The data set comprises of 607 images from one micrograph, and the corresponding ones from the other. The defocus values and Gaussian fit parameters determined from each set are as follows:

Micrograph	Defocus (Å)	M1	M2	M3	M4
5881	8190	6237.4	41169	-0.19584	0.25781
5882	14640	6047.5	43077	-0.2101	0.26578

The in-plane rotation between the two micrographs is 0.166° (ie, rotate 5882 counter-clockwise).

---

## 3. CTFMIX & single datasets – phase correction

In the early stages of analysis, defocus pairs of images cannot yet be combined since no orientation information is known (in particular, the phase origins). The raw images may be used for initial orientation determination, but if data beyond the first CTF zero is to be used, then usually the phase reversals in the raw images need to be corrected. To do this, the only information required is the defocus of the micrograph, and other microscope details used in CTFZEROS, and a ‘dummy’ PFT-format orientation file which should be constructed ‘by-hand’. An example run of CTFMIX is shown below with the closer-to-focus dataset (5881) of the example pair.

```
CTFmix v1.5
=====
James F. Conway, compiled Oct 18 2000, 17:21:48
-----
Enter parameters
=====
EM> What is the spherical aberration constant (in mm) ? 2
EM> What is the accelerating voltage (in keV) ? 120
CTF> Correct for partial lateral/temporal coherence (exp decays) (Y/n) ? no
CTF> Deconvolve PCTF ? (Y/n) no
CTF> Avoid PCTF scaling at low freqs ? (y/n/fraction from 0-1) yes
CTF> Apply smoothing to phase flip when |PCTF| below what level (0-1) ? 0.3
CTF> Enter CTF amplitude lower limit (Weiner replacement) (0-1) ? 0.3
```

```

CTF> Enter low pass limit and width for Gaussian profile (Angstroms) ? 16.5,0.5
GENERAL> Image file format:
      1 or B: Series of BP files (only on VMS systems)
      2 or P: Packed PIF file
      3 or M: Series of MRC files (not yet available)
Enter choice > pif
GENERAL> Are input images flipped in y-axis relative to PFT coords (y/N) no
GENERAL> Apply constant background mask to output images ? (Y/n) yes
GENERAL> Enter radius for onset of mask (pixels) > 182
GENERAL> Enter width of fade-in ring (pixels, 0 up) > 10
Enter output PIF file filename >> 5881_flip.pif
-----
Summary of parameters:
=====
Spherical aberration constant = 2 mm
Electron voltage      = 120 kVolts
Electron wavelength = 0.0335087 A
Partial lateral and temporal coherences not corrected
PCTF will not be deconvolved
PCTF not applied to low freqs
Smoothing applied when |PCTF| < 0.3
CTF amplitude lower limit (Weiner replacement) is 0.3
Low pass limit = 16.5 A, width = 0.5 A
File format is PIF
Input images will not be flipped in y-axis
Background mask applied, radius = 182 pixels, fade ring = 10 pixels
Output PIF image filename is 5881_flip.pif
-----
Enter parameters specific to each micrograph in the series
=====
Micrograph 1 parameters:
EMPFT-format filename ? 5881.dat
0 lines of header, 607 particles read from '5881.dat'
What is the digitization rate (in Angstroms/pixel) ? 1.842
Enter input PIF file filename >> ../x3d/5881.pif
Enter defocus (+ve) (Angstroms) ? 8190
What is the proportion of amplitude contrast (0-1) ? 0.07
Rotationally align images to dataset 1 (y/N) no
Decay Correction - none, exponential or gaussian decay (N/e/g) ? none
Another micrograph (y/n) ? no
-----
Summary of micrograph parameters
=====
Proportion amplitude contrast 1 = 0.07
Defocus 1 = 8190 A; zeros (pure phase contrast) at:
      1 - 16.61 A
      2 - 11.77 A
      3 -  9.63 A
      4 -  8.36 A

Omega - no alignment
No Exponential or Gaussian decay correction
Weight factor = 1
-----
Opening PIF file...
- Image data is big-endian (eg, Unix)
- Created from 'JFC-PIF 1.4.8 of Sep 25 2000'
Doing particle #1 of 607 (id=1)
Doing particle #2 of 607 (id=2)
.....
Doing particle #607 of 607 (id=607)
Finished processing

```

## Notes

1. Coherence corrections need not be done for any but the highest resolution reconstruction attempts. In fact, I have never used them.

2. “Deconvolve PCTF” – this question is critical for choosing between correcting only for phase reversals (answer: no) or performing a deconvolution of the phase-contrast transfer function (answer: yes), which includes the phase reversal correction. I use the phase-flip to generate images with consistent contrast but no increase in noise – these will be used in orientation determination. The deconvolution I apply to images destined for reconstruction, and these will usually be focal pairs that are corrected and combined (see section 5 below).
3. “Avoid PCTF scaling at low freqs” – this has little effect on the phase-reversal correction. See section 5 below for a description of its purpose.
4. “Apply smoothing to phase flip when |PCTF| below what level (0-1)” allows the data in the vicinity of a phase reversal to be attenuated. For example, when the amplitude of the ideal phase-contrast transfer function is below 0.3 (which I typically use) then an attenuation is applied which drops to zero at the node. Since the signal-to-noise ratio will be very low in this region, suppressing the data may be beneficial (but I have not proven it!).
5. “Enter CTF amplitude lower limit” prevents division by small numbers. In fact, the effect of this is also negligible and is intended more for the correction described in section 5.
6. “Enter low pass limit and width for Gaussian profile” are parameters for a smooth low-pass edge. The high-frequency data has too little signal-to-noise ratio to be useful in refinement of orientations (remember, images have to be refined independently of each other) and may be dispensed with (especially the noise) to as to improve the contrast of the image. In the example, “16.5, 0.5” specifies an edge starting at 16.75Å and smoothly falling to zero at 16.25Å.
7. “Are input images flipped in y-axis relative to PFT coords” is an old compatibility issue. If you don’t use OpenVMS systems, and also don’t use PIC and BP images, then you will need only ever answer this “no”. If you do have data that was once in the BP format, and if that data was converted to PIF-format using a program called “PURDUE”, then the images were inverted during conversion. That would not matter, provided that the images were inverted once, and remained that way for all the steps of analysis. However, such was not always the case. If in doubt, contact me.
8. “Apply constant background mask to output images” is intended to suppress ripples that extend into the background. The circular mask imposed should probably match that used initially on the raw images (such as by the X3dPreprocess particle-picking program). The radius of the mask, and width of the fade ring should be the same as used in X3dPreprocess: the fade ring is additional radii that extend from the mask radius and within which the fade from image values to the constant mask is done.
9. “EMPFT-format filename” is a dummy file in this instance. It is required, but only useful when combining images (see sections 4 and 5 below). What is important here is to supply a file in the PFT format (header lines are optional) with one line for each image, as follows:

```

1, 0.0, 0.0, 0.0, 218.0, 218.0
2, 0.0, 0.0, 0.0, 218.0, 218.0
3, 0.0, 0.0, 0.0, 218.0, 218.0
...
607, 0.0, 0.0, 0.0, 218.0, 218.0

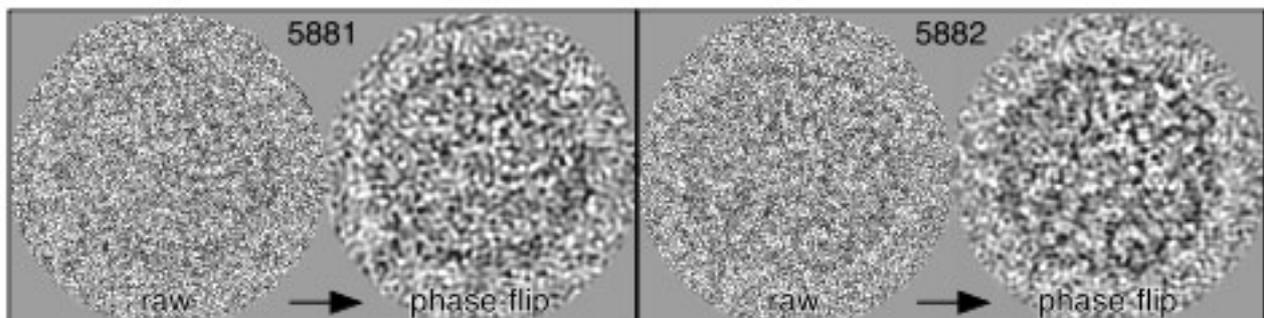
```

The first number on each row is important – it specifies the ID of images in the image database (PIF file, or series of BP or MRC files). The order of ID's is also important – the resulting images will be written in the order specified by this EMPFT.DAT file. I find that it is essential to keep the numbers straight – always number in order from 1, and always use all images (so that the ID's correspond to the position of images in the image database). The remaining numbers are not used, but are theta, phi, omega, X and Y.

10. “What is the proportion of amplitude contrast” should be the same as used in CTFZEROS for fitting the defocus. Here it is a proportion, and 0.07 (or 7%) is a figure from the Unwin and Toyoshima papers.
11. “Rotationally align images to dataset 1” is not done in this case, because we are not combining images, and in any case, this dataset is the first dataset.
12. “Decay Correction” is also recommended for section 5.

## Results

The resulting images should be useful in orientation determination (eg, with PFT), in particular to estimate good phase origins so that the images can be aligned with their defocus partners for correcting-and-mixing, as described in the following two sections. For the curious, the output images can be analyzed with SumPS, and the averaged power spectra compared before and after phase-flipping. There should be little difference, since the primary correction is to the phases and not the amplitudes. However, the effects of the low-pass filter (step 6) should be evident in the output, as should any damping near the positions of phase reversal (step 4). I strongly recommend that images are compared visually to ensure that the sequence of images has been maintained correctly (eg, that an image has not been accidentally omitted) and also to check that the results look reasonable for confirming that the input parameters did not contain any serious errors.



An example of a defocus pair is shown above. The improvement in contrast is largely a result of the low-pass filter and subsequent renormalization of image intensities.

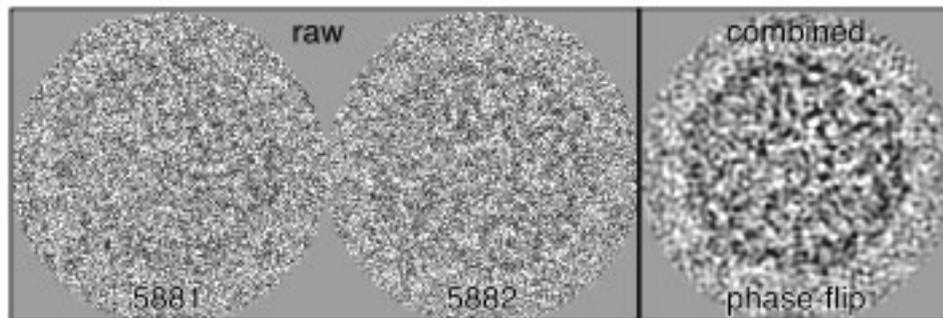
---

## 4. CTFMIX & defocus pairs – phase correction

There are several differences with the single dataset correction described in section 3. Obviously, two (or more) datasets are used, each with the same particles but from micrographs imaged at different defocus values. In the program dialog above, the final prompt for another micrograph will be answered “yes” and details of the second dataset will be entered. To avoid decaying the high spatial frequencies of the closer-to-focus data, I avoid interpolation on this dataset by entering it first, and any further-from-focus datasets will be entered subsequently. These additional images are translationally and

rotationally aligned to the first, and thus suffer some inevitable attenuation of detail as a result of interpolation.

The PFT.DAT files determined independently for each dataset are used for translational alignment. Only the X and Y columns are needed unless a particular method of rotational alignment is specified, in which case the Omega column will also be used. For in-plane rotational alignment I recommend determining the angle from the micrographs (for example, the difference in slope between lines connecting two particles at opposite corners) and using this for the whole dataset. Be careful of the sense of rotation – if the images are flipped about an axis, the sense of rotation may be reversed. When the angle is large enough, you could confirm visually that the superposition of images is correct. As before, I recommend a quick visual inspection as a check for gross errors.



An example of a corrected-and-combined image is above. The resulting images should be refined further (in PFT) and the closer-to-focus orientation file can be used as a starting point. Note that discrepancies in the angles have been ignored for combining images, so start with PFT mode 2 which allows a global search of the asymmetric unit using only the origins. Remember to change the header of the PFT.DAT file to reflect the new corrected-and-mixed input image file.

---

## 5. CTFMIX & defocus pairs – deconvolution

The dialog is similar to that of section 4, but the question “Deconvolve PCTF” will now be answered “yes” and, as a consequence, a couple of prompts will change. Gaussian correction can also be applied here (I usually do this). An example of a complete dialog follows.

```
CTFmix v1.5
=====
James F. Conway, compiled Oct 18 2000, 17:21:48
-----
Enter parameters
=====
EM> What is the spherical aberration constant (in mm) ? 2
EM> What is the accelerating voltage (in keV) ? 120
CTF> Correct for partial lateral/temporal coherence (exp decays) (Y/n) ? no
CTF> Deconvolve PCTF ? (Y/n) yes
CTF> NIH Weighting is Sum(I/c), MRC is Sum(cI/cc)
CTF> --Do NIH or MRC style weighting (N/m) ? NIH
CTF> Avoid PCTF scaling at low freqs ? (y/n/fraction from 0-1) yes
CTF> Enter CTF amplitude lower limit (Weiner replacement) (0-1) ? 0.3
CTF> Enter low pass limit and width for Gaussian profile (Angstroms) ? 7.75, 0.5
GENERAL> Image file format:
          1 or B: Series of BP files (only on VMS systems)
          2 or P: Packed PIF file
          3 or M: Series of MRC files (not yet available)
```

```

Enter choice > PIF
GENERAL> Are input images flipped in y-axis relative to PFT coords (y/N) no
GENERAL> Apply constant background mask to output images ? (Y/n) yes
GENERAL> Enter radius for onset of mask (pixels) > 182
GENERAL> Enter width of fade-in ring (pixels, 0 up) > 10
Enter output PIF file filename >> 5881-5882deconv.pif
-----

```

Summary of parameters:

```

=====
Spherical aberration constant = 2 mm
Electron voltage      = 120 kVolts
Electron wavelength = 0.0335087 A
Partial lateral and temporal coherences not corrected
PCTF will be deconvolved
Weighting method is NIH
PCTF not applied to low freqs
CTF amplitude lower limit (Weiner replacement) is 0.3
Low pass limit = 7.75 A, width = 0.5 A
File format is PIF
Input images will not be flipped in y-axis
Background mask applied, radius = 182 pixels, fade ring = 10 pixels
Output PIF image filename is 5881-5882deconv.pif
-----

```

Enter parameters specific to each micrograph in the series

=====

Micrograph 1 parameters:

```

EMPFT-format filename ? 5881.dat_019
2 lines of header, 607 particles read from '5881.dat_019'
What is the digitization rate (in Angstroms/pixel) ? 1.842
Enter input PIF file filename >> 5881.pif
Enter defocus (+ve) (Angstroms) ? 8190
What is the proportion of amplitude contrast (0-1) ? 0.07
Rotationally align images to dataset 1 (y/N) no
Decay Correction - none, exponential or gaussian decay (N/e/g) ? gaussian
--Gaussian decay: a+b*exp(-(x-c)^2/(d*d))
--Note - units for curve fit are 1/A
--What is term a ? 6237.4
--What is term b ? 41169
--What is term c ? -0.19584
--What is term d ? 0.25781
Another micrograph (y/n) ? yes

```

Micrograph 2 parameters:

```

EMPFT-format filename ? 5882.dat_019
2 lines of header, 607 particles read from '5882.dat_019'
What is the digitization rate (in Angstroms/pixel) ? 1.842
Enter input PIF file filename >> 5882.pif
Enter defocus (+ve) (Angstroms) ? 14640
What is the proportion of amplitude contrast (0-1) ? 0.07
Rotationally align images to dataset 1 (y/N) yes
--Rotationally align by Individual omega's
      Average omega
      User input omega (A/i/u) ? user
      Note: +ve for CCW, -ve for CW (same as PIC's CA output)
      Enter constant omega for rotating this dataset > 0.166
Decay Correction - none, exponential or gaussian decay (N/e/g) ? gaussian
--Gaussian decay: a+b*exp(-(x-c)^2/(d*d))
--Note - units for curve fit are 1/A
--What is term a ? 6047.5
--What is term b ? 43077
--What is term c ? -0.21010
--What is term d ? 0.26578
Another micrograph (y/n) ? no
-----

```

Summary of micrograph parameters

```

=====
Proportion amplitude contrast 1 = 0.07
Defocus 1 = 8190 A; zeros (pure phase contrast) at:

```

```

1 - 16.61 A
2 - 11.77 A
3 - 9.63 A
4 - 8.36 A

Omega - no alignment
Gaussian decay (a+b*exp[-(x-c)^2/(d*d)]) will be corrected
-- Decay parameter a = 6237.4
-- Decay parameter b = 41169
-- Decay parameter c = -0.19584
-- Decay parameter d = 0.25781
Weight factor = 1
-----
Proportion amplitude contrast 1 = 0.07
Defocus 1 = 14640 A; zeros (pure phase contrast) at:
1 - 22.17 A
2 - 15.69 A
3 - 12.82 A
4 - 11.11 A

Omega - alignment by constant user-entered difference = 0.166000
Gaussian decay (a+b*exp[-(x-c)^2/(d*d)]) will be corrected
-- Decay parameter a = 6047.5
-- Decay parameter b = 43077
-- Decay parameter c = -0.2101
-- Decay parameter d = 0.26578
Weight factor = 1
-----
Opening PIF file...
- Image data is big-endian (eg, Unix)
- Created from 'JFC-PIF 1.4.8 of Sep 25 2000'
Opening PIF file...
- Image data is big-endian (eg, Unix)
- Created from 'JFC-PIF 1.4.8 of Sep 25 2000'
Doing particle #1 of 607 (id=1)
Doing particle #2 of 607 (id=2)
.....
Doing particle #607 of 607 (id=607)
Finished processing

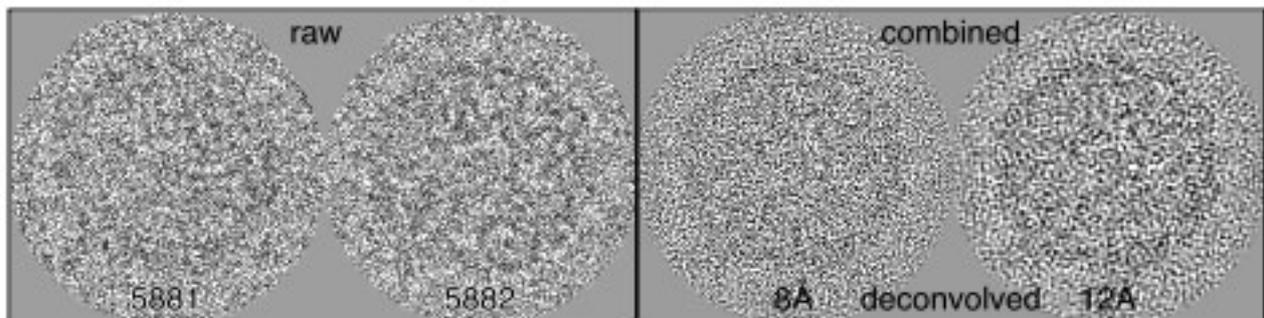
```

## Notes

- 1 “Do NIH or MRC style weighting” chooses between the NIH method (ref [1]) where only the deconvolution is done, and the MRC (ref [11]) method, where the image is weighted by its CTF and the deconvolution is by the CTF-squared. This only effects the combination of images, and not single dataset usage. I have found little difference in comparing the two methods.
- 2 “Avoid PCTF scaling at low freqs” is a difficult area to give firm recommendations about. Because there is strong signal at low spatial frequencies, this part of the spectrum does not follow the phase-contrast imaging mechanism which would suggest low or no contrast at the spatial frequency origin. I use “yes” to avoid the scaling of low-frequency amplitudes, which means that the CTF correction begins at the first maxima of the CTF (usually between 25-40 Å, depending on defocus). Other studies find that some partial correction is better, especially for spherical virus capsids which are full of nucleic acid (eg, ref [12]).
- 3 “Decay Correction - none, exponential or gaussian decay” is chosen to be Gaussian, although an alternative is a simpler exponential decay. The parameters m1-m4 from the Kaleidagraph fit (see the document “SumPS Analysis”) correspond to a-d. Although parameters a and b are not essential, they provide some relative scaling between the datasets that may prove useful (but this has not been established).

## Results

The deconvolved images are very noisy, primarily as a result of the deconvolution which boosts frequencies with low signal at and near the points of phase reversal. The Gaussian decay correction has rather less affect. An example of a corrected-and-combined image is below, in one case with the low-pass filter at 8Å, the other at 12Å.



Do not despair! The strong high-frequency noise can be overcome with sufficient averaging. Make sure to set the low-pass filter limit to as low a value as you can (ie, to limit the inclusion of high frequency noise) according to how high the resolution you expect to achieve. This will avoid unnecessary noise contamination and concomitant loss of contrast. The 8Å limit shown above is rather aggressive. Ultimately, a 12Å 3D structure would be acceptable for this data, and the same image with a 12Å limit (above) shows improved contrast.

---

## 6. References

- 1 Conway JF & Steven AC (1999) Methods for reconstructing density maps of “single” particles from cryoelectron micrographs to subnanometer resolution. *J Struct Biol* **128**, 106-118.  
<[http://www.ncbi.nlm.nih.gov/80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list\\_uids=10600565&dopt=Abstract](http://www.ncbi.nlm.nih.gov/80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10600565&dopt=Abstract)>
- 2 Baker TS & Cheng RH (1996) A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy, *J Struct Biol* **116**, 120-130.
- 3 Fuller SD, Butcher SJ, Cheng RH & Baker TS (1996) Three-dimensional reconstruction of icosahedral particles - The uncommon line. *J. Struct. Biol.* **116**, 48-55.
- 4 Conway JF, Cheng N, Zlotnick A, Wingfield PT, Stahl SJ & Steven AC (1997) Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature* **386**, 91-94.
- 5 Conway JF, Cheng N, Zlotnick A, Stahl SJ, Wingfield PT, Belnap DM, Kanngießer U, Noah M & Steven AC (1998) Hepatitis B virus capsid: Localization of the putative immunodominant loop (residues 78-83) on the capsid surface, and implications for the distinction between c- and e-antigens. *J Mol Biol* **279**, 1111-1121.
- 6 Conway JF, Cheng N, Zlotnick A, Stahl SJ, Wingfield PT & Steven AC (1998) Localization of the N-terminus of hepatitis B virus capsid protein by peptide-based difference mapping from cryo-electron microscopy. *Proc Natl Acad Sci USA* **95**, 14622-14627.

- 7 Cheng N, Conway JF, Watts NR, Hainfeld JF, Joshi V, Powell RD, Stahl SJ, Wingfield PT & Steven AC (1999) Tetra-iridium, a 4-atom cluster, is readily visible as a density label in 3-dimensional cryo-em maps of proteins at 10-25 Å resolution. *J Struct Biol* **127**, 169-176.
- 8 Yang F, Forrer P, Dauter Z, Conway JF, Cheng N, Cerritelli ME, Steven AC, Plückthun A & Wlodawer A (2000) Novel fold and capsid-binding properties of gpD, the phage display platform protein. *Nature Struct Biol* **7**, 230-237.
- 9 Hewat EA, Neumann E, Conway JF, Moser R, Ronacher B, Marlovits TC & Blass D (2000) The cellular receptor to human rhinovirus 2 binds around the 5-fold axis and not in the canyon: a structural view. *EMBO J* **19**, 6317-6325.
- 10 Conway JF, Wikoff WR, Cheng N, Duda RL, Hendrix RW, Johnson JE & Steven AC (2000) Virus maturation via large subunit rotations and local refolding. Submitted.
- 11 Bottcher B, Wynne SA & Crowther RA (1997) Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature* **386**, 88-91.
- 12 Belnap DM, Filman DJ, Trus BL, Cheng N, Booy FP, Conway JF, Curry S, Hiremath CN, Tsang SK, Steven AC & Hogle JM (2000) Molecular tectonic model of virus structural transitions: the putative cell-entry states of poliovirus. *J Virol* **74**, 1342-1354.