High fidelity DNA replication and repair: new structures and mechanisms using cryogenic electron microscopy

Borsellini, A.

Introduction

Background on DNA

In all living organisms, deoxyribonucleotide acid (DNA) stores the genetic information that is needed to perform every biological process inside a cell. To utilise this genetic information, the DNA is first transcribed into messenger RNA, which is in turn translated into proteins that perform diverse functions in the cell. The flow of information from DNA to RNA to protein is known as the central dogma in biology. Moreover, the DNA in the cells is copied and divided between daughter cells during cell division, through a very accurate process called DNA replication. As suggested by Watson and Crick in their 1953 paper [1], the ability of the DNA to store and transmit the genetic information is a direct result of its structure.

The DNA is a right handed double helix consisting of two anti-parallel strands of deoxyribonucleotides[2] (Fig. 1). Each deoxyribonucleotide consists of a deoxyribose sugar, a base and a phosphate group, and the individual nucleotides are covalently bonded together through the linking of the 3' hydroxyl group of the sugar on one nucleotide to the 5' phosphate group of the adjacent nucleotide through phosphodiester bonds. These 5'-3' linkages form the sugar-phosphate backbone that is located on the outside of the DNA molecule. Within the DNA molecule, the genetic information is stored in the bases that group together to form sequences of instructions for the cell. The DNA is made of four bases, adenosine (A), guanine (G), thymine (T) and cytosine (C), that are arranged perpendicularly to the phosphate backbone.

The bases form pairs in a specific manner, where A pairs with T and G pairs with C as described by Watson and Crick. The structure of DNA is further stabilized through π-π interactions between the adjacent vertically stacked base-pairs (Fig. 1). The specific base-pairing ensures that if the sequence on one strand is known the sequence on the other strand can be automatically determined.

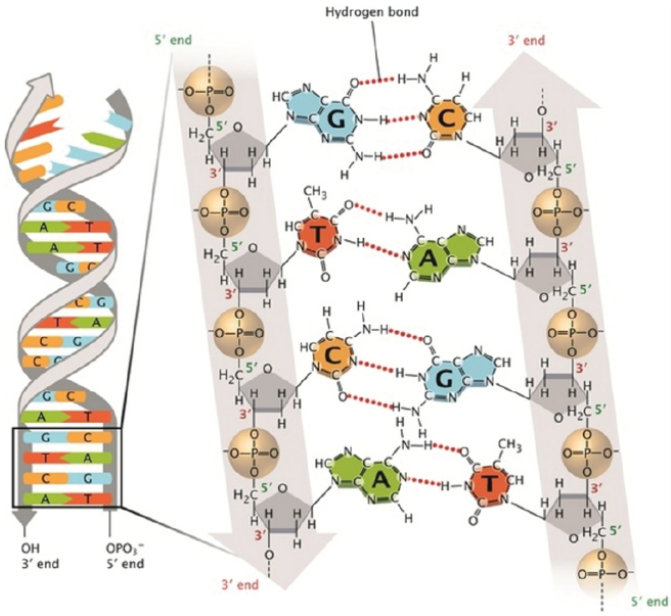

Figure 1. The structure of DNA helix | Schematic representation of the anti-parallel double helical structure of DNA showing the sugar-phosphate backbone and complementary base pairing between bases. Figure taken from Praym 2008.

DNA replication

When DNA is replicated, each strand of the double helix acts as the template for the synthesis of a new strand through Watson and Crick base-pairing. Enzymes called DNA polymerases catalyse the reaction between the 3' hydroxyl of the ribose ring and the triphosphate group of the incoming nucleotide [3]. Due to the unique geometry of this reaction, DNA synthesis can only take place in one direction, which is commonly called 5' to 3'. This poses a challenge to the replication of the DNA since the two strands of the double helix run in antiparallel direction. To overcome this problem, nature has evolved a multi-protein molecular machinery called the replisome that is able to synthetize DNA on both strands simultaneously while moving in only one direction [4] (Fig. 2a). In order to do so, the replisome synthetizes one strand continuously (leading strand) and one strand discontinuously (lagging strand) in short ~1 kb fragments called Okazaki fragments [5] (Fig. 2b), which are than ligated together by the action of other enzymes. All the necessary events for DNA replication are collectively orchestrated within the replisome, which efficiently syntheses leading and lagging strand ensuring that the two copies of the DNA are made simultaneously and divided between mother and daughter cells.

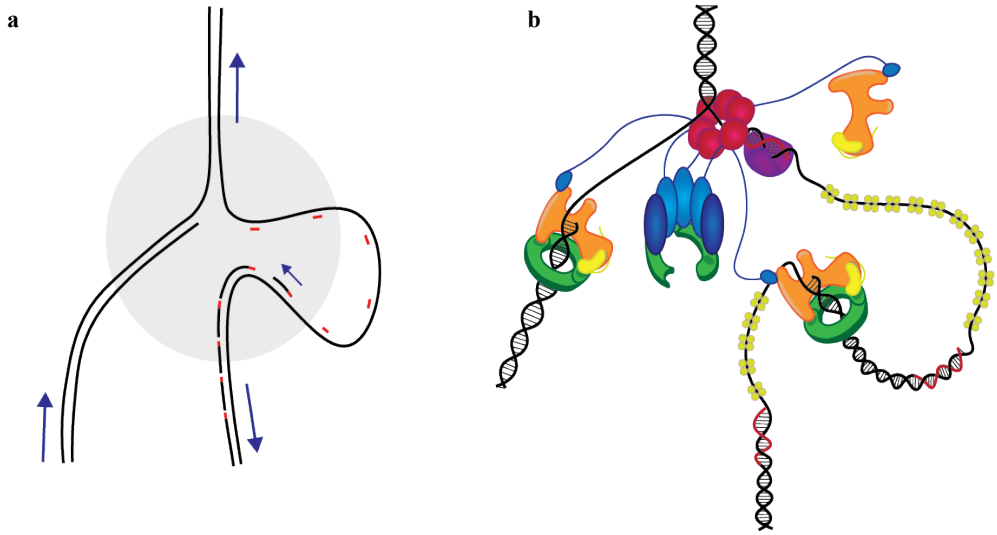

Figure 2 | The replisome synthetizes both DNA strands simultaneously. a, Schematic representation of leading and lagging strand DNA synthesis. Grey circle represents the position of the replisome. **b,** Cartoon representation of the *E. coli* replisome. The helicase (red) unwinds the DNA. The primase (purple) synthetises small RNA primers for the discontinuous synthesis on the lagging strand. Both strands are replicated by a DNA polymerase (orange) that is bound to an exonuclease and a clamp (yellow and green). The polymerases (orange) synthetize DNA starting from the RNA primers. The clamps (green), loaded on the DNA by the clamp loader (blue), hold the polymerases on the DNA increasing their processivity. SSB (yellow) coats the ssDNA preventing DNA from forming secondary structures.

Replication fidelity

Importantly, DNA replication is a very accurate process, with an average of one wrong nucleotide incorporated per 10^8 –10^{10} nucleotides polymerized. The fidelity of DNA replication relies on the ability of replicative polymerases to select the correct nucleotide, together with the activity of proofreading exonucleases and DNA mismatch repair that remove mismatches during and after DNA replication, respectively [6,7]. All three depend on the ideal geometry between base pairs (Fig. 3). First, DNA polymerases, have an high insertion fidelity, and only make a mistake every 10^5-10^6 incorporated nucleotides [7]. The low rate at which replicative DNA polymerases introduce mismatches in the daughter strand is a result of their narrow active site, that only favours the reaction with the complementary nucleotides that respect Watson-Crick geometry [8]. In the rare occasions in which a mismatch is introduced, the distorted geometry of the base pair prevents further synthesis of the daughter DNA strand [3]. Therefore, to continue DNA synthesis, all high-fidelity DNA polymerases have a 3'-5' exonuclease domain that can excise the wrong nucleotide. It has been shown that the presence of a mismatch on the primer strand increases

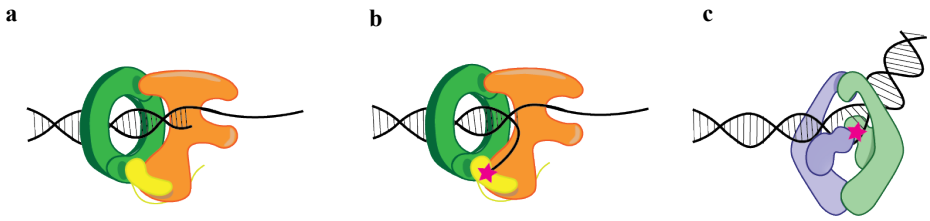

Figure 3 | Three mechanisms to ensure high fidelity DNA replication. a, DNA polymerases only make a mismatch every 10^4-10^5 incorporated nucleotides. **b,** DNA polymerase in editing mode. The exonuclease domain (yellow) removes mis-incorporated nucleotides during DNA synthesis. **c,** The MutS protein is able to detect mismatches among millions of perfectly matched base pairs.

the fraying of the DNA, and favours the transition of the primer terminus from the polymerase active site to the exonuclease active site [9,10,11]. The presence of the exonuclease domains reduces the error rate of DNA replication to one mismatch every 10^7-10^8 incorporated nucleotides [6]. Finally, the DNA mismatch repair pathway is able to detect and remove mismatches left behind by the replication machinery, further reducing the error rate to one mismatch every 10^9-10^{10} incorporated nucleotides. The activity of DNA mismatch repair proteins also depends on the incorrect pairing of the DNA bases. Mismatches on the DNA have a distorted geometry due to the non-ideal pairing of the nucleobases. The protein MutS is able to recognize the distortion of a mismatched base pair and start the repair cascade that removes the wrong nucleotide [12], safeguarding the genome from the incorporation of potentially harmful mutations.

In the DNA helix the complementary bases lie at a distance which is determined by the hydrogen bonds between bases. Only when a base pair does not respect the Watson-Crick pairing than the activity of exonucleases or mismatch repair enzymes is favoured. Although it is the proteins that catalyse the different reactions which ensure high fidelity DNA replication, it is the DNA itself that governs and determines their activity.

DNA mismatch repair

DNA mismatch repair (MMR) is an essential safeguard of genomic integrity that corrects mismatches and insertion/deletion loops generated during DNA replication [13,14]. MMR is conserved from prokaryotes to eukaryotes and its deficiency leads to increased mutation rates, antibiotic drug resistance in bacteria, and Lynch syndrome (LS) in humans, also known as hereditary non-polyposis colon cancer (HNPCC) [15].

The *Escherichia coli* mismatch repair pathway is the most studied, and it is used as a model to understand the cascade of enzymatic activities that ensure the removal of the mis-incorporated nucleotide from the newly synthesized strand (Fig. 4). This process is initiated when the dimeric MutS protein binds to a mismatched or looped out base [16]. MutS then

undergoes a large conformational change in an ATP-dependent manner that results in the recruitment of the second repair protein MutL [17]. MutL activates the endonuclease MutH that selectively nicks the newly synthesized strand at hemi-methylated GATC, sites which are temporarily present on the DNA after DNA replication [18,19]. The single stranded nick then forms the entry point for the UvrD helicase, that with the aid of a 5'-3' or 3'-5' exonuclease will excise the newly synthesized strand [20]. The removed stretch of DNA will then be resynthesized by a DNA polymerase and the remaining nick sealed by a DNA ligase [21,22].

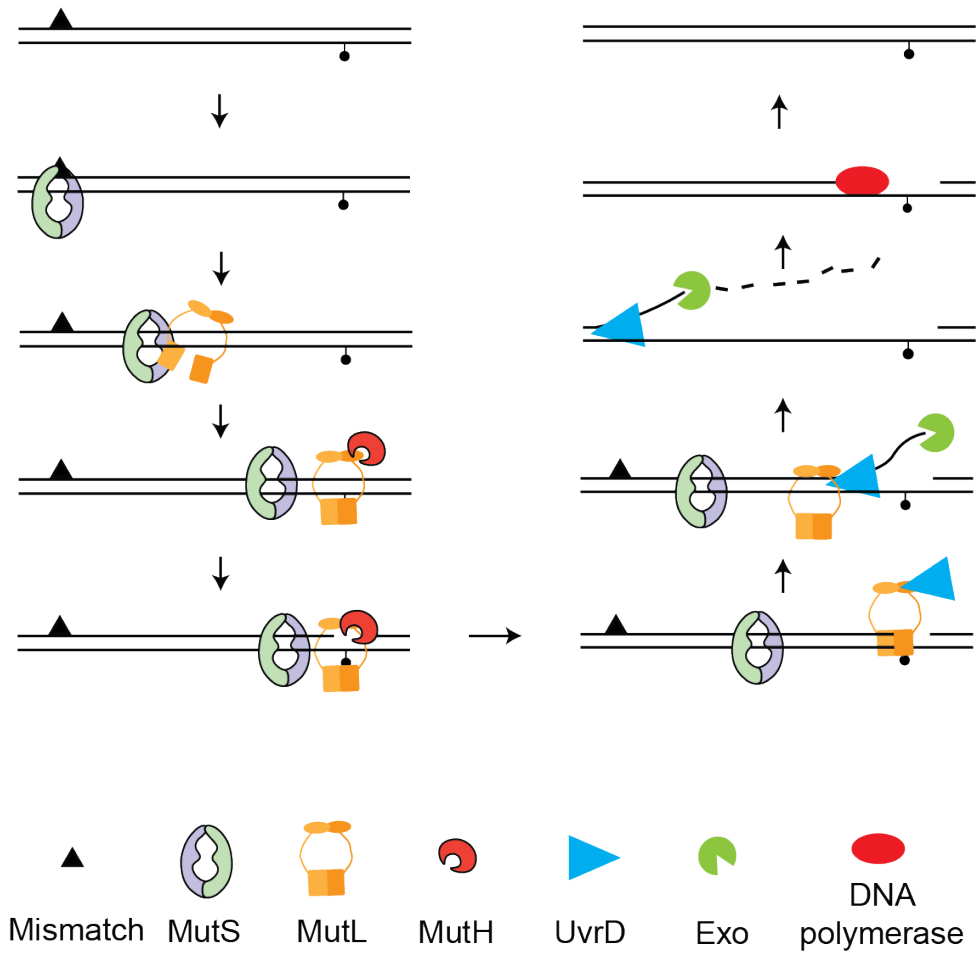

Figure 4 | Schematic representation of the *E. coli* Mismatch repair. MutS recognizes the mismatch and recruits MutL on the DNA. MutL activates MutH to nick the DNA and UvrD to unwind it. An exonuclease excises the ssDNA and a polymerase resynthesises the DNA once the mismatch is removed.

The MutS protein

Although the mismatch repair reaction requires the action of multiple proteins, the master coordinator of this process is the MutS protein, which can sense the presence of a mismatch, activate the signal to start the repair cascade and de-activate it once the mismatch is removed. In order to do so, it uses ATP to change conformation in response to different signal on the DNA molecule.

The MutS homologues are dimeric proteins, belonging to the ABC family of ATPases [23]. Characteristic of this ATPase family are the two ATP hydrolysis active sites which are at the interface of the dimer. In order to hydrolyse ATP, the two domains need to come close together and adopt a conformation which is suitable for hydrolysis. Once ATP is hydrolysed and the product is released, the two domains can open up and reset the protein for another round of ATP binding and hydrolysis. The consecutive rounds of ATP binding and hydrolysis generate an open-close movement in the ATPase domain, which is used by multiple proteins, such us membrane transporters or chromatin remodelling enzymes, to perform different tasks inside the cell [24,25].

Thanks to the crystal structure published in the year 2000, the recognition mode of a DNA mismatch by the MutS protein is well characterized [26,27](Fig. 5). "The MutS dimer is an asymmetric molecule that looks like a pair of praying hands, with the wrists close together, the thumbs coming close and the fingers lightly touching. The DNA is held between thumbs and fingers" [26]. The "thumbs" which are called mismatch binding domains, interact with the DNA from the bottom with only one of the two binding specifically to the mismatch. The "fingers" are called clamp and lever domains and engage the DNA from the top, trapping it between the monomers. The "thumbs" and "fingers" are connected through the connector and core domains which are the "palm of the hand". The DNA in this structure is kinked of 60°, thanks to the mismatch binding domain that

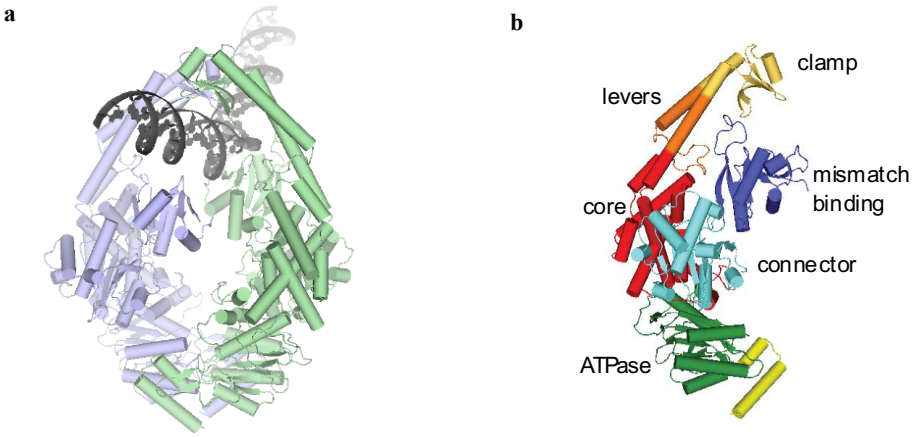

Figure 5 | Structure of MutS. a, Structure of *E. coli* MutS bound to a mismatched DNA. Monomer A, B and the DNA are coloured in pale green, light blue and grey, respectively. **b**, The domains of MutS marked in different colours and indicated by name.

protrudes into the minor groove and bends the DNA structure. The ATPase domains lie at the bottom of the protein, and although they do not directly interact with the DNA, their role is crucial for the conformational changes required after mismatch recognition.

Conformational changes of MutS during DNA mismatch repair

Recent cryo-EM structures of MutS at different stages of the repair process reveal how a small conformational change from matched (homoduplex) to mismatched (heteroduplex) bound acts as a licencing step that triggers a dramatic conformational change in MutS and subsequent activation of the repair process [28]. When MutS scans the DNA for mismatches the DNA structure acts as a steric block which can only be overcome through kinking of the DNA at a mismatch. After mismatch binding, MutS uses ATP to rearrange in the sliding clamp conformation which can only dissociate from the DNA at open DNA ends or single-stranded DNA stretches [29,30]. It is only in this sliding clamp conformation that MutS can recruit MutL and start the repair cascade. It is well known that ATP binding is essential for MutS to switch between the mismatch bound and sliding clamp state[31] yet the molecular basis of this transition and how it is regulated remain unresolved. Moreover, it has been shown that once it is transformed into the sliding clamp conformation, MutS remains on the DNA for a long time before dissociating unless ssDNA is present [29,30]. It was hypothesized that ATP hydrolysis was suppressed in the clamp state, and the MutS dimer was not able to return to its open -post hydrolysis- conformation. However, recent observation that Taq MutS can hydrolyse ATP in the clamp [32] state suggests is not ATP that keeps the dimer closed, yet it is unclear what is the mechanism that traps MutS on the DNA.

How the DNA and ATP work together to regulate the conformational changes in MutS during DNA mismatch repair is not well understood. **Chapter 2** will address how the DNA molecule acts as a modulator of the ATP hydrolysis cycle in MutS. When MutS is scanning for mismatches, the DNA keeps the ATPase dimer open and prevents ATP hydrolysis. Once MutS has detected the mismatch and is transformed in the sliding clamp conformation, the does allow ATP hydrolysis. However, the DNA keeps the ATPase dimer closed and prevents release of MutS from the DNA. As the repair process continues, ssDNA is generated which allows the opening of the ATPase domains and the release of the MutS clamps from the DNA. Hence, different signals on DNA modulate the ATP hydrolysis cycle of MutS, allowing the initiation and termination of the mismatch reaction.

The MutL protein

While the MutS protein is responsible for sensing the presence of a mismatch and initiate or terminate the repair cascade, the MutL protein plays a role in activating the nicking and the unwinding of the mismatch containing strand [19,33,34]. The MutL homologues are dimers (homodimers in prokaryotes and heterodimers in eukaryotes), where every

monomer is composed of two domains connected by a flexible linker of ~30 amino acids (Fig. 6). The C-terminal domains form the primary dimerization interface, while the N-terminal domains contain ATP-binding sites that dimerize upon ATP binding [35]. Thanks to its flexible linkers and the ATP induced dimerization domains, the MutL protein is able to adopt multiple conformations and connect the mismatch repair activation signal given by the presence of MutS sliding clamps on the DNA, to downstream effector of the repair process.

First, MutL is recruited to the DNA by the interaction of one of its N-terminal domain with the MutS sliding clamp [17]. It was demonstrated that MutS and MutL can remain bound and diffuse along the DNA, or dissociate and travel on the DNA as individual proteins [36,37]. Indeed, the ATP induced dimerization of the N-terminal domains in MutL traps the DNA helix in the central channel between the N- and C- terminal domains. This way MutL can travel long distances on the DNA in search for the strand discrimination signal [38].

While the N-terminal domains of MutL are essential for the recruitment on the DNA and the formation of a mobile clamp, the C-terminal domains are important for the interaction with other proteins to continue the repair process. In *E. coli*, the C-terminal domains of MutL interact with the endonuclease MutH and the helicase UvrD, promoting the nicking and the resection of the mismatched DNA strand [39,40,18,19]. In other species including eukaryotes, the C-terminal MutL domains themselves, nick the mismatched DNA strand thanks to the interaction with the processivity clamp [41,42].

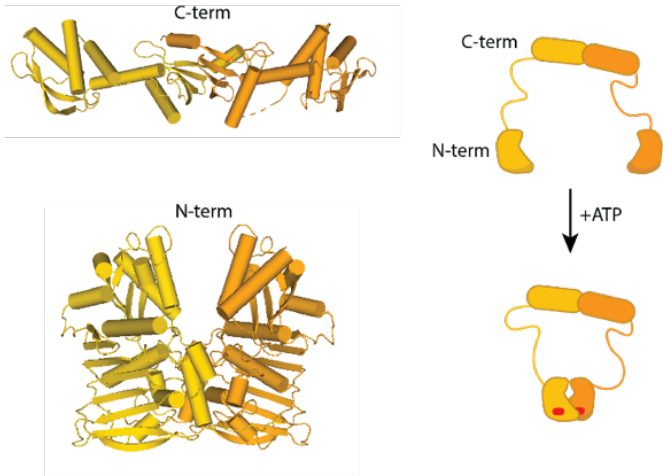

Figure 6. Structure of MutL | a, Crystal structure of *E. coli* MutL C- and N-terminal domains. **b**, Cartoon representation of the dimerization of MutL N-terminal domains induced by ATP binding.

The interplay between excision and resynthesis

During DNA mismatch repair in *E. coli*, removal of the mismatch containing strand is performed by the helicase UvrD that unwinds the DNA and by 5'-3' or 3'-5' exonucleases

that digest the unwound strand [43]. During excision, MutL acts as a processivity factor for the helicase UvrD by direct binding of the MutL C-terminal domains with the 2B domain of UvrD [44,40]. However, the increase of UvrD processivity is rather moderate, going from ~15 to ~40 unwound bases per binding event. Analysis of the activity, processivity and unwinding speed of UvrD in presence and absence of MutL suggests that strand separation is a discontinuous process which requires multiple turnover events. This poses a question about the interchange between excision and resynthesis: if the excision is a discontinuous process, then multiple times during the repair reaction 3' resected DNA ends will be present that are the substrate for any DNA polymerase. What stops DNA polymerases from binding to their natural substrate and prevents them from prematurely re-synthetising the ssDNA gap? Discoveries of the interaction between MutS, MutL and the β clamp suggested that the mismatch repair proteins could recruit the replisomal proteins at the end of DNA mismatch repair [45,46]. However, this does not explain how the resynthesis of the ssDNA gap is delayed until removal of the mismatch on the DNA.

Chapter 3 describes a new role of the mismatch repair protein MutL that prevents premature DNA synthesis during MMR, by binding to the 3' resected DNA ends and inhibiting DNA polymerases activity. Only when the mismatch is removed and all MMR proteins dissociate from the DNA, DNA polymerases can re-synthetise the ssDNA gap.

DNA replication as a novel target for antibiotics

Antibiotic resistance is a worldwide problem that is a threat to global health. The issue arises from the ability of bacteria to become resistant to antibiotics. Resistance is achieved through several mechanisms, such as clearance by membrane pumps, enzymatic inactivation of the antibiotic, or mutation of the target site to reduce affinity of the inhibitor [47]. As a result, bacteria have developed ways to overcome common treatments. Hence, finding new targets for the development of new antibiotics has become an essential task to combat the spreading of multidrug resistance, and improve the treatment of resistant strains. Current antibiotics can be divided into four main categories based on their target [48,49]: (i) inhibition of protein synthesis by targeting the ribosome, (ii) inhibition of RNA synthesis by targeting RNA polymerases, (iii) inhibition of RNA and DNA synthesis by targeting DNA gyrases and topoisomerases, and (iv) inhibition of cell membrane by targeting enzymes in the membrane biosynthetic pathway.

With the exception of gyrases and topoisomerases, there is no antibiotic targeting any of the DNA replication proteins [50]. However, DNA polymerases have emerged as an attractive potential target for antibacterial drug discovery, due to the structural differences between eukaryotic and prokaryotic replicative polymerases [51].

DNA polymerases can be classified into different families (A, B, C, D, X, Y) based on their sequence and structural organization. All bacterial replicative polymerases belong to the C-family DNA polymerases and are only found in the bacterial kingdom [52,53]. Eukaryotic replicative polymerases in contrast, belong to the B-family polymerases [54]. Finally, Archaeal replicative DNA polymerases are from both B- and D-family [55]. The remaining families are conserved throughout different species with different roles during DNA replication or participate in other DNA repair pathways and Translesion DNA synthesis. All DNA polymerase families share a common overall architecture that can be associated to a right hand which threads the DNA between its palm, fingers and thumb domain [56] (Fig. 7b). The main differences between Eukaryotic and Prokaryotic replicative DNA polymerases lie in the structural arrangements of the catalytic residues in the palm domain, in the size of the finger domains, and in the presence of an N-terminal

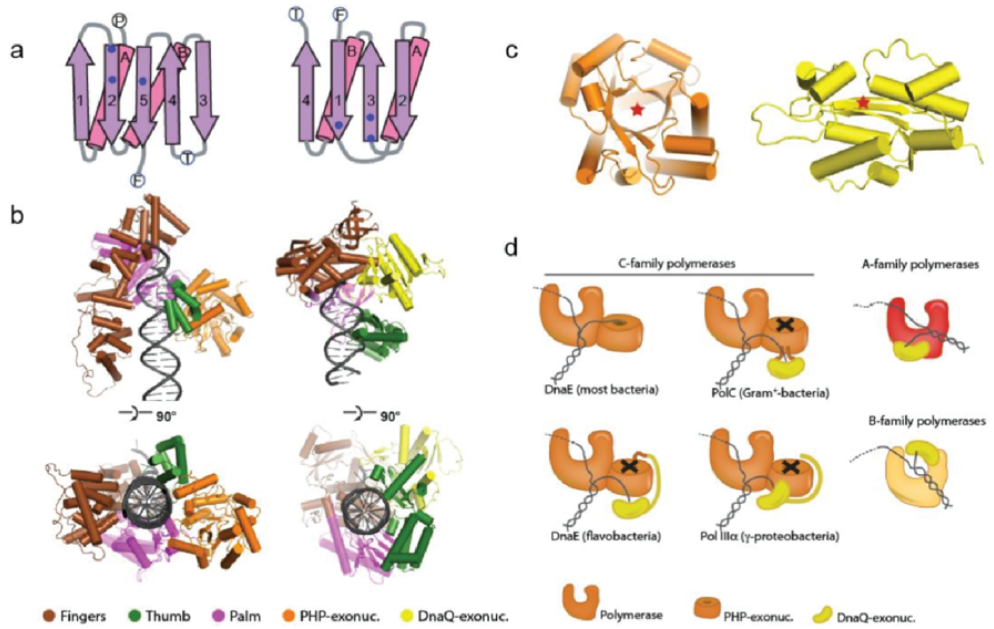

Figure 7 | Unique features of bacterial replicative DNA polymerases. a,. The active site of the bacterial replicase consists of a 5-stranded parallel/antiparallel β-sheet. The active site of the eukaryotic replicative DNA polymerase contains 4-stranded, all antiparallelβ-sheet. Catalytic site residues, marked with blue dots, are located at opposite site of theβ-sheet. **b,** Comparison of the bacterial replicative DNA polymerase (left, PDB: 5FKW, EMDB: 3202) and the eukaryotic replicative DNA polymerase δ (right, PDB: 3IAY). Domains are indicated with different colours. The C-terminal tail of bacterial polymerase is omitted for clarity. **c**, The Polymerase and Histidinol Phosphatase (PHP)-exonuclease (left) and DnaQ-type exonuclease (right) are structurally distinct. The red star indicates the location of active site (**d**). Schematic representation of the different polymerase families, highlighting the different locations of the PHP- and DnaQ-exonuclease domains. Black cross marks inactive PHP-exonucleases. Figure taken from Santos, J. A. & Lamers, M. H. *Antibiotics* 2020 [51].

"Polymeraseand Histidinol Phosphatase" (PHP) domain that is unique to bacteria, and a few isolated fungi [57] (Fig. 7a-c).

Finally, the 5' to 3' exonucleases which correct mismatches during DNA replication are also diverse among the different domains of life. Most bacterial species, use the intrinsic exonuclease activity of the PHP domain in the polymerase chain, although in some cases an additional exonuclease domain (DnaQ) is present which replaces the role of the PHP domain [58,59]. A DnaQ-type exonuclease is also present in the Eukaryotic A and B-family DNA polymerases, however, its location within the polymerase structure is different compared to the C-family.

The many structural differences among replicative DNA polymerases, makes them an attractive target for developing new antibiotics which could target selectively bacterial DNA replication, and not its eukaryotic counterpart.

Novel antibiotics targeting bacterial replicative DNA polymerases

Several inhibitors of bacterial replicative DNA polymerases have been discovered in the last two decades. Among those, Ibezapolstat that inhibits the replicative polymerase of *C. difficile* [60] is currently in Phase 2 clinical trials. Another example is the compound Nargenicin, a recently discovered natural product isolated from the bacterium *Nocardia devorans*, which is found to be potent inhibitor of the DnaE polymerase from Gram-positive bacteria such as *S. aureus* and *Micrococcus luteus* [61]. **Chapter 4** describes the bactericidal activity of the compound Nargenicin against *Mycobacterium tuberculosis* together with a biochemical and structural characterization of its mode of action against the *Mtb* replicative polymerase DnaE1.

An air-blades based plunge freezer for improved cryo-EM grid preparation

The results presented in the chapters of this thesis are obtained with different biochemical techniques combined with structural analysis using cryo-EM. Cryo-EM is a technique used to image biomolecules such as proteins, embedded in a thin layer of vitreous water. The sample immobilized into the vitreous water is exposed to a transmission electron beam which generates 2D projections. The 2D images of the sample are than analysed and used to reconstruct a 3D volume. Thanks to the technological advances of the last decades, which brought developments in microscopes, direct electron detectors, and sophisticated data-processing algorithms, this technique has become a powerful tool to obtain high resolution protein structures.

While the hardware and software for cryo-EM analysis have become particularly reliable, the bottleneck of this technique remains sample preparation. The common way is to apply a few microliters of sample onto a cryo-EM grid, blot away the excess of liquid

buffer using filter paper, and plunge the grid into liquid ethane at ~ -182 degrees. In this process called plunge freezing, the particles ideally are evenly distributed and randomly oriented into the thin layer of vitrified water [62]. Since its invention by Jacques Dubochet in 1988, many protein structures have been solved preparing the sample using the plunge freezing method, and this technique is still the most used one to prepare cryo-EM samples. However, there are a series of problems such as aggregation, denaturation, dissociation of multiprotein complexes and preferred orientation, which represent a barrier to high-resolution imaging of some plunge frozen proteins. In most cases, these problems arise from the hydrophobic patches of proteins that interact with the air-water interface generated during the blotting and plunge freezing procedure [63]. The issue arises from the time passing

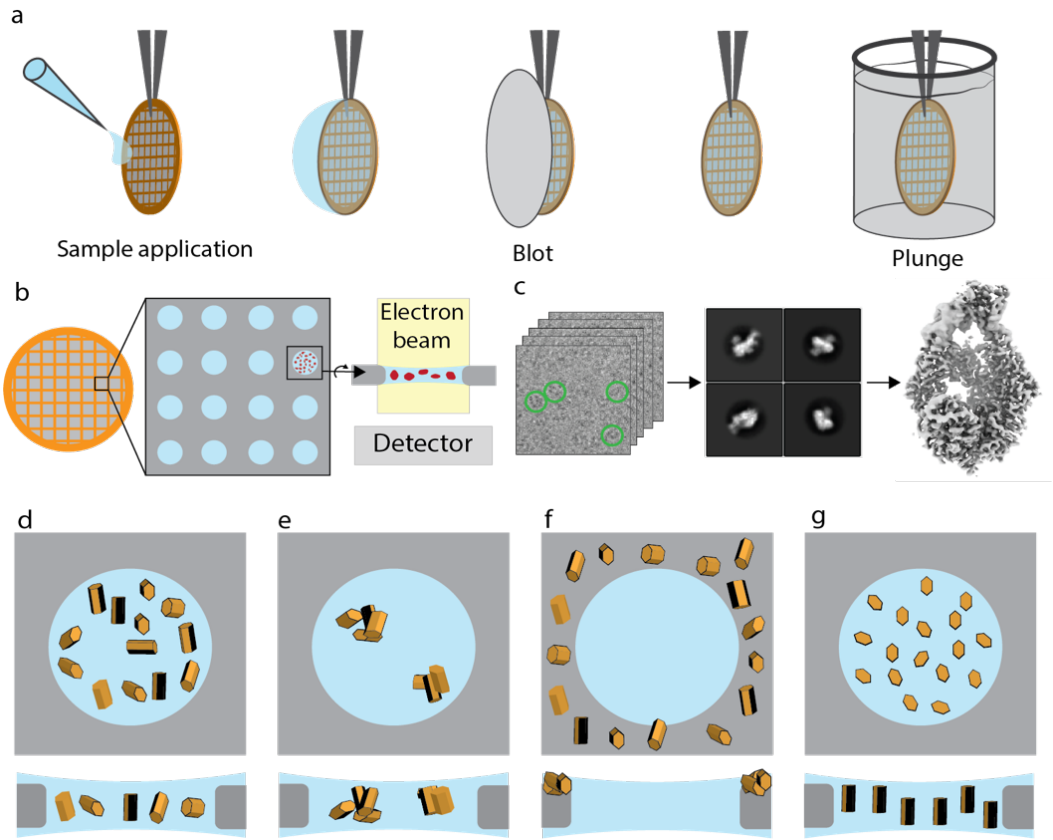

Figure 8. Cryo-EM structure determination pipeline | a, Graphical representation of the blotting/plunge freezing procedure. **b**, Illustration of the vitrified water layer of the cryo-EM grid. The sample is vitrified into the holes of the carbon support, and imaged with an electron beam. **c**, Schematic workflow of single particle structure determination. First, individual 2D projections are extracted from the original images and classified in 2D. After that, they are used to reconstruct a high resolution 3D volume. **d**, Ideal distribution of the sample into the vitreous water layer. The particles are homogeneously distributed and have multiple orientation. **e**, Aggregation, (**f**) water exclusion and (**g**) preferred orientation of the particles in the water layer.

between blotting and plunge freezing, which in common instruments is in the range of 1 to 5 seconds. During this interval of time, proteins which are normally thumbling in solution have the chance to accumulate and get trapped at the air-water interface.

Chapter 5 of this thesis describes the development of a new cryo-EM grid preparation device which allows formation of a thin layer of protein sample embedded in liquid buffer in the milliseconds time scale. By reducing the time that proteins can interact with the air-water interface we open up the possibility to prepare better samples of proteins that have a tendency to accumulate in a preferred orientation at the air-water interface when prepared with common plunge freezers. Moreover, the new plunge freezer offers the possibility to rapidly mix two samples on the cryo-EM grid, which together with the fast-freezing capabilities, allows the study of conformational changes which occur in the millisecond time scale.

Thesis outline

The scientific research of this thesis is built on three main categories: 1) fundamental research, 2) drug discovery and 3) method development. All three categories fall in the context of structural biology research, particularly focused on cryo electron microscopy and high-fidelity DNA replication and repair.

Chapter 2 and **3** provide new insights in the molecular mechanisms of the DNA mismatch repair enzymes MutS and MutL using a combination of cryo-EM and biochemistry. **Chapter 4** focuses on a novel structure of the replicative DNA polymerase from *M. tuberculosis* bound to an inhibitory compound, which highlights the importance of cryo-EM in drug discovery research. **Chapter 5** describes the development of a new instrument to that prepares better samples for cryo-EM and is able to perform time-resolved cryo electron microscopy.

References

1. WATSON, J. . C. F. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature 171, 737–738 (1953). doi:https://doi.org/10.1038/171737a0
2. O'Donnell, M., Langston, L. & Stillman, B. Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb. Perspect. Biol.* **5**, (2013).
3. Steitz, T. A. DNA polymerases: Structural diversity and common mechanisms. *Journal of Biological Chemistry* **274**, 17395–17398 (1999).
4. Johnson, A. & O'Donnell, M. Cellular DNA replicases: Components and dynamics at the replication fork. *Annual Review of Biochemistry* **74**, 283–315 (2005).
5. Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K. & Sugino, A. *MECHANISM OF DNA CHAIN GROWTH, I. POSSIBLE DISCONTINUITY AND UNUSUAL SECONDARY STRUCTURE OF NEWLY SYNTHESIZED CHAINS*.
6. Lindahl T. The Croonian Lecture, 1996: endogenous damage to DNA. Philos Trans R Soc Lond B Biol Sci. 1996 Nov 29;351(1347):1529-38. doi: 10.1098/rstb.1996.0139. PMID: 8962440.
7. Kunkel, T. A. DNA Replication Fidelity. *Journal of Biological Chemistry* **279**, 16895–16898 (2004).
8. Johnson, K. A. The kinetic and chemical mechanism of high-fidelity DNA polymerases. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1804**, 1041–1048 (2010).
9. Fernandez-Leiro, R., Conrad, J., Scheres, S. H. W. & Lamers, M. H. Cryo-EM structures of the E. Coli replicative DNA polymerase reveal its dynamic interactions with the DNA sliding clamp, exonuclease and τ. *Elife* **4**, 1–16 (2015).
10. Fernandez-Leiro, R. *et al.* Self-correcting mismatches during high-fidelity DNA replication. *Nat. Struct. Mol. Biol.* **24**, 140–143 (2017).
11. Dodd, T. *et al.* Polymerization and editing modes of a high-fidelity DNA polymerase are linked by a well-defined path. *Nat. Commun.* **11**, (2020).
12. Modrich, P. & Lahue, R. *MISMATCH REPAIR IN REPLICATION FIDELITY, GENETIC RECOMBINATION, AND CANCER BIOLOGY*. *Annu Rev. Biochem. 19%* **65**, (1996).
13. Li, Z., Pearlman, A. H. & Hsieh, P. DNA mismatch repair and the DNA damage response. *DNA Repair* **38**, 94–101 (2016).
14. Jiricny, J. Postreplicative mismatch repair. *Cold Spring Harbor Perspectives in Biology* **5**, 1–23 (2013).
15. Lynch_JMG_1999.

16. Shin-San Su & Modrich, P. Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5057–5061 (1986).
17. Groothuizen, F. S. *et al.* MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. *Elife* **4**, 1–24 (2015).
18. Welsh, K. M., Lu, A. L., Clark, S. & Modrich, P. Isolation and characterization of the Escherichia coli mutH gene product. *J. Biol. Chem.* **262**, 15624–15629 (1987).
19. Junop, M. S., Yang, W., Funchain, P., Clendenin, W. & Miller, J. H. In vitro and in vivo studies of MutS, MutL and MutH mutants: Correlation of mismatch repair and DNA recombination. *DNA Repair (Amst).* **2**, 387–405 (2003).
20. Kunkel, T. A. & Erie, D. A. DNA mismatch repair. *Annual Review of Biochemistry* **74**, 681–710 (2005).
21. Dzantiev, L. *et al.* A defined human system that supports bidirectional mismatch-provoked excision. *Mol. Cell* **15**, 31–41 (2004).
22. Lahue, R. S., Au, K. G. & Modrich, P. DNA mismatch correction in a defined system. *Science (80-.).* **245**, 160–164 (1989).
23. Hopfner, K. P. & Tainer, J. A. Rad50/SMC proteins and ABC transporters: Unifying concepts from high-resolution structures. *Curr. Opin. Struct. Biol.* **13**, 249–255 (2003).
24. Melby, T. E., Ciampaglio, C. N., Briscoe, G. & Erickson, H. P. *The Symmetrical Structure of Structural Maintenance of Chromosomes (SMC) and MukB Proteins: Long, Antiparallel Coiled Coils, Folded at a Flexible Hinge*. *The Journal of Cell Biology* **142**, (1998).
25. Moody, J. E. & Thomas, P. J. Nucleotide binding domain interactions during the mechanochemical reaction cycle of ATP-binding cassette transporters. *Journal of Bioenergetics and Biomembranes* **37**, 475–479 (2005).
26. Lamers, M. H. *et al.* The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature* **407**, 711–7 (2000).
27. Obmolova, G., Ban, C., Hsieh, P. & Yang, W. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature* **407**, 703–710 (2000).
28. Fernandez-Leiro, R. *et al.* The selection process of licensing a DNA mismatch for repair. *Nat. Struct. Mol. Biol.* **28**, 373–381 (2021).
29. Heo, S. D., Cho, M., Ku, J. K. & Ban, C. Steady-state ATPase activity of E. coli MutS modulated by its dissociation from heteroduplex DNA. *Biochem. Biophys. Res. Commun.* **364**, 264–269 (2007).
30. Jeong, C. *et al.* MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat. Struct. Mol. Biol.* **18**, 379–385 (2011).
31. Hingorani, M. M. Mismatch binding, ADP-ATP exchange and intramolecular signaling during mismatch repair. *DNA Repair (Amst).* **38**, 24–31 (2016).

32. Hao, P. *et al.* Recurrent mismatch binding by MutS mobile clamps on DNA localizes repair complexes nearby. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 17775–17784 (2020).
33. Hall, M. C. & Matson, S. W. The Escherichia coli MutL protein physically interacts with MutH and stimulates the MutH-associated endonuclease activity. *J. Biol. Chem.* **274**, 1306–1312 (1999).
34. Yamaguchi, M., Dao, V. & Modrich, P. MutS and MutL activate DNA helicase II in a mismatch-dependent manner. *J. Biol. Chem.* (1998). doi:10.1074/jbc.273.15.9197
35. Ban, C., Junop, M. & Yang, W. Transformation of MutL by ATP Binding and Hydrolysis. *Cell* **97**, 85–97 (1999).
36. Liu, J. *et al.* Cascading MutS and MutL sliding clamps control DNA diffusion to activate mismatch repair. *Nature* **539**, 583–587 (2016).
37. Lee, J. B. *et al.* Single-molecule views of MutS on mismatched DNA. *DNA Repair (Amst).* **20**, 82–93 (2014).
38. Mardenborough, Y. S. N. *et al.* The unstructured linker arms of MutL enable GATC site incision beyond roadblocks during initiation of DNA mismatch repair. *Nucleic Acids Res.* **47**, 11667–11680 (2019).
39. Ordabayev, Y. A., Nguyen, B., Niedziela-Majka, A. & Lohman, T. M. Regulation of UvrD Helicase Activity by MutL. *J. Mol. Biol.* **430**, 4260–4274 (2018).
40. Matson, S. W. & Robertson, A. B. The UvrD helicase and its modulation by the mismatch repair protein MutL. *Nucleic Acids Res.* **34**, 4089–4097 (2006).
41. Pluciennik, A. *et al.* PCNA function in the activation and strand direction of MutL endonuclease in mismatch repair. *Proc. Natl. Acad. Sci.* **107**, 16066–16071 (2010).
42. Pillon, M. C., Miller, J. H. & Guarné, A. The endonuclease domain of MutL interacts with the β sliding clamp. *DNA Repair (Amst).* **10**, 87–93 (2011).
43. Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S. T. & Modrich, P. In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 6765–70 (2001).
44. Ordabayev, Y. A., Nguyen, B., Kozlov, A. G., Jia, H. & Lohman, T. M. UvrD helicase activation by MutL involves rotation of its 2B subdomain. *Proc. Natl. Acad. Sci.* **116**, 201905513 (2019).
45. López De Saro, F. J., Marinus, M. G., Modrich, P. & O'Donnell, M. The β sliding clamp binds to multiple sites within MutL and MutS. *J. Biol. Chem.* **281**, 14340–14349 (2006).
46. Pluciennik, A., Burdett, V., Lukianova, O., O'Donnell, M. & Modrich, P. Involvement of the β clamp in methyl-directed mismatch repair in vitro. *J. Biol. Chem.* **284**, 32782–32791 (2009).
47. C Reygaert, W. An overview of the antimicrobial resistance mechanisms of

bacteria. *AIMS Microbiol.* **4**, 482–501 (2018).
48. Bbosa, G. S., Mwebaza, N., Odda, J., Kyegombe, D. B. & Ntale, M. Antibiotics/antibacterial drug use, their marketing and promotion during the post-antibiotic golden age and their role in emergence of bacterial resistance. *Health (Irvine. Calif).* **06**, 410–425 (2014).
49. Kapoor, G., Saigal, S. & Elongavan, A. Action and resistance mechanisms of antibiotics: A guide for clinicians. *Journal of Anaesthesiology Clinical Pharmacology* **33**, 300–305 (2017).
50. Sanyal, G. & Doig, P. Bacterial DNA replication enzymes as targets for antibacterial drug discovery. *Expert Opinion on Drug Discovery* **7**, 327–339 (2012).
51. Santos, J. A. & Lamers, M. H. Novel antibiotics targeting bacterial replicative dna polymerases. *Antibiotics* **9**, 1–14 (2020).
52. Ito, J. & Braithwaite, D. K. *Compilation and alignment of DNA polymerase sequences*. *Nucleic Acids Research* **19**, (Oxford University Press, 1991).
53. Delarue, M., Poch, O., Tordo, N., Moras, D. & Argos, P. *An attempt to unify the structure of polymerases*. (Protein Engineering, 1990).
54. Doublié, S. & Zahn, K. E. Structural insights into eukaryotic DNA replication. *Frontiers in Microbiology* **5**, (2014).
55. Greenough, L., Menin, J. F., Desai, N. S., Kelman, Z. & Gardner, A. F. Characterization of Family D DNA polymerase from Thermococcus sp. 9°N. *Extremophiles* **18**, 653–664 (2014).
56. Bebenek, K. & Kunkel, T. A. Functions of DNA polymerases. *Adv. Protein Chem.* **69**, 137–165 (2004).
57. Aravind, L. & Koonin, E. V. *Phosphoesterase domains associated with DNA polymerases of diverse origins*. *Nucleic Acids Research* **26**, (1998).
58. Rock, J. M. *et al.* DNA replication fidelity in Mycobacterium tuberculosis is mediated by an ancestral prokaryotic proofreader. *Nat. Genet.* **47**, 677–681 (2015).
59. Timinskas, K., Balvočiute, M., Timinskas, A. & Venclovas, Č. Comprehensive analysis of DNA polymerase III α subunits and their homologs in bacterial genomes. *Nucleic Acids Res.* **42**, 1393–1413 (2014).
60. Xu, W. C., Silverman, M. H., Yu, X. Y., Wright, G. & Brown, N. Discovery and development of DNA polymerase IIIC inhibitors to treat Gram-positive infections. *Bioorganic and Medicinal Chemistry* **27**, 3209–3217 (2019).
61. Painter, R. E. *et al.* Elucidation of DnaE as the antibacterial target of the natural product, nargenicin. *Chem. Biol.* **22**, 1362–1373 (2015).
62. Schultz, P. Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* **21**, 129–228 (1988).
63. Taylor, K. A. & Glaeser, R. M. Retrospective on the early development of

cryoelectron microscopy of macromolecules and a prospective on opportunities for the future. *J. Struct. Biol.* **163**, 214–223 (2008).