UNIVERSITY OF LEEDS

Article:
Johnson, R, Rawson, S orcid.org/0000-0002-2973-2630, McPhillie, M et al. (2 more authors) (2017) The Growing Role of Electron Microscopy in Anti-parasitic Drug Discovery. Current Medicinal Chemistry, 24. ISSN 0929-8673

https://doi.org/10.2174/0929867324666171109120526

The Growing Role of Electron Microscopy in Anti-parasitic Drug Discovery

Johnson, R. M.[a,b,c], Rawson, S.[a,c], McPhillie, M. J.[b,c], Fishwick, C. W. G.[b,c], Muench, S. P.[a,c]*

[a] School of Biomedical Sciences, Faculty of Biological Sciences, University of Leeds, Leeds, U.K. LS2 9JT. [b] School of Chemistry, University of Leeds, Leeds, U.K. LS2 9JT.[c] Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, U.K. LS2 9JT.

Abstract: Parasitic diseases are a huge burden on human health causing significant morbidity and mortality. However, parasitic structure based drug discovery programmes have been hindered by a lack of high resolution structural information from parasitic derived proteins and have largely relied upon homology models from mammalian systems. The recent renaissance in electron microscopy (EM) has caused a dramatic rise in the number of structures being determined at high resolution and subsequently enabled it to be thought of as a tool in drug discovery. In this review, we discuss the challenges associated with the structural determination of parasitic proteins. We then discuss the reasons behind the resurgence in EM, how it may overcome some of these challenges and provide examples of EM derived parasitic protein structures. Finally, we discuss the challenges which EM needs to overcome before it is used as a mainstream technique in parasitic drug discovery.

Keywords: Electron Microscopy, Structure based drug design, Plasmodium falciparum, Parasitic disease.

1. Introduction

Parasitic protozoa are responsible for causing a wide range of devastating diseases, predominantly in the developing world[1,2]. Diseases such as malaria, leishmaniasis, toxoplasmosis and sleeping sickness are a significant global burden with over ~500 million people affected each year, and in 2015, it was estimated that 429,000 people died from malaria alone[3]. Current therapies are limited, exacerbated by drug resistance for many anti-malarial compounds and no effective treatments for the

bradyzoite form of Toxoplasma[4]. Therefore, there is an urgent need for new medicines. There have been many different approaches to designing new therapeutics for parasitic targets including high throughput screens (HTS), phenotypic screens and structure based drug design (SBDD)[5]. Drug discovery campaigns can be significantly enhanced when the 3D structural information of the protein target is known, especially if the protein is bound to an inhibitor allowing the mode of binding to be established[6]. This information can be utilised by medicinal chemists in the design of new compounds which can optimise the interactions to the target protein, subsequently improving the binding affinity[7]. Furthermore, structural information for the target parasitic protein and comparisons to its homologous proteins can be exploited to design new compounds which are selective for their target. Structural information has underpinned many different anti-parasitic drug discovery projects, which have been reviewed elsewhere[8,9]. A recent example is the work by Philips et al. who targeted the Plasmodium falciparum dihydroorotate dehydrogenase (DHODH) enzyme. After a structure-based drug discovery programme the lead compound, DSM265, has reached phase II clinical trials emphasising the success of this approach[10]. However, there is a low percentage of structures deposited in the Protein Data Bank from parasitic organisms which can significantly hinder SBDD projects when homology models are not available. The challenges of obtaining parasitic structures are described below.

2. Challenges Associated with Anti-Parasitic Structure Guided Drug Discovery

Structure based inhibitor design is explicitly reliant upon structural information which has been typically provided by X-ray crystallography and nuclear magnetic resonance (NMR)[11]. In the absence of a structure of the target protein, homology models can be generated which will approximate the target binding pocket. This is dependent on two things, the first is an accurate starting model on which to base the modelling and the second is sufficient sequence identity/similarity which provides confidence in the resulting model[12]. For parasitic targets this can provide a significant challenge in terms of pre-existing structures and sequence similarity with many targets being unique to the parasite family. Moreover, a common feature of many apicomplexan parasites is the presence of low complexity inserts which can vary significantly in length[13]. For

example, enoyl reductase is the target of potent inhibitors, such as triclosan[14]. Within the apicomplexan family there is a low complexity insert situated around the inhibitor binding site[15]. A similar insert is seen with the M. tuberculosis family, which results in a significant enlargement of the inhibitor binding pocket[16]. For the apicomplexan family, crystallographic studies have shown that this insert does not alter the inhibitor binding pocket, as seen within the M. tuberculosis family, with the apicomplexan binding pocket adopting a very similar architecture to homologues without the insert (Figure 1)[15]. The role for the low complexity insert is currently unknown but could facilitate protein complex formation.

In comparison, Birkholtz et al[17] have shown that parasite-specific inserts in the P. falciparum S-adenosylmethionine decarboxylase/orinithine decarboxylase (PfAdoMetDC/ODC) are important for enzyme activity. These decarboxylase enzymes are critical in the synthesis of polyamines. There are four inserts in total which have effectively doubled the size of the AdoMetDC/ODC complex in comparison to its mammalian homologues. Each of the individual inserts were deleted and the activity of both decarboxylase enzymes was reduced. Thereby suggesting the insertions are important in mediating how the two enzymes interact and carry out their catalytic function. Another example of low complexity parasitic inserts being important to the function of the enzyme is shown by Yuvaniyama et al[18]. P. falciparum dihydrofolate reductase-thymidylate synthase (PfDHFR-TS) is an important therapeutic target for anti-malarial drugs and crystal structures have revealed three additional helical inserts in the enzyme. The function of two of these inserts has been determined from structural studies. One insert is composed of ten amino acids and could represent a recognition site for additional interactions which is important to the function of the enzyme. The other insert (containing four amino acids) is thought to be involved in both dimerization and domain organisation of the two different domains, DHFR and TS. As this sequence is unique to the parasitic enzyme it represents an exciting therapeutic target as small molecule inhibitors could be developed to disrupt this interaction.

In other instances, there are significant differences between homologues from different species both at the sequence and structural level. For example, the F-ATPases are large rotary machines which produce ATP in the final step of the

mitochondrial electron transport chain[19]. There are many crystal and EM derived structures from mammalian and bacterial systems which show that the catalytic domain of the complex exists as a hexameric ring of alternating □□□subunits[20–22]. Within Trypanosoma brucei (T. brucei) it has been shown that the complex contains 22 subunits, of which 14 are unique to the parasite organism, as they have no recognizable eukaryotic homologues. Through electron tomography the structure of the T. brucei F-ATPase was determined at ~30 Å resolution revealing that the parasitic $(□□)_3$ subunit exists in a pyramidal shape which is significantly different to its bovine homologue[23]. This suggests that the mechanism of ATP synthesis is not as highly conserved as previously thought and may occur by a different mechanism in parasitic protozoa. It also highlights how despite the strong architectural conservation of this fundamental protein complex within bacteria, plants and mammals, there are significant differences in the parasitic homologues, which are not easily predicted through computational modelling. Furthermore, the structure of the ATP synthase revealed that a key subunit involved in the production of ATP (subunit □) was split into two unique fragments[23]. Despite being cleaved the complex is still able to synthesize ATP and it is not known how this affects the mechanism of ATP synthesis in parasitic organisms. Moreover, in P. falciparum the ATP synthase does not have homologues for key subunits such as the membrane bound subunit a, which is critical for proton translocation across the membrane, and the peripheral stalk subunits b and d[24]. Due to the apparent absence of these subunits, it is difficult to accurately model the structure of the complex as it is unknown how the subunits interact with one another within the complex based on our current structural understanding. Therefore, gaining high resolution structural information about the parasitic complex via EM would allow the mechanism of action to be established.

Many therapeutic targets, for both mammalian and parasitic systems, are membrane proteins (eg GPCRs and ion channels) which are notoriously difficult to study[25] . This has resulted in a lack of structural information for many classes of membrane proteins, not just in parasites but all species. Recent developments in EM (discussed below), improved crystallisation methods such as LCP (lipid cubic phase)[26] and improved membrane protein screens have started to address this problem but a scarcity of information remains[27]. This results in difficulties when generating reliable homology models which can be validated through biochemical means, for example

mutational analysis or crosslinking. An improvement in our basic structural understanding of parasitic membrane proteins is essential if we are to begin to develop new therapeutics against this important clade.

If there is no structural information, including a reliable model, for the target protein available experimental methods such as X-ray crystallography, NMR or EM are conducted to probe the inhibitor binding site. A pre-requisite for structure determination is the ability to obtain sufficient quantities of pure protein. For crystallography, milligrams of very high purity sample are typically needed to produce crystals that diffract to high resolution. A significant reduction in protein quantity is seen for EM which typically uses microgram quantities, but even this can be a challenge for many systems. Sufficient protein quantities can be achieved in two ways, the first is through cloning the target protein(s) and overexpressing in a suitable host, for example E. coli. The second route is where the protein is extracted, direct from the native source. Traditionally structural biology has been underpinned by the ability to overexpress proteins in more amenable hosts, such as bacteria or yeast which can be easily cultured into high volumes[28]. There are significant hurdles for the overexpression of many parasitic proteins, especially those for Plasmodium where the AT composition of the genome is unusually high (~80%) making it difficult to express the genes in the traditional hosts[29]. Moreover, the composition of the genome also results in a change in the codon usage which results in a limitation on certain tRNA's in more traditional expression hosts. This can be overcome by using the RIG plasmid expression system which allows codons commonly used by Plasmodium to be over expressed thereby fixing the codon bias[30]. This involves expressing the rare tRNA's, such as Arg, Ile and Gly, which are rarely used in E. coli, in parallel to your gene of choice, resulting in high yields of Plasmodium proteins[30]. A significant limitation of this route is in the study of larger protein complexes which can be a significant hurdle due to the requirement of specific chaperones and complex formation pathways which cannot always be compensated for in other host organisms.

For some systems, the preferred route is often native source as this can mitigate any concerns with post translational modifications or the need for chaperones which might not be replicated in the expression host. This is exemplified by the vacuolar ATPase (V-ATPase) where in some species such as S. cerevisiae, the RAVE complex is involved in the assembly of the complex after it has dissociated in response to being

deprived of glucose[31]. It has been shown that in the absence of RAVE, the V-ATPase complex shows both functional and structural defects, highlighting the importance of the chaperone in the assembly of the complex[31]. Obtaining protein from the native source is a viable option for some systems. For example, the ability to obtain significant quantities of bovine heart muscle has allowed sufficient quantities of the proteins involved in the mitochondrial electron transport chain to be extracted from the mitochondria. This is shown in the crystal structures of bovine complex I[32], cytochrome bc_1 (complex III)[33] and the F-ATPase (complex V)[34] where the proteins were extracted from the heart muscle. Furthermore, the structure of the porcine respiratory supercomplex ($I_1III_2IV_1$), was determined using cryo-EM providing us with a unique insight into how the individual components of the electron transport chain interact with one another in the native tissue[35]. This approach is reliant on a natural abundance of the protein and an appropriate method of purification that does not rely on engineered tags, such as His or FLAG tags. For parasitic organisms obtaining sufficient quantities of protein is challenging as the parasites are relatively small and are difficult to culture. Therefore, obtaining large (litres) volumes of culture becomes unfeasible or too costly for most projects.

For a number of parasitic small protein targets, X-ray crystallography and to a lesser extent NMR has been a powerful approach in structure determination. In total, there are 504 structures for 'P. falciparum only' in the PDB (as of June 2017) which can be broken down into 480 structures solved using X-ray crystallography, 19 using NMR and just 5 using EM (Figure 2a). Three of the most common entries are for triose-phosphate isomerase (29), M1 family aminopeptidase (20) and L-lactate dehydrogenase (16), which have a variety of different ligands bound. Interestingly, of the 480 X-ray deposited structures, ~95% of these have a molecular weight lower than 250 kDa (Figure 2b). Therefore, for larger protein complexes and membrane proteins which are typically difficult to express in high abundance, alternative approaches are needed. This review describes how recent advances in the EM field can address many of these challenges and provide a new alternative to structure determination to drive SBDD programs against parasitic targets.

3. Resurgence in Electron Microscopy

Electron microscopy has recently undergone a 'resolution revolution' resulting in an increase in the number of high resolution protein structures being determined (Figure 3)[36]. Recently its impact has been recognised through the award of the Nobel Prize in chemistry (2017) to Jacques Dubochet, Joachim Frank and Richard Henderson for "developing cryo-electron microscopy for the high-resolution structure determination *of biomolecules in solution"*. The developments in the field have not just improved the resulting resolution but it can now study much smaller systems (~100kDa) than would have typically been studied previously. This is highlighted by the large increase in the total number of structures from EM being deposited into the protein data bank each year; 65 in 2012 compared to 408 deposited in 2016 alone. The reasons behind this increase is developments made to microscope stability, the detector technology and image processing algorithms[37].

EM is no longer considered as a technique which can only be used for large symmetrical systems or large protein complexes such as viruses or ribosomes, respectively. A significant increase in the number of structures has been in the membrane protein field, which has traditionally proven to be highly challenging for crystallographic techniques[38]. For instance, high resolution structures of the membrane proteins TRPV1[39], Slo2.2[40], $Ca_v1.1$[41] and gamma secretase[42] have been determined using EM approaches, with these targets previously evading crystallographic studies. By negating the need for highly ordered protein crystals which is often hindered by a lack of crystal contacts within detergent enclosed membrane proteins, EM is a powerful alternative for structure determination[43]. Given the number of parasitic membrane protein targets this is an exciting opportunity to open up new avenues in SBDD against previously intractable systems. Furthermore, EM is a powerful technique for determining the structure of large protein complexes which can enable individual proteins to be studied in a more native environment than their isolated individual structures. One such example is the respiratory supercomplex which has been determined at ~7 Å[35] and more recently 4.0 Å[44] by EM. The cryo-EM structures have enabled the overall architecture of the supercomplex to be understood despite there already being crystal structures of the individual components highlighting the complementary nature of the two techniques.

Over the past 5 years there has been an increasing number of smaller proteins (<250 kDa), studied using EM which is shown by there being ~29 unique, sub nm resolution maps deposited in the EMDB. Smaller systems can now be studied with the use of phase plates which increases the signal to noise ratio in the images, subsequently boosting the contrast which is needed to see the protein[45,46]. For instance, the smallest structure determined by cryo-EM is the structure of haemoglobin (~60 kDa) which was determined using the Volta phase plate at 3.2 Å resolution[47]. Previously the smallest protein determined was isocitrate dehydrogenase (93 kDa) which achieved a resolution of 3.8 Å[48]. This highlights the step-change EM has undergone in recent years as targets which were previously thought to be too small for EM are now being determined at high resolution.

Despite the "resolution revolution" in EM the expected resolution is still typically ~3-4 Å. Although not all side chains can be unambiguously assigned at this resolution, it has proven to be sufficient to characterise the inhibitor biding site in TRPV1 (2.9 Å)[39] and the proteasome (3.6 Å)[49]. Higher resolution has been achieved in the case of β-galactosidase (2.2 Å)[50] which reportedly resolved not only the bound inhibitor but also co-ordinating waters. The ability to resolve inhibitor density and see how compounds interact with their target protein, means that EM has the potential to be used as a tool in SBDD[48,51]. Because of this improvement in resolution, it is now feasible that EM can play a role in elucidating structural information for parasitic proteins which have so far evaded structural characterisation.

One significant limitation with current SBDD approaches is the use of "static" structures, whereby the target protein is determined in only one or a handful of conformational states. However, EM has become a powerful technique to study different conformational states and the dynamic nature of large protein complexes, with EM accounting for the flexibility of different systems in the data processing steps. For instance, the 3.4 Å structure of the human γ-secretase identified a flexible active site which could change conformation upon substrate binding[42]. Additionally, the vacuolar ATPase has been shown to have inherent flexibility with the soluble, catalytic V_1 domain moving in relation to the membrane bound V_0 domain[52]. Moreover, the structure of the full Manduca V-ATPase was originally determined at 9.2 Å but using

the same data set, masking of the individual V_1 domain improved the resolution to 8.2 Å with the local resolution showing further improvements to ~6 Å[53]. By accounting for the flexibility of the system, the resolution was substantially improved which could reveal new insights into the structure of dynamic systems. This can be advantageous over X-ray crystallography because upon formation of the crystal lattice, the protein can become trapped into one particular conformation. Comparing the different conformational states, which can be identified and separated using EM, will allow new inhibitors to be designed against the most populated, and potentially the most physiologically relevant state. For instance, obtaining structures of the F-ATPase via cryo-EM revealed the complex existed in three distinct conformations which could be sub-divided into seven different sub-states[22]. Furthermore, Hite et al have determined the structure of the Slo2.2 potassium channel in both an open and closed state via cryo-EM thus revealing the channel's mechanism of action switching between the two states, which is vital information when targeting particular states which predominate in the native environment[54].

The recent resurgence in EM has been primarily underpinned by the introduction of direct electron detectors which has enabled the signal to noise ratio of the data to be improved resulting in higher contrast in the images[55,56]. Moreover, instead of taking single images, the detectors record movies containing a number of individual frames which have only been exposed to a small proportion of the total dose in a single exposure. This is advantageous because during a typical two second exposure, radiation damage to the sample and/or mechanical drift in the microscope can cause blurring of the image. Computational programmes can subsequently correct for this movement to further enhance the signal and quality of the images thus allowing the high-resolution information to be retained. Before the introduction of these detectors images were recorded on film or CCD cameras which do not record movies but instead take an average image of the movement during an exposure[57]. This resulted in the loss of the higher resolution information as the particle would appear to be blurred, significantly hindering the structural determination of the sample. Recording the exposures as movies instead of single images also allows one of the main limitations of EM to be mitigated; radiation damage[58,59]. As the sample is irradiated by the electron beam it undergoes significant amounts of radiation damage which can affect side chains by differing amounts. For instance, whilst taking a single micrograph, the

sample is typically exposed to between 20-100 $e^-/Å^2$. However, it has been shown that doses as low as ~$3e^-/ Å^2$ can have a serious negative impact on some charged amino acids resulting in some high resolution information to be lost[60]. By removing the later frames, where the sample has been exposed to the highest amount of electrons, or by weighting the frames by the dose they have received, the radiation damage to some extent can be accounted for[60].

Cryo-electron tomography is another technique which could be exploited to identify new parasitic drug targets within cells as individual cells can now be plunge-frozen and imaged[61]. This is particularly advantageous when the target protein cannot be extracted from its environment therefore traditional single particle analysis cannot be performed. In tomography, images are generally obtained as tilt-series which results from the specimen being imaged at a series of different angles in the microscope (typically between ± 65°with an image being taken at 2°increments). Comp utationally these images can then be recombined using subtomogram averaging methods, to generate a 3D model of the target macromolecular complex in situ[62]. The resolution achieved and the thickness of the sample being imaged are the main limitations of this technique. As multiple images are taken at the same position, low electron doses are used to prevent losing all of the biological information to radiation damage. Therefore, in order for the electrons to be detected the sample needs to be less than ~500 nm in thickness otherwise there would be no signal in the images[63]. Traditionally, resolutions obtained from cryo-tomography have been ~1-2 nm allowing for the organisation and globular shape of the protein to be understood. However, more recently the resolution of tomographic reconstructions has significantly improved as exemplified by the structure of the HIV-capsid-SP1 at 3.9 Å[64]. This suggests that in the future, high resolution structures derived from tomography could play an important role in elucidating how drugs can bind to their target protein in situ.

4. Examples of Parasitic Structures determined using EM

Currently the EMDB contains many high-resolution EM structures of parasitic proteins including a number of ribosomal subunits from different parasitic species such as Trypanosoma brucei (5.6 Å)[65], Trypanosoma cruzi (2.5 Å) (Figure 4A)[66] and two from Leishmania donovani (2.8 Å and 2.9 Å)[67,68]. The most commonly studied parasitic

protein in the EMDB is the P. falciparum 80S ribosome, with three different structures. One of which has revealed the dynamic features of the ribosome during translation, identified through 3D classification of the data[69]. This involves sorting the particles into classes based upon their different 3D structures. Therefore, if the protein exists in two distinct conformational states the particles which make up these states will be separated into two different classes thereby producing two 3D models. The other two structures are of the ribosome with two different anti-malarial drugs bound, Emetine[70] and Mefloquine[71], highlighting the power of EM as a tool for visualising how ligands are bound to their target protein in anti-parasitic drug discovery.

Recently Wong et al have provided one of the first examples of a parasitic EM structure being used to drive anti-malarial inhibitor design[71]. By determining the structure of the P. falciparum 80S ribosome in the presence of Mefloquine, the study not only identified the mechanism of action of the compound, they also de novo designed a series of Mefloquine analogues which had improved anti-malarial activity based on the EM derived structure. The structure of the 80S ribosome was determined at 3.2 Å and clearly showed density for one enantiomer of the compound at the GTPase-associated centre of the large ribosomal subunit a. The binding site for the compound was probed to reveal the amino acid residues involved in binding to the compound which was used to design a series of derivatives which showed improved activity.

EM can play an important role in designing selective anti-parasitic inhibitors. This is highlighted by Li et al who determined the structure of the P. falciparum proteasome at 3.6 Å in the presence of an inhibitor (Figure 4B)[72]. The structure showed clear density for the peptide inhibitor which was covalently linked to the protein. Furthermore, the structures of Plasmodium and human proteasomes were compared to reveal the parasitic inhibitor binding site was more open than the human orthologue, which demonstrates how modelling alone based on sequence similarity cannot always account for the more subtle structural differences. This information can subsequently be utilised in the design of new compounds which are highly selective for the parasite protein thus enhancing the chances of creating new therapeutic drugs.

Lower resolution structures obtained by electron cryo-tomography have also provided unique insights into the structure of parasitic proteins and how they interact with their host. One such example is by Akhouri et al. who have used tomography to identify how the P. falciparum Erythrocyte membrane protein 1 (PfEMP1) interacts with human IgM in parasitized red blood cells (pRBC)[73]. They found that PfEMP1 clusters the human IgM on the pRBC surface which enables them to interact with uninfected red blood cells. In doing so, the parasites are able to avoid IgM mediated cell lysis. Another example by Lemgruber et al used cryo-electron tomography to study Plasmodium apicoplasts[74]. Previous to their work, it was unknown whether the Plasmodium apicoplast had three of four membranes. A tilt-series of images were taken and a 3D reconstruction showed that there were clearly four membranes in the apicoplast which was consistent with a number of different apicomplexan parasites. This highlights the importance of the technique in providing key information about an important organelle within a parasitic organism.

5. Challenges for Electron Microscopy to Overcome

Despite the vast improvements in EM over the past five years, X-ray crystallography still remains the gold standard technique for structure based drug design[75]. The main advantage for X-ray crystallography over EM is the resolution of the structures obtained. For instance, high resolution structures of ~1-2 Å are routinely solved by crystallography. At this resolution, the position of water molecules and coordinating metal ions can be established which provides atomic level detail which is rarely seen in EM. This is further utilised when analysing modes of inhibitor binding as high resolution crystal structures can highlight the non-covalent interactions made to the protein. Whilst the number of high resolution structures being produced by EM is increasing, there are still numerous hurdles, such as resolution, for EM to overcome in the future.

Where protein is readily available, stable and homogeneous then X-ray crystallography is often the primary route for structural determination. An advantage of X-ray crystallography over EM is the speed of data collection as many different crystals, which could be complexed with a variety of different inhibitors, can be screened and processed within a day. For EM, the number of particles which are needed to obtain high resolution structural information can vary between ~50,000 to

1,000,000 particles depending on the protein sample[56]. This requires thousands of individual images to be collected over several days before data processing can begin. Large data sets can be generated using software which allows for automated data acquisition. The current speed of the detectors can vary between 50-150 images per hour depending on the camera which is used and the mode of data collection. For instance, integrated mode on the Falcon III can take up to ~150 images per hour but when it is on the more sensitive counting mode the number of exposures per hour is reduced to ~40.

Once all of the data has been collected the time taken to go from raw micrograph to 3D model can be between ~1-4 weeks depending on the amount of data which has been collected. The speed of data processing is improving and has been significantly enhanced via graphical processing unit (GPU) cards[76]. This has greatly increased the speed of the initial processing steps such as motion correction, CTF correction and auto-picking whilst also improving the speeds of the classification and refinement steps of the processing[77]. Furthermore, there will be more emphasis on the initial on-the-fly processing which will increase the speed of the image processing as the individual movies will be motion-corrected and CTF corrected as they are being taken. Another hurdle for EM to overcome is the amount of data generated per data collection. For instance, if each individual movie file is ~1.5 GB in size, a data collection of ~3,000 micrographs will result in ~4.5 TB of raw data being produced. Processing the data will further increase the amount of data generated therefore the computational power and subsequent storage of the data represents a huge challenge in the EM community as the appropriate computational infrastructure needs to be in place. Most funding bodies in the UK also require data to be stored for at least ten years which means that both short and long term storage plans need to be executed.

One of the biggest challenges in the EM field is radiation damage which occurs when the specimen is exposed to the electron beam. It has been described that sufficient exposure to the electron beam in order to obtain an image has the same effect on the specimen as standing next to a nuclear detonation, therefore the level of sample irradiation needs to be kept to a minimum in order to minimise the loss of high resolution structural information[78,79]. When the beam of electrons first hits the specimen, it causes the biggest movement of the particles in the ice, which is thought

to be caused by charging or contraction between the carbon support and the metal grid bars on the grid[80]. This beam-induced movement of the particles causes some of the highest resolution information to be lost as the particles have not yet suffered from radiation damage. If this information can be restored then the resolution which EM could achieve will be further increased. However, fully understanding and limiting the factors that cause the initial movement of the specimen during initial beam exposure will be a significant challenge.

Another limitation which currently hinders EM is the amount of contrast in the individual micrographs. This means that small samples, such as ~100 kDa proteins, are difficult to see without using high levels of defocus which subsequently limits the high-resolution information in the image[81]. This reduction in signal to noise for smaller samples is problematic for downstream processing of the data for example when picking particles and/or aligning and classifying the data. In order to improve the contrast in the images, without applying the high levels of defocus, phase plates can be used. For instance the structure of haemoglobin and Prx3 which were determined at 3.2 Å and 4.4 Å, respectively, using the phase plate which significantly improved the contrast in the raw micrographs[45,82]. This technology is still in its infancy and further work needs to be carried out before it becomes routinely used.

6. Future prospects

Despite using less protein than X-ray crystallography, EM still requires ~3-5 µL of pure sample per grid (typically at ~1mg/ml) which can be challenging to obtain from parasitic organisms. Currently EM grids are prepared by applying ~3 µL of sample onto the grid, blotting the majority of this away using filter paper and then plunge-freezing this into liquid ethane leaving the sample suspended in a thin layer of vitreous ice. The disadvantages of this approach are that it can be unreliable, inconsistent and as most of the sample is blotted away it can waste precious material. Therefore, more efficient ways of grid preparation are being developed which involve spraying pico-litres of sample directly onto the grid[83]. Using this method, it would also be possible to have numerous protein samples applied onto different areas of the same grid which would enable more efficient ways to screen numerous samples, ie protein with different ligands bound.

Moreover, on-grid purification techniques are currently being developed which would also significantly reduce the amount of purified protein needed for grid preparation. Two different approaches have been developed for this. The first involves creating a Nickel monolayer on the surface of the grid thus extracting any His-tagged proteins directly onto the grid[84]. The second approach involves using immobilised antibodies onto the surface of the grid so the target protein becomes attached to the surface[85]. This is particularly useful for proteins where obtaining high quantities from traditional purification columns can be challenging. By purifying directly onto the grid, less protein is needed, therefore this could prove a vital tool in determining parasitic-derived structures.

As already discussed, EM is a useful tool for studying dynamic systems which are not constrained by the crystal lattice. These conformational changes, which occur on a 10-1000 ms timescale, can be monitored using EM via flash freezing the sample at different time-points in a biological reaction[86]. This could be useful in monitoring any conformational changes which occur upon formation of a large macromolecular complex or upon binding of a ligand. So far, time-resolved EM has provided useful insights into the catalytic cycle of the ribosome[87]. Currently developments are being made to the rapid mixing/spraying of the sample onto the grid before it is plunge-frozen[88]. By being able to trap biological systems at different time points, key insights into the mechanism of action of many complexes and the flexibility of inhibitor binding sites can be gained which will enhance future drug discovery projects.

7. Concluding Comments

The rapid advances in EM has given rise to a huge increase in the number of structures routinely being determined to high resolution. Due to the improved resolution of the structures being determined, EM can now be considered as a tool in structure based drug design. By moving away from the static snapshots provided by X-ray crystallography, EM can be used to study dynamic and flexible systems. This will aid SBDD programmes as the dynamic nature of inhibitor binding sites can be probed thus allowing inhibitors to be designed against the most physiologically relevant states. Despite the large improvements in EM, the technique is still in its infancy and future developments within the field are inevitable. In addition to established techniques such as X-ray crystallography and NMR, EM can now be used

to elucidate the structural information of parasitic protein targets thereby enabling a new generation of inhibitors to be designed.

CONFLICT OF INTEREST

No authors declare no conflict of interest.

ACKNOWLEDGEMENTS

RMJ is a Wellcome Trust funded PhD student (109158/B/15/Z) and SR is a Wellcome Trust funded ISSF fellow (105615/Z/14/Z).

REFERENCES

1. Pink, R., Hudson, A., Mouriès, M.-A. & Bendig, M. Opportunities and challenges in antiparasitic drug discovery. Nat. Rev. Drug Discov. **4,** 727–740 (2005).
2. Guido, R. V. C. & Oliva, G. Structure-based drug discovery for tropical diseases. Curr. Top. Med. Chem. **9,** 824–843 (2009).
3. World Health Organisation. Malaria Fact Sheet. (2017). Available at: http://www.who.int/mediacentre/factsheets/fs094/en/. (Accessed: 31st May 2017)
4. McPhillie, M., Zhou, Y., El Bissati, K., Dubey, J., Lorenzi, H., Capper, M., Lukens, A. K., Hickman, M., Muench, S., Verma, S. K., Weber, C. R., Wheeler, K., Gordon, J., Sanders, J., Moulton, H., Wang, K., Kim, T.-K., He, Y., Santos, T., et al. New paradigms for understanding and step changes in treating active and chronic, persistent apicomplexan infections. Sci. Rep. **6,** 29179 (2016).
5. Fernandez-Leiro, R. & Scheres, S. H. W. Unravelling biological macromolecules with cryo-electron microscopy. Nature **537,** 339–346 (2016).
6. Agarwal, A. K. & Fishwick, C. W. G. Structure-based design of anti-infectives. Ann. N. Y. Acad. Sci. **1213,** 20–45 (2010).
7. Rao, V. S. & Srinivas, K. Modern drug discovery process : An in silico approach. J. Bioinforma. Seq. Anal. **2,** 89–94 (2011).
8. Gordon, J., Fishwick, C. W. G. & Mcphillie, M. J. New Opportunities in the Structure-based Design of Anti-Protozoan Agents. Curr Top Med Chem **17,** 79–90 (2017).

9. Hol, W. G. J. Three-dimensional structures in the design of therapeutics targeting parasitic protozoa: Reflections on the past, present and future. Acta Crystallogr. Sect. FStructural Biol. Commun. **71,** 485–499 (2015).
10. Phillips, M. A., Lotharius, J., Marsh, K., White, J., Dayan, A., White, K. L., Njoroge, J. W., Mazouni, F. El, Lao, Y., Kokkonda, S., Tomchick, D. R., Deng, X., Laird, T., Bhatia, S. N., March, S., Ng, C. L., Fidock, D. A., Wittlin, S., Lafuente-Monasterio, M., et al. A long-duration dihydroorotate dehydrogenase inhibitor (DSM265) for prevention and treatment of malaria. Sci. Transl. Med. **7,** 296ra111 (2016).
11. Ferreira, L. G., Santos, R. N., Oliva, G. & Andricopulo, A. D. Molecular Docking and Structure-Based Drug Design Strategies. Molecules **20,** 13384–13421 (2015).
12. Krieger, E., Nabuurs, S. B. & Vriend, G. Homology modeling. Struct. Bioinforma. 507–521 (2003).
13. Pizzi, E. & Frontali, C. Low-Complexity Regions in Plasmodium falciparum Proteins. Genome Res. **11,** 218–229 (2001).
14. Levy, C. W., Roujeinikova, A., Sedelnikova, S., Baker, P. J., Stuitje, A. R., Slabas, A. R., Rice, D. W. & Rafferty, J. B. Molecular basis of triclosan activity. Nature **398,** 383–384 (1999).
15. Muench, S. P., Sean, T., Mcleod, R., John, B., Kirisits, M. J., Roberts, C. W., Mui, E. J. & Rice, D. W. Studies of Toxoplasma gondii and Plasmodium falciparum enoyl acyl carrier protein reductase and implications for the development of antiparasitic agents. Acta Crystallogr. Sect. D **D63,** 328–338 (2007).
16. Kuo, M. R., Morbidoni, H. R., Alland, D., Sneddon, S. F., Gourlie, B. B., Staveski, M. M., Leonard, M., Gregory, J. S., Janjigian, A. D., Yee, C., Musser, J. M., Kreiswirth, B., Iwamoto, H., Perozzo, R., Jacobs, W. R., Sacchettini, J. C. & Fidock, D. A. Targeting Tuberculosis and Malaria through Inhibition of Enoyl Reductase. J. Biol. Chem. **278,** 20851–20859 (2003).
17. Birkholtz, L.-M., Wrenger, C., Joubert, F., Wells, G. A., Walter, R. D. & Louw, A. I. Parasite-specific inserts in the bifunctional S-adenosylmethionine decarboxylase/ornithine decarboxylase of Plasmodium falciparum modulate catalytic activities and domain interactions. Biochem. J. **377,** 439–448 (2004).
18. Yuvaniyama, J., Chitnumsub, P., Kamchonwongpaisan, S., Vanichtanankul, J.,

Sirawaraporn, W., Taylor, P., Walkinshaw, M. D. & Yuthavong, Y. Insights into antifolate resistance from malarial DHFR-TS structures. Nat. Struct. Biol. **10,** 357–365 (2003).

19. Althoff, T., Mills, D. J., Popot, J. & Ku, W. Arrangement of electron transport chain components in bovine mitochondrial. EMBO J. **30,** 4652–4664 (2011).
20. Gibbons, C., Montgomery, M. G., Leslie, A. G. W. & Walker, J. E. The structure of the central stalk in bovine F1-ATPase at 2.4 Å resolution. Nat. Struct. Biol. **7,** 1055–1061 (2000).
21. Morales-rios, E., Montgomery, M. G., Leslie, A. G. W. & Walker, J. E. Structure of ATP synthase from Paracoccus denitrificans determined by X-ray crystallography at 4 . 0 Å resolution. PNAS **112,** 1–6 (2015).
22. Zhou, A., Rohou, A., Schep, D. G., Bason, J. V, Montgomery, M. G., Walker, J. E. & Grigorieff, N. Structure and conformational states of the bovine mitochondrial ATP synthase by cryo-EM. Elife **4,** 1–15 (2015).
23. Mühleip, A. W., Dewar, C. E., Schnaufer, A., Kühlbrandt, W. & Davies, K. M. In-situ structure of trypanosomal ATP synthase dimer reveals unique arrangement of catalytic subunits. PNAS **114,** 992–997 (2017).
24. Nina, P. B., Morrisey, J. M., Ganesan, S. M., Ke, H., Pershing, A. M., Mather, M. W. & Vaidya, A. B. ATP synthase complex of Plasmodium falciparum: Dimeric assembly in mitochondrial membranes and resistance to genetic disruption. J. Biol. Chem. **286,** 41312–41322 (2011).
25. Arinaminpathy, Y., Khurana, E., Engelman, D. M. & Gerstein, M. B. Computational analysis of membrane proteins : the largest class of drug targets. Drug Discov. Today **14,** 1130–1135 (2009).
26. Bull, S. C. & Doig, A. J. Properties of Protein Drug Target Classes. PLoS One 1–44 (2015). doi:10.1371/journal.pone.0117955
27. Cherezov, V. Lipidic cubic phase technologies for membrane protein structural studies. Curr. Opin. Struct. Biol. **21,** 559–566 (2011).
28. Consortium, S. G., Structural, B., Consortium, G., Innovation, F., Structural, I., Structural, N., Consortium, G., Protein, O., Facility, P., Production, P. S. & Initiative, P. Protein production and purification. Nat. Methods **5,** 135–147 (2008).
29. Weber, J. L. Analysis of sequences from the extremely A + T-rich genome of PIasmodium falciparum. Gene **52,** 103–109 (1987).

30. Baca, A. M. & Hol, W. G. J. Overcoming codon bias : A method for high-level overexpression of Plasmodium and other AT-rich parasite genes in Escherichia coli. Int. J. Parasitol. **30,** 113–118 (2000).
31. Smardon, A. M., Tarsio, M. & Kane, P. M. The RAVE complex is essential for stable assembly of the yeast V-ATPase. J. Biol. Chem. **277,** 13831–13839 (2002).
32. Baradaran, R., Berrisford, J. M., Minhas, G. S. & Sazanov, L. A. Crystal structure of the entire respiratory complex I. Nature **494,** 443–448 (2013).
33. Xia, D., Yu, C. a, Kim, H., Xia, J. Z., Kachurin, a M., Zhang, L., Yu, L. & Deisenhofer, J. Crystal structure of the cytochrome bc1 complex from bovine heart mitochondria. Science **277,** 60–66 (1997).
34. Menz, R. I., Walker, J. E. & Leslie, A. G. W. Structure of Bovine Mitochondrial F 1 -ATPase with Nucleotide Bound to All Three Catalytic Sites : Implications for the Mechanism of Rotary Catalysis. Cell **106,** 331–341 (2001).
35. Gu, J., Wu, M., Guo, R., Yan, K., Lei, J., Gao, N. & Yang, M. The architecture of the mammalian respirasome. Nature **537,** 639–643 (2016).
36. Kuehlbrandt, W. The Resolution Revolution. Science (80-.). **343,** 1443–1444 (2014).
37. Bai, X., Mcmullan, G. & Scheres, S. H. W. How cryo-EM is revolutionizing structural biology. Trends Biochem. Sci. **40,** 49–57 (2015).
38. Rawson, S., Davies, S., Lippiat, J. D. & Muench, S. P. The changing landscape of membrane protein structural biology through developments in electron microscopy. Mol. Membr. Biol. **33,** 12–22 (2016).
39. Gao, Y., Cao, E., Julius, D. & Cheng, Y. Article TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. Nature **534,** 347–351 (2016).
40. Hite, R. K., Yuan, P., Li, Z., Hsuing, Y., Walz, T. & MacKinnon, R. Cryo-electronmicroscopy structure of the Slo2.2 Na1-activated K1 channel. Nature **527,** 198–203 (2015).
41. Wu, J., Yan, Z., Li, Z., Qian, X., Lu, S., Dong, M., Zhou, Q. & Yan, N. Structure of the voltage-gated calcium channel Cav1.1 at 3.6 Å resolution. Nat. Publ. Gr. **537,** 191–196 (2016).
42. Bai, X., Yan, C., Yang, G., Lu, P., Sun, L., Zhou, R., Scheres, S. H. W. & Shi, Y. An atomic structure of human y-secretase. Nature **525,** 212–217 (2015).
43. Kühlbrandt, W. Cryo-EM enters a new era. Elife **3,** 1–4 (2014).

44. Meng Wu, Jinke Gu, Runyu Guo, Yushen Huang, M. Y. & Correspondence. Structure of Mammalian Respiratory Supercomplex Article Structure of Mammalian Respiratory. Cell **167,** 1598–1609 (2016).
45. Danev, R. & Nagayama, K. Phase Plates for Transmission Electron Microscopy. Methods in Enzymology **481,** (Elsevier Masson SAS, 2010).
46. Khoshouei, M., Radjainia, M., Baumeister, W. & Danev, R. Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate. Nat. Commun. **8,** 1–6 (2017).
47. Danev, R. & Baumeister, W. Cryo-EM single particle analysis with the Volta phase plate. Elife **5,** 1–14 (2016).
48. Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M. I., Pragani, R., Boxer, M. B., Earl, L. A., Milne, J. L. S. & Subramaniam, S. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. Cell **165,** 1698–1707 (2016).
49. Huang, X., Luan, B., Wu, J. & Shi, Y. An atomic structure of the human 26S proteasome. Nat. Publ. Gr. **23,** 778–785 (2016).
50. Bartesaghi, A., Merk, A., Banerjee, S., Matthies, D., Wu, X., Milne, J. L. S. & Subramaniam, S. 2.2 Å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor. Science (80-.). **348,** 1147–1151 (2015).
51. Rawson, S., Mcphillie, M. J., Johnson, R. M., Fishwick, C. W. G. & Muench, S. P. The potential use of single-particle electron microscopy as a tool for structure-based inhibitor design research papers. Acta Crystallogr. Sect. D Biol. Crystallogr. **73,** 1–7 (2017).
52. Song, C. F., Papachristos, K., Rawson, S., Huss, M., Wieczorek, H., Trinick, J., Harrison, M. A. & Muench, S. P. Flexibility within the Rotor and Stators of the Vacuolar H + - ATPase. PLoS One **8,** 1–12 (2013).
53. Rawson, S., Iadanza, M. G., Ranson, N. A. & Muench, S. P. Methods to account for movement and flexibility in cryo-EM data processing. Methods **100,** 35–41 (2016).
54. Hite, R. K. & Mackinnon, R. Structural Titration of Slo2.2, a Na+-Dependent K+ Channel. Cell **168,** 390–399 (2017).
55. Bammes, B. E., Rochat, R. H., Jakana, J., Chen, D. & Chiu, W. Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3 / 4 Nyquist frequency. J. Struct. Biol. **177,** 589–601 (2012).

56. Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P. & Ranson, N. A. An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. Methods **100,** 3–15 (2016).
57. Grigorieff, N. Direct detection pays off for electron cryo-microscopy. Elife **2,** 1–3 (2013).
58. Baker, L. A. & Rubinstein, J. L. Radiation Damage in Electron Cryomicroscopy. Methods in Enzymology **481,** (Elsevier Masson SAS, 2010).
59. Karuppasamy, M., Nejadasl, K. & Vulovic, M. Radiation damage in single-particle cryo-electron microscopy : effects of dose and dose rate. J. Synchotron Radiat. **18,** 398–412 (2011).
60. Grant, T. & Grigorieff, N. Measuring the optimal exposure for single particle cryo-EM using a 2.6 A reconstruction of rotavirus VP6. Elife **4,** 1–19 (2015).
61. Lučić, V., Rigort, A. & Baumeister, W. Cryo-electron tomography: The challenge of doing structural biology in situ. J. Cell Biol. **202,** 407–419 (2013).
62. Irobalieva, R. N., Martins, B. & Medalia, O. Cellular structural biology as revealed by cryo-electron tomography. J. Cell Sci. **129,** 469–476 (2016).
63. Briggs, J. A. G. Structural biology in situ — the potential of subtomogram averaging. Curr. Opin. Struct. Biol. **23,** 261–267 (2013).
64. Schur, F. K. M., Obr, M., Hagen, W. J. H., Wan, W., Jakobi, A. J., Kirkpatrick, J. M., Sachse, C., Kräusslich, H.-G. & Briggs, J. A. G. An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. Science (80-.). **353,** 506–508 (2016).
65. Hashem, Y., Georges, A., Fu, J., Buss, S. N., Jossinet, F., Jobe, A., Zhang, Q., Liao, H. Y., Grassucci, R. A., Bajaj, C., Westhof, E., Madison-antenucci, S. & Frank, J. High-resolution cryo-electron microscopy structure of the Trypanosoma brucei ribosome. Nature **494,** 385–389 (2013).
66. Liu, Z., Gutierrez-vargas, C., Wei, J., Grassucci, R. A., Ramesh, M. & Espina, N. Structure and assembly model for the Trypanosoma cruzi 60S ribosomal subunit. PNAS **113,** 1–6 (2016).
67. Shalev-benami, M., Zhang, Y., Matzov, D., Jaffe, C. L., Yonath, A., Skiniotis, G., Shalev-benami, M., Zhang, Y., Matzov, D., Halfon, Y., Zackay, A. & Rozenberg, H. 2.8-A Cryo-EM Structure of the Large Ribosomal Subunit from the Eukaryotic Parasite Leishmania. CellReports **16,** 288–294 (2016).
68. Zhang, X., Lai, M., Chang, W., Yu, I., Ding, K., Mrazek, J., Ng, H. L., Yang, O.

O., Maslov, D. A. & Zhou, Z. H. Structures and stabilization of kinetoplastid-specific split rRNAs revealed by comparing leishmanial and human ribosomes. Nat. Commun. **7,** 1–10 (2016).

69. Sun, M., Li, W., Blomqvist, K., Das, S., Hashem, Y., Dvorin, D. & Frank, J. Dynamical features of the Plasmodium falciparum ribosome during translation. Nucleic Acids Res. **43,** 10515–10524 (2015).
70. Wong, W., Bai, X., Brown, A., Fernandez, I. S., Hanssen, E., Condron, M., Tan, Y. H., Baum, J. & Scheres, S. H. W. Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. Elife **3,** 1–20 (2014).
71. Wong, W., Bai, X.-C., Sleebs, B. E., Triglia, T., Brown, A., Thompson, J. K., Jackson, K. E., Hanssen, E., Marapana, D. S., Fernandez, I. S., Ralph, S. A., Cowman, A. F., Scheres, S. H. W. & Baum, J. Mefloquine targets the Plasmodium falciparum 80S ribosome to inhibit protein synthesis. Nat. Microbiol. **2,** 17031 (2017).
72. Li, H., Donoghue, A. J. O., Linden, W. A. Van Der, Xie, S. C., Yoo, E., Foe, I. T., Tilley, L., Craik, C. S., Fonseca, P. C. A. & Bogyo, M. Structure- and function-based design of Plasmodium-selective proteasome inhibitors. Nature **530,** 233–236 (2016).
73. Erythrocyte, P., Protein, M., Akhouri, R. R., Goel, S., Furusho, H., Skoglund, U. & Wahlgren, M. Architecture of Human IgM in Complex with P . falciparum Erythrocyte Membrane Protein 1. CellReports **14,** 723–736 (2016).
74. Lemgruber, L., Kudryashev, M., Dekiwadia, C., Riglar, D. T. & Baum, J. Cryo-electron tomography reveals four-membrane architecture of the Plasmodium apicoplast. Malar. J. **12,** 1–8 (2013).
75. Zheng, H., Hou, J., Zimmerman, M. D., Wlodawer, A. & Minor, W. The future of crystallography in drug discovery. Expert Opin. Drug Discov. **9,** 125–137 (2015).
76. Forsberg, O., Scheres, S. H. W. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife **5,** 1–21 (2016).
77. Zhang, K. Gctf : Real-time CTF determination and correction. J. Struct. Biol. **193,** 1–12 (2016).
78. Taylor, A. & Glaeser, R. M. Radiation damage relative to transmission electron microscopy of biological specimens at low temperature : a review. J. Microsc.

112, 127–138 (1978).

79. Orlova, E. V & Saibil, H. R. Structural Analysis of Macromolecular Assemblies by Electron Microscopy. Chem. Rev. **111,** 7710–7748 (2011).
80. Brilot, A. F., Chen, J. Z., Cheng, A., Pan, J., Harrison, S. C., Potter, C. S., Carragher, B., Henderson, R., Grigorieff, N. & Em, H. Beam-induced motion of vitrified specimen on holey carbon film. J. Struct. Biol. **177,** 630–637 (2012).
81. Glaeser, R. M. & Hall, R. J. Reaching the Information Limit in Cryo-EM of Biological Macromolecules : Experimental Aspects. Biophys. J. **100,** 2331–2337 (2011).
82. Khoshouei, M., Radjainia, M., Phillips, A. J., Gerrard, J. A., Mitra, A. K., Baumeister, W. & Danev, R. Volta phase plate cryo-EM of the small protein complex Prx3. Nat. Commun. **7,** 1–5 (2016).
83. Razinkov, I., Dandey, V. P., Wei, H., Zhang, Z., Melnekoff, D., Rice, W. J., Wigge, C., Potter, C. S. & Carragher, B. A new method for vitrifying samples for cryoEM. J. Struct. Biol. **195,** 190–198 (2016).
84. Kelly, D. F., Abeyrathne, P. D., Dukovski, D. & Walz, T. The Affinity Grid : A Pre-fabricated EM Grid for Monolayer Purification. J. Mol. Biol. **382,** 423–433 (2008).
85. Yu, G., Li, K. & Jiang, W. Antibody-based affinity cryo-EM grid. Methods **100,** 16–24 (2016).
86. Chen, B. & Frank, J. Two promising future developments of cryo-EM : capturing short-lived states and mapping a continuum of states of a macromolecule. Microscopy **65,** 69–79 (2016).
87. Dashti, A., Schwander, P., Langlois, R., Fung, R., Li, W., Hosseinizadeh, A. & Liao, H. Y. Trajectories of the ribosome as a Brownian nanomachine. PNAS **111,** 17492–17497 (2014).
88. Feng, X., Fu, Z., Kaledhonkar, S., Jia, Y., Shah, B., Jin, A., Liu, Z., Sun, M., Chen, B., Grassucci, R. A., Ren, Y., Jiang, H., Frank, J. & Lin, Q. A Fast and Effective Microfluidic Spraying-Plunging Method for High-Resolution Single-Particle Cryo-EM. Structure **25,** 663–670 (2017).

Figures

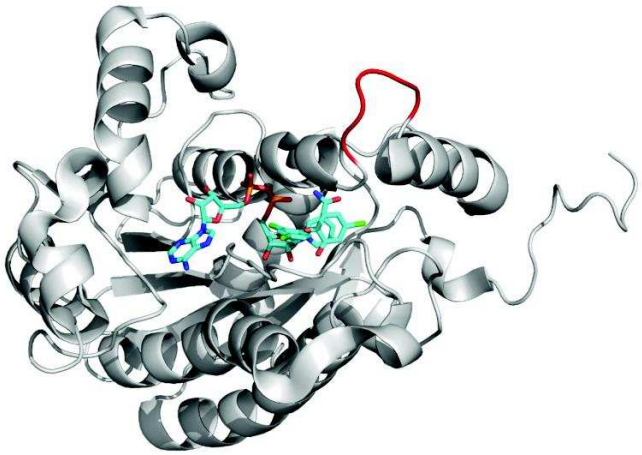

Figure 1: Low complexity insert in T. gondii enoyl reductase (ENR). The parasitic insert can vary between ~5-40 amino acid residues. In T. gondii the parasitic insert, consisting of 5 amino acid residues, is shown in red and is in close proximity to the inhibitor (cyan) binding site (pdb 2o2s)[15].

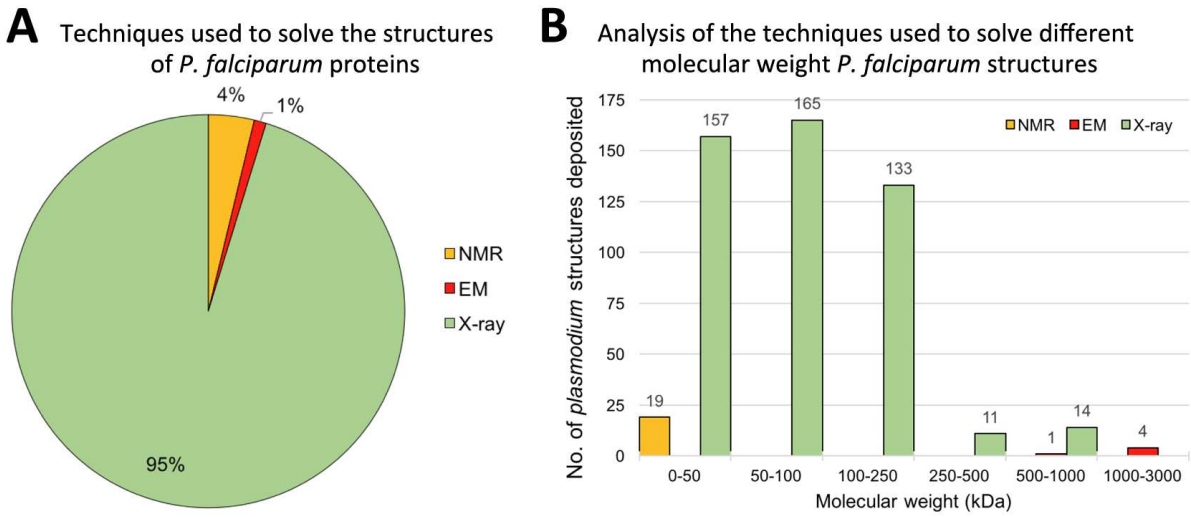

Figure 2: The number of P. falciparum structures deposited in the protein data bank (pdb). A) A pie chart highlighting the percentage of structures deposited into the pdb according to the technique used. B) The proteins deposited into the pdb according to technique used and molecular weight. This shows that EM is an important tool for proteins with a molecular weight greater than 250 kDa as there are fewer X-ray crystal structures solved for higher molecular weights. The numbers were accurate as of June 2017.

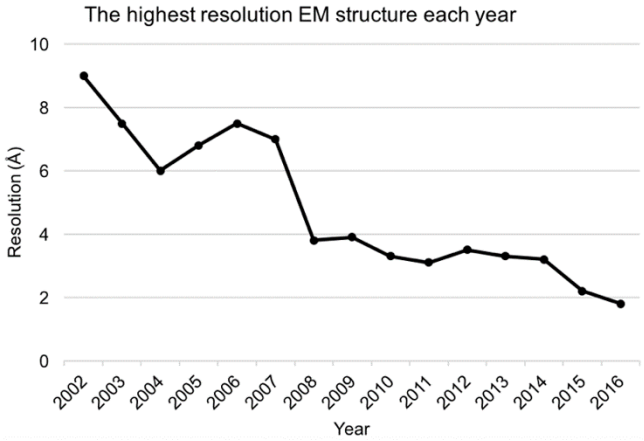

Figure 3: The trend showing the highest resolution EM map each year. The highest resolution map deposited in the EMDB each year between 2002-2017.

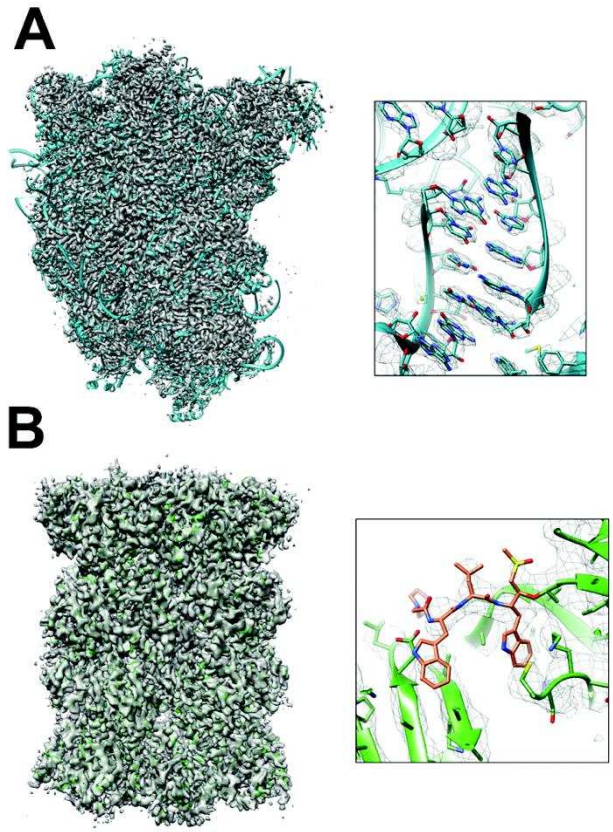

Figure 4: Examples of parasitic structures determined by EM. A) The highest resolution structure determined from a parasitic organism of the T. cruzi ribosome at 2.5 Å (EMDB 8361, pdb 5t5h)[66] with example density of DNA (right). B) The 3.6 Å structure of the P. falciparum proteasome (EMDB3231, pdb 5fmg)[72] with density for the covalently bound peptide inhibitor (right).