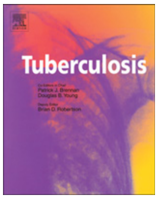
Tuberculosis 91 (2011) 533–543

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: http://intl.elsevierhealth.com/journals/tube



REVIEW

Base excision and nucleotide excision repair pathways in mycobacteria

Krishna Kurthkoti [a], Umesh Varshney [a,b,*]

[a] Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560012, India
[b] Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560064, India

A R T I C L E I N F O

Article history:
Received 15 April 2011
Received in revised form
1 June 2011
Accepted 12 June 2011

Keywords:
Hypoxia
BER
NER
Ung
UdgB

SUMMARY

About a third of the human population is estimated to be infected with *Mycobacterium tuberculosis*. The bacterium displays an excellent adaptability to survive within the host macrophages. As the reactive environment of macrophages is capable of inducing DNA damage, the ability of the pathogen to safeguard its DNA against the damage is of paramount significance for its survival within the host. Analysis of the genome sequence has provided important insights into the DNA repair machinery of the pathogen, and the studies on DNA repair in mycobacteria have gained momentum in the past few years. The studies have revealed considerable differences in the mycobacterial DNA repair machinery when compared with those of the other bacteria. This review article focuses especially on the aspects of base excision, and nucleotide excision repair pathways in mycobacteria.

1. DNA repair in mycobacteria and the scope of the article

Cellular DNA is continually exposed to reactive radicals of intracellular or extracellular origin often leading to irreversible changes in the genetic blue print. If left unattended, such changes lead to mutations. Some of the well studied intracellular DNA damaging agents such as the reactive oxygen species (ROS) and reactive nitrogen intermediates (RNI) are generated during cellular metabolism. As the pathogenic bacteria have to establish and perpetuate within the host, the study of DNA repair in these bacteria forms an interesting model system. In fact, *Mycobacterium tuberculosis* which preferentially infects macrophages, the host's first line of defence in the immune system (which produce ROS and RNI), should possess a robust DNA repair machinery to ensure maintenance of the integrity of its genome to survive within the host and to establish a successful infection.[1–3] While as a part of its survival strategy the bacterium possesses an elaborate mechanism to detoxify the ROS and RNI,[4] any residual levels of these agents can be detrimental to the genomic integrity. A detailed knowledge of DNA repair mechanisms in pathogens may offer a basis to design therapies or prophylactics against them.

Availability of the genome sequences of some of the important human pathogens such as *Haemophilus influenzae*, *Helicobacter pylori*, *M. tuberculosis* has offered a new approach to advance our knowledge of DNA repair mechanisms in these organisms.[5–7] There is a growing body of evidence suggesting that pathogens often behave differently from *Escherichia coli* and *Bacillus subtilis* models in having a distinct representation of DNA repair enzymes.[8] For example, in *Campylobacter jejuni* that causes gastrointestinal infections, enzymes of direct repair pathway and base excision repair pathway are under-represented. In *Streptococcus pneumoniae* that colonizes upper respiratory tract, not only the direct and base excision repair proteins but also the members of recombinational repair are under-represented.

The genus mycobacterium consists of members that continue to have a great impact on the human society as important pathogens.[9] The prominent members include *M. tuberculosis*, the causative agent of tuberculosis and *Mycobacterium leprae* that causes leprosy. The genome of *M. tuberculosis* (~4.2 MB) has a high $G + C$ content (~66%). Thus, considering a highly reactive environment of macrophage, there is a greater risk of cytosine deamination to uracil and guanine oxidation to 7,8-dihydro-8-oxoguanine (8-oxoG) or other oxidized intermediates. The *M. tuberculosis* genome sequence revealed,[7,10,11] that it contains many of the base excision and nucleotide excision repair genes found in *E. coli*. However, it lacks homologues of mismatch repair pathway genes.[7,10–12] Interestingly, the bacterium possesses ERCC3 (XPB) and Mpg enzymes, which were until then recognized exclusively in the mammalian cells. The bacterium also encodes homologues of the

---

* Corresponding author. Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560012, India. Tel.: +91 80 2293 2686; fax: +91 80 2360 2697.
E-mail addresses: varshney@mcbl.iisc.ernet.in, uvarshney@gmail.com (U. Varshney).

1472-9792/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tube.2011.06.005

nonhomologous end joining pathway (NHEJ), which has been established as a major pathway for repairing double strand DNA breaks in eukaryotes. Proteins such as Ku 70, Ku 80, DNA ligase D that function in this pathway have been identified in *M. tuberculosis* and other bacterial members.[13,14] Expression of Ku (Rv0937c) increases in *M. tuberculosis* in the infected human samples.[15] Recently, two members of the Y family of polymerases (which perform translesion DNA synthesis) from *M. tuberculosis* DinP (Rv3056) and DinX (Rv1537), referred to as DinB2 and DinB1, respectively were also reported.[16]

DNA repair studies in mycobacteria make extensive use of *Mycobacterium smegmatis*, a non-pathogenic and a relatively fast growing member of the group, as a model. Similar to *M. tuberculosis*, *M. smegmatis* has a G + C rich genome (~67%), lacks mismatch repair genes, and the DNA repair proteins in *M. tuberculosis* and *M. smegmatis* share a high degree of similarity.[10,11] In addition, the availability of relatively more tractable genetic methods for *M. smegmatis* render it useful to obtain the first information report of the distinct features of DNA repair in mycobacteria.

As the members of mycobacteria do not encode homologues of the mismatch repair enzymes,[11] the base excision repair (BER) and the nucleotide excision repair (NER) pathways may play a major role in maintaining the integrity of DNA in these bacteria. Furthermore, because of the G + C richness of the genome, it is not surprising that there is more emphasis on enzymes that are involved in the excision of uracil and the oxidized guanine bases. In the recent past, several reviews[11,17] have addressed general aspects of DNA repair in mycobacteria. More recently, role of DNA repair in *M. tuberculosis* pathogenesis has also been reviewed.[18] In this article, we restrict ourselves to the salient features of BER and NER pathways in mycobacteria. A list of various proteins involved in these pathways is provided in Table 1.

2. BER pathway

The BER pathway involves excision of modified bases in DNA by a class of enzymes called DNA glycosylases which are highly evolved to specifically recognize the modified/damaged base(s) in the context of DNA backbone to initiate BER. A number of DNA glycosylases are monofunctional enzymes which hydrolyze the N-glycosidic bond between the base and the sugar and result in the formation of abasic (AP) sites.[19] However, the bifunctional DNA glycosylases, in addition to hydrolyzing the N-glycosidic bond, continue to cleave the phosphodiester backbone at the AP sites by their lyase activity. Presence of unprocessed AP sites can be more detrimental than the modified bases as they impede the essential cellular processes such as replication and transcription. The accumulation of AP sites in DNA is both mutagenic and cytotoxic.[20–22] A class of enzymes called AP endonucleases (APE) act on AP sites to cleave the phosphodiester bond resulting in a nick in the DNA. The nicks so generated require further processing before the ends of the DNA can be used by the downstream processing enzymes. A major pathway for processing of the AP sites in *E. coli* employs exonuclease III or endonuclease IV to hydrolyse the phosphodiester bond 5′ to the abasic deoxyribose sugar to generate a 3′ hydroxyl, and a 5′ deoxyribose ends. A deoxyribosephosphodiesterase (dRpase) activity (*e. g.* RecJ) is then utilized to cleave the deoxyribose and generate a 5′ phosphate end.[23,24] The 5′ phosphate end can also result from β-elimination of the deoxyribose, a reaction promoted by Fpg.[25] The single nucleotide gap surrounded by 3′ hydroxyl and 5′ phosphate ends is filled in by DNA polymerase I and sealed by DNA ligase to restore the original sequence.[23,24] The lyase activity of the bifunctional DNA glycosylases or the AP lyases (*e. g.* endonuclease III) may also process the AP sites by cleaving 3′ to the abasic sugar to generate a 5′ phosphate end. Although such a reaction bypasses the requirement of dRpase, it requires further processing of the 3′ end (*e. g.* by exonuclease III or endouclease IV) to convert the 3′ unsaturated aldehyde (indicated by 3′ PA in Figure 1) to a 3′ hydroxyl end to serve as primer for the DNA polymerase.[26,27] In an alternate pathway, even though the 5′-3′ exonuclease activity of DNA polymerase I is unable to remove the 5′ deoxyribose, alternate DNA polymerase(s) (*e. g.* as known in eukaryotes) may carry out the fill in reaction (from the 3′ hydroxyl end) by replacement synthesis. The 5′ flank containing the deoxyribose residue may then be removed as a part of a DNA oligomer by a structure specific endonuclease activity (*e. g.* FEN-1) followed by sealing of the nick (containing 3′ OH and 5′ PO_4) by DNA ligase.[24] Thus, while the predominant BER pathway results in single nucleotide repair patch (Figure 1, *left panel*), the alternate pathway would lead to a multiple nucleotide repair patch (Figure 1, *right panel*). A recent study shows a repair size of more than one nucleotide in mycobacteria.[28] However, it would be interesting to investigate the mechanism of this longer size patch repair in mycobacteria.

2.1. DNA glycosylases involved in uracil excision in mycobacteria

Uracil DNA glycosylases (UDGs) excise uracil formed either due to deamination of cytosine or misincorporation of dUTP (as dUMP against A) during replication.[29] Inability to repair uracil arising from cytosine deamination leads to G:C to A:T transitions. Although, direct incorporation of dUMP is not mutagenic in itself, it may impact protein binding to DNA. The incorporation of dUMP in genome is kept to a minimum by dUTPase (encoded by *dut*) which hydrolyses dUTP to minimize the cellular pools of dUTP and thereby reducing the chances of its misincorporation by DNA polymerases.[30,31] Loss of dUTPase leads to increased incorporation of uracil in DNA and may be lethal or result in elevated mutation rates.[32,33] *M. tuberculosis* and the other members of mycobacteria have been shown to contain *dut*.[11] *M. tuberculosis* Dut (Rv2697c) has been shown to be a bifunctional protein possessing dCTPase and dUTPase activities,[34] and the *dut* gene has been identified as an essential gene in a transposon mutagenesis screen.[35]

A number of UDG activities have been identified in various organisms. Of these, the Ung proteins (family 1 UDGs) are the most conserved, highly efficient and by far the best characterized UDGs. The family 1 UDGs excise uracil from both the single-stranded and double-stranded DNAs, and possess two highly conserved amino acid sequence motifs, A and B (GQDPY and HPSPLS, respectively).[36–39] The Ung proteins are potently inhibited by a *B. subtilis* phage protein, Ugi.[40] The family 2 UDGs (also known as dsUDGs, MUG or DUG) which excise uracil or thymine from G:U/T mismatches in dsDNA, possess GINPG and MPSSSAR as motifs A and B, respectively.[41] The single stranded DNA selective monofunctional UDGs (SMUG, family 3) possess GMNPG and HPSPRNP as motifs A and B, respectively.[42] SMUG also acts on dsDNA but requires AP endonuclease for the product release. Two more UDGs, UdgA (family 4) containing GEAPG and HPAAVLR, and UdgB (family 5) having GLAPA and HPSPLNV as motifs A and B, respectively have been characterized.[43–46].

Mycobacteria possess family 1 and family 5 UDGs (Table 1). Family 1 UDG from *M. smegmatis*,[47] though similar to *Eco*Ung in its activities, possesses certain novel characteristics. For example, compared to *Eco*Ung, *Msm*Ung is more efficient in excising uracils from the loop substrates. This property may be of particular significance in mycobacteria because of high G + C contents of its genome, which may lead to formation of more stable looped DNA structures. The three dimensional structure of *Mtu*Ung with Ugi has been determined.[48] The structure of the central core of *Mtu*Ung is similar to the Ung from other sources. However, its N- and C-terminal tails show variability and its DNA-binding region is rich in arginine residues. Whether these structural features are