(International Tables for Crystallography Volume F, 2010, in press)

Four-Dimensional Cryo Electron Microscopy at Quasi Atomic Resolution: "IMAGIC 4D"

Marin van Heel [1], Rodrigo Portugal [1], Alexis Rohou [1, 2], Charlotte Linnemayr [3], Cecilia Bebeacua [1], Ralf Schmidt [4], Tim Grant [1], and Michael Schatz [4]

[1]) Faculty of Natural Sciences, Division of Molecular Biosciences, Imperial College London, SW7 2AZ London, UK

[2]) current address: Rosenstiel Center, Brandeis University, 415 South St. Waltham MA 02454-9110, USA

[3]) Division of Internal Medicine, Inflammation Research, University Hospital Zürich, Switzerland

[4]) Image Science Software GmbH, Gillweg 3, D-14193 Berlin, Germany

ABSTRACT

The traditional tools of the structural biologist seeking to understand macro-molecules and their complexes are X-ray crystallography and NMR spectroscopy. Single-particle cryo electron microscopy ("cryo-EM") has established itself as a new structural biology technique over the last 15 years. Spectacular insights into the functioning of macromolecular complexes have been achieved especially from combining cryo-EM with the earlier approaches. The resolution levels achieved improved over the last decade from ~10Å to sometimes better than ~4Å, meaning that a *de-novo* structure determination based on single-particle cryo-EM studies alone is now feasible. More challenging is the new perspective that cryo-EM brings: sorting heterogeneous populations of molecules into individual 3D conformers resulting in sequences of related 3D structures, or short: "4D cryo-EM". Thanks to these developments, single-particle cryo-EM has become the technique of choice to shed light on the functioning of many a complex biological system. The design of the software instrumentation for 4D cryo-EM is crucial. In this paper we elaborate on organisation issues for single-particle cryo-EM software, as exemplified by recent developments in the "IMAGIC 4D" software system.

INTRODUCTION

During the nineteen fifties and sixties the electron microscope became a routine research instrument. Many biological structures we now take for granted, such as the ribosome (Chemistry Nobel prize 2009; see: www.nobel.se) were first discovered using the electron microscope. An avalanche of such discoveries emerged after the appropriate preparation techniques for biological specimens were developed like the contrasting of molecular complexes with heavy-metal salts [Brenner 1959]. This dry "negative-stain" approach, however, can also create structural deformations of the sample hampering its detailed analysis. The introduction of the vitreous water ("vitreous ice") specimen preparation technique by Jacques Dubochet and co-workers [Adrian 1984], represented a huge improvement in structural preservation for biological specimens in the harsh vacuum environment of the electron microscope. Notwithstanding the contributions made to the cryo-EM field by the data-processing methodology and software – the subject of this paper – it was the vitreous-ice specimen preparation approach that marked the future success of single-particle cryo-EM. For all practical purposes, vitreous ice is water and the first electron images of virus particles trapped in their native solution state were spectacular [Adrian 1984]. The vitreous ice preparations can trap the molecular complexes in full action, in different conformational states, and may represent a window into the detailed functioning of the complexes in their natural environment.

Structural analysis by single particle cryo-EM is therefore very different to that carried out by traditional X-ray crystallography. With the latter technique, the data collected are diffraction intensities, necessarily reflecting already *averaged* contributions from all unit cells in the crystal. Any information relating to the behaviour of individual molecules is thus lost during data collection. Single-particle cryo-EM, in contrast, is unique in that it preserves the identity of each individual molecular image registered in whatever structural state it happens to be. The marked difference with X-ray crystallography is that the information about each individual molecule or complex is explicitly available in digital form, albeit strongly deteriorated by noise. This also means that in cryo-EM all averaging takes place in the computer; the computational requirements of cryo-EM therefore exceed those of X-ray crystallography by orders of magnitude.

The intrinsic heterogeneity of cryo-EM samples was originally known as the curse of cryo-EM because it degrades the quality of the resulting average structures. Many structures simply would not refine to a reasonable resolution (for reviews of the single particle methods see: [van Heel, 2000; Frank 2006]). In recent years, however, heterogeneity is seen more as a blessing, since, with the appropriate advanced data processing technology, it is now possible to separate the different conformations of the complex into individual 3D structures. Each of these 3D structures may reveal a different state of the biological complex as was first shown with a mixture of functional state of the *E. coli* 70S ribosome in complex RF3 [Klaholz 2004]. For recent reviews on the challenges posed by heterogeneous biological samples, see [Leschziner 2007; Spahn 2009]. These fascinating new developments build upon almost four decades of advances in software instrumentation, that started in the mid nineteen sixties [DeRosier 1968]. We here review recent developments of the IMAGIC system aimed at optimally processing large heterogeneous "4D" cryo-EM data sets to high resolution.

The IMAGIC software system

The IMAGIC software system [van Heel 1981; van Heel 1996] is the result of over 30 years of continuous developments, including various major redesign phases. It was one of the first software packages in electron microscopy and many now standard procedures were pioneered in IMAGIC including: surface rendering, MSA, automatic classification, projection matching, angular reconstitution, automatic particle selection, etc. The basic file format of IMAGIC consists of a header file, to hold image parameters and processing information, and a separate file holding all actual image data. This "stack" format can hold millions of individual molecular images and all IMAGIC programs are designed to loop over all the images in the stack [van Heel 1981]. Users are not burdened with the task of formulating loops over the 2D images needing the same treatment.

The IMAGIC philosophy is to create specialized high-level programs for all more complicated task. High-level programs allow much better interactive guidance for the user than do image processing scripts (which are operating-system dependent). A unique feature of all IMAGIC programs is that all user interaction is concentrated in a so-called User Interaction Block (UIB). Whether interactively processing images or creating a large "batch job" for later processing, the user always communicates with the program's UIB with all interactive guidance available (each question comes with its own interactive help). The IMAGIC system is available for most popular current operating systems including: Linux, Windows XP/Vista/7, and Mac OS-X.

IMAGIC “4D” Processing / Data Format

Recently, another major overhaul of the system was undertaken in order to better handle 3D structural heterogeneity. The relevant IMAGIC programs can now loop over entire 3D volumes (and not just 2D images) allowing programs to tackle problems on a much higher level of complexity. About 25 years ago a radical redesign of the system allowed small 2D images to reside “in core” (the main memory of the computer) rather than on disk and to facilitate all alignment algorithms. The new redesign allows whole 3D volumes to remain in core and to be manipulated efficiently by single subroutine calls without I/O overhead. The file format for “4D” processing has hardly needed any updating other than defining some new header parameters. The reason is that a stack of 3D volumes (a “4D” data format) is still just a stack of 2D images, the traditional IMAGIC data format (**Fig. 1**). This “4D” upgrade, however, did require a radical overhaul of all programs, since all relevant programs needed an extra loop over a set of 3D volumes. At the same time these 3D programs have all become much faster, more compact and maintainable because they rely on the new incore-3D libraries and are now freed of all excess I/O calls.

Software Parallelisation

In cryo-EM all averaging and information extraction from all individual molecular images takes place "in silico"; the computational requirements are thus huge. Computers are never fast enough for the most demanding single-particle approaches.

Over the last 15 years much emphasis was placed on the parallelisation of critical IMAGIC code [van Heel 2000] mainly using "MPI" (Message Passing Interface, [Gropp 1994]) to take advantage of modern "cluster" computing environments. Other software packages have since followed the same parallelisation path [Smith 2008]. The IMAGIC software is implemented such that the same code will run on all machines from a single-CPU notebook computer up to large cluster systems with hundreds of CPUs.

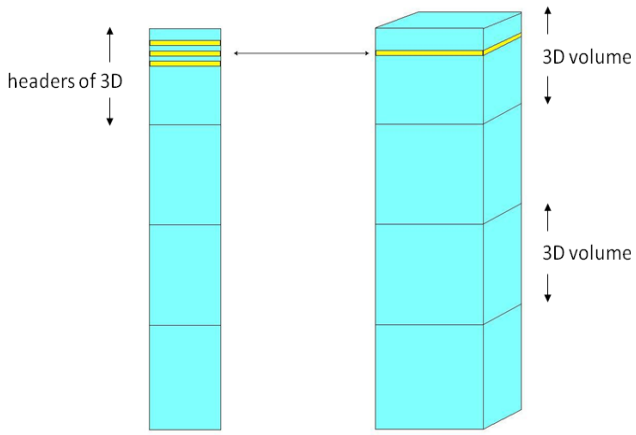

Figure 1: IMAGIC 4D file format
The 4D file format is not much different from the original file format with one header record per 2D image in a stack [van Heel 1981]. The main difference is that specific header locations now indicate the additional organisation of a sequence of 2D images in the stack belonging together in a 3D volume. Substantial changes were required in the software, however, to make programs loop over the 3D volumes in one 4D IMAGIC file, the way the programs looped over the 2D images in a stack. For full format details see: www.imagescience.de

Recently, a "GPU" library (see: www.nvidia.com/object/cuda_home.html) has been implemented in IMAGIC to exploit the properties of cheap graphic processors. However, for most "standard library" operations, like 2D FFTs, which can be almost directly linked to existing programs, most time gained in GPU processing is still lost in transporting data to and from the limited memory of the GPUs. The question is whether the time invested in software development for exploiting specific properties of any specific parallel computing hardware will pay off on the longer term for the specific needs of cryo-EM. In the case of the GPUs, the amount of memory typically available per GPU core is rather low.

Full 2D (parallel) astigmatic CTF Correction

The IMAGIC CTF estimation and correction programs [van Heel 2000] have recently been upgraded to work fully in 2D, enabling the accurate detection of all CTF parameters including astigmatism. This has been carried out in order to accomplish two goals. Firstly, the programs are now capable of operating on an entire dataset of CCD images or patches of micrographs. This now allows the use of the parallel MSA programs (see below) to classify sets of amplitude spectra (and create class averages thereof) prior to the precise determination of the defocus and astigmatism parameters. This enables the simple and largely automatic determination of CTF parameters, and subsequent CTF correction via phase flipping of entire dataset-stacks. Secondly, the ability to accurately detect even extreme levels of astigmatism enables the use of highly astigmatic images, unlocking extremely close to focus defoci in order to push the achievable resolution. These programs have been used for the processing of a highly astigmatic dataset of the *Limulus polyphemus* hemocyanin, resulting in a ~4Å reconstruction from only 15,000 raw molecular images, corresponding to 60,000 asymmetric units for this C2 pointgroup-symmetry structure (Grant *et al*. submitted for publication).

Parallel Automatic Particle Picking

Parallel processing is of increasing importance for collecting the raw large data sets required for 4D cryo-EM. If we are interested in a structure that represents 1% of all molecular complexes, we need to increase the size of the data set one-hundred fold in order to achieve the same resolution we had for a mono-disperse dataset. This implies we need to be able to rapidly process data sets of the order of 1Tbyte in size. One of the fastest automatic particle selection procedures is still one of the oldest ones based on the local variance [van Heel 1982] in the raw micrographs or CCD images. It is applied immediately after the full-data set CTF correction discussed above. Very important for the calculation of the variance image is the choice of the frequency range used for discrimination the presence of a particle with respect to the background. The sensitivity of the approach is approximately as good as that of particle searching using the CCF function [Saxton 1977]. However, like all CCF alignments (see below) the CCF particle picking requires templates and may tend to bias the particle selection towards the references [Boekema 1986; Stewart 2004]. Modulation image particle picking is a variant of the original variance image detection [van Heel 1982] which is based on the local standard deviation rather that the local variance and avoids squaring of amplitudes. This new algorithm is often the method of choice for an unbiased automatic particle selection. Once a reliable first 3D structure has been calculated, the particle picking program can then look for all possible views in all possible orientations in an extensive stack of input images (the massive calculations involved exploit MPI parallelisation.

Parallel Multi-Reference Alignments and Reference Bias

Large multi-reference alignment (MRA) schemes [van Heel 1985], have been accelerated by some orders of magnitude with the introduction of MPI parallelisation on computer clusters [van Heel 2000] and MRA remain the largest source of CPU needs in single particle cryo-EM. However, with the recent focus on discriminating between different functional states, the issue of reference bias resurfaces. With the renewed emphasis on avoiding reference bias [Boekema 1986, Stewart 2004], “alignment by classification” of centered molecular images (Dube 1993) has taken a new emphasis. An especially useful variant is the rotational alignment of the data set with respect to the main symmetry-related eigenvector. This alignment does not bias the dataset with respect to specific references, yet concentrates much of the variance in the data set into the lower eigenimages [van Heel 2009] and thus yields better class averages while using the same total number of eigenvectors. Such new unbiased approaches are still in full development.

MRAs are now used as a tool in 4D analysis by generating MRA references from a set of 3D structures rather than just a single one [Klaholz 2004; Spahn 2009]. Although not different from the 2D versions of MRA, the new IMAGIC "4D" programs make it simpler to explicitly track a "3D membership" indicating to which 3D volume the best alignment was achieved. The overall administration of the 4D procedures are greatly facilitated especially during iterative refinement rounds (**Fig. 2**)

MSA and its parallelisation

Multivariate Statistical Analysis ("MSA") was originally introduced to electron microscopy to handle different views of a macromolecule [van Heel 1981b; Borland 1990; van Heel 2000; Frank 2006; for an extensive review: van Heel 2009]. MSA approaches are the techniques of choice to create a structure in complex using heterogeneous data sets. Today, a wealth of information on MSA techniques is readily available on the internet. The highly efficient IMAGIC MSA programs are specifically optimised for large data sets: the CPU time required for calculating the main eigenvectors is directly proportional to the total size of the data set. The naïve use of standard MSA libraries in most EM software leads to eigenvector/eigenvalue calculations that are proportional to the *square* of the number of images in a data set, making it impossible to process more than a few thousand images [van Heel 2009].

In the first "4D cryo-EM" paper revealing different functional states of the same structure [Klaholz 2004], the MSA separation of the dataset into the two states was achieved by a manually-supported separation of the two conformers in eigenvector space, each representing a different functional states of the 70S ribosome in complex with RF3. Each functional state of a complex represents a "manifold" in factor space, and one seeks to separate the different manifolds in an unambiguous way, possibly by weighing the eigenvectors [van Heel 1984, Elad 2008]. An alternative is to use the MSA programs to explicitly analyse large sets of 3D volumes the way one traditionally analyses large sets of 2D images (see [van Heel 2009]).

The ever-increasing size of the data sets posed a new problem: the MSA processing of large data sets containing, say, a million images (totalling say ~1 Tbyte) take months to process on a single CPU. An MPI parallelisation project was undertaken to speed up the eigenvector/eigenvalue calculations. This required parallelisation of not only the pure computations but also the Input/Output operations since the algorithm reads the large input data sets many times (~30 times) in order to iteratively refine towards a stable answer leading to I/O bottlenecks (Portugal *et al*., publication in preparation; [van Heel 2009]). Like all other programs mentioned in this paper, the parallel MSA program is now part of the normal IMAGIC distribution.

Handling multiple 3D reconstructions in parallel

A new parallel 3D reconstruction program was created based on the earlier ex-core program implementing the "exact filter" algorithm [Harauz 1986]. With all 3D reconstructions now organised in in-core subroutines, the new program is organised to loop over many 3D volumes and to generate whole series of 3D reconstructions using new options. It is an ideal tool to generate (large) 4D data sets as needed, for example, for the 4D MSA approaches. It is a perfect match for the 4D MRA (above) and the 4D Angular Reconstitution programs (**Fig. 2**).

The new program, apart from the exact filter algorithm, now also features a novel 3D deconvolution algorithm based on the idea that one can predict the three dimensional point-spread function (PSF) for any 3D reconstruction geometry. The correct 3D reconstruction is then calculated by the 3D deconvolution of the unfiltered 3D reconstruction with the PSF. This 3D deconvolution algorithm scales proportionally to the number of projections used to generate a 3D reconstruction (the exact filter

algorithm scales as the *square* of the number of projections) and yields results of the same overall quality as the classical exact filter algorithm (Grant *et al.*, publication in preparation).

Angular Reconstitution 4D refinements

With the angular reconstitution approach [van Heel 1987; van Heel 2000] one typically assigns an Euler angle orientation to a 2D image by finding its best overall sinogram correlation peaks with respect to an "anchor set" of 2D projections from one single 3D structure. The IMAGIC "Euler" program, which performs the angular reconstitution orientational search, has now been extended to loop over multiple anchor sets and has thus been upgrade to a "4D" level. As a consequence, the Euler program is now capable of competitively choosing the best anchor-set among anchor sets generated from different 3D structures. The program can now thus also be used for refining towards multiple 3D structures as was already the case with the "MRA" program in IMAGIC.

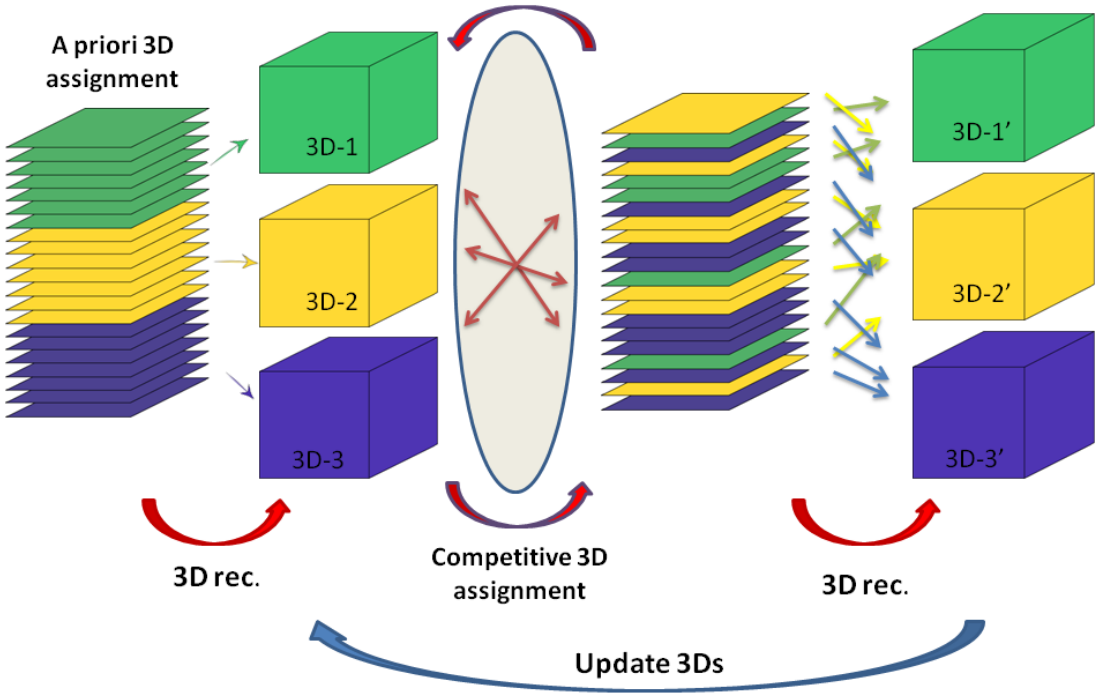

Figure 2. Generic 4D refinement scheme: one starts with the (*a priori*) assignment of the images in a stack (original images or class averages) to a number of 3D volumes (named 3D-1 to 3D-3). This assignment can be of various origins: 2D MSA manifold separation; 3D MSA classification; or even a random number generator, depending on the strategy pursued. Once different 3Ds have been generated, various "competitive" 3D membership assignment schemes can be applied to the 2D images in the stack. A classical one is multi-reference alignment with respect to re-projections of the different 3D volumes. Another approach is the multi-anchor set Euler angle assignment by “angular reconstitution” as described in the main text. The resulting new 3Ds (named 3D-1’ to 3D-3’) take the place of the earlier 3Ds in this generic iterative refinement scheme. The iterations are stopped once convergence criteria like the Fourier shell correlation (FSC) resolution of the 3Ds reach stability [Harauz 1986; van Heel 2005].

Discussion

The potential benefits of 4D cryo-EM at atomic resolution are obvious. With some further developments, 4D cryo EM may become the structure-biology method *par excellence*. The growing importance of software in cryo-EM has been the subject of various editorials and special issues of journals [Carragher 1996; Smith 2008]. The processing of 4D data sets, however, requires solving a large set of 3D structures simultaneously and that is necessarily a complex matter. The researcher will have to understand the basic principles of the approach and its pitfalls and will thus have to go through a steep learning curve. The basic methodology (and the use of IMAGIC) is taught in courses such as the EMBO course for 3D cryo EM, or the Brazil school for single particle cryo-EM (www.single-particles.org/school/). The IMAGIC 4D system is aimed at accompanying the user at the appropriate level of complexity, without forcing the user to become an expert script programmer. Throughout the IMAGIC 4D system, the software will directly understand the 4D frame of thinking of the user.

There is a growing understanding in the field that continuity of development, maintenance and support of the software for cryo-EM are important issues for the success of a project. One needs to safeguard the human investment in mastering the learning curve, and one needs continuity of support over the duration of a (long-term) project. Software continuity is at least as important as it is to continue maintaining electron microscopes for data collection in cryo-EM. An example of a successful community effort to consolidate software instrumentation in the field of X-ray crystallography is the maintenance of the CCP4 suite of programs [CCP4 1994]. One must realise however, that this model cannot directly be transferred to the cryo-EM field because each of the handful of packages in use in single-particle cryo-EM is at least as complex as the CCP4 system, largely due to the necessity of parallelisation.

One further remark about using scripts for running existing software: scripts tend to lead a life of their own and are often exchanged among users; combinations of scripts are sometimes even distributed as a "new" package. However, scripts are operating system dependent, and any change in the original programs requires updating of all associated scripts in all different operating systems. The lack of "backward" compatibility with existing scripts actually impedes the development of the original software since many users may refuse to update the core software to avoid having to modify their (borrowed) scripts. Creating wrappers (Python) to interface with original programs has a similar effect of possibly hampering software evolution.

Another compatibility issue is that of the formats used in cryo-EM. The original 2D stack format of IMAGIC [van Heel 1981] is one of the richest formats in use cryo-EM and therefore it is possible to convert from all file formats in use cryo-EM into IMAGIC without loss of information. The popular EM2EM conversion program (www.imagescience.de) for converting all cryo-EM formats into each other is an IMAGIC program that converts formats without any unnecessary loss of header data. The popular CCP4 density format [CCP4 1994] has been adopted as a standard for the current EM databases [Smith 2008]. This format is a 3D format, with one header record for the full 3D volume. It can also hold a single 2D image, but it cannot hold a stack of images, each with its own header record, as is required for single particle cryo-EM. For 4D cryo-EM applications a 4D format will be required for the

structure data bases. There is some urgency for the cryo-EM and X-ray community to agree on a more elaborate joint density format data for our structure data bases.

Atomic-resolution structures (~3Å) have hitherto been elucidated mainly by X-ray crystallography, where the biological molecules are confined to the rigidity of a 3D crystal. Single-particle cryo-EM gives a direct window into the solution, revealing a plethora of different views, of different complexes, in different functional states albeit at a resolution level still somewhat inferior to that achieved by X-ray crystallography and NMR spectroscopy. Single-particle cryo-EM techniques are now approaching resolution levels previously only achievable by X-ray crystallography. For a better understanding of biological processes, it is essential to see the sequence of conformational changes and interactions that molecules undergo during their functional cycle. A primary challenge in structural biology today is to generate "4D movies" of biological complexes at (quasi) atomic resolution. The new IMAGIC 4D software has been tailored for such "4D analysis" at atomic resolution.

Acknowledgements:

We acknowledge recent financial support from the EU/NOE (NOE - PE0748), from the Dutch ministry of economic affairs (Cyttron Project: BIBCR_PX0948) and from the BBSRC (Grant: BB/G015236/1). We thank all users of the IMAGIC system for constructive feedback and past contributors for their code and ideas. We also acknowledge the use of the Eclipse/SVN software suite for helping to keep the IMAGIC system coherent with contributors operating from three different continents.

REFERENCES

Adrian M, Dubochet J, Lepault J, & McDowall AW. **Cryo-electron microscopy of viruses**, *Nature* **308 (**1984) 32-36.

Boekema EJ, Berden JA & van Heel M. **Structure of mitochondrial F1-ATPase studied by electron microscopy and image processing**, *Bioch. Biophys. Acta* **851** (1986) 353-360.

Brenner S & Horne RW. **A negative staining method for high resolution electron microscopy of viruses**, *Biochim Biophys Acta* **34 (1959)** 103-10.

Borland L & van Heel M. **Classification of image data in conjugate representation spaces**, *J. Optic. Soc. Am*. A, **7** (1990) 601-610.

Carragher B & Smith PR. **Advances in computational image processing for microscopy**, *J Struct Biol* **116** (1996) 2–8.

CCP4: Collaborative Computational Project #4. **The CCP4 Suite: Programs for Protein Crystallography**, Acta Cryst. **D50** (1994) 760-763

DeRosier DJ & Klug A. **Reconstruction of three-dimensional structures from electron micrographs,** *Nature* **217** (1968) 130-134

Dube P, Tavares P, Lurz R & van Heel M. **Bacteriophage SPP1 portal protein: a DNA pump with 13-fold symmetry**, *EMBO J.* **12** (1993) 1303-1309.

Elad N, Clare DK, Saibil HR & Orlova EV. **Detection and separation of heterogeneity in molecular complexes by statistical analysis of their two-dimensional projections,** *J. Struct. Biol.* **162** (2008) 108–120,

Frank J. **Three-Dimensional Electron Microscopy of Macromolecular Assemblies**, Oxford University Press (2006)

Gropp W, Lusk E; Skjellum A. **Using MPI: portable parallel programming with the message-passing interface**. MIT Press, In Scientific and Engineering Computation Series, Cambridge, MA, USA. (1994).

Harauz G, van Heel M. **Exact filters for general geometry three dimensional reconstruction**, *Optik* **73** (1986) 146-156.

Klaholz BP, Myasnikov AG & van Heel M. **Release factor 3 seen on the ribosome during termination of protein synthesis,** *Nature* **427** (2004) 862-865.

Leschziner AE & Nogales E. **Visualizing flexibility at molecular resolution: Analysis of heterogeneity in single-particle electron microscopy reconstruction**s, *Annu. Rev. Biophys. Biomol. Struct.* **36** (2007) 43-62.

Saxton WO & Frank J. **Motif detection in quantum noise-limited electron micrographs by cross-correlation**, *Ultramicroscopy* **2** (1977) 219-227.

Smith R & Carragher B. **Software tools for molecular microscopy**, *J Struct Biol* **163** (2008) 224–228.

Spahn CMT & Penczek PA. **Exploring conformational modes of macromolecular assemblies by multiparticle cryo-EM**, Curr. Opin. in Struct. Biol. 19 (2009) 1–9

Stewart A & Grigorieff N. **Noise bias in the refinement of structures derived from single particles,** *Ultramicroscopy* **102** (2004) 67-84.

van Heel M & Stöffler-Meilicke M. **The characteristic views of *E. coli* and *B. stearothermophilus* 30S ribosomal subunits in the Electron Microscope**, *EMBO J.* **4** (1985) 2389-2395.

van Heel M & Frank J. **Use of multivariate statistics in analyzing the images of biological macromolecules**, *Ultramicroscopy* **6** (1981b) 187-194.

van Heel M & Keegstra W. **IMAGIC: a fast flexible and friendly image analysis software system**, *Ultramicroscopy* **7** (1981) 113-130.

van Heel M. **Detection of objects in quantum noise-limited images**, *Ultramicroscopy* **7** (1982) 331-342.

van Heel M. **Multivariate Statistical Classification of Noisy Images (Randomly Oriented Biological Macromolecules)**, *Ultramicroscopy* **13** (1984) 165-183.

van Heel M. **Angular reconstitution: *a posteriori* assignment of projection directions for 3D reconstruction**, *Ultramicroscopy,* **21** (1987) 111-124.

van Heel M, Portugal R & Schatz M. **Multivariate Statistical Analysis in Single Particle (Cryo) Electron Microscopy**. in: **An electronic text book: Electron microscopy in Life Science**, 3D-EM Network of Excellence, Ed.: Verkley A & Orlova EV (2009).

van Heel M, Harauz G, Orlova EV, Schmidt R & Schatz M. **A new generation of the IMAGIC image processing system**, *J. Struct. Biol.* **116** (1996) 17-24.

van Heel M, Gowen B, Matadeen R, Orlova EV, Finn R, Pape T, Cohen D, Stark R, Schmidt R, Schatz M, & Patwardhan A. **Single-particle cryo electron microscopy: towards atomic resolution**, *Quart. Rev. Biophys*.**33** (2000) 307-369.

van Heel M & Schatz M. **Fourier Shell Correlation Threshold Criteria,** *J. Struct. Biol.* **151** (2005) 250-262.