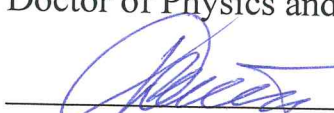Ministry of Education and Science of the Russian Federation
NATIONAL RESEARCH
TOMSK STATE UNIVERTSITY (NR TSU)
Faculty of Physics
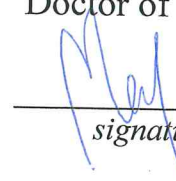Department of General and Experimental Physics

ADMIT TO THE DEFENCE AT SEB
Director of the Master's Programme
Doctor of Physics and Mathematics,
Professor
_____ V.P. Demkin
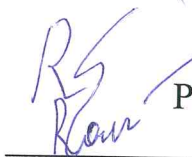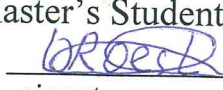_____18. 06._____2018

MASTER'S THESIS

BREAKING THE 100 KDA BARRIER USING PHASE PLATE CRYOGENIC ELECTRON MICROSCOPY

Within the Basic Educational Programme of Master's Degree
subject area 03.04.02 – Physics

Roest Natascha

Research supervisor
Doctor of Physics and Mathematics,
Professor
_____ V.N. Cherepanov
signature
_____11. 06._____2018

Research supervisor
PhD, Assistant Professor
_____ R.B.G. Ravelli
signature
_____11.06_____2018

Author
Master's Student of group 526 MB
_____N. Roest
signature

Tomsk-2018

ABSTRACT

Introduction: Tuberculosis (TB) is the leading infectious bacterial disease worldwide, caused by *Mycobacterium Tuberculosis*. The virulence of *Mycobacterium Tuberculosis* is dependent on several ESX secretion systems, also known as type VII secretion systems (T7SS), which are responsible for the transport and secretion of several effector proteins. *Mycobacterium Tuberculosis* encodes for five ESX systems, and three of these systems have been shown to be active in secretion. Apart from the fact that the ESX-1 system is poorly characterized, the secreted proteins are the target of intense research due to their involvement in the pathogenesis and virulence. Each ESX system is responsible for the secretion of several secreted proteins. In the case of ESX-1, these include EspA, EspC, EspE, EspF, EspJ, EspK & EspB. EspB is the most studied secreted protein from the ESX-1 system and is compulsory for the death of the host-cell. Despite the fact that EspB is an essential element of the ESX-1 secretion system and its role in the regulation of ESX-1 virulence, the function of EspB is poorly defined. Further, structural data of EspB or any other Esp protein in the ESX-1 secretion system is lacking.

Aim of the research: Currently, several research groups have demonstrated that the VPP is able to yield high-resolution (near-atomic) cryo-EM structures. During this study, we hypothesize to break the earlier mentioned barrier of 100kDa with near atomic resolution, by implementing the phase plate in conventional cryo-EM. We will study the cut-on periodicity using simulations and collect data of two secreted proteins from *Mycobacterium tuberculosis*; EspB & EspK.

Methods: RELION is software used for cryo-EM structure determination. It consists of several steps to completely process raw data into a 3D structure. For the 2D class averages of EspB, the particles were subjected to 2D classification (20 classes, regularization parameter 4, 50 iterations and a mask diameter of 120 Å). The best 2D classes were selected (10 classes) and contained 61618 particles (38.44% of the whole data set). Reported resolutions of the individual classes were below 4 Å. For the determination of the C-terminal domain of EspB, a smaller particles diameter was used.

The best 2D classes were selected (5 classes) and contained 11719 particles (85.87% of the whole data set). Reported resolutions of the individual classes were all below 4 Å.

Results: The mean phase shift is 115.8 degrees, with a standard deviation of 23.7 degrees. A defocus of 0.5 μm (5000 Å) was requested. The average defocus value is 0.49 μm (4984 Å) with a standard deviation of 1247 Å. The mean astigmatism was 193 Å with a standard deviation of 139 Å. 2D class averaging of the EspB oligomer resulted in 10 satisfactory classes, all with reported resolution below 4 Å. Same was observed in the 2D class averaging of the C-terminal domain of the oligomer; 5 satisfactory classes were obtained, all with resolution below 4 Å. Data processing on the 2D class averages of EspK are currently running, so cannot be incorporated in the present work.

Discussion: During this research multiple logistic and technological problems were addressed with. Currently, most problems remain present in the sample preparation. Getting an amorphous ice layer with a steady ice thickness remains one of the most problematic for current cryo-EM research. Furthermore, problems are encountered with the air-water interface which is present on the grid and responsible for the preferred orientation present in the case of EspB. Another difficulty during data processing was astigmatism.

Conclusion: In this current work, the aim was to break the 100 kDa barrier in cryogenic electron microscopy by implementing a phase plate in the back-focal plane of the microscope, to enhance phase contrast without adjusting the defocus. The ultimate aim was to generate a 3D reconstruction of two biological relevant proteins, EspB & EspK, secreted proteins from *Mycobacterium tuberculosis*. Unfortunately, due to several limitations in sample, technique and time restrictions, a 3D reconstruction of the mentioned proteins was not reachable within the period of time in which this research was conducted. Too much problems with the protein, the sample preparation, the vitrification and air-water interface remain to have a good reconstruction at the present time.

In the future, local refinement strategies can be used to analyse small structures, heterogeneous samples and flexible samples. Furthermore, further research can be conducted into other strategies to «tumble» the protein out of its preferred orientation. During this research, small investigations was with crosslinking EspB, however, due to limiting time there was no success with this so far.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Cryo-EM   Cryogenic Electron Microscopy

CTD   C-terminal domain

DWF   Debye-Waller Factor

FPLC   Fast Protein Liquid Chromatography

LB   Luria-Bertani

MycP1   Mycosin-1

OD   Optical Density

PMSF   Phenyl Methane Sulfonyl Fluoride

PS   Phase Shift

RELION   Regularized Likelihood Optimization

RPM   Rotations Per Minute

SNR   Signal-to-Noise Ratio

T7SS   Type VII secretion system

TB   Tuberculosis

VPP   Volta Phase Plate

WHO   World Health Organization

INTRODUCTION

Tuberculosis (TB) is the leading infectious bacterial disease worldwide, caused by *Mycobacterium Tuberculosis*. In the Netherlands, tuberculosis is not frequently occurring. In 2016, thousand cases of TB were reported (680 cases among men & 320 cases among women) of which thirty-eight cases of mortality were reported. However, the incidence of TB worldwide is much higher (Figure 1). In 2016, 10.4 million cases of TB were reported to the World Health Organization (WHO), and 1.3 million TB deaths were reported in 2016 [1].

Figure 1 – Estimated Tuberculosis (TB) incidence rates in 2016. Incidence is expresses as the number of cases per 100,000 population, per year.
Data source: Global tuberculosis report 2017 [1]

The virulence of *Mycobacterium Tuberculosis* is dependent on several ESX secretion systems, also known as type VII secretion systems (T7SS), which are responsible for the transport and secretion of several effector proteins. *Mycobacterium Tuberculosis* encodes for five ESX systems, and three of these systems have been shown to be active in secretion [2, 3]. The most studied secretion system is ESX-1, because of its involvement in the pathogenesis of TB. However, despite the critical role of ESX-1 in the virulence of *M. tuberculosis*, the structure of the ESX-1 secretion

system remains poorly identified [4]. Apart from the fact that the ESX-1 system is poorly characterized, the secreted proteins are the target of intense research due to their involvement in the pathogenesis and virulence. Each ESX system is responsible for the secretion of several secreted proteins. In the case of ESX-1, these include EspA, EspC, EspE, EspF, EspJ, EspK & EspB [4]. EspB is the most studied secreted protein from the ESX-1 system and is compulsory for the death of the host-cell [5]. The full-length EspB protein, 60 kDa, is cleaved at the C-terminal domain (CTD) by a membrane-bound protease, mycosin-1 (MycP1). The cleavage by MycP1 yields the mature 50 kDa isoform of the EspB protein [3]. The purpose of the cleavage by MycP1 is unknown, nevertheless is required for the pathogenesis of *M. tuberculosis*.

The cleavage of MycP1 has been described to modulate the secretion of proteins from the ESX-1 and potentially regulates phospholipid binding by EspB [4]. Normally, phospholipids are presented on the surface of apoptotic cells. The presence of EspB disrupts the normal phosphatidylserine signalling and can therefore influence membrane-mediated innate immune responses [6].

Despite the fact that EspB is an essential element of the ESX-1 secretion system and its role in the regulation of ESX-1 virulence, the function of EspB is poorly defined. Further, structural data of EspB or any other Esp protein in the ESX-1 secretion system is lacking [3, 4].

Within the area of structural biology, the structure of biological macromolecules is determined in order to describe and understand its function [7]. Methods which are widely used in the current period are X-ray crystallography and cryogenic electron microscopy (cryo-EM). X-ray crystallography is a powerful technique which can achieve atomic resolutions (~3 Å) in the three-dimensional analysis of proteins. However, X-ray crystallography poses problems with the analysis of large biological macromolecules; The analysis becomes labour-intensive and the radiation used in the analysis is likely to damage the biological material [7]. In cryo-EM, a small amount of a solution, containing the substance of interest, is immediately frozen in liquid nitrogen to avoid the formation of ice crystals. The particles of interest are present in the solution in random orientations, and a considerable number of particles allows the

reconstruction of a three-dimensional model of the protein of interest. Currently, limitations in cryo-EM are mainly due to technical limitations [8], but there is a lot of potential for the application of single particle cryo-EM in the structural analysis of proteins.

Currently, several research groups have demonstrated that the VPP is able to yield high-resolution (near-atomic) cryo-EM structures [9]. The purpose of this research is to break the earlier mentioned barrier of 100kDa with near atomic resolution, by implementing the phase plate in conventional cryo-EM. We will study the cut-on periodicity using simulations and collect data of two secreted proteins from *Mycobacterium tuberculosis*; EspB & EspK.

The research specific tasks include:

1. Determination of cut on periodicity based on the findings of *Danev et al* [21]. This will be simulated using the computer program Matlab, where an existing script will be modified to allow implementation of the cut-on periodicity.

2. Creation of 2D class averages of the proteins of interest, EspB & EspK, to gain more knowledge about the structure of the protein, allowing better interpretation about the function of the protein.

3. Discuss current bottlenecks, of cryogenic electron microscopy, and fetch solutions to address these issues currently present. New processing methods, such as focused classification will be described as well.

1 Transmission Electron Microscopy

In conventional transmission electron microscopy, further referred to as TEM, a thin specimen is irradiated with an electron beam of uniform current density. The acceleration voltage of routine instruments is 100-200kV. Medium-voltage instruments, 200-500kV, yield better transmission and resolution and in high-voltage electron microscopy (HVEM) the acceleration voltage reaches 500kV – 3MV. Electrons are emitted in the electron gun by thermionic or field emission. Field emission guns are used when a high gun brightness and coherence are required. The condenser lens system permits variation of the illumination aperture and the area of the specimen illuminated. The electron-intensity distribution behind the specimen is imaged with a projector lens system, onto the detector in place. The image can be recorded digitally via a fluorescent screen coupled by a fibre-optic plate to a charge-couple device (CCD) camera.

1.1 Particle optics of Electrons

The relevant properties of an electron in particle optics are the rest mass m_0 and the charge $- e$ (Table 1). In an electric field E and a magnetic field B, electrons experience the Lorentz force

$$F = -e \, (E + v \, x \, B), \tag{1}$$

where F (N) is acting on a particle with electric chare e (C) with instantaneous velocity v (m/s), due to an external electric field E (N · C^{-1}) and magnetic field B (T). In an electron gun, the electrons leave the cathode of the latter as a result of thermionic or field emission. The cathode is held at a negative potential $\Phi_C = -U$ in which U (V) is the acceleration voltage, relative to the anode, which is grounded, $\Phi_A = 0$.

The electrode potentials create an electric field E in the vacuum between the cathode and anode, which can also be characterized by equipotential $\Phi = constant$. The electric field is the negative gradient of the potential

$$E = -\nabla\Phi = -\left(\frac{\partial\Phi}{\partial x}, \frac{\partial\Phi}{\partial y}, \frac{\partial\Phi}{\partial z}\right), \tag{2}$$

where E (N · C^{-1}) is the electric field and Φ (V) is the potential gradient. The existence of a potential implies that force $F = -e{\cdot}E$ is conservative and that the law of energy conservation

$$E + V = constant, \tag{3}$$

where E is kinetic energy and V is potential energy, can be applied. The kinetic energy of the cathode equals $E = 0$, where the potential energy V is zero at the anode.

The potential energy of the cathode can be obtained from the work W that is needed to move an electron from the anode to the cathode against the force F

$$V = -W = -\int_A^C F \cdot ds = e \int_A^C E \cdot ds = -e \int_A^C \nabla\Phi \cdot ds = -e(\Phi_C - \Phi_A) = eU,$$
$$\tag{4}$$

where V is potential energy, W is work, F is force, E (N · C^{-1}) is the electric field, e (C) is electric charge, Φ_C (V) is potential of the cathode, Φ_A (V) is the potential of the anode and U (V) is acceleration voltage.

In the reverse direction, the electrons acquire the amount of eU (V) of kinetic energy of the anode. This implies the gain of kinetic energy $E = eU$ of an accelerated electron depends only on the potential difference U, irrespective of the real trajectory between cathode and anode. An electron acquires kinetic energy $E = 1.602$ x 10^{-19} J if accelerated through a potential difference $U = 1V$ because in SI units
1 C V = 1 A V s = 1 W s = 1 J.

This energy of 1 eV = 1.602 x 10^{-19} J is used as a new unit, 'one electron volt'.

Table 1 – Properties of the electron

Rest mass	$m_0 = 9.1091 \times 10^{-31}$ kg
Charge	$e = -1.602 \times 10^{-19}$ C
Kinetic energy	$E = e \cdot U$
	$1\ eV = 1.602 \times 10^{-19}$ J
Velocity of light	$c = 2.9979 \times 10^{8}$ m·s^{-1}
Rest energy	$E_0 = m_0 \cdot c^2 = 511$ keV
Spin	$s = h/4\pi$
Planck's constant	$h = 6.626 \times 10^{-34}$ J · s

	Nonrelativistic $(E \ll E_0)$	Relativistic $(E \sim E_0)$	
Newton's law	$F = \dfrac{\mathrm{d}p}{\mathrm{d}\tau}$	$F = \dfrac{\mathrm{d}}{\mathrm{d}\tau}(mv)$	(5)
Mass	$m = m_0$	$m = m_0/\sqrt{1 - v^2/c^2}$	(6a)
Energy	$E = e \cdot U =$ $0.5 m_0 v^2$	$mc^2 = m_0 c^2 + eU = E_0 + E$ $m = m_0\,(1 + E/E_0)$	(7) (6b)
Velocity	$v = \sqrt{2E/m_0}$	$v = c\sqrt{1 - \dfrac{1}{(1 + E/E_0)^2}}$	(8)
Momentum	$p = m_0 v$ $= \sqrt{2m_0 E}$	$p = \sqrt{2\,m_0 E(1 + E/2E_0)}$ $p = \dfrac{1}{c}\sqrt{2EE_0 + E^2}$	(9)
Wavelength	$\Lambda = \dfrac{h}{p}$ $= h/\sqrt{2m_0 E}$	$\Lambda = h/\sqrt{2m_0 E(1 + E/2E_0)}$ $\Lambda = hc/\sqrt{2EE_0 + E^2}$	(10)

The most important law of relativistic mechanics is the equivalence of energy and mass: $E = mc^2$. The total energy of mc^2 of an accelerated electron is the sum of the rest energy $E_0 = m_0 c^2$ and the kinetic energy $E = e \cdot U$, from equation (7). $E_0 = m_0 c^2$ corresponds to an energy of 0.511 MeV. The relativistic increase of the mass m can be formulated not only in terms of equation [6a] but also in energy described in equation

(6b), which follows directly form equation (7). The mass, therefore, increases linearly with increasing energy E; it reaches three times the rest mass m_0 at $E = 2E_0 \cong 1$ MeV.

The velocity v (8) cannot exceed the velocity of light c and can be obtained by comparing the right-hand sides of equation (6a) and (6b). At 100 keV, the electron velocity v reaches 1.64×10^8 m · s⁻¹, which is more than half of the velocity of light c. The electron momentum p from equation (9) is important because the conservation of both energy and momentum has to be considered in electron collisions. The radius of an electron trajectory in a homogenous magnetic field B and the Broglie wavelength Λ (10) also depend on the value of the momentum (10).

1.2   Wave optics of Electrons

Broglie has showed that an electron can be treated as a quantum of electron wave and that the relation $E = hv$ for light quanta should also be valid for electrons. Consequently, he postulated that the momentum $p = mv$ is also related by $p = hk$ to wave factor k, the magnitude of which (wave number) may be written $|k| = 1 / \Lambda$, in which Λ is the wavelength in meters. This implies that $\Lambda = h/p$ (10) with the relativistic momentum p (9). Substitution of the constants in (10) results in the formula

$$\Lambda = \frac{h}{mv} = \frac{1.226}{[U(1+0.9788*10^{-6}U]^{1/2}},$$ (11)

where Λ is the wavelength (nm) and U the potential (V) (11).

1.3   Elements of a Transmission Electron Microscope
1.3.1 Thermionic emission

The conduction electrons in metals or compounds have to overcome the work function ϕ_W if they are to be emitted from the cathode into the vacuum. The potential energy $V(z)$ of an electron in front of a conducting surface at a distance z larger than the atomic diameter can be calculated by considering the effect of a mirror charge with opposite sign behind the surface. With an electric field E, the potential energy $V = -e|E|z$ is superposed on that of the mirror charge, giving

$$V(z) = \phi_W - \frac{e^2}{16\pi\epsilon_0}\frac{1}{z} - e|E|z, \tag{12}$$

where $V(z)$ is the potential energy and ϕ_W is the work function.

Increasing the cathode temperature leads to a broadening of the Fermi distribution $f(E)$ at the Fermi level E_F, and for high temperatures, electrons in the tail of the Fermi distribution acquire enough kinetic energy to overcome the work function. The current density j_c (A · m^{-2}) of the cathode emission can be estimated from Richardson's law,

$$j_c = AT_C^2 \exp\left(-\frac{\phi_W}{kT_c}\right), \tag{13}$$

where k is Boltzmann's constant (k = 1.38 x 10^{-23} J · K^{-1}), T_C is the cathode temperature (K) and A is a constant that depends on the cathode material ($A \approx$ 12 x 10^5 A · K^{-2} · m^{-2}).

Most metals melt before they reach a sufficiently high temperature for thermionic emission. An exception is tungsten, which is widely used at a working temperature T_c of 2500-3000 K (melting point T_m = 3650 K). Lanthanum hexaboride (LaB$_6$) cathodes with T_c = 1400-2000 K are also employed because their work function is lower (Table 2). The tungsten metal evaporates continuously during operation, limiting the lifetime of the filament, which decreases from ~ 200 hours to 5 hours if T_c increases from 2500 K to 2900 K (12).

1.3.2 Field emission

The width b of the potential barrier at the metal-vacuum boundary decreases with increasing E; for $|E| \geq 10^9$ V m^{-1}, using a tip radius r $\leq$ 0.1 μm, the width b becomes less than 10nm, and electrons can penetrate the potential barrier by quantum-mechanical tunnelling effect. This means that the electron waves near the Fermi energy are reflected at the potential barrier but penetrate with an exponential decrease of their

amplitude ψ into the barrier. When the width b of the barrier is small, the amplitude at the vacuum side of the barrier is still appreciable and the

Table 2 – Parameters of thermionic and field-emission cathodes at $E = 100$ keV

Characteristic parameters:	
Cathode temperature T_c	Tip radius r of pointed cathodes
Work function ϕ_W	Diameter d of source
Emission current density j_c	Operating vacuum p
Gun brightness β at $E = 100$ keV	Field strength $\|E\|$ at cathode
Energy spread ΔE	
Thermionic cathodes: Tungsten hairpin	**Pointed LaB$_6$ rod**
$T_c = 2500\text{-}3000$ K	$T_c = 1400\text{-}2000$ K
$\phi_W = 4.5$ eV	$\phi_W = 2.7$ eV
$j_c \approx (1\text{-}3) * 10^4$ A/m^2	$j_c = (2\text{-}5) * 10^5$ A/m^2
$\beta = (1\text{-}5) * 10^9$ A/m^2 sr	$\beta = (1\text{-}5) * 10^{10}$ A/m^2 sr
$\Delta E = 1.5\text{-}3$ eV	$\Delta E = 1\text{-}2$ eV
$d = 20\text{-}50$ μm	$d = 10\text{-}20$ μm
$p \leq 10^{-3}$ Pa	$p \leq 10^{-4}$ Pa
$\|E\| \approx 10^6$ V/m	
Point-source cathodes: Field emission	
$T_c = 300$ K or ≈ 1500 K	$\Delta E = 0.2\text{-}0.7$ eV
$\phi_W = 4.5$ eV	$r \leq 0.1$ μm
$j_c \approx 10^9 \text{-} 10^{10}$ A/m^2	$d \approx 2.5$ nm
$\beta = 2 * 10^9 – 2 * 10^{13}$ A/m^2 sr	$p \leq 10^{-8}$ Pa
$\|E\| \approx 5 * 10^9$ V/m	

probability of tunneling across the barrier is proportional to $\psi\psi^*$. The emitted electrons do not need to overcome the potential barrier and it is not necessary to heat the cathode, whereas this is essential for thermionic emissions. If a field-emission

source is heated, it is mainly to prevent the absorption of gas molecules. The current density of field emission can be estimated for the Fowler-Nordheim formula

$$j = \frac{k_1 \, |E|^2}{\phi_W} \exp\left(-\frac{k_2 \phi_W^{3/2}}{|E|}\right), \tag{14}$$

where j is the current density, E is the electric field (V), ϕ_W is the work function and constants k_1 and k_2 are weakly depended on $|E|$ and ϕ_W (12).

1.3.3 Illumination system of TEM

One of the lens systems in a transmission electron microscope is the condenser lens-system and is responsible for the following things:

− focusing of the electron beam at the specimen, so a high intensity is maintained, even at high magnification;

− irradiation of the specimen area that corresponds with the viewing screen with uniform current density.

Transmission electron microscopes are equipped mostly with two condenser-lens-systems to make sure that the mentioned requirements are met.

The electron-gun system of the microscope can be adjusted onto the axis of the condenser lens system by the tilting/shifting of the gun system. This is also required to bring the electron beam onto the axis of both the objective-lens system and the magnification-lens systems.

The alignment procedure needed to bring the electron beam in the right position involves a mechanical shift and tilt of the condenser-lens system [12].

1.3.4 Specimens

The maximum thickness of the sample depends on the type of interactions occurring between the electron beam and specimen, together with the mode of operation. For high resolution, it is of high importance to apply phase-contrast methods. The image contrast observed is then caused by the interference of the

scattered waves with the unscattered incident electron beam. The phase contrast reduces when the thickness of the specimen increases.

For medium and low-resolution mode, most work on amorphous specimens rely on scattering contrast. In the bright-field mode, the intensity of the image is depending on the electron dose applied to the specimen.

Another example of gaining contrast is the negative stain method. In biological specimens, the scattering contrast is increased by staining thin sections with heavy metals. These will form a barrier surrounding the specimen. On the image, the background will appear black, while the protein will appear white on the image.

Another way to look at biological specimens is by mounting them on copper grids with different meshes, mostly 100-200 μm meshes. The meshes are covered with a supported film which needs to possess the following properties:

 − low atomic number to reduce scattering of the primary beam;

 − high mechanical strength;

 − resistance to electron irradiation and heating;

 − electrical conductivity to avoid charging;

 − low granularity for high resolution;

 − easy preparation.

For high resolution, the granularity of carbon films is useful for investigating the contrast transfer function (CTF).

The specimen grids are mounted in a specimen cartridge, which can be inserted in an airlock of the TEM. The specimen position is in a strongly excited objective lens.

The principle methods of specimen manipulation for imaging can be observed in table 3. Specimen rotation about an axis parallel to the electron beam can be applied to get the specimen in convenient positions for the final image. Tilting of the specimen can produce stereo pairs for quantitative measurements and observation of the three-dimensional specimen structure.

Specimen cooling devices operating at temperatures below -150 °C can be used to reduce contamination of the specimen and protect against irradiation damage [12].

Table 3 – Specimen manipulation

Specimen manipulation	Applications
1. <u>Specimen rotation</u> Rotation on an axis parallel to the electron beam	Orientation of specimen structures or diffraction patterns relative to the edges of the final screen
2. <u>Specimen tilt</u> Tilt about an axis in the specimen plane	Determination of orientation Favorable orientation of biological sections
3. <u>Staining devices</u> Staining of the specimen by movement of two clamps by mechanical effects	Staining of metals and high polymers
4. <u>Specimen heating</u> Direct heating of the grid or indirect heating	Recovery and recrystallization Precipitation and transition phenomena
5. <u>Specimen cooling</u> Cooling between -100°C and -150 °C degrees Cooling with liquid nitrogen	Temperature sensitive specimens Decrease of specimen contamination Direct observations of cryo-sections Decrease of radiation damage Superconducting states

1.3.5 Imaging system of TEM

In the different lens systems, most is demanded from the objective lens. The objective lens is responsible for the main magnification of the image. The objective lens must have several mechanical tolerances to perform this task. There can be several aberrations present in the lenses:

- spherical aberration;
- astigmatism;

19

- field curvature;

- distortion;

- coma.

The spherical aberration has the effect of reducing the focal length for electrons passing through outer zones of the lenses. Every lens has its own spherical aberration coefficient used to correct during processing. For most objective lenses, the coefficient falls within the range of 0.5-2.0 mm. The spherical-aberration coefficients of the objective lens do not only influence the resolution of the image but can also be observed when looking at crystalline specimens. The value of the coefficient is also necessary for the determination of the contrast transfer function (CTF).

To get a good image, the electron beam should be circular when interacting with the specimen. Once the probe beam is in the form of an eclipse, then we speak of astigmatism. The presence of an astigmatic beam can pose difficulties in focusing of the beam.

In order to achieve high resolution images, the astigmatism of an objective lens should be so small that the stigmator can correct for these distortions in the probe beam. Resolution limiting factors such as spherical and chromatic aberrations are only of importance to the objective lens systems because a magnification M decreases the apertures for the following lenses to $\alpha = \alpha_0/M$. Even for a small magnification (20-50x) at the first intermediate image, the aperture becomes so small that the spherical aberration of the intermediate and subsequent lenses can be neglected.

Any of three or four diaphragms of 20-200 μm diameter can be inserted in the focal plane of the objective lens, therefore allowing the objective aperture α_0 to be changed. The nature of the objective diaphragm is important for the quality of the image. The diaphragm has to be of heat resisting material, capable of tolerating the largest possible current density in the focal plane [12].

1.3.6 Image recording & Electron detection

The final image of a transmission electron microscope can be observed on a fluorescent screen consisting of *ZnS* or *ZnS/CdS* powder. The light intensity of a fluorescent screen is proportional to the electron current density. The light-generating efficiency decreases when the electron range exceeds the thickness of the fluorescent layer, which can be a problem in high-voltage electron microscopy.

For digital imaging processing and electron microscope alignment it is highly interesting to record two-dimensional arrays of pixels directly. A charge-coupled-device (CCD) image sensor consist of an array of silicon-based photodiodes (pixels). The absorbed quanta of light will generate electron-hole pairs, which are separated in the depletion layer of the diodes, and the electrons are accumulated in the potential wells of the diodes during storage. By applying sequences, the charges can be transferred to a serial shift register, after which a built-in amplifier and ADC transfers the signal to an external buffer store [12].

1.4   Limitations of single-particle cryo-EM

The fact that cryo-EM would pose some limitations already became clear several years ago. In 1995, *Henderson* provided a rough estimation of the smallest-size structure based on a hypothetical model. Using several parameters, *Henderson* showed that, in principle, structures can be determined for any molecule with a size larger than ~100 kDa [13]. Nevertheless, many proteins structures, including EspB, are smaller than the proposed barrier. The average human protein length is around 375 amino-acids which corresponds to a protein weight of ~ 42 kDa [14]. This demonstrates the need for technological improvements for cryo-EM.

After the initial theory in 1995, *Rosenthal & Henderson* described the constraint of cryo-EM, using a Guinier plot [15]. The Guinier plot, based on the initial theory of *Henderson*, shows the natural logarithm of the average structure factor *F* as a function of resolution, $1/d^2$ (Figure 2). The structure factor describes how the material scatters incident radiation and is critical in the interpretation of the scattering patterns obtained in electron experiments. As observed in *figure 2*, two regions were identified by *Rosenthal & Henderson*; a sharp declining region at low resolution ($d > 10Å$) and a high-resolution region ($d < 10Å$) which describes the atom scattering.
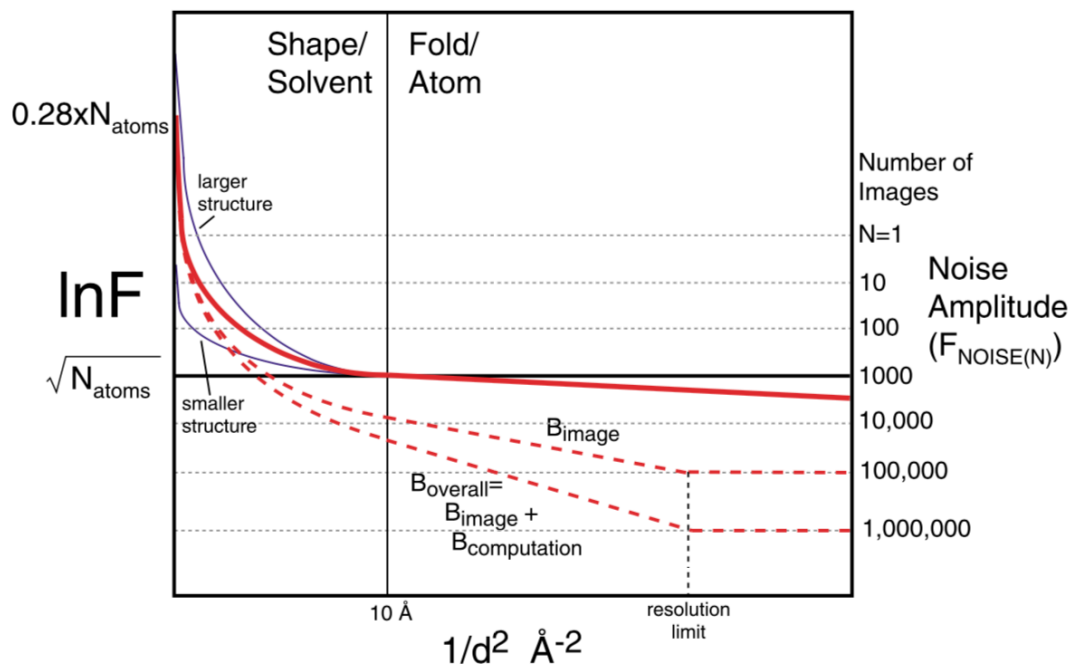

Figure 2 – Guinier plot, based on the theory of Henderson [13]. The graph shows the natural logarithm of the spherically averaged structure amplitude (F) for a protein against the resolution ($1/d^2$), where d is the resolution ($Å^2$). The red line consists of a low-resolution region, determined by the shape of the molecule, and a high-resolution region, defined by the scattering of the atoms. The noise amplitude is described by the dotted lines $F_{NOISE\,(N)}$. The resolution limit, is indicated, where the structure curve (F) equals the noise level. Transcribed from [15]

The noise level of an acquired image limits the ability to record a high-resolution map. Images of identical particles can be averaged, resulting in a higher signal-to-noise ratio (SNR), and therefore the high-resolution information embedded in the image can be accessed. The number of images required can be deducted from the resolution limit

in figure 2, whereas the single temperature factor B-overall models the structure factor amplitude decay curve (*F*). The resolution that can be observed is determined by the intersection of the average structure factor amplitude (*F*) and the noise level, and the noise level is determined by the number of particles. That the resolution is dependent on the number of particles can be discovered in figure 3. In Figure 3, the black line with a temperature factor of $B = 1000$ Å^2, is plotted against the resolution *1/d²,* where the structure factor amplitude, *F* in figure 2, is expressed as particle sum [15]. The temperature factor *B*, is known in literature as Debye-Waller Factor (DWF). The DWF describes the effect of thermal vibrations of atoms on the beam diffraction [16]. As seen in figure 3, a lower DWF factor requires less particles to achieve the same resolution. The atoms with a lower *B* factor are less flexible, the structure is well ordered and therefore the diffraction of the beam is more coherent.
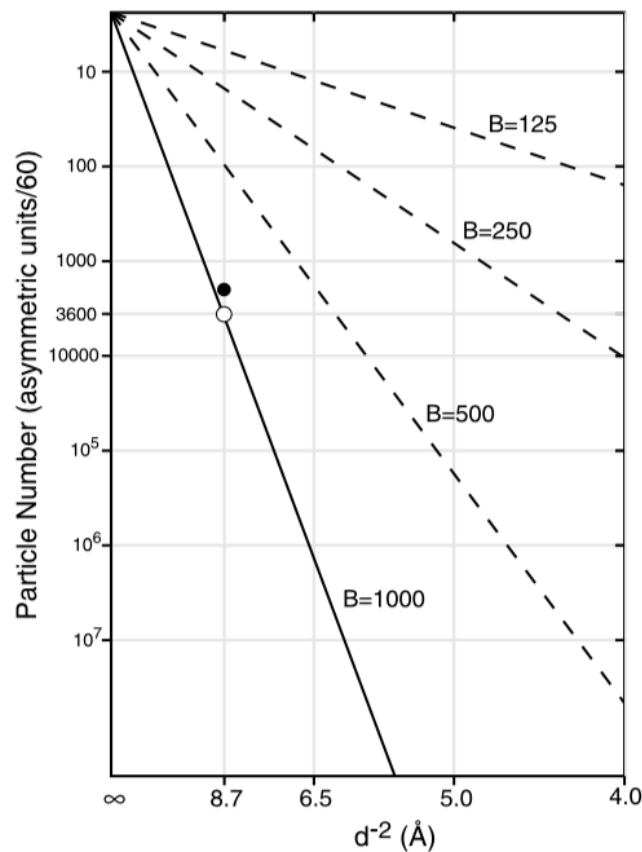

Figure 3 – Estimation the number of identical particles required to attain a high-resolution map, as a function of resolution (1/d2). The black continues line, B =1000 Å2, is the average amplitude decay function (F) expressed in particle numbers. The dotted lines describe F as well, but with a different temperature factor, respectively 500, 250 and 125 Å2. Transcribed from [15]

1.5   Volta Phase Plate (VPP)

The current obstacle in the structural determination of biological complexes by means of cryo-EM is the amount of data required for a three-dimensional reconstruction [9]. The application of phase plates has a steadily growing interest over the last couple of years and the technology has the ability to improve the structural determination of protein complexes by cryo-EM [17, 18].

In-focus images, namely biological samples have a low signal-to-noise ratio (SNR) and have therefore low contrast [9]. A phase plate is thin carbon film, placed in the back-focal plane of the microscope, which induces phase contrast by introducing a phase shift between the scattered and unscattered electron beams. Ideally, the phase shift should be $\pi/2$ (90 degrees). However, in practice the phase shift is never optimal and lies within the range of $\pi/4 - 3\pi/4$, respectively [17, 19].

The application of phase plate contrast methods has been developed over the past fifteen years. The most promising phase plate was the Zernike phase plate. The Zernike phase plate consists of a thin carbon-film with a thickness which is depending on the voltage of the microscope. Furthermore, the Zernike plate has a small hole in the centre for the central beam of unscattered electrons [17]. Even though it induced sufficient phase contrast, some practical issues became apparent in the practical use of the Zernike plate. The most limiting factor was the short lifespan ($\sim$ 7 days) of the Zernike plate and the creation of artefacts in the images [20].

As a successor of the Zernike phase plate, the Volta Phase Plate (VPP) was developed in the last couple of years. The VPP facilitates in-focus phase contrast in cryo-EM. It is the first phase plate where near-atomic resolutions were reported [21]. The VPP consists of a thin amorphous carbon film, around $\sim$ 10nm thick. The carbon film is continuously heated to a temperature of $\sim$ 200°C to prevent contamination caused by the electron beam [9, 17]. There are two types of phase plates; A positive phase plate, which advances the phase of the scattered beam, relative to the unscattered beam. Consequently, the negative phase plate retards the phase of the scattered beam, relative to the unscattered beam [19]. In the case of a negative phase plate, which is

most applied in current research, the induced phase shift (PS) gradually increases while images are acquired. However, the gradual increase in phase shift is not a limiting factor; It is a parameter to take into account during automated data collection. The fact that no hole is required, has a great practical advantage since a phase shift can be created at any spot of the VPP and data acquisition can be completed automatically. The phase shift contributes to the contrast transfer function (CTF) and can be included as parameter in the CTF correction during processing [9, 17].

1.6 Cut-on periodicity of VPP

As mentioned above, during data acquisition a gradual increase in phase shift can be observed. Data is generally acquired between a low phase shift of $0.2\ \pi$ and a high phase shift of $0.8\ \pi$. The growing phase shift is beneficial for single particle analysis; the phase shift changes the position of the zeros in the CTF, therefore change of defocus is not necessary during data acquisition. In practice, the slower the phase shift evolves, the more images can be acquired at each position on the VPP [21].

After the phase shift has progressed to $0.8\ \pi$, it is desirable to move to a new position on the VPP. It has been previously established that a higher PS can cause deformations in the CTF and therefore weaken the quality of the data collected [21]. The inclusion of the PS in the CTF gives the following formula, transcribed from *Danev, 2017* [21]:

$$|CTF\ (k)| = \left| \sin\left[-\varphi \left(1 - e^{\frac{-k^2}{2s^2}} \right) + \pi \left(-\Delta z \lambda k^2 + 0.5\ C_s \lambda^3 k^4 \right) \right] \right|, \quad (15)$$

where φ is the phase shift in radians, k is the spatial frequency, s is the radius of the VPP spot, Δz is the defocus, λ is the electron wavelength and C_s is the spherical aberration coefficient.

The first term of the formula describes the so-called cut-on frequency of the phase plate. The cut-on frequency is the spatial frequency at which the VPP induces

the phase shift in the CTF and is inversely related to the spot size of the VPP, *s* in formula (15) [19].

When the spot size becomes infinite small, the PS is considered ideal (Figure 4). In practice, a PS of 0.5 π (dotted lines), shows a CTF profile which is similar to the increase in PS. When the PS is high (0.9 π), the CTF develops quickly and saturates quickly. This leads to the fact that the first zero of the CTF is shifted towards the lower spatial frequencies, resulting in blurry images [21]. The black dotted line, which describes the ideal phase shift, does not take the cut-on periodicity of the phase plate into account. Once the cut-on periodicity is added to formula (15), the curve changes into red line (Figure 4). The difference between the ideal situation and practical findings indicate that images with a high PS are not dealt with properly during data processing [21].
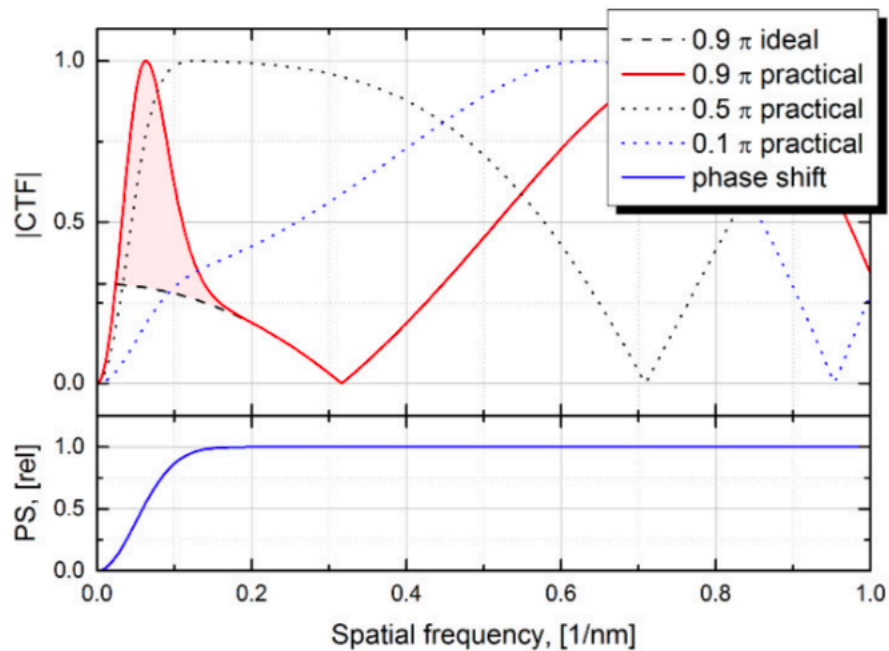

Figure 4 – Contrast transfer function (CTF) with different phase shifts. The developing phase shift is depicted in the bottom halve (blue line). Black dotted line represents the ideal PS, with a PS of 0.9 π. Red line represents the practical CTF, with a phase shift of 0.9 π. Transcribed from [21]

1.7   RELION workflow

RELION is software used for cryo-EM structure determination. It consists of several steps to completely process raw data into a 3D structure. After importing the recorded movies, beam induced Motion correction can be applied. Beam-induced motion correction uses the MOTIONCOR2 [22]. It allows whole movie alignment. It used parameters such as pixel size, B-Factor. The movies are divided in patches, defined by the user. The movement of the patches during the whole movie is tracked and corrected for in the end. The movies can also be dose corrected. In order to correct for the dose applied at the sample, the voltage of the microscope and the dose per frame (e/A^2) should be inserted. The estimated beam-induced shifts can be visualized in the created file, which looks like figure 5.
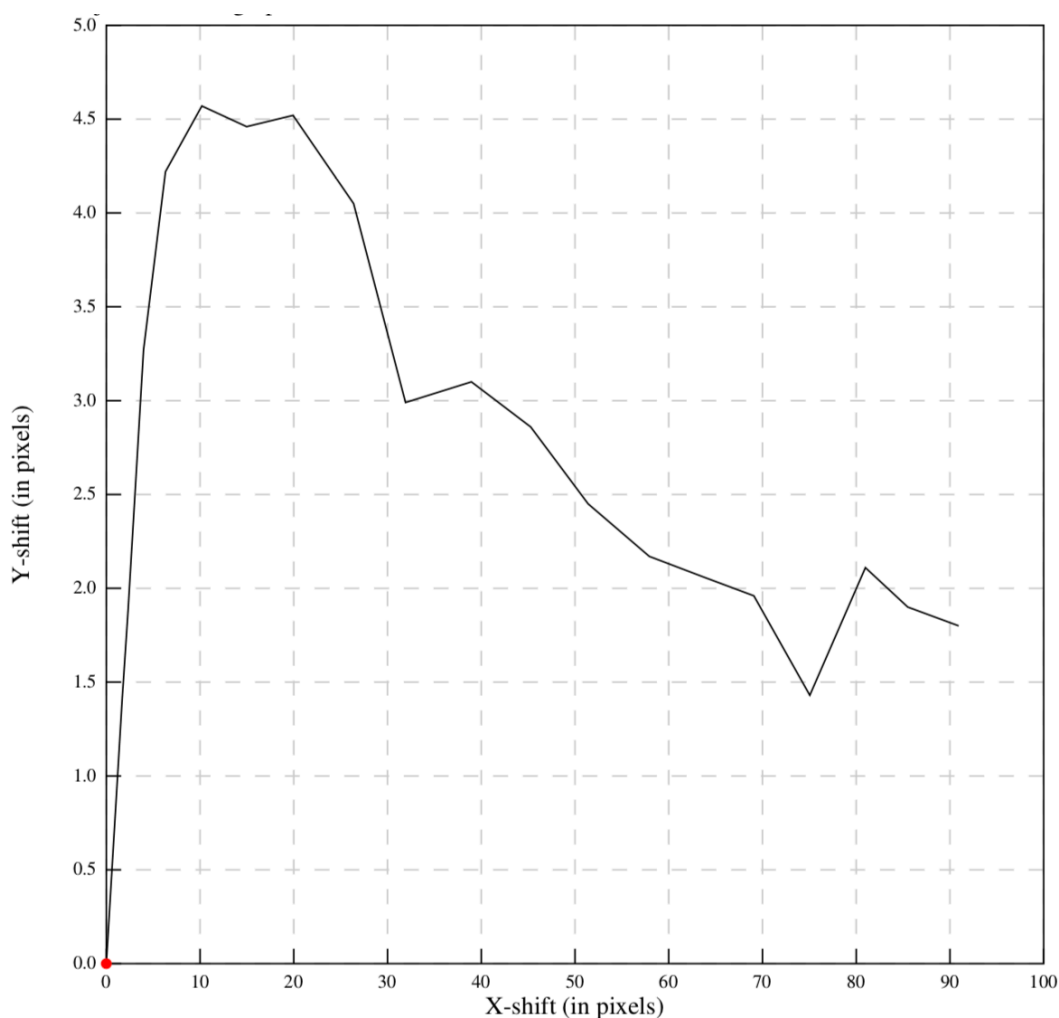

Figure 5 – Example of beam-induced correction using Motioncor2

After beam-induced motion correction, the next step in the processing of the data is CTF correction. During this step, the CTF parameters will be estimated for each corrected micrograph. The CTF uses the GCTF function by Zhang [23]. It requires the following parameter from the user:

- spherical aberration (mm);
- voltage (kV);
- amplitude contrast;
- pixel size;
- astigmatism;
- minimum & Maximum resolution;
- defocus values;
- estimation of phase shifts, when phase plate was used during data acquisition.

The GCTF function also allows for several additional options such as:
- local refinement;
- phase flipping of the data set;
- B-factor.

After finding the best parameters for CTF correction, unsatisfactory micrographs can be discarded by selecting only the ones which have good Thon-ring fitting. After selected the satisfactory micrographs, manual particle picking can be carried out.

During manual picking, a couple of micrographs is selected. Manual picking is best to do to get acquainted with the collected data. The picked micrographs can then be subjected to reference-free 2D classification, where the averages calculated can be used for automated particle picking. Parameters required for manual particle picking include: Particle diameter, low-pass filter (which can help with noisy images) and the pixel size.

One a coordinate file for each micrograph is created (through either manual or automated particle picking) the particles can be extracted from the micrographs. The extracted particles can be used to create a template for automated particles.

Template of the manual picked particles can be created via 2D classification of the extracted particles. Parameters which are required are:

- number of classes;
- regularisation Parameter T;
- number of iterations;
- mask Diameter (Å), the diameter should be longer then the longest dimension of the picked particle;
- limit resolution E-step (Å): To create templates it is often not required. However, by setting a value, frequencies beyond that value will not be included in the averaging and can be useful for difficult classifications (for instance with noisy data).

Selection of suitable templates is again carried out in the subset selection job type. After selecting the satisfactory class averages, these can be used for automated particle picking. Automated particle picking is dependent on several parameters including:

- pixel size;
- low-pass filter references (Å);
- mask diameter;
- picking threshold;
- minimum inter-particle distance (Å).

Then the particles can be extracted once more. After extraction two thing can be done: 1) Particles can be subjected for 2D classification or 2) can first be sorted. The RELION particle sorting option is based on the difference between each particle and their aligned, CTF-corrected reference. Particles can then be sorted based on their Z-score and also be subjected to 2D classification as well.

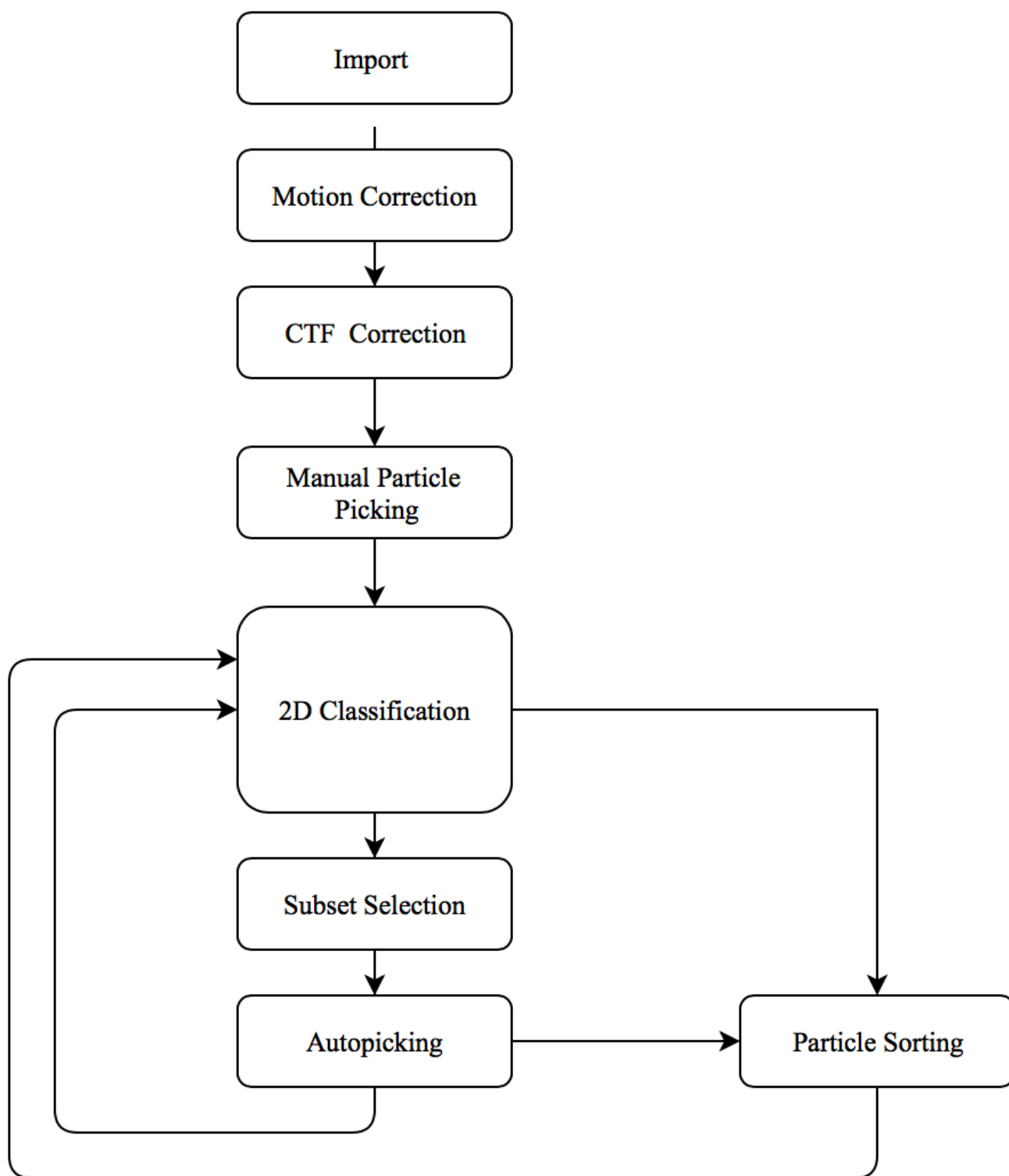Figure 6 gives a general overview of the RELION workflow.

Figure 6 – General RELION workflow with different step in the process of data reconstruction

2 Materials & Methods

2.1   Expression of EspB

Day 1: From the primary culture, one colony was taken and inoculated with 5 mL Luria-Bertani (LB) medium, supplemented with 100 μg/mL of ampicillin and 25 100 μg/mL of chloramphenicol. The colony was incubated for a period of 18 hours, at 37°C with 250 rotations per minute (RPM) shaking. The optical density (OD) was measured at 600nm by diluting 50 μL in 950 μL of medium.

Day 2: From the secondary culture, after 18 hours, the primary culture was inoculated with 500mL Overnight Express$^{\text{TM}}$ Instant LB medium (supplemented with 100 μg/mL of carbenicillin and μg/mL of chloramphenicol) to an OD_{600} of 0.05. Secondary culture was incubated for 60 hours, at 25°C with 250 rpm shaking. The optical density (OD) was measured at 600nm by diluting 50 μL in 950 μL of medium. Cell harvesting was carried out by centrifuging at 5,000 rpm, at 4°C. Pellet was used immediately or otherwise stored at -20°C.

2.2   Purification of EspB

The pellet was resuspended in 70 mL of 20 mM of Tris pH 8, 300 mM NaCl, 1 mM phenyl methane sulfonyl fluoride (PMSF) and 10 μL of Benzonase.   Cell lysis was carried out with EmulsiFlex-C3. One cycle was carried out without pressure and four cycles with a pressure of 5000 and 10000 psi. The solution was centrifuged at 5,000 rpm/4°C for 30 minutes. The pellet was discarded, and the supernatant was centrifuged at 40k rpm at 4°C for one hour. The supernatant was mixed with Ni-NTA resin (20 mM Tris pH 8, 300 mM NaCl) and shacked in a tube roller overnight, at 4°C.

The next day, shacking was stopped, and supernatant was taken away. The resin (which is still present in the other container) was added to a clean column and the collection of the flow through the column was started. The supernatant was added little by little, until everything has been loaded in the column. Before the column turned dry, 20 mM Tris pH 8, 300 mM NaCl supplemented with 20 mM Imidazole was added (30

mL). After a wash of the column, the EspB was eluted with 20 mM Tris pH 8, 300 mM NaCl, 400 mM Imidazole (around 20 mL).

The sample was diluted 6-fold with 20 mM Tris pH 8 buffer. The diluted sample was loaded into the Biorad Fast Protein liquid chromatography (FPLC) and the method was started. The sample was loaded continuously, to load the entire sample. Right before the final volume of the sample, 20-30 mL of 20 mM Tris pH 8, 10 mM NaCl was added. The 96-wells plate was placed in place to collect the fractions generated by the FPLC.

Once the ion exchange was finished, the concentration of the fractions was started. This was done from the first till the second peak (generated by the FPLC) with Amicon Ultra 15 mL Centrifugal Filter of EspB. The sample was concentrated to 500 µL and then diluted with 3 mL of 20 mM Tris pH 5.5, 150 mM NaCl buffer. After one hour, the concentration of the sample was finished. The sample was centrifuged at 10k/4°C for five minutes. The sample was injected in the FPLC once more and run. The 96-plate was placed in position to collect the different fractions. Finally, a sample was run of every purification step on a 12% SDS-PAGE gel.

2.3   Sample preparation

1.5 μL of the EspB sample (with a concentration of 0.5 mg/mL) was placed on a Quantifoil R1.2/1.3 Cu grid, with a 300 mesh. The grid was vitrified using the VitroBot (FEI). The Vitrobot offers automated vitrification, which is easy to complete. Parameters used for the Vitrobot: humidity of 95%, a blot force of 5, with a single blot. The blot time was three seconds, before plunge-frozen in liquid ethane ($\sim$ -150°C) before stored in liquid nitrogen.

2.4 Data acquisition

The data was collected on an FEI Tecnai Arctica (FEI) cryogenic TEM, operated at 200kV and equipped with a Falcon II detector (FEI). The data was acquired using the following parameters: magnification x110,000, 50 μm C2 aperture, spot size 4, 42.158% C2 lens, 35.121% diffraction lens, pixel size 0.935 Å. The data was acquired by the automatically with EPU software (FEI). The phase plate PhP3 was aligned and started at position 74. Requested defocus -0.5 μm. The periodicity of the phase plate was set to 50 exposures, with an activation time of 20 s. 12 squares were selected on the grid, containing 1291 holes selected for data acquisition. In total, 966 holes were acquired successfully.

2.5 Data processing of EspB

After initial screening of the dataset, the VPP collected dataset contained 704 images. The movies were aligned using MotionCorr2. Parameters were 5 x 5 patches, dose-weighting 2 e-/Å2, pixel size 0.935 and a b-factor of 100 was applied. The CTFs were fitted on the dose-weighted micrographs with Zhang's Gctf (23). Gctf is a GPU accelerated program for real-time CTF determination and correction. Following parameters were used: spherical aberration 2.7 mm, voltage 200kV, amplitude contrast 0.1, minimum resolution 20 Å, maximum resolution 4 Å, minimum defocus 3000 Å, maximum defocus 7000 Å (step size 300 Å), minimum phase shift 40 degrees, maximum phase shift 150 degrees with a step size of 10 degrees. Additionally, the astigmatism was 200 Å and the additional validation option was used. A B-factor of 150 was applied during CTF correction. After visual inspection, 391 CTF-corrected images were used for the selection of particles.

Particle picking was carried out manually using the GPU-accelerated beta version of RELION 2.1 [24], with a particle diameter of 100 Å, pixel size 0.935 Å. A selection of the 391 images was used to carry out the particle picking. Manual picked particles were extracted with a box size 192, based on the particle diameter and pixel

size. From the selected images, 458 particles were extracted and were subjected to 2D-classification using the following parameters: 10 classes, with regularization parameter of 2, 40 iterations and a mask diameter of 120 Å. After 2D classification, the good 2D classes were selected to use as a reference during the automated particle picking. Automated particle picking was carried out on the 391 selected images, with a particle diameter of 100 Å, pixel size 0.935 Å and an inner-particle distance of 80 Å. 160308 particles were extracted and sorted. Sorted particles with a Z-mas score below 0.8 were selected. These particles were again subjected to 2D classification (20 classes, regularization parameter 4, 50 iterations and a mask diameter of 120 Å). The best 2D classes were selected (10 classes) and contained 61618 particles (38.44% of the whole data set). Reported resolutions of the individual classes were below 4 Å.

2.6   Data processing of EspB C-terminal domain

Apart from generating 2D classes for the oligomer of EspB, additional research was carried out trying to visualize the C-term of EspB. It was carried out on a selection of the 391 images (40 in total). Automated particle picking was carried out on those 40 images with a particle diameter of 50 Å (in the center of the oligomer), pixel size 0.935 Å. A total of 13647 particles were extracted, with a box size of 92. The particles were subjected to 2D classification using the following parameters: 10 classes, regularization parameter 2, 50 iterations and a mask diameter of 70 Å. The particles were sorted and particles with a Z-mas score below 0.8 were again selected. The selected particles were subjected to 2D classification (10 classes, regularization parameter 4, 100 iterations, a mask diameter of 80 Å and limit resolution E-step to 4 Å). The best 2D classes were selected (5 classes) and contained 11719 particles (85.87% of the whole data set). Reported resolutions of the individual classes were all below 4 Å.

3   Results
3.1   Simulation of the cut-on frequency

As described earlier, the VPP contains a cut-on periodicity in which the phase shift needs to gradually increase. As observed in figure 4, the ideal phase shift of 0.5π leads to a direct increase of the CTF, so maximal information can be acquired from the image, in both the low and high spatial frequency regions. In figure 7, a Matlab simulated CTF is displayed, with a phase shift of 0.5π. The blue line represents the CTF from equation [15] without the cut-on periodicity, while the red line describes the CTF function from equation [15] including the cut-on periodicity.
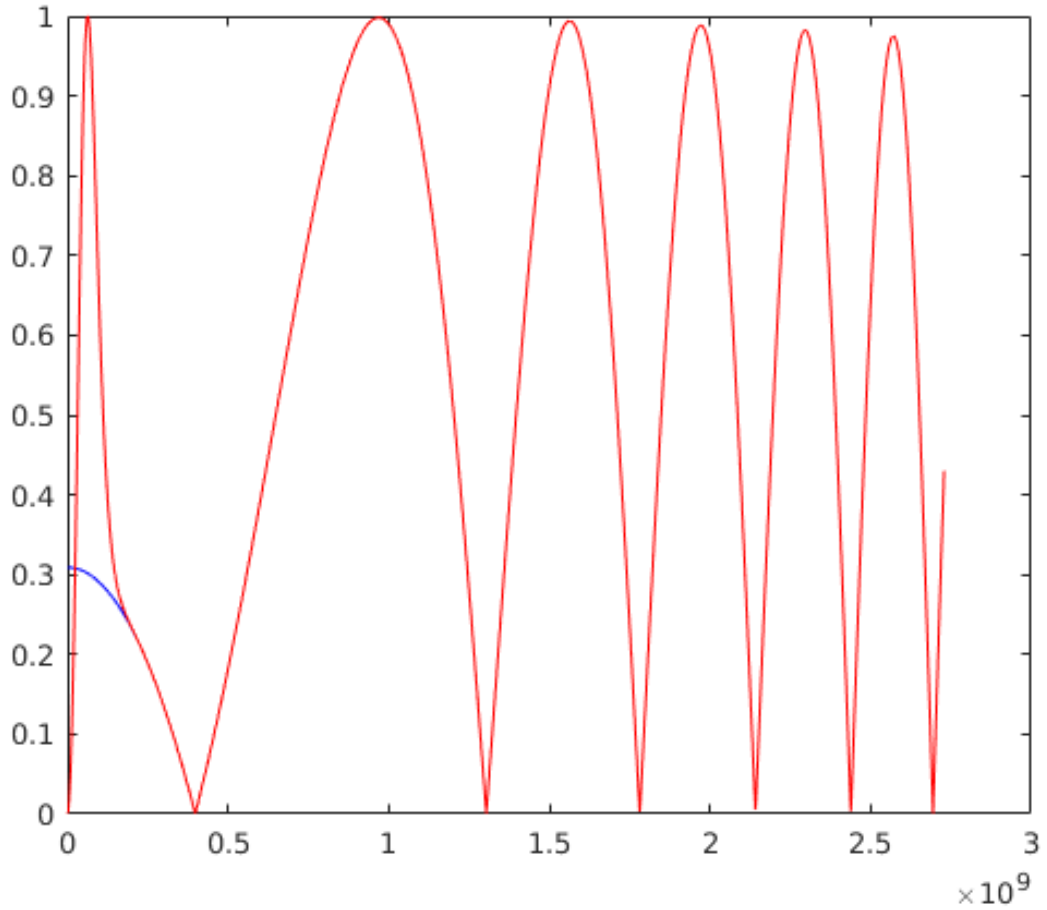

Figure 7 – Matlab simulation of the contrast transfer function, including the cut-on periodicity in red, and without the cut-on periodicity in blue. CTF is based on equation (15) and has a phase shift of 0.5 π

Similar results are observed in figure 8 & 9, whereas figure 8 simulates the CTF with a high phase shift of $0.9\,\pi$ and figure 9 the CTF simulates with a low phase shift, of $0.1\,\pi$.

Figure 8 – Matlab simulation of the contrast transfer function, including the cut-on periodicity in red, and without the cut-on periodicity in blue. CTF is based on equation (15) and has a phase shift of $0.9\,\pi$

From these simulations we can observe the following; The simulations correspond to the CTF function described in equation (15) based on the findings of *Danev et al*. Even though there seems to be little difference in the red and blue line, the implementation of the cut-on periodicity is important during further data processing to maintain the quality of the data acquired.
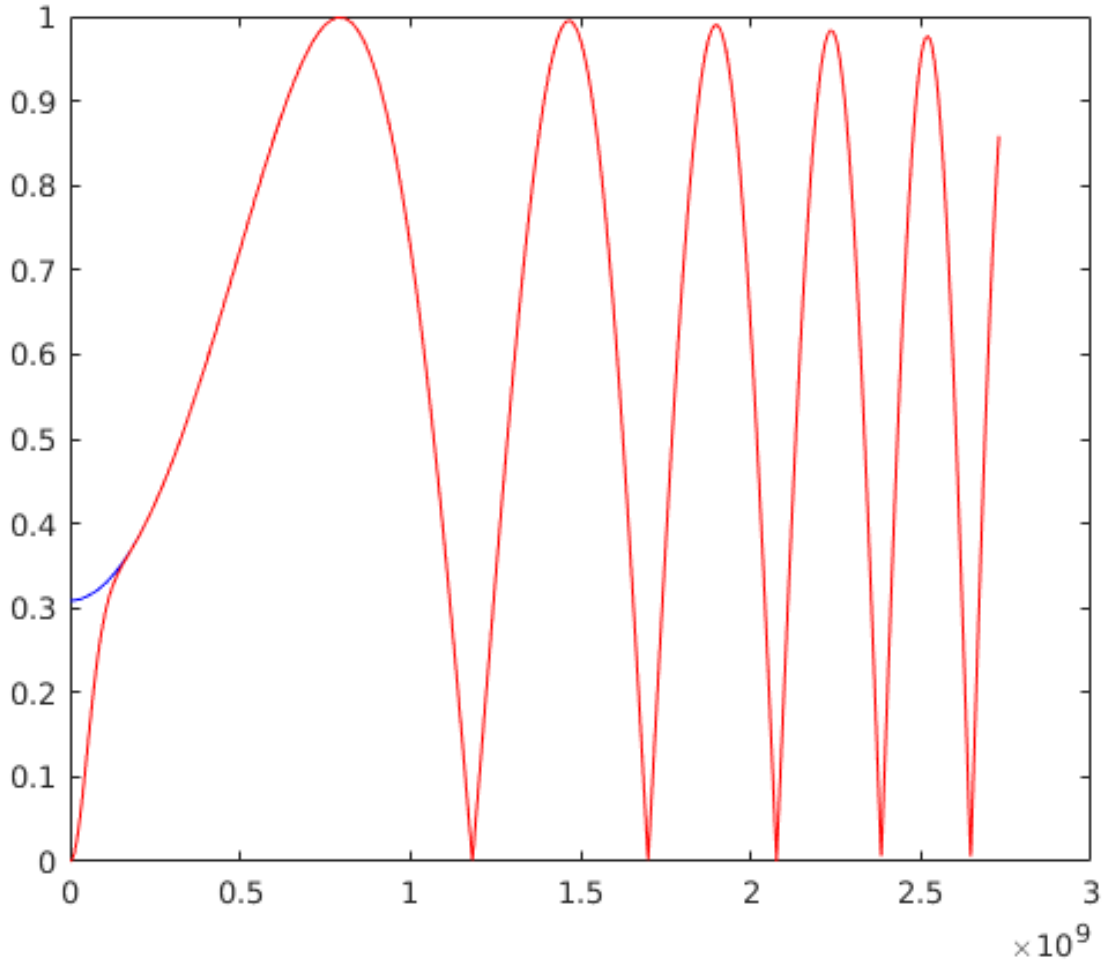
Figure 9 – Matlab simulation of the contrast transfer function, including the cut-on periodicity in red, and without the cut-on periodicity in blue. CTF is based on equation (15) and has a phase shift of 0.1 π

3.2   Phase shift, Astigmatism and defocus of the dataset

As described earlier, a phase shift of 0.5π is considered ideal, corresponding to 90 degrees. However, in practice the phase shift builds up when the VPP is exposed to the electron beam. In figure 10, the phase shift built-up for the data set (containing 391 images) is displayed. From the figure, it is clear to see when the VPP was moved towards a new position; It corresponds to a drop in the phase shift. The mean phase shift is 115.8 degrees, with a standard deviation of 23.7 degrees. The minimum phase shift is 25 degrees and the maximum phase shift is 160 degrees.
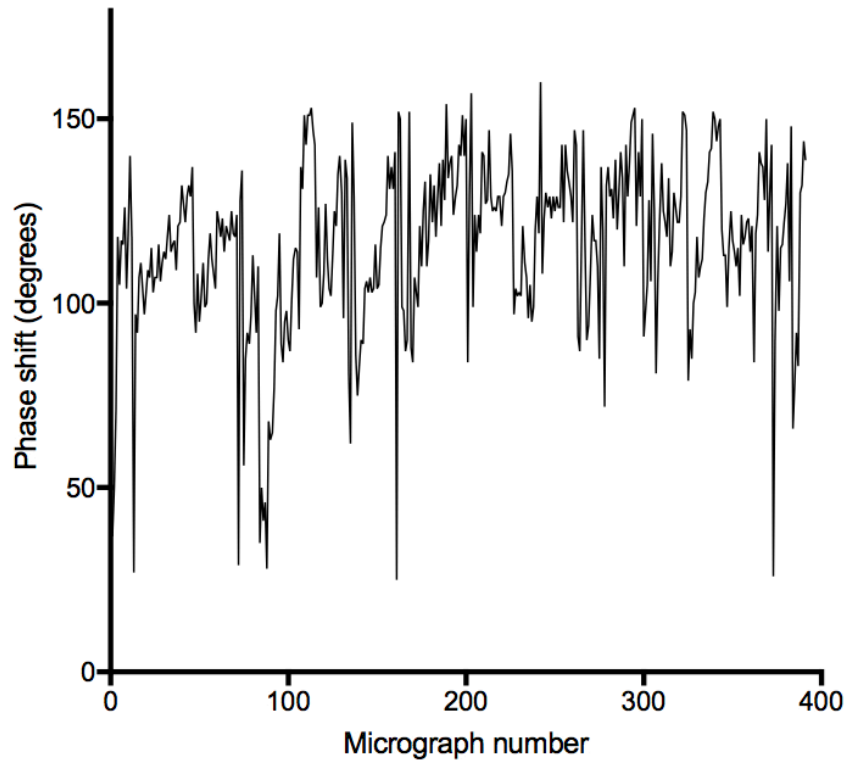
Figure 10 – Phase shift development for the entire dataset, including 391 images. Mean phase shift is 116 degrees, with a standard deviation of 24 degrees. The minimum is 25 degrees and the maximum is 160 degrees

In figure 11, the built-up of the defocus during the dataset is displayed. During setting up data acquisition, a defocus of 0.5 μm (5000 Å) was requested. From figure 11 it becomes clear that the defocus is not steady, it gradually increases and decreases as result of beam-induced movements and the movement of the phase plate. However, the average defocus value is 0.49 μm (4984 Å) with a standard deviation of 1247 Å. Most of the defocus values lie within the range applied during CTF correction, 3000-7000 Å.
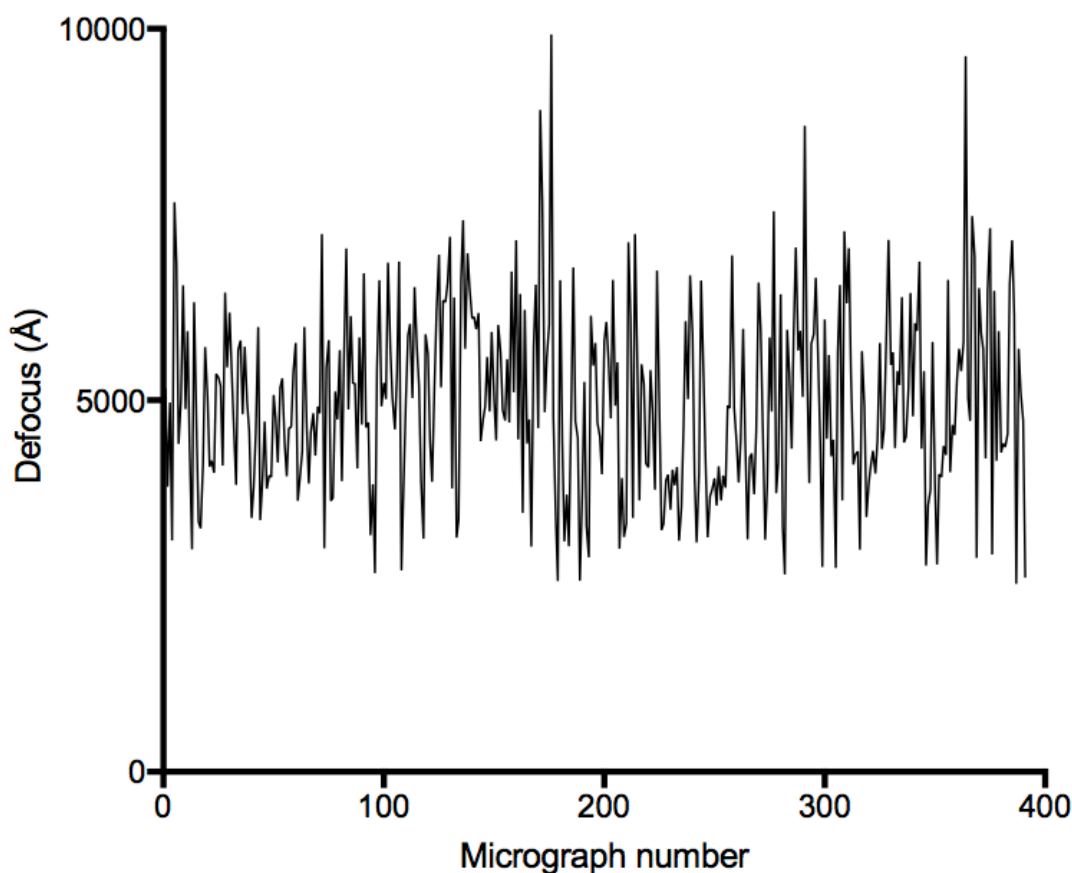
Figure 11 – Development of defocus during the data collection. Mean defocus is 4984 Å, with a standard deviation of 1247 Å

Figure 12 describes the amount of astigmatism of the beam during data collection. From figure 12 we can observe that there is quite some variation in the amount of astigmatism. During CTF correction a final value of 200 Å was used for correction. Even though there is quite some variation, the mean astigmatism was 193 Å with a standard deviation of 139 Å; A minimum of 4 Å and a maximum of 868 Å was present in the entire data set.
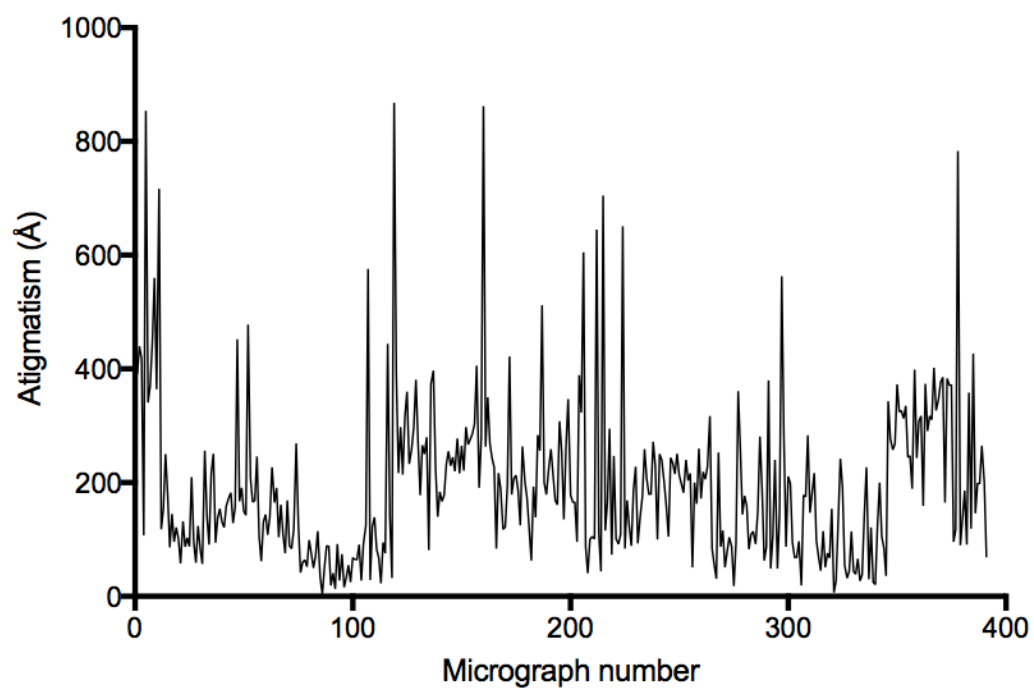
Figure 12 – Astigmatism in the entire data set. Mean of 193 Å and a standard deviation of 139 Å

3.3   2D classification of EspB

From literature and previous experiments conducted by the research group, it became clear that EspB forms an oligomer, which can be built-up from either 6, 7 or 8 monomers. Figure 13 shows the final 2D class averages of the subjected particles. 2D class averages built from 6 monomers can be observed in figure *13H* and *13I,* while figure *13A-C, E-G, J* can be categorized as 2D class averages built from 7 EspB monomers. The oligomers consisting 8 monomers are more difficult to spot, but we hypothesize that figure *13F* shows 8 monomers since it has a rounder shape, which implicates more monomers. A detailed description of the individual classes can be found in table 4.

Table 4 – Detailed description of 2D class averages, from figure 13

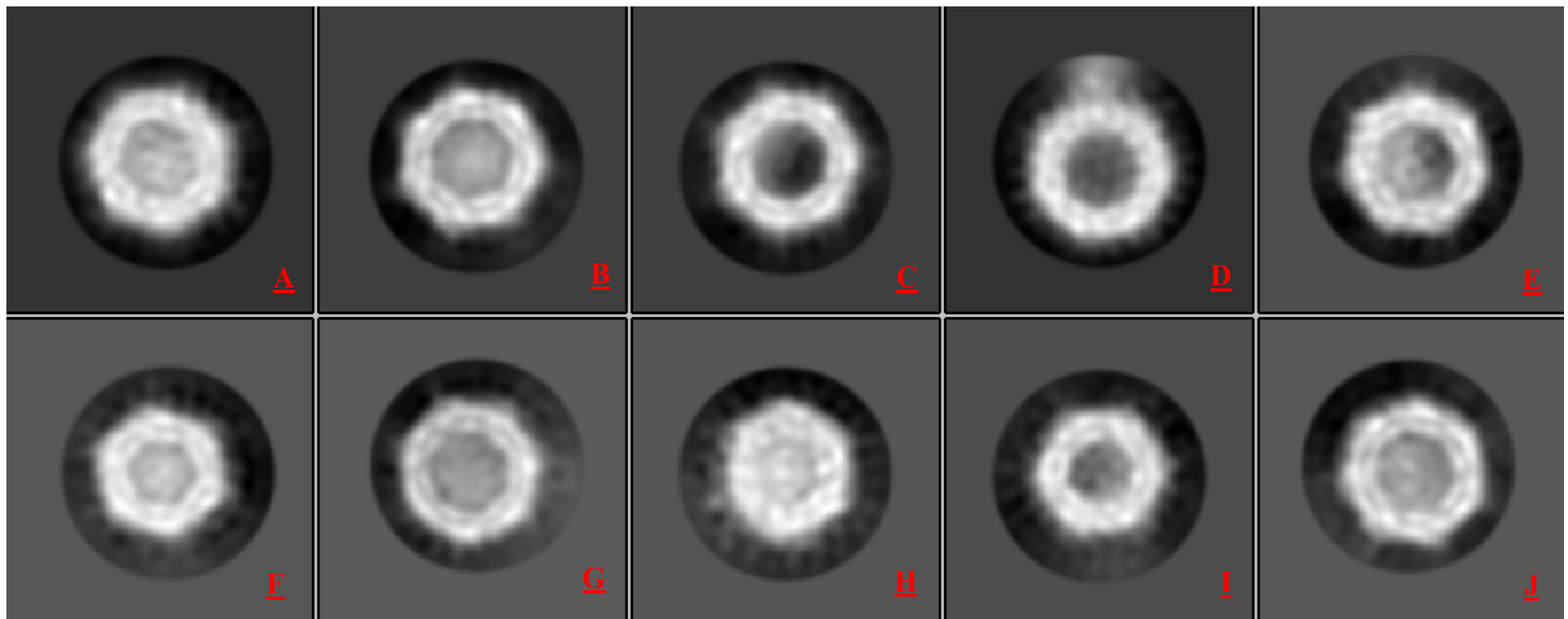| Class	Number of particles	Percentage of total particles (%)	Resolution (Å)
A	14483	23.5	3.39
B	10726	17.4	3.32
C	8904	14.4	3.39
D	8726	14.2	3.45
E	3936	6.4	3.45
F	3379	5.5	3.45
G	3204	5.2	3.45
H	2965	4.8	3.59
I	2809	4.6	3.59
J	2486	4.0	3.42

Figure 13 – 2D class averages by RELION.13A-13C, 13D-G and J describe the structure of the EspB oligomer composed of 7 monomers. 13H and 13I describe the structure of an EspB oligomer composed of 6 monomers, while 13H describes the oligomer composed of 8 EspB monomers.

3.4   2D classification of the C-terminal domain of EspB

Figure 13 clearly describes the top-view of EspB. However, in many of the class averages, and also in individual particles, 'something' can be observed. Thus far, no explanation has been given for the found structure in the middle of the oligomer. From earlier models, the structure of the monomer EspB has become clear, described in figure 14. From the figure a couple of things can be concluded:

− the alpha helixes describe a so-called PE-PPE protein structure which are linked with a so-called 'linker'. The PE-PPE protein structure form the N-terminal domain and are very well defined in the 2D class averages;

− the structure on the left of the figure, is the C-terminal domain of the monomer.

As seen in figure 14, the C-terminal domain is very flexible, and therefore very hard to define using the 2D classification. However, there are some tricks that can be applied to have a more detailed look at the C-terminal domain. For the determination of the C-terminal domain, figure 15, a smaller particle diameter was used during particle picking, centred in the middle of the oligomer, where we hypothesize the C-terminal domain is most likely positioned. Details of the 2D class averages can be observed in table 5.

Table 5 – Detailed description of 2D class averages of the C-terminal domain of EspB, from figure 15

Class	Number of particles	Percentage of total particles (%)	Resolution (Å)
A	4684	34.3	1.87
B	4206	30.8	1.87
C	1073	7.9	3.58
D	957	7.0	2.96
E	799	6.0	3.31

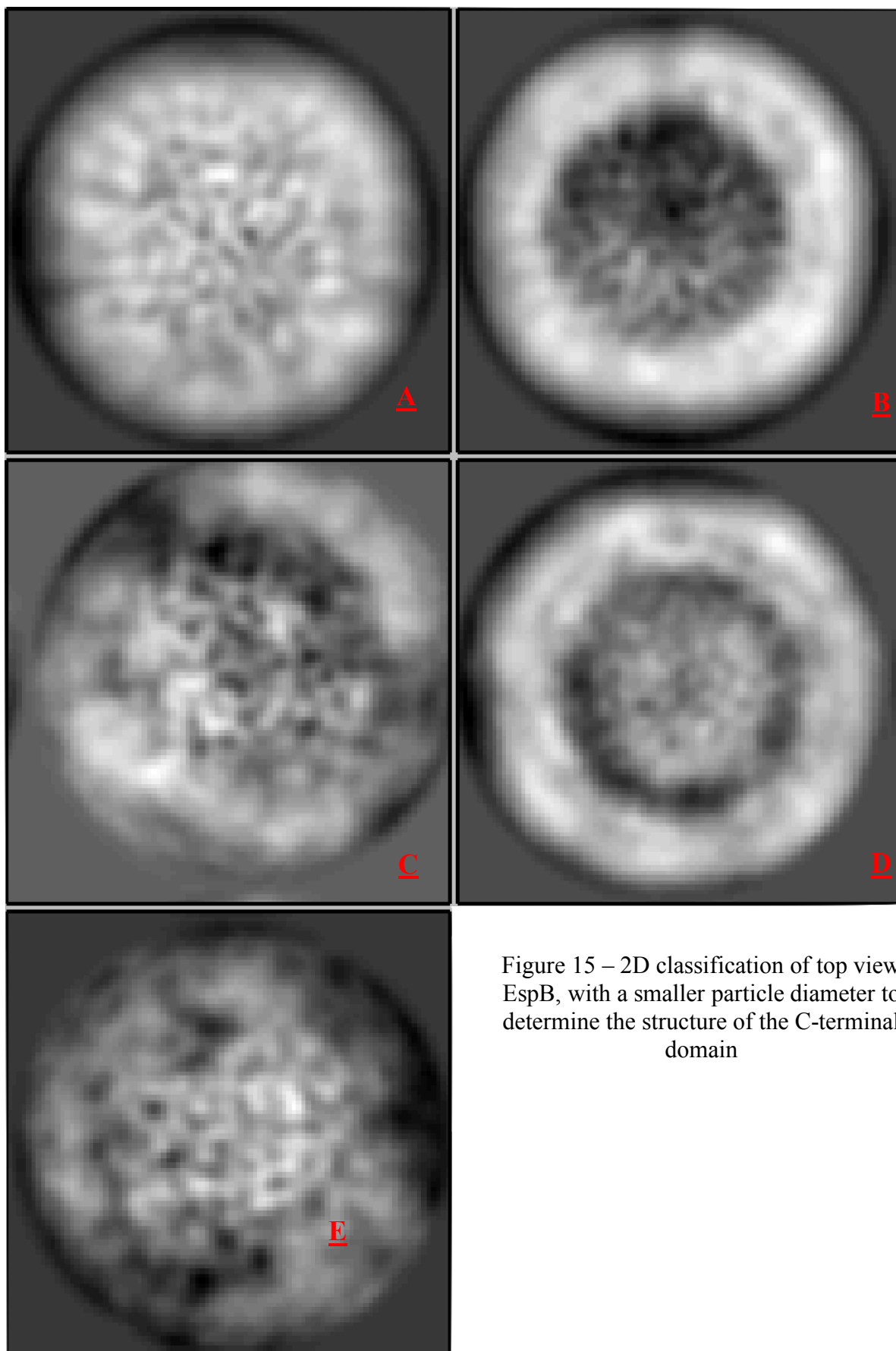Figure 14 – Model of the EspB monomer

Figure 15 – 2D classification of top view EspB, with a smaller particle diameter to determine the structure of the C-terminal domain

3.5 2D classification of EspK

EspK is another secretory protein secreted by the ESX1 secretion system. The presence of EspK has been established years ago. However, thus far no research group has reported a definite structure of EspK.

The structure of the monomer EspK is presented in figure 16. It has been hypothesized that the monomer EspK also forms an oligomer, but the composition of the oligomer is unknown.

During this research, we attempted to get 2D classifications of EspK. Unfortunately, due to time restrictions, results have not been implemented in current work since data processing is still running. There were some unexpected difficulties since the EspK monomer is very flexible and therefore hard to locate in the micrographs.
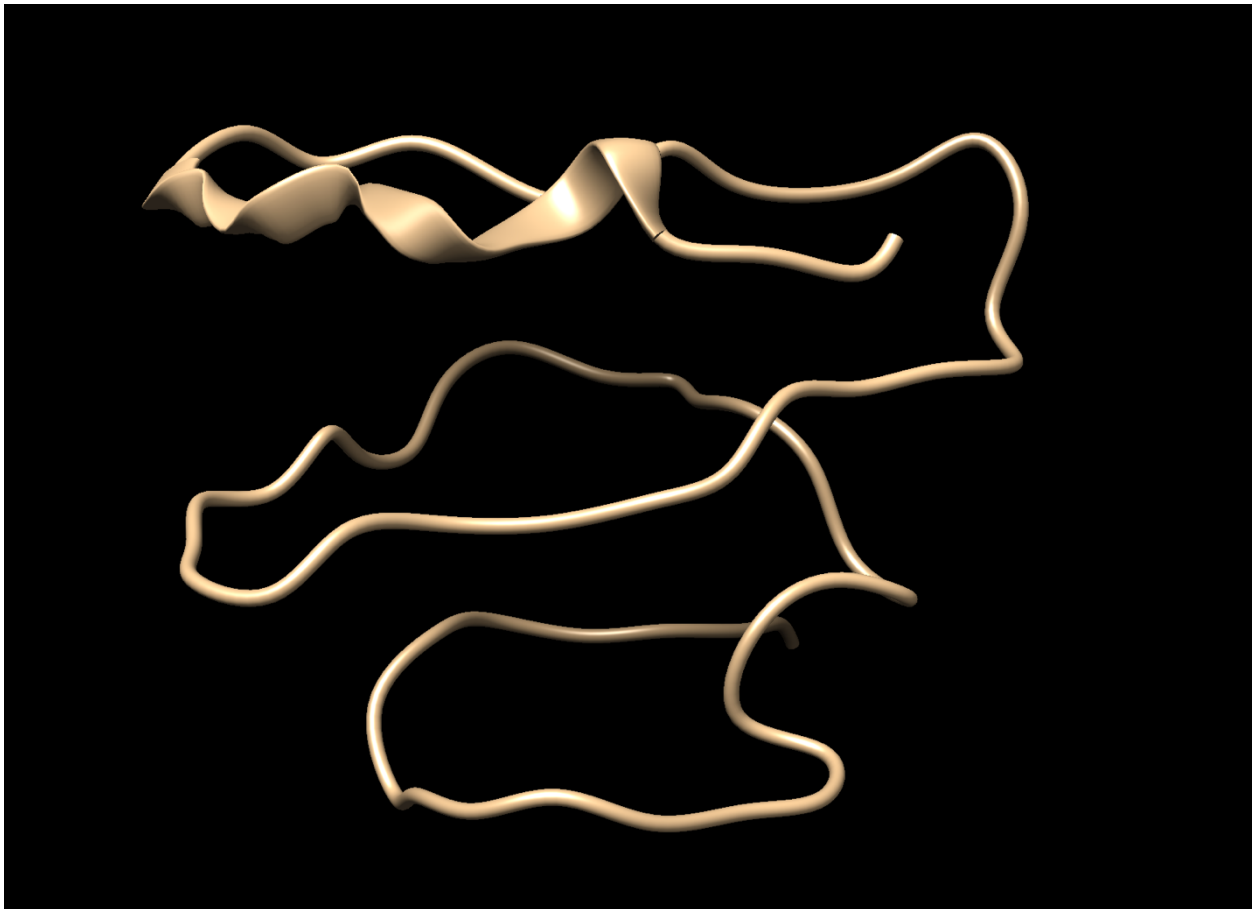

Figure 16 – Proposed model of the EspK monomer

4 Discussion

During this research, there have been several technical and logistic problems. During the first couple of months the equipment generated a lot of problems with as result that in that time no data collection was possible. During this period of time I practiced the RELION workflow on free assessable datasets on the internet, to make the data processing smoother, and less time consuming.

Although I was not able to get a 3D reconstruction of the proteins of interest, EspB & EspK, there were still novel results: For the first time, a research group attempted to look at the C-terminus of the EspB protein. This is not carried routinely since the protein is very flexible and therefore very hard to average during 2D classification.

The bottleneck in the current research field is the difficulties that are being dealt with during sample preparation. There can be poor consistency of the ice thickness, within same grid or between different grids. Currently, samples are made by blotting excess liquid from grids with filter paper. The automated control of parameters such as humidity, blotting force and blot time and the interval between blotting and plunge freezing, have drastically improved the quality of the samples currently used in cryo-EM, but there is still a lot to gain from optimal sample preparation. Achieving a thickness about 100 nm becomes highly important when the resolution increases. The problems with ice thickness are often contributed to the fact that the support film on the EM grid is continuous and highly hydrophilic.

It is known that thin aqueous films are greatly stabilized against rupture when there is a surfactant at the air-water interface. Some research groups have hypothesized that Marangoni flow (the bulk transfer of a liquid due to a gradient in surface tension) opposed any fluctuations in layer thickness. The idea of forming a locally thinner area, occurrence of a nascent de-wetting event is accompanied by an increase in surface area, and therefore a local decrease in surface coverage by the surfactant. Nevertheless, a surfactant layer is hypothesized to prevent rupture of the thin layer [25].

In cryo-EM, it is nevertheless a requirement that one makes very thin films, which have extremely high values of the surface-to- volume ratio. Macromolecules diffusing freely within such thin films, perhaps 100 nm thick, are bound to collide with the air–water interface 1000 times per second [26]. As a result, even though the interval between blotting and vitrification may be kept as short as possible, there is plenty of time for something to go wrong. It is well known, for example, that some types of particles exhibit preferential orientation in cryo-EM images.

A likely explanation is that one or more hydrophobic patches on the surface may allow a particle to bind to the air–water interface, in much the same way as it would do to a hydrophobic interaction chromatography bead. Little or no activation energy may be required for this type of binding to occur. Although such binding need not cause a structural change in the particle, it can still make the grid useless for three-dimensional data collection because of preferential orientation of the particles.

More worrisome, of course, is the concern that the particles of interest may become denatured immediately after they adsorb to the air–water interface. The rate at which a denatured-protein layer forms may, in fact, be limited only by the rate of diffusion from the underlying solution to the air–water interface. It is fairly common practice to add detergents or other surfactants to samples that are used to make cryo-EM grids. This is a requirement when working with solubilized membrane proteins, of course.

In addition, it is often something that is done with soluble proteins, especially those that are initially found to be difficult to prepare as cryo-EM specimens. When the mass of added surfactant is in great excess relative to that of the particle of interest, a monolayer of surfactant will form at the air–water interface before most of the proteins ever get there. Surfactant monolayers can be much more structure friendly to proteins than is the air–water interface, and even arguably more structure friendly—as support films—than are glow-discharge treated carbon films. This is because surfactant monolayers are expected to act as a barrier that reduces access to the hydrophobic side of an air–water interface. Nevertheless, proteins may still be

able to diffuse into the monolayer and bind via a hydrophobic interaction, unless the (surfactant) surface pressure is unusually high [27].

During the research, several attempts have been carried out with a detergent called OPOE. We tested the detergent in different concentrations and different environmental factors, such as changing the ph. However, none of those experiments yielded usable grid, due to vitrification difficulties. This stresses the fact that sample preparation remains a difficulty even more.

Apart from the difficulties encountered during sample preparation, other limiting factors are present in the cryo-EM research field, including:

− electron-beam tilt;

− inaccurate determination of defocus values;

− focus gradient through particles;

− multiple scattering of electrons.

Electron beam tilt

To obtain high resolution reconstruction with cryo-EM, the incident electron beam should be perfectly parallel to the optical axis, as any beam tilt introduces phase shift to images. Beam tilt can originate from three sources in electron microscopy: the overall beam tilt relative to the optical axis ('direction tilt'), beam divergence by over-focus of the C2 condenser lens ('divergence tilt'), and the spiraling trajectory of electrons in the inhomogeneous magnetic lens ('spiral tilt'). The latter two tilts are also known as local beam tilts since they are position-dependent.

Since the phase shift induced by beam tilt increases dramatically with spatial frequency, correcting its effect computationally or experimentally is critical for achieving atomic resolution by single particle cryo-EM. Therefore, to reach atomic resolution, the beam tilt must be reduced to such low values. Also, a higher accelerating voltage (e.g., an increase from 200 kV to 300 kV) can reduce the wavelength of the electron and in turn reduce the phase shift due to beam tilt for

each structure. However, in practice, higher accelerating voltage has one disadvantage of reducing image contrast.

Methods to reduce beam tilt include use of beam-deflecting coils to physically align the beam to the optical axis and use of a collimating lens to produce a parallel beam. In theory, it is also possible to compensate for the effect of beam tilt with image processing software, but the efficacy of this approach is limited to some extent by the low signal-to-noise ratio (SNR) of high resolution cryo-EM images. Most significantly, the effect of total beam tilt can be completely eliminated by use of a hardware device, a spherical aberration corrector [28].

Inaccurate determination of defocus values

Astigmatism generates variation in defocus values as a function of direction (azimuth angle) within the image. With no astigmatism, defocus values would be the same in all directions, and CTF could be visualized in the Fourier transform of the image as circular rings. With astigmatism, CTF rings takes the form of ellipses, and variation in the defocus values can be described using three parameters: defocus 1 (Δf_1), defocus 2 (Δf_2) and the angle (β) between the long axis of the ellipse and X axis. Therefore, accurate determination of astigmatism is also essential.

For two reasons, an accurate determination of these values is difficult. First, since the defocus and astigmatism values of images are usually determined by fitting theoretical CTF curves into the power spectrum of cryo-EM images, the accuracy of both derived defocus and astigmatism values is inherently compromised by the poor SNR of cryo-EM images. Second, since the thickness of embedding ice is usually greater than the diameter of particles, particles in an image may be located at different heights in the ice, leading to defocus variation among different particles [28].

Focus gradient through particles

Most 3D reconstruction methods are based on the Central Projection Theorem, which states that the Fourier transform of a 2D projection image is

equivalent to the central section of the 3D Fourier transform of the particle being imaged. The Central Projection Theorem is based on the assumption that there is no change in the focus values from the top of a particle to the bottom. The assumption is valid only for small or thin particles. For large particles, the effect of a large gradient in focus values from slice to slice is to add different phase shifts to the electron wave fronts emerging from different slices of the particle. Failure to take this effect into consideration in current software packages based on the Central Projection Theorem thus creates a resolution barrier for the single particle cryo-EM [28].

Multiple scattering of electrons

Electrons are scattered much more strongly than X-ray photons by specimens, making it possible to collect structural information for a single biological complex as small as 38 kDa by cryo-EM [13]. However, the stronger scattering of electrons also makes the dynamic (multiple) scattering effects of electrons much more severe than that of X-ray photons. Because biological specimens are primarily composed of light elements and are thus weak phase objects, neglect of the effects of dynamic scattering is likely to be valid when the specimen is thin and at low resolution. With an increase in specimen thickness and accompanying increase in scattering, and with improvements in targeted resolution, the effect of dynamic scattering becomes significant, making compensation for the effects of dynamic scattering necessary [28].

Astigmatism

One of the major physical problems encountered during this research astigmatism. To allow accurate imaging, the electron beam (probe) should be circular in cross section when it reaches the specimen. The probe cross section can be distorted to form an ellipse, due to a range of factors such as level of machining accuracy and the material of the pole-piece, imperfections in the casting of the iron magnets and the copper winding. This distortion is called astigmatism and causes

focus difficulties. Bad or "gross" astigmatism can be seen as "streaking" in the image in an X direction that changes to the Y direction as the image passes through focus from under focus to over focus. At exact focus the streaking vanishes and focus can be correctly achieved if the spot size is suitable.

Good alignment of the microscope, such as the gun, condensers, apertures, beam tilt, coma, astigmatism, etc., is a prerequisite to reaching optimal illumination lens and imaging lens conditions for high resolution TEM imaging. Currently, microscope alignments are still performed manually and rely on visual, qualitative feedback. Moreover, manual microscope alignment requires extensive training and experience but still often suffers from suboptimal efficiency and quality. Minimizing astigmatism of the objective lens is an indispensable daily instrument alignment task essential for high resolution TEM imaging.

Astigmatism of the objective lens represents the angular dependency of defocus. 2-fold astigmatism is the major type of astigmatism relevant to cryo-EM, which results in the elliptical elongation of Thon rings in the power spectra of TEM images. Currently, many microscopists visually correct astigmatism at large magnifications and small defocuses, then switch to a lower magnification to collect data and intentionally vary defocus to sample the entire reciprocal space and even-out the zero-nodes of the contrast transfer function (CTF). Similarly, many microscopists also use a magnification in the focus mode larger than that of the exposure mode during low-dose imaging. The implicit assumption of these strategies is that the astigmatism is invariant to the change of magnification and defocus. Conversely, other microscopists prefer to use the same magnification for astigmatism correction and in the focus mode as that used in the exposure mode during data collection.

The astigmatism of TEM images has been shown to vary with changes in imaging conditions (e.g. defocus, magnification), indicating that correction of astigmatism at high magnification and near-focus conditions by the current approach will not be optimal after switching to different conditions for data acquisition. What's worse, the dependence of astigmatism on the imaging conditions varies from

time to time, so that astigmatism cannot be reliably compensated by pre-calibration of the instrument. Although post-imaging computational CTF correction has now become a common practice, the community-wide CTF Challenge has found that it is still more challenging to reliably and accurately fit the astigmatism than the average defocus [29].

To aim at best quality results for each step and to avoid accumulation of issues in early steps that must be rescued by later steps, it is thus desirable to minimize the astigmatism during data collection. Based on systematic measurements and analyses the work of *Yan et al*, they suggest that (1) the magnification used for astigmatism correction during instrument alignment should be the same as the one used for data collection; (2) the defocus used for correction of astigmatism during instrument alignment should be set at the median defocus of the defocus range intended for subsequent data collection; (3) the focus-mode in the search-focus-exposure iterations of low-dose imaging should use the same magnification that is used for the exposure mode. Additionally, there are other factors such as stage and sample grid position that also lead to the variation of astigmatism, especially when the conductivity of materials in the imaging area is poor and local charging is induced. To optimally correct the astigmatism for images taken at different defocuses for different sample areas, a fast, accurate and automated method needs to be developed to avoid the defocus-dependent astigmatism by adaptively correcting the astigmatism at all focuses [29].

Because of the high astigmatism observed in the current work, half of the acquired data had to be excluded from analysis because it was not usable during processing. Even though during data acquisition, we had corrected for astigmatism, it was not sufficient to maintain the image quality.

A possible explanation for the high astigmatism is de fact that the grid was moved around quite a lot during data acquisition. A lot of grid squares were selected, so the grid moved quite some distance during data acquisition. In theory, this should not be a problem since there is a period of time in which the grid movement 'relaxes'

but maybe in this case, that time was not sufficient enough to relax the stage completely and minimize the effect of astigmatism.

CONCLUSION

In this current work, the aim was to break the 100 kDa barrier in cryogenic electron microscopy by implementing a phase plate in the back-focal plane of the microscope, to enhance phase contrast without adjusting the defocus. The ultimate aim was to generate a 3D reconstruction of two biological relevant proteins, EspB & EspK, secreted proteins from Mycobacterium tuberculosis.

Unfortunately, due to several limitations in sample, technique and time restrictions, a 3D reconstruction of the mentioned proteins was not reachable within the period of time in which this research was conducted. Too much problems with the protein, the sample preparation, the vitrification and air-water interface remain to have a good reconstruction at the present time.

However, in this research several results were achieved. The top view, the preferred orientation of EspB was addressed with more detail and yielded near-atomic resolutions by implementation of the VPP. Furthermore, the hypothesized C-terminal domain, in the center of the EspB oligomer was investigated further. Although it is still speculation, the results imply that the C-terminal domain could be present inside the oligomer since there is clearly something present within the oligomer.

For the analysis of EspK, there was not enough time to incorporate it in the current report. We encountered more problems than anticipated with the data processing since the protein is poorly defined and extremely flexible.

Luckily there are opportunities. The field continues to evolve regarding sample preparation, sample vitrification, phase plates and new processing strategies. One of the most recent developments in the data processing is 'localized reconstruction'. Localized reconstruction can be used for solving structures which are highly heterogeneous, or flexible. The method allows symmetry-mismatched and/or flexible subunits to be treated as isolated single particles separate from the rest of the complex, facilitating their classification, refinement and reconstruction [30].

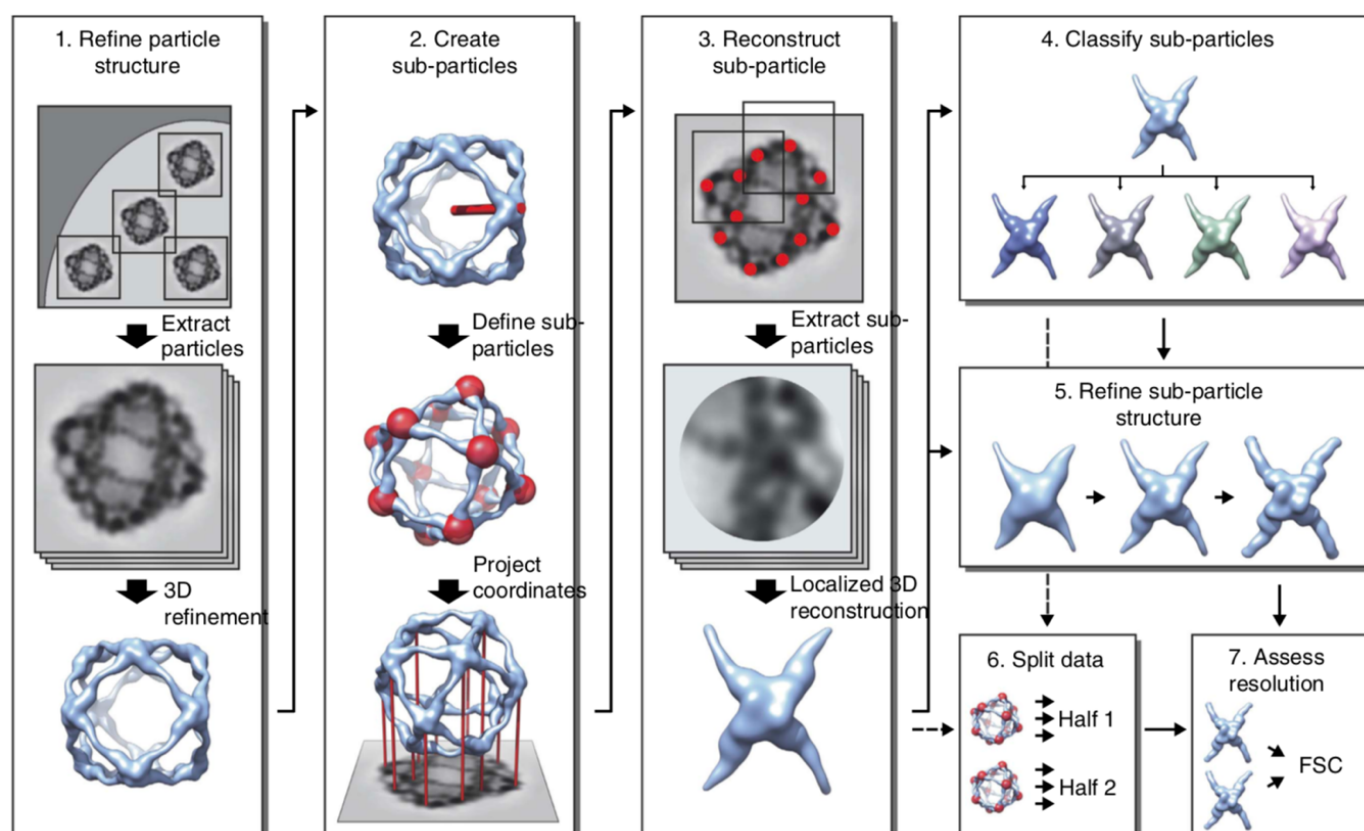For local refinement the general workflow is described in figure 17.

Figure 17 – Localized reconstruction of sub-particles from images of macromolecular complexes. Schematic diagram of the workflow for localized reconstruction. First the structure of the macromolecular complex is solved using conventional 3D refinement (1), after which the locations of the sub-structures (red spheres) are calculated based on the particle orientation, a symmetry operator and a vector defining one sub-structure relative to the particle model (red stick; 2). After extracting the sub-particles (red dots) from the particle images, a localized 3D reconstruction is calculated (3). This reconstruction can be used as a starting model for further classification (4) and 3D refinement (5) of sub-particles to improve the structure. Finally, two independent sets of data (6) are compared by Fourier shell correlation (FSC) to assess the resolution of the reconstruction (7). See text for a detailed description. Transcribed from [30]

Besides solving the high-resolution structure of each compositional rigid parts of a complex, the segmentation algorithm provides additional information of the spatial relationship between the rigid parts within each individual particle image [31].

Unfortunately, due to limitations in time, this technique cannot be implemented into current work. However, this is definitely a technique which can help in further research with solving difficult structures, which are flexible, heterogeneous and small structures.

Future research should implement local refinement strategies but should also continue research about sample preparation and other manners to get the protein out of its preferred orientation. Crosslinking is such a way. By crosslinking the protein to another protein, it has a different preferred orientation and absorption by the air-water interface changes. During the current research, there was little research done about the effect of crosslinking. However, it did not yield satisfactory results so far. Therefore, research should also be continued to find out the optimal parameters.

REFERENCES

1. (WHO) WHO. Global Tuberculosis Report 2017. . Geneva: World Health Organization (WHO); 2017.

2. Houben EN, Korotkov KV, Bitter W. Take five - Type VII secretion systems of Mycobacteria. Biochim Biophys Acta. 2014;1843(8):1707-16.

3. Korotkova N, Piton J, Wagner JM, Boy-Rottger S, Japaridze A, Evans TJ, et al. Structure of EspB, a secreted substrate of the ESX-1 secretion system of Mycobacterium tuberculosis. J Struct Biol. 2015;191(2):236-44.

4. Solomonson M, Setiaputra D, Makepeace KA, Lameignere E, Petrotchenko EV, Conrady DG, et al. Structure of EspB from the ESX-1 type VII secretion system and insights into its export mechanism. Structure. 2015;23(3):571-83.

5. Ohol YM, Goetz DH, Chan K, Shiloh MU, Craik CS, Cox JS. Mycobacterium tuberculosis MycP1 protease plays a dual role in regulation of ESX-1 secretion and virulence. Cell Host Microbe. 2010;7(3):210-20.

6. Wong KW. The Role of ESX-1 in Mycobacterium tuberculosis Pathogenesis. Microbiol Spectr. 2017;5(3).

7. Miyaguchi K. Direct imaging electron microscopy (EM) methods in modern structural biology: overview and comparison with X-ray crystallography and single-particle cryo-EM reconstruction in the studies of large macromolecules. Biol Cell. 2014;106(10):323-45.

8. Zanotti G. Cryo-EM and X-ray Crystallography: Complementary or Alternative Techniques? . Nano World Journal. 2016;2(2):22-3.

9. von Loeffelholz O, Papai G, Danev R, Myasnikov AG, Natchiar SK, Hazemann I, et al. Volta phase plate data collection facilitates image processing and cryo-EM structure determination. J Struct Biol. 2018.

10. Reimer L, Kohl H. Transmission Electron Microscopy. Transmission Electron Microscopy Physics of Image Formation 5ed. New York: Springer-Verlag; 2008. p. 17-43.

11. Reimer L, Kohl H. Transmission Electron Microscopy. Transmission Electron Microscopy Physics of Image Formation. New York: Springer-Verlag; 2008. p. 44-77.

12. Reimer L, Kohl H. Transmission Electron Microscopy Physics of Image Formation New York: Springer-Verlag; 2008. p. 77-140.

13. Henderson R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. Q Rev Biophys. 1995;28(2):171-93.

14. Brocchieri L, Karlin S. Protein length in eukaryotic and prokaryotic proteomes. Nucleic Acids Res. 2005;33(10):3390-400.

15. Rosenthal PB, Henderson R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J Mol Biol. 2003;333(4):721-45.

16. Peng LM. Electron atomic scattering factors, Debye-Waller factors and the optical potential for high-energy electron diffraction. J Electron Microsc (Tokyo). 2005;54(3):199-207.

17. Danev R, Baumeister W. Cryo-EM single particle analysis with the Volta phase plate. Elife. 2016;5.

18. Danev R, Buijsse B, Khoshouei M, Plitzko JM, Baumeister W. Volta potential phase plate for in-focus phase contrast transmission electron microscopy. Proc Natl Acad Sci U S A. 2014;111(44):15635-40.

19. Danev R, Nagayama K. Optimizing the phase shift and the cut-on periodicity of phase plates for TEM. Ultramicroscopy. 2011;111(8):1305-15.

20. Danev R, Glaeser RM, Nagayama K. Practical factors affecting the performance of a thin-film phase plate for transmission electron microscopy. Ultramicroscopy. 2009;109(4):312-25.

21. Danev R, Tegunov D, Baumeister W. Using the Volta phase plate with defocus for cryo-EM single particle analysis. Elife. 2017;6.

22.   Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods. 2017;14(4):331-2.

23.   Zhang K. Gctf: Real-time CTF determination and correction. J Struct Biol. 2016;193(1):1-12.

24.   Kimanius D, Forsberg BO, Scheres SH, Lindahl E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife. 2016;5.

25.   Glaeser RM, Han BG, Csencsits R, Killilea A, Pulk A, Cate JH. Factors that Influence the Formation and Stability of Thin, Cryo-EM Specimens. Biophys J. 2016;110(4):749-55.

26.   Taylor KA, Glaeser RM. Retrospective on the early development of cryoelectron microscopy of macromolecules and a prospective on opportunities for the future. J Struct Biol. 2008;163(3):214-23.

27.   Glaeser RM, Han BG. Opinion: hazards faced by macromolecules when confined to thin aqueous films. Biophys Rep. 2017;3(1):1-7.

28.   Zhang X, Zhou ZH. Limiting factors in atomic resolution cryo electron microscopy: no simple tricks. J Struct Biol. 2011;175(3):253-63.

29.   Yan R, Li K, Jiang W. Defocus and magnification dependent variation of TEM image astigmatism. Sci Rep. 2018;8(1):344.

30.   Ilca SL, Kotecha A, Sun X, Poranen MM, Stuart DI, Huiskonen JT. Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. Nat Commun. 2015;6:8843.

31.   Zhou Q, Zhou N, Wang HW. Particle segmentation algorithm for flexible single particle reconstruction. Biophys Rep. 2017;3(1):43-55.

Отчет о проверке на заимствования №1

Автор: Печерицын Алексей Анатольевич pecher@phys.tsu.ru / ID: 449043
Проверяющий: Печерицын Алексей Анатольевич (pecher@phys.tsu.ru / ID: 449043)

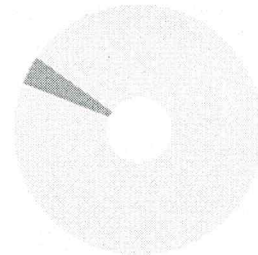Отчет предоставлен сервисом «Антиплагиат»- http://www.antiplagiat.ru

ИНФОРМАЦИЯ О ДОКУМЕНТЕ

№ документа: 24
Начало загрузки: 18.06.2018 19:57:35
Длительность загрузки: 00:00:01
Имя исходного файла: NRoest-diss
Размер текста: 4168 кБ
Символов в тексте: 80963
Слов в тексте: 12501
Число предложений: 507

ИНФОРМАЦИЯ ОБ ОТЧЕТЕ

Последний готовый отчет (ред.)
Начало проверки: 18.06.2018 19:57:37
Длительность проверки: 00:00:04
Комментарии: не указано
Модули поиска:

ЗАИМСТВОВАНИЯ	ЦИТИРОВАНИЯ	ОРИГИНАЛЬНОСТЬ
3,92%	0%	96,08%

№	Доля в отчете	Доля в тексте	Источник	Ссылка	Актуален на	Модуль поиска	Блоков в отчете	Блоков в тексте
[01]	0,5%	0,95%	не указано	http://ddd.uab.cat	05 Июл 2017	Модуль поиска Интернет	7	14
[02]	0,47%	0,92%	http://etheses.whiterose.ac...	http://etheses.whiterose.ac.uk	08 Янв 2018	Модуль поиска Интернет	3	6
[03]	0,11%	0,54%	http://doras.dcu.ie/16094/1...	http://doras.dcu.ie	06 Июл 2017	Модуль поиска Интернет	1	7

Еще источников: 14
Еще заимствований: 2.85%

Отв. за проверку на „Антиплагиат"
ст. преп. Печерицын А.А.