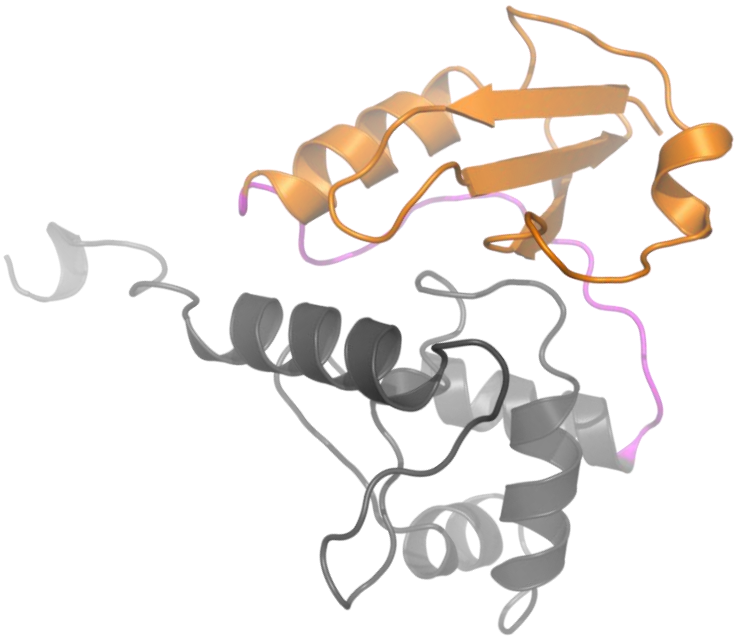
Università degli Studi del Piemonte Orientale
"Amedeo Avogadro"

Dipartimento di Scienze del Farmaco

Dottorato di Ricerca in Biotecnologie Farmaceutiche ed Alimentari
XXVI ciclo a.a. 2010-2013

BIOCHEMICAL AND STRUCTURAL STUDIES OF THE *Mycobacterium tuberculosis* O^6-METHYLGUANINE METHYLTRANSFERASE AND MUTATED VARIANTS



Riccardo Miggiano

Università degli Studi del Piemonte Orientale
"Amedeo Avogadro"

Dipartimento di Scienze del Farmaco

Dottorato di Ricerca in Biotecnologie Farmaceutiche ed Alimentari
XXVI ciclo a.a. 2010-2013

BIOCHEMICAL AND STRUCTURAL STUDIES OF THE *Mycobacterium tuberculosis* O^6-METHYLGUANINE METHYLTRANSFERASE AND MUTATED VARIANTS

Riccardo Miggiano

Supervised by
Dott.ssa Franca Rossi
Prof. Menico Rizzi

PhD program co-ordinator Prof. Menico Rizzi

A Mattia, Martina, Gioia e Jenny

Contents

Tuberculosis

Tuberculosis (TB) is a global public health emergency ranking as a leading cause of death from a single infectious agent worldwide, second only to Acquired Immune Deficiency Syndrome (AIDS).

According to a World Health Organization report, in 2012 there were almost 9 million new TB cases and 1.3 million TB deaths. Most of the estimated number of cases in 2012 occurred in Asia (58%) and Africa (27%); a smaller proportions of cases occurred in the Eastern Mediterranean Region (8%), the European Region (4%) and the Region of the Americas (3%) (Figure 1) (1).

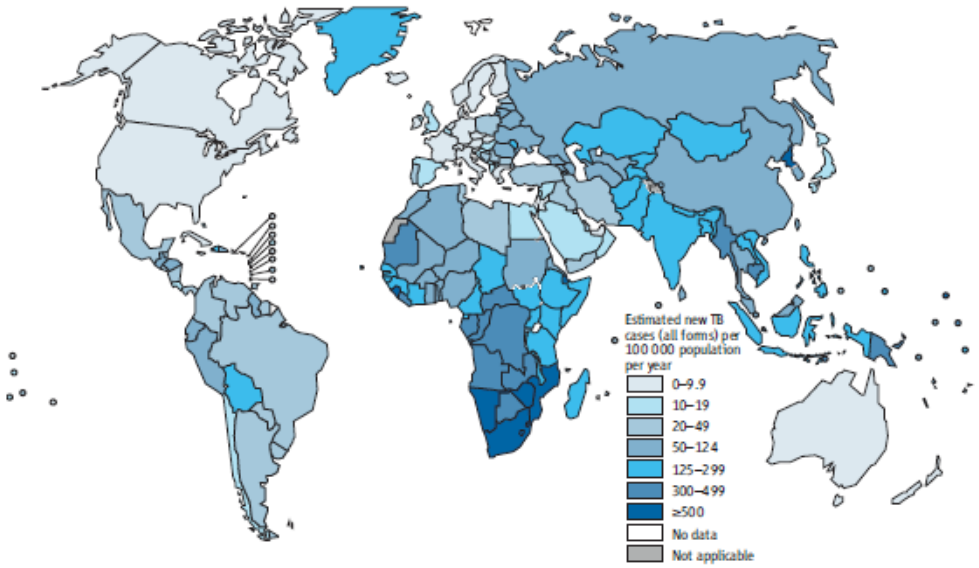


Figure 1- ***Estimated TB incident rates, 2012 (1).***

The majority of TB cases and deaths occur in underdeveloped countries, mainly due to less health care access as well as exposure to unhealthy living condition, malnutrition, HIV infection, homelessness, alcohol misuse and drug consumption (2).

TB is an airborne disease caused by *Mycobacterium tuberculosis* (MTB), the tubercle bacillus; although MTB can infect almost any tissue or organ of the body, most commonly the disease results in a chronic, spreading lung infection that, in untreated cases, leads to progressive destruction of the lung tissue (3).

At present anti-mycobacterial chemotherapy consists in a combination of different drugs; however, treatments outcomes are far from satisfactory. Long-term therapy makes patient's compliance one of the most difficult issue to overcome for chemotherapy success. The coupling of noncompliance with the nature of MTB's infectious process itself, resulted in widespread emergence of multi-drug and extensively-drug resistant MTB strains (MDR-TB and XDR-TB, respectively) (4). Treatment of MDR-TB and XDR-TB – there were approximately 0.5 million cases in 2012 – is more challenging, requiring the use of second-line drugs that are more expensive, cause more severe side effects, and must be taken for up to two years; moreover, healing rates for MDR-TB are lower: overall only 48% of patients were successfully treated (1). TB new face and the global scourge of MDR-TB and XDR-TB is reaching epidemic proportions.

Mycobacterium tuberculosis

Domain: Bacteria; Phylum: Actinobacteria; Class: Actinobacteria; Order: Actinomycetales; family: Mycobacteriaceae; Genus: Mycobacterium

Mycobacterium tuberculosis forms a complex with other closely related bacteria (MTB complex) that includes several other obligate human pathogens, most importantly *M. leprae* (the etiologic agent of leprosy), *M. africanum* (common cause of TB in West-African countries characterized by a lower pathogenicity with respect to MTB), and *M. bovis* (the causative agent of TB in a range of animal species, mainly in cattle). Similarly to the other members of this *genus*, MTB is a large, non-motile, rod-shaped bacterium distantly related to the Actinomycetes. The bacilli are 2-4 μm in length and 0.2-0.5 μm in width. As an obligate aerobe MTB grows most successfully in highly oxygenated tissues such as the lungs and it is characterized by a slow growth rate (i.e. generation time of 12 to 18 hours) mainly related to the tough cell wall that resists to the permeation of nutrients into the cell and inhibits waste products to be excreted out of the cell. This peculiar physiological characteristic contribute to mycobacterial virulence. The bacterium is a facultative intracellular parasite that does not seem to fit the Gram-positive category from an empirical standpoint (in general, it does not retain well the crystal violet stain). For this reason it is classified as an acid-fast Gram-positive bacterium due to the lack of an outer cell

membrane and its unique feature to be impermeable by certain dyes and stains. Despite this, once stained, acid-fast bacteria will retain dyes when heated and treated with acidified organic compounds. The structure of MTB external envelope is very complex and unique among prokaryotes, consisting of a plasma membrane, a cell wall and a capsule-like outermost layer (Figure 2). By means of advanced electron microscopy techniques at least four layers, possessing significantly different properties, can been identified: i) an inner plasma membrane; ii) an electron-dense layer corresponding to the peptidoglycan and arabinogalactan coat; iii) an electron transparent zone that corresponds to the mycolic acid layer; and iv) an outer electron dense layer, the capsule-like layer (5).

The mycolic acids assemble through powerful interactions, to form a lipid shell surrounding the cell (6); this feature is responsible for the "lipid barrier" that confers resistance to many drugs. Mycolic acids are threalose estherified to form threalose-6,6-dimycolate, a structure named "cord factor" (7) that is responsible for a unique feature of MTB known as "cording"; this phenomenon consists in the formation of ropelike tangles of laterally associated bacilli that has long been thought to be associated with bacterial virulence (8).

As supported by the identification of a porin-encoding gene, small hydrophilic molecules diffusion in MTB is mediated by porin-dependent mechanism (9).

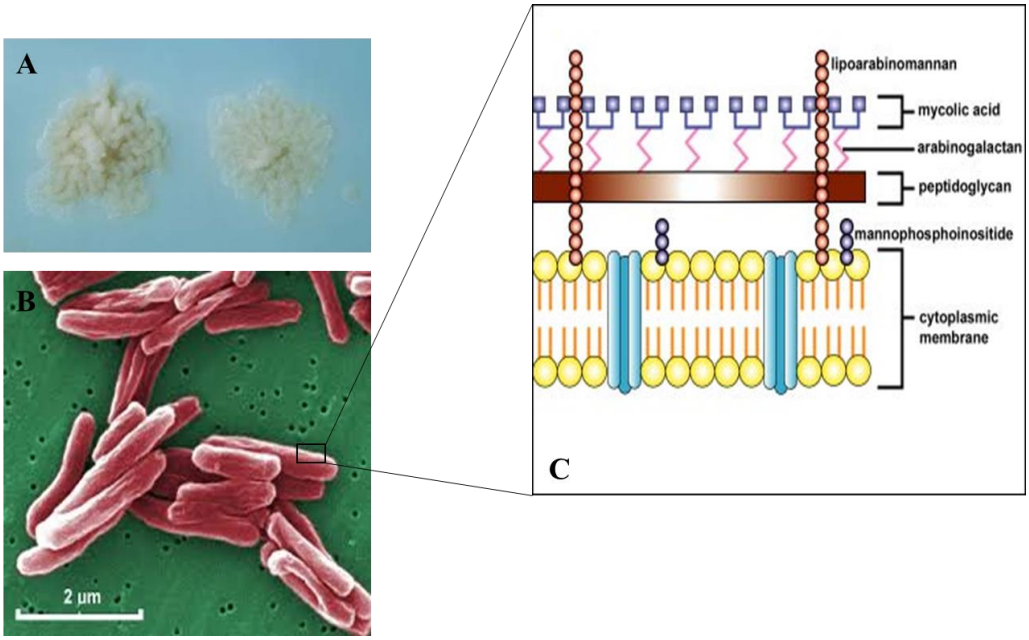


Figure 2- ***(A) Colonies of M. tuberculosis on Lowenstein-Jensen egg-based-medium. (B) Scanning electron microscope image of M. tuberculosis. (C) Mycobacterial cell wall structure (6).***

Transmission and pathogenesis of TB

MTB is carried in airborne particles, called droplet nuclei, of 1-5 microns in diameter. Infectious droplet nuclei are generated when persons who have pulmonary or laryngeal active TB disease cough, sneeze, shout, or sing. Depending on the environment, these tiny particles can remain suspended in the air for several hours. Transmission occurs when a person inhales MTB-infected droplet nuclei, which then traverse the mouth or nasal passages, upper respiratory tract, and bronchi to reach the alveoli of the lungs (Figure 3).

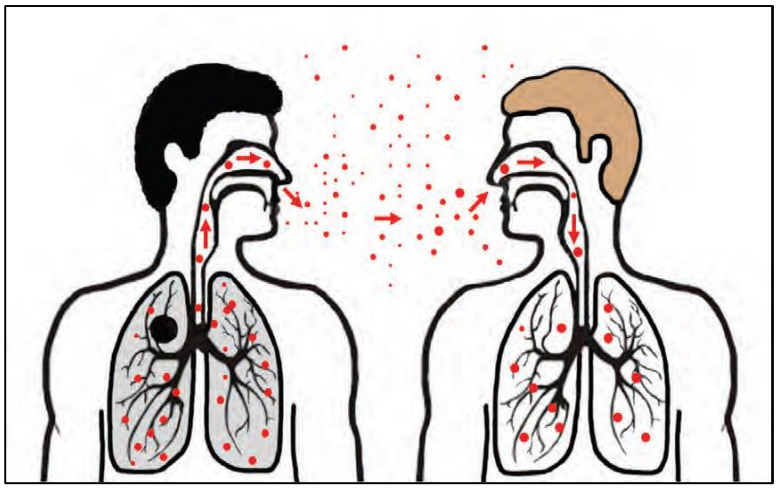


Figure 3- ***Airborne horizontal transmission of MTB (10).***

The infectiousness of a person with TB disease is directly related to the number of tubercle bacilli expelled into the air (10); it has been estimated that a droplet containing 1-10 bacteria is sufficient to initiate infection, establishing the so called primary infection. Primary infection can be i) either eradicated by the host's immune system; ii) or contained as latent infection; iii) or it can progress towards the active disease (11) that typically develops and spreads as caseating granulomas to regional lymph nodes (Figure 4). When the host's immune response and MTB virulence are in balance, MTB is contained at the primary site of infection by the immune cells. Subsequently to phagocytosis of MTB bacilli by alveolar macrophages, lymphocytes and in particular T-cells, are recruited to the infected site, processing and presenting MTB antigen to Major Histocompatibility Complex Class II (MHC II). This results in T-cell activation and release of cytokines, including Interferon Gamma (IFN-γ), that activate macrophages recruited to the site of infection.

MTB is able to survive and replicate inside macrophages. The ability of pathogenic mycobacteria to adapt to the hostile environment of macrophage has been instrumental in their success as pathogens. Once inside the macrophages MTB interferes with the host intracellular trafficking pathways by modulating events in endosomal/phagosomal maturation pathway to create a protected niche for itself: the mycobacterial phagosome (12). Phagocytosed pathogens experience increasing acidification, highly reactive oxidative and nitrosative species (ROS and RNS), hydrolytic enzymes and cationic antimicrobial peptides (13). ROS and RNS - generated by the phagosomal enzymes NADPH phagocyte oxidase and inducible nitric oxide synthase (iNOS) - damage captured microbes by modification of their DNA, lipids, proteins and active centers of metal-dependent proteins (14). The final step of bacterial destruction and clearance requires phagolysosome fusion.

Pathogenic mycobacteria have evolved a unique strategy to survive intracellularly within macrophages: they modulate phagosome maturation to prevent the fusion with lysosomes (15). The block of phagolysosome biogenesis plays a central role in the pathogenesis of mycobacteria. By blocking its delivery to lysosomes, MTB is able to avoid the acidic proteases of the lysosomes, to avoid exposure to the bactericidal mechanisms that operate within lysosomes, preventing degradation and hence processing and presentation of mycobacterial antigens to the immune system (16).

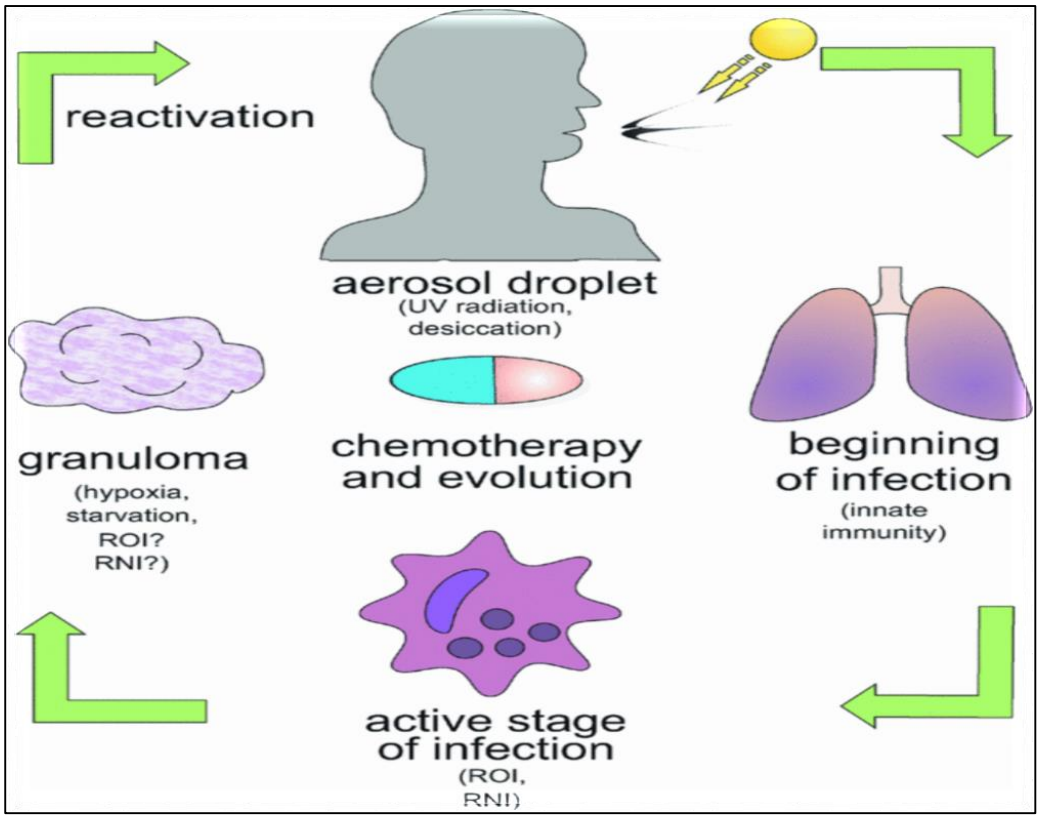


Figure 4- ***Overview of the MTB infection in humans (4).***

Granuloma

A peculiar histopathologic trait of mycobacterial infection is the formation of granuloma. This structure is maintained and stabilized by events mediated by both the host and the pathogen. It is beneficial for the host as it helps to contain the infection to localized regions. On the other hand, bacteria can live for prolonged time within the granuloma (16), while the bacterial spreading to other areas of the organism is restricted (17). The tuberculosis granuloma is a compact, organized

aggregate of epithelioid macrophages that have undergone a specialized transformation to have tight contacts with adjacent cell membranes. Bacteria are most commonly present in the central necrotic areas in which dead and dying macrophages can be detected. Many other cell types also populate the granuloma, such as neutrophils, dendritic cells, B and T cells, natural killer (NK) cells, fibroblasts and cells that secrete extracellular matrix components (Figure 5) (18).

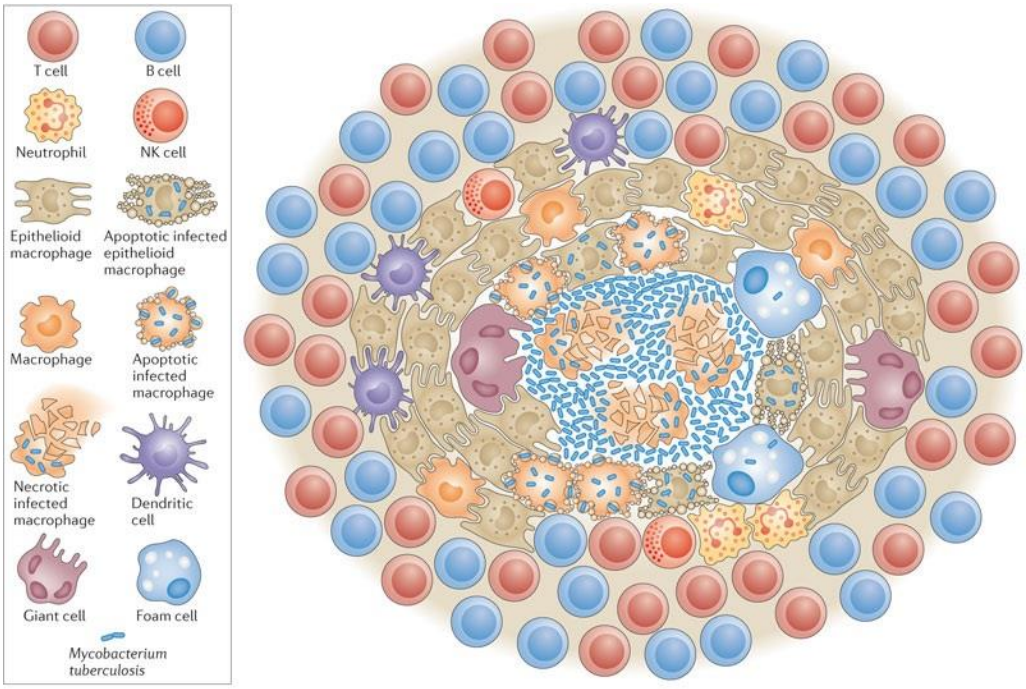


Figure 5- ***Structure and cellular constituents of the tuberculosis granuloma (18).***

In humans, the granuloma shows a high plasticity and three major types can be distinguished. These are solid granulomas, characteristic of the containment phase of the infection; necrotic granulomas, typical for early stages of active TB; and caseous granulomas during end-stage or

severe TB. These different stages are not distinct entities but form a continuum.

Unfavourable conditions inside the granuloma, such as nutrient limitation and low oxygen tension, trigger the metabolic downshift of subpopulations of MTB to dormancy (11).

Since the central region of granuloma does not contain blood vessels, it becomes hypoxic (19) and it has been demonstrated that hypoxia leads to DNA damage (20) with several DNA repair genes undergoing changes in their expression profile (21).

When tubercle bacilli overcome the immune system and multiply, it results in progression from latent TB infection (LTBI) to TB disease that represents the active phase of the disease. Persons who have TB disease are usually infectious and may spread the bacteria to other people. The progression from LTBI to TB disease may occur at any time, from soon to many years later after infection.

Chemotherapy and mycobacterial evolution

Since the control measures for TB, such as Bacillus-Calmette-Guèrin vaccination and chemoprophylaxis, appear to be unsatisfactory, treatment with anti-tubercular drug represents the only option available. Long-term treatment with specific combination of antitubercular drugs is required (22). As indicated by WHO guidelines (23) treatment of TB and drug-resistant cases requires:

1- An initial intensive phase consisting in the administration of rifampicin (RIF), isoniazid (INH) and pyrazinamide (PYZ), and ethambutol (ETB) daily for two months
2- A continuation phase consisting in administration of RIF and INH for a further four months, either daily or three times *per* week.

INH eradicates most of the rapidly replicating bacilli in the first two weeks of treatment, together with ETB. Thereafter, RIF and PYZ have an important role in the sterilisation of lesions by eradicating the microorganisms; these two drugs are crucial for successful six-month treatment regimens. RIF kills low or non-replicating bacteria and the high sterilising effect of PYZ serves to act on semi-dormant bacilli located in sites resistant to the penetration and action of the other anti-TB drugs (24). Anti-TB drugs sites of action are showed in Figure 6.

Selective antibiotics promote the occurrence of mutations in the pathogen's genome that could lead to drug-resistance (25). Consistent with these findings is the occurrence of drug-resistant strains of MTB, which have been divided into two main groups: i) the multi-drug resistant (MDR) strains that are resistant to two front line drugs, isoniazid and rifampicin, and they cover 5.3% of all new cases of TB worldwide (26); and the extensively drug resistant (XDR) strains that are resistant to isoniazid and rifampicin plus one of the fluoriquinolones and at least one second-line injectable drug (amikacin).

Without any evidence about MTB plasmids (27) or other functional apparatus, it is believed that drug resistance determinants are

chromosomally encoded and they are the consequence of spontaneous mutations (28) or gene inactivation by a mobile genetic element (29).

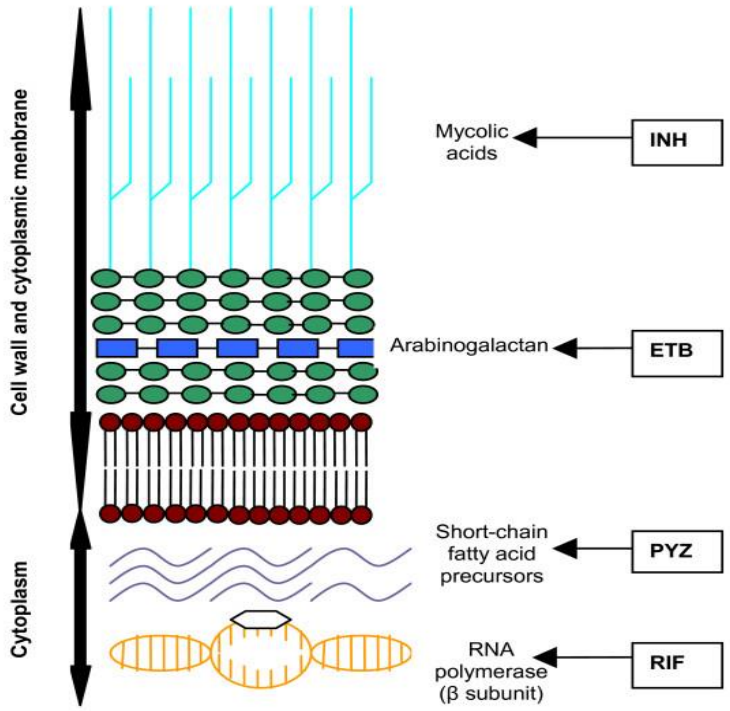


Figure 6- ***Sites of action of the anti-TB drugs (24).***

As in all organisms, DNA repair systems directly influence the rates and types of mutation in mycobacteria. When they perform properly their function, these systems keep mutation rates at a low level, althought in several circumstances they can increase the prevalence of specific types of mutations. Indeed, DNA repair systems might not be efficient enough in repairing large amounts of DNA damage, which can therefore lead to mutations; secondly, some DNA damage repair systems are inherently error-prone, possibly introducing mutations during their action; finally,

mutated DNA repair proteins can lead to an increased mutation rate, as has been observed upon gene deletions experiments resulting in the absence of key enzymes in DNA repair systems of MTB (4).

DNA repair in *M. tuberculosis*

At all stages of its complex life cycle MTB is exposed to several physical and chemical potential DNA-damaging stresses. Before establishing a primary infection mycobacteria must survive inside the aerosol droplets in which survival rate falls to 50% after six hours (30). They experience both desiccation and UV radiation, possibly resulting in DNA damage; in fact it has been demonstrated that intracellular dehydration leads to double-strand breaks of DNA (31); moreover, UV radiation induces pyrimidine dimers as photoproducts. Subsequently to inhalation of MTB droplets by a healthy individual, mycobacteria continue to be exposed to DNA damaging assaults consisting in both oxidative and nitrosative stresses generated by the activated macrophages (32), leading to DNA alkylation through N-nitrosoamines formation (33). As described above, in the last stage of its life MTB survives in the granulomatous tissue characterized by a hypoxic environment as well as starvation, two conditions strictly related to DNA damage (34) (35). As also suggested by the induction of key enzymes involved in prevention and repair of DNA damage during *in vitro* stationary phase, mycobacteria survival in the dormant phase could be related to a proper deployment of DNA

repair systems (4). The slow-growing feature of MTB makes investigations on DNA repair machineries more difficult compared to other bacteria. As a result, most assumptions about MTB DNA repair are still based on homology rather than on functional studies (36). Genes encoding proteins involved in nucleotide excision repair (NER), base excision repair (BER), recombinational and SOS repair have been identified in MTB genome (Table 1). In particular, a full set of genes known to be directly involved in the repair of oxidative and alkylation damage are present in MTB. In contrast, homologs of gene involved in mismatch repair (MMR) have not been detected (37). As observed in *Helicobacter pilory*, missing MMR repair system could potentially contribute to a high level of genetic diversity (38); in contrast MTB display a remarkably low level of genetic diversity among isolates (39), signalling that the other systems could compensate for this lack.

DNA repair pathway	**Repair strategy**
Base excision repair (BER)	Removal of small, non-helix-distorting base lesions from the genome.
Nucleotide excision repair (NER)	Removal of short single-stranded DNA segment that contains bulky helix-distorting lesions
Recombinational repair	Repair of chromosomal site-specific double-strand breaks. In homologous recombination, a second intact copy of the broken segment serves as a template for DNA synthesis. Non-homologous end joining operate when only one chromosomal copy is available.
SOS repair and mutagenesis	The SOS response is a global response to DNA damage in which the cell cycle is arrested and DNA repair and mutagenesis are induced. It is an error-prone system.

Table 1- ***MTB multi-step DNA repair pathways.***

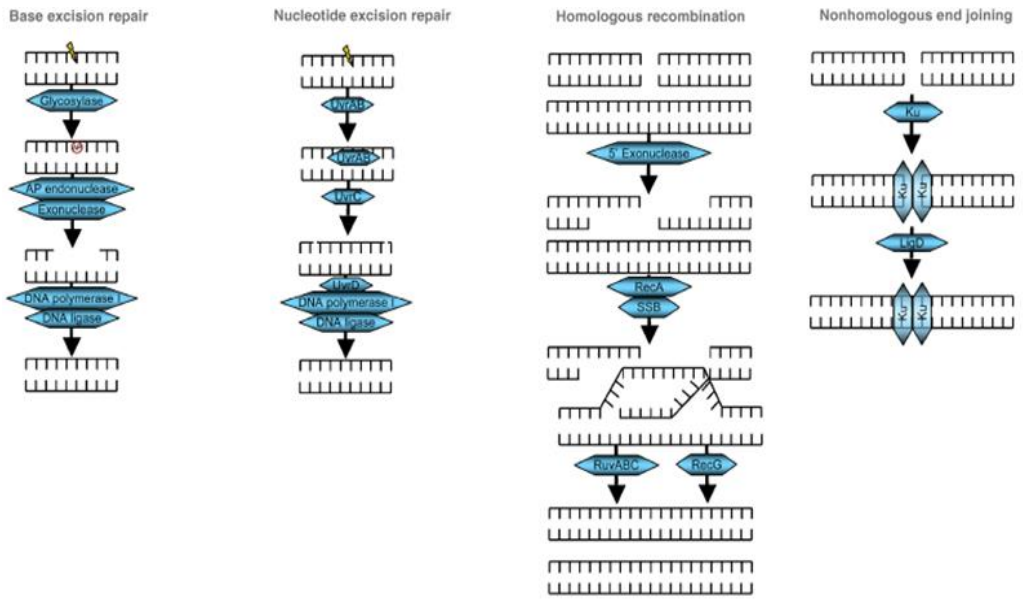


Figure 7- ***Major DNA repair pathways in MTB, [adapted from (37)]***

Alkylation damage and the Ada operon in MTB

MTB is expected to sustain significant levels of nitrosative and oxidative stresses that form potent DNA alkylating compounds during the infection process (4). Alkylation damage is both genotoxic and pro-mutagenic. Similarly to what observed in other organisms MTB can repair alkylated DNA by using multi-enzymatic systems (i.e. Nucleotide Excision Repair and Base Excision Repair whose mechanisms are summarized in table 1 and Figure 7) or by mounting an adaptive response, (Ada response) (40,41).

In *Escherichia coli* four genes are induced by Ada response: *ada, alkA, alkB* and *aidB.* In this system the methyltransferase AdaA acts as a positive regulator of the operon; AlkA is a 3-methyladenine DNA glycosylase whose expression level is 10-fold increased after exposure to sublethal doses of alkylating agents; the AlkB protein is Fe-(II)-dependent dioxygenase responsible for the repair of 1-methyladenine and 3-methylcytosine; finally AidB is a DNA binding protein predicted to catalyze direct repair of alkylated DNA (44).

E. coli Ada consists of two different domains: the N-terminal domain displaying an "AdaA" fold and the C-terminal domain displaying an AdaB fold. The 20 kDa AdaA domain demethylates innocuous methylphosphotriesters in DNA by methyl transfer on to the cysteine residue of its active site (i.e. cysteine at position 38). Self-methylation converts Ada into a transcriptional activator with high DNA binding affinity to genes containing the ada operator sequence in their promoter, including the ada gene itself (42,43). The 19 kDa AdaB domain acts as an O^6-alkylguanine DNA alkyltransferase by performing the transfer of alkyl groups from the mutagenic and cytotoxic bases O^6-methylguanine (O^6-meG) and O^4-methylthymine (O^4-meT) to a strictly conserved cysteine of its active site (i.e. cysteine at position 321).

Genome wide sequencing revealed that the genes coding for players of the Ada response are highly conserved, although characterized by different domain combinations and genomic organization in different bacterial species. In MTB, a fusion of *adaA* with *alkA* constitutes an

operon with the *adaB* gene encoding an O^6-methylguanine DNA methyltransferase (Figure 8) (44). Searches in genome databases reveal that the AdaA-AlkA fusion seems to be as common as AdaA with AdaB in bacteria (42).

Yang *et al.* have recently characterized the *ada* operon of MTB by biochemical analysis and functional complementation assays. Activity assays with purified recombinant proteins demonstrated that both AdaA-AlkA and AdaB possess DNA methyltransferase activity, whereas no 3-methyladenine DNA glycosylase activity could be detected. Furthermore, trans-complementation of an OGT-defective *E.coli* strain (KT233) with MTB *adaB* gene rescues the hypermutator phenotype while the expression of the MTB AdaA-AlkA in the same system does not produce the same effect. These results suggest that in MTB AdaB counteracts the mutagenic effect of O^6-meG, thus behaving as a genuine O^6-methylguanine DNA methyltransferase (OGT), whereas AdaA-AlkA possesses a methyltransferase activity that removes methyl groups from innocuous methylphosphotriesters (44).

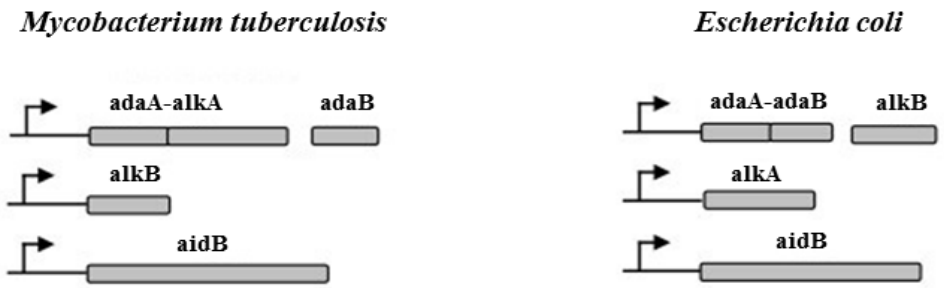


Figure 8- ***Schematic representation of chromosomal organization of the Ada response genes in E.coli and structural homologs in MTB (44).***

MTB mutant strains lacking the *ada* operon, including *adaA-alkA* and *adaB/ogt* showed a 100-fold increase in mutation frequency as compared to wild-type cells after treatment with the alkylating agent N-methyl-N′-nitro-N-nitrosoguanidine (MNNG). Moreover this mutant displayed hypersensitivity to MNNG comparable to that of *adaB*/*ogt*-defective *E.coli* strains (45). Analyzing the phenotypic effect of the absence of *ada* operon, mice infected with wild type and mutant MTB strain did not show significant differences in bacillary loads present in the lungs, liver, and spleen, when they were measured at a relatively early stage of stationary infection. Therefore it has been proposed that in MTB the Ada response to alkylation damage is restricted to counteract mutagenesis without producing any evident effect onto mycobacterial viability *in vivo* (44, 45). These findings lead to the hypothesis that a defective Ada response in mutant MTB strains could produce a selective advantage in terms of adaptation to the human host and environmental changes. Notably, the analysis of MDR W-Beijing strains revealed that nine out of fiftyfive analysed strains had nonsynonymous single-nucleotide polymorphisms (nsSNP) on codon 37 of the OGT-encoding gene, leading to the Arg to Leu substitution (R37L). nsSNP on codon 15, leading to Thr to Ser substitution was observed in twentyfour of twentynine strains of the Harlem family (46). Further sequencing of the *ada* operon of hundreds Central African Republic (CAR) MDR and non-MDR MTB strains revealed other SNPs in their Ada operon genes (47) (Figure 9 and 10), suggesting that certain MTB strains characterized by

a defective DNA-alkylation repair could be particularly prone to acquire mutations that may result in drug resistance and in a better adaptation to the host.

Strains	Genotype	No. isolates	Country of isolation	Group	*mutT2*	*mutT4*	*ogt*
ZA20/65	W-Beijing	2	Spain	1	wt	Arg CGG 48 GGG Gly	Arg CGC 37 Leu CTC
ZA67-69	W-Beijing	3	Spain	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
ZA11/16	Haarlem	2	Spain	2	wt	wt	Thr ACC 15 Ser AGC
ZA12-14/17	other	4	Spain	nd	wt	wt	wt
ZA19	*M. bovis*	1	Spain		wt	wt	wt
ZA15	other	1	Spain	nd	wt	wt	wt
ZA60-62	W-Beijing	3	Spain	1	wt	Arg CGG 48 GGG Gly	Arg CGC 37 Leu CTC
CDC1551		1	USA	2	wt	wt	wt
H37Rv		1	USA	3	wt	wt	wt
MT210	W-Beijing	1	USA	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
20	W-Beijing	1	Mongolia	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
30	W-Beijing	1	South Africa	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
34	W-Beijing	1	Malaysia	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
43	W-Beijing	1	China	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
44	W-Beijing	1	Thailand	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
45	W-Beijing	1	Malaysia	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
91/102-6	W-Beijing	6	Vietnam	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
110/116/119/124-5/140-2	W-Beijing	8	the Netherlands	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
133	W-Beijing	1	South Africa	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
W4/10/126/129	W-Beijing	4	USA	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
W99	W-Beijing	1	Singapore	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
W147	W-Beijing	1	Russia	1	Gly GGA 58 CGA Arg	Arg CGG 48 GGG Gly	Gly GGG 12 Gly GGA
94	W-Beijing	1	Vietnam	1	wt	Arg CGG 48 GGG Gly	Arg CGC 37 Leu CTC
111	W-Beijing	1	South Korea	1	wt	Arg CGG 48 GGG Gly	Arg CGC 37 Leu CTC
115	W-Beijing	1	the Netherlands	1	wt	Arg CGG 48 GGG Gly	Arg CGC 37 Leu CTC
5107(HG1)	W-Beijing	1	USA	1	wt	Arg CGG 48 GGG Gly	Arg CGC 37 Leu CTC
114, 139	W-Beijing	1	the Netherlands	1	wt	Arg CGG 48 GGG Gly	wt
166(HD6)	W-Beijing	1	USA	1	wt	Arg CGG 48 GGG Gly	wt
165(001)	W-Beijing	1	USA	1	wt	wt	Arg CGC 37 Leu CTC
107(LB)	W-Beijing	1	USA	1	wt	wt	wt
113	W-Beijing	1	the Netherlands	1	wt	wt	wt
122(CI1)	W-Beijing	1	USA	1	wt	wt	wt
IK/KY/LB2/DV/DU2/HI	W-Beijing	6	Russia	1	wt	wt	wt
N16	W-Beijing	1	USA	1	wt	wt	wt
AM	W-Beijing	1	USA	1	wt	wt	wt

[a]wt, wild-type alleles (identical to H37Rv strain); nd, not determined

Figure 9- ***Characteristics of MDR W-Beijing strains originating from different countries (46).***

Strains	Genotype	No. isolates	Country of isolation	Group	*mutT2*	*mut T4*	*ogt*
AU	Haarlem	1	USA	2	wt	wt	Thr ACC 15 Ser AGC
3,5,22,32,39,48,50 ,52-3,55	Haarlem	10	Argentina	2	wt	wt	Thr ACC 15 Ser AGC
8	Haarlem	1	Vietnam	2	wt	wt	Thr ACC 15 Ser AGC
13/28	Haarlem	2	Sri Lanka	2	wt	wt	Thr ACC 15 Ser AGC
51	Haarlem	1	the Netherlands	2	wt	wt	Thr ACC 15 Ser AGC
57/59	Haarlem	2	Czech republic	2	wt	wt	wt
84	Haarlem	1	Czech Republic	2	wt	wt	Thr ACC 15 Ser AGC
86/143/145	Haarlem	3	Bolivia	2	wt	wt	Thr ACC 15 Ser AGC
87	Haarlem	1	USA	2	wt	wt	Thr ACC 15 Ser AGC
99	Haarlem	1	Italy	2	wt	wt	Thr ACC 15 Ser AGC
123	Haarlem	1	Czech Republic	2	wt	wt	Thr ACC 15 Ser AGC
144/146-7	Haarlem	3	Bolivia	2	wt	wt	wt
Apr-35	Africa	2	Rwanda	2	wt	wt	wt
37	Africa	1	Uganda	2	wt	wt	wt
40/120	Africa	2	Burundi	2	wt	wt	wt
72	Africa	1	Central African Republic	2	wt	wt	wt
97	Africa	1	Uganda	2	wt	wt	wt
121	Africa	1	Central African Republic	2	wt	wt	wt
2	BCG	1	the Netherlands	1	wt	wt	wt
6/47/73/130	*M. bovis*	4	the Netherlands	1	wt	wt	wt
12	Other	1	Tunisia	3	wt	wt	wt
15/31	Other	2	Iran	2	wt	wt	wt
16	Other	1	Canada	2	wt	wt	wt
17	Other	1	Greenland	2	wt	wt	wt
18	Other	1	USA	2	wt	wt	wt
19/36/74	Other	2	India	1	wt	wt	wt
25/62	*M. microti*	2	UK	1	wt	wt	wt
26	Other	1	Zimbabwe	2	wt	wt	wt
27	Other	1	Ethiopia	2	wt	wt	wt
38/42	Other	2	Tahiti	2	wt	wt	wt
41/46	Other	2	Chile	2	wt	wt	wt
49	Other	1	Tanzania	1	wt	wt	wt
56	Other	1	Curacao	2	wt	wt	wt
64	Other	1	Honduras	2	wt	wt	wt
65/112	Other	2	the Netherlands	1	wt	wt	wt
71	BCG	1	Japan	1	wt	wt	wt
76/101/126	*M. bovis*	3	Argentina	1	wt	wt	wt
83	BCG	1	Russia	1	wt	wt	wt
89/95	Other	2	Spain	2	wt	wt	wt
96	Other	1	the Netherlands	3	wt	wt	wt
98	Other	1	Ecuador	2	wt	wt	wt
100	*M. africanum*	1	the Netherlands	1	wt	wt	wt
108	Other	1	China	2	wt	wt	wt
118	Other	1	Honduras	2	wt	wt	wt

[a]nd, not determined; wt, wild-type alleles (identical to H37Rv strain).

Figure 10- ***Characteristics of MTB strains from Harlem family originating from different countries (46).***

Structural and functional aspects of alkylated-DNA repair enzymes

As mentioned above, the alkylation of exocyclic oxygens of the DNA nitrogen bases, and in particular the O^6 position of guanine and to a lesser extent the O^4 position of thymine are promutagenic. While alkylation at the O^6 position of guanine typically constitutes a small fraction of the damage produced by exogenous alkylating stresses, it is principally responsible for the GC→AT transition mutations associated with mutagenic effects of many alkylating agents (48).

Alkylguanine-DNA alkyltransferases (AGT) are evolutionary conserved DNA repair proteins that represent the main cellular mechanism responsible for alkylguanine repair.

All members of this protein family described till now (including the mycobacterial O^6-methylguanine-DNA methyltransferase), invariably act through a suicidal mechanism by performing the stoichiometric transfer of the O^6-alkyl group from the modified base to a strictly conserved cysteine buried in their active site (a S_N2-like reaction) (Figure 11) (49), resulting in a permanently inactivated protein probably addressed to degradation (50, 51).

Sequences encoding AGTs are readily recognized in genomic data bases since virtually all AGT proteins contain the consensus sequence -(I/V)PCHR(V/I)(I/V)- surrounding the cysteine residue that serves as the alkyl acceptor.

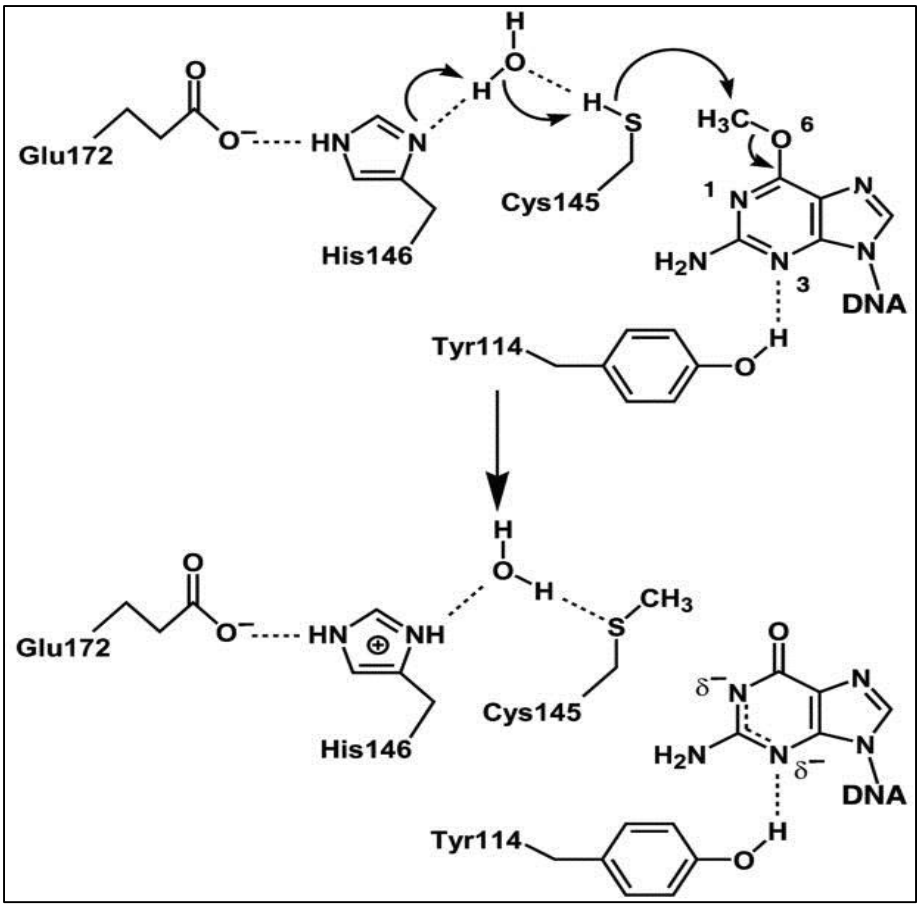


Figure 11- ***Reaction mechanism of human AGT by activation of catalytic cysteine via interaction with glutamate-histidine-water-cysteine (49).***

These proteins, which consist of a small single polypeptide chain, have been found in most bacteria, archea, fungi and animals but not in plants. Notably, an AGT ortolog is also absent from the fission yeast *Schizosaccharomyces pombe* (49). Structural studies (52, 53, 54, 55, 56, and 57) revealed that these proteins show a similar overall structure despite limited sequence identity, which is restricted to the DNA binding domain.

The three-dimensional structure consists of two roughly globular domains connected by a loop (Figure 12). The highly structurally conserved C-terminal domain contains the active site pocket that houses the catalytic cysteine within the PCHR consensus sequence, adjacent to the DNA-binding Helix-Turn-Helix motif. The poorly functionally characterized N-terminal domain displays a low degree of primary sequence and structural conservation between different species. X-ray structures of human AGT bound to dsDNA were solved (54, 55). These studies revealed that the helix-turn-helix motif of the C-terminal domain unusually binds the DNA at the level of its minor groove. The recognition helix is small and contains hydrophobic residues allowing it to pack closely within the DNA minor groove with minimal sequence-specificity (54, 58). The binding alter the DNA structure, bending the chain sugar-phosphate by about 15°, widening the minor groove and displacing the target modified base into the active site pocket. This displacement is caused by a rotation promoted by specific residues (54), and is stabilized by a conserved arginine finger that replaces the flipped-out base in the regular base stacking.

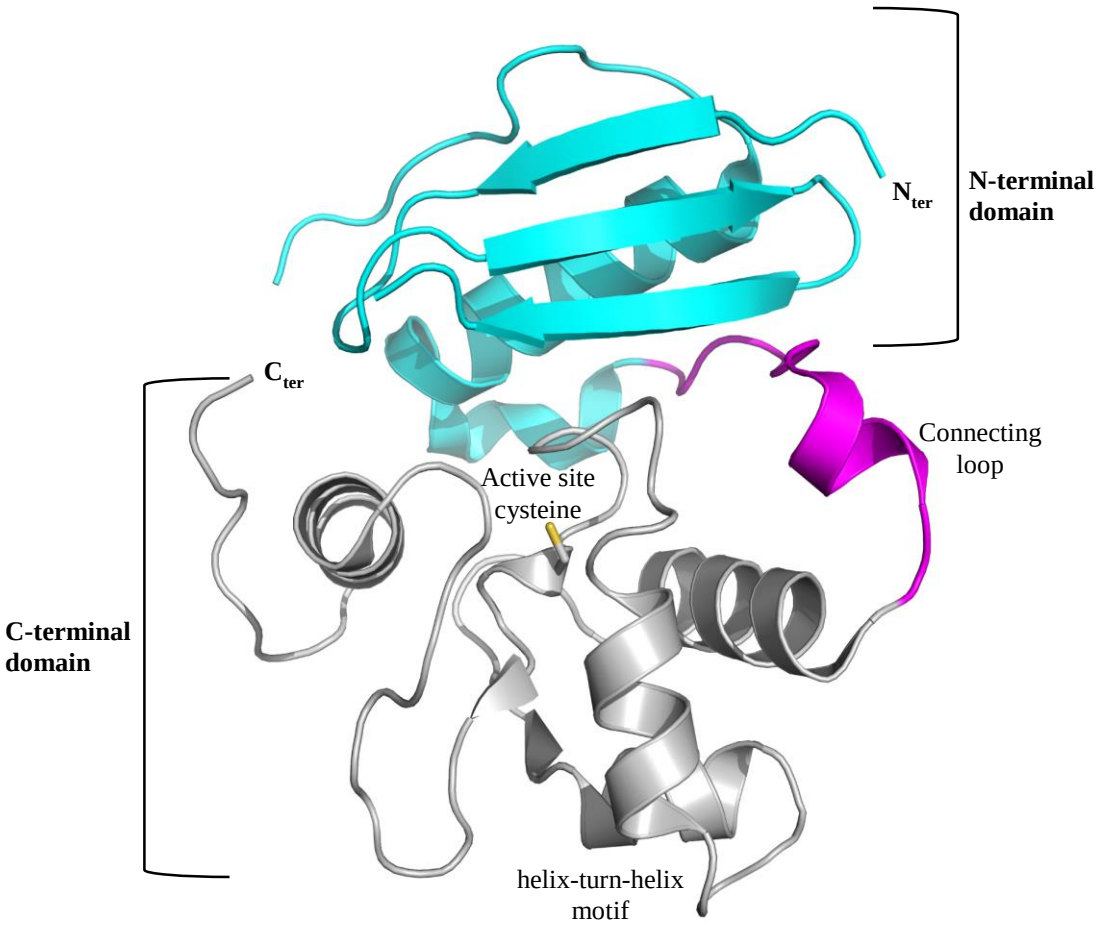


Figure 12- ***Molecular architecture of human AGT (53).***

It is known that human AGT binds the target DNA in a cooperative fashion (59). It has been proposed that cooperative binding could facilitate both directional scanning and the efficient repair of lesion contained within chromatin structure. A model, predicting quaternary structure topology, demonstrated that cooperative assemblies contain a 3-start helical array of proteins with dominant protein-protein

interactions between the N-terminal face of protein *n* and the C-terminal face of protein *n*+3 (Figure 13) (59). The occluded binding-site sizes in cooperative complexes are significantly smaller compared to the DNA surface occupied by an AGT monomer, corresponding to four nucleotides for both double-strand and single strand DNA. As mentioned above, differently from other helix-turn-helix DNA binding proteins, AGT interacts with DNA through a minor groove-binding mechanism. It is well know that major groove binding allows base-specific hydrogen bonds and seems optimal for sequence-specific recognition, whereas the DNA minor groove binding performed by AGT seems advantageous for both sequence-independent binding and nucleotide flipping (54). Considering substrate specificity, competition studies with oligodeoxynucleotides containing O^6-benzylguanine (O^6-bG) and O^6-meG in the same sequence clearly showed that O^6-bG is the preferred substrate for mammalian OGTs, that *E. coli* OGT has no preference and *E. coli* Ada-B greatly prefers O^6-meG (60). Finally, it is fully established that the DNA adduct, O^4-methylthymine, which is a very minor product of the reaction of DNA with methylating agents, can also be repaired by OGTs from different species, althought at different rates (61).

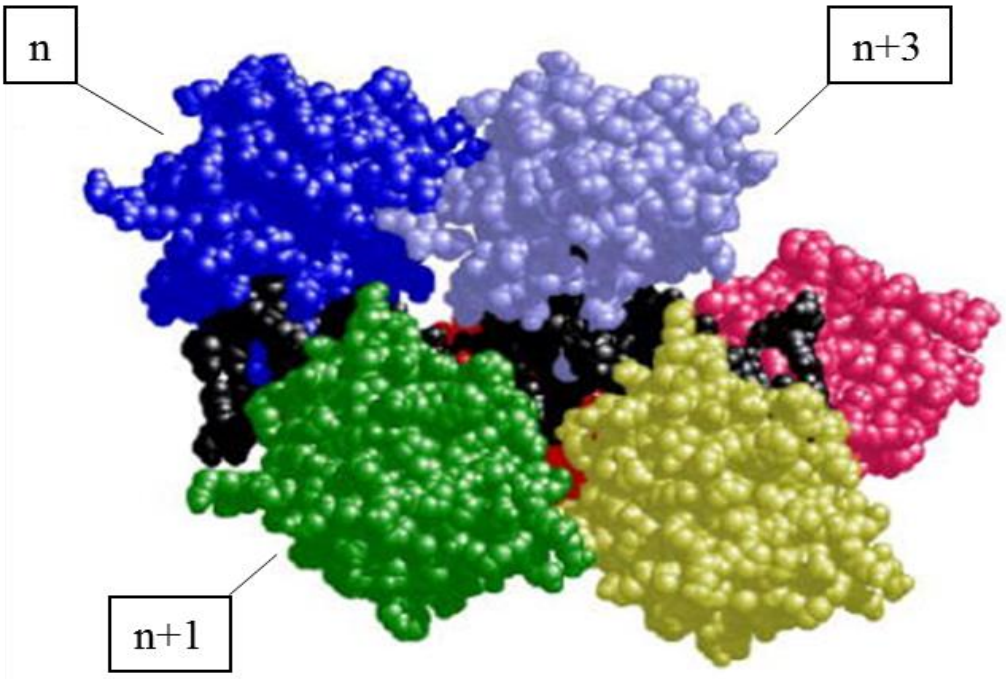


Figure 13-*A model of human OGT-DNA quaternary structure resulting in a three-start helical array. [Adapted from (59)].*

Bibliography

1- **World Health Organization -** Global Tuberculosis report 2013

2- **Lönnroth K., Castro K. G., Chakaya J. M., Chauhan L.S., Floyd K, Glaziou P., Raviglione M.C.** (2010) Tuberculosis control and elimination 2010-50: cure, care, and social development. Lancet; **375**:1814-29

3- **Flynn J.L., Chan J.** (2001) Tuberculosis: latency and reactivation. Infect Immun.; **69**:4195-4201.

4- **Gorna A.E., Bowater R.P., Dziadek J.** (2010) DNA repair systems and the pathogenesis of Mycobacterium tuberculosis: varying activities at different stages of infection. Clin. Sci.; **119**:187-202.

5- **Paul T.R., Beveridge T.J.** (1992) Reevaluation of envelope profiles and cytoplasmic ultrastructure of mycobacteria processed by embedding and freeze-substitution protocols. J. bacteriol.; **174**:6508-6517.

6- **Brennan P.J., Nikaido H.** (1995) The envelope of mycobacteria. Annu. Rev. Biochem.; **64**:29-63

7- **Goren M.B., Brennan P.J.** (1979) Mycobacterial lipids: chemistry and biological activities. Tuberculosis pp. 63-93

8- **Glickman M.S., Cox J.S., Jacobs W.R. Jr.** (2000) A novel mycolic acids cyclopropane synthetase is required for coding, persistence and virulence of Mycobacterium tuberculosis. Mol. Cell; **5**: 717-727

9- **Trias J., Jarlier V., Benz R.**, (1992) Porins in the cell wall of mycobacteria. Science; **258**:1479-1481

10- **Center for Disease and Prevention** - TB pathogenesis.

11- **Gengenbacher M. and Kaufmann S.H**. (2012) Mycobacterium tuberculosis: Success through dormancy. FEMS Microbiol Rev.; **36**:514–532.

12- **Houben E.N., Nguyen L., Pieters J.** (2006) Interaction of pathogenic mycobacteria with the host immune system. Curr Opin Microbiol.; **9**:76-85

13- **Flannagan R.S., Cosio G., Grinstein S. (**2009) Antimicrobial mechanisms of phagocytes and bacterial evasion strategies. Nat Rev Microbiol.; **7**:355–366

14- **Fang F.C.** (2004) Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. Nat Rev Microbiol.; **2**:820–832

15- **Pieters J.** (2001) Evasion of host cell defense mechanisms by pathogenic bacteria. Curr. Opin. Immunol.; **13**:37–44

16- **Grosset J.** (2003) Mycobacterium tuberculosis in the extracellular compartment: an underestimated adversary. Antimicrob. Agents Chemother.; **47**:833–836

17- **Ulrichs T., Kaufmann S.H.** (2006) New insights into the function of granulomas in human tuberculosis. J. Pathol.; **208**:261–269

18- **Ramakrishnan L.** (2012) Revisiting the role of the granuloma in tuberculosis. Nature Reviews Immunology; **12**:352-366

19- **Tsai M. C., Chakravarty S., Zhu, G., Xu, J., Tanaka, K., Koch, C., Tufariello, J., Flynn, J. and Chan, J.** (2006) Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension. Cell. Microbiol.; **8**: 218-232

20- **Moller, P., Loft, S., Lundby, C. and Olsen, N. V.** (2001) Acute hypoxia and hypoxic exercise induce DNA strand breaks and oxidative DNA damage in humans. FASEB J.; **15**:1181-1186

21- **Hampshire, T., Soneji, S., Bacon, J., James, B. W., Hinds, J., Laing, K., Stabler, R. A., Marsh, P. D. and Butcher, P. D.** (2004) Stationary phase gene expression of Mycobacterium tuberculosis following a progressive nutrient depletion: a model for persistent organisms? Tuberculosis; **84**:228-238

22- **Fox W., Ellard G.A., Mitchison D.A.** (1999) Studies on the treatment of tuberculosis undertaken by the British Medical Research. Council tuberculosis units, 1946–1986, with relevant subsequent publications. Int J Tuberc Lung Dis; **3**:231-279.

23- **World Health Organization. Communicable Diseases Cluster**: Fixeddose combination tablets for the treatment of tuberculosis. (1999)

24- **Du Toit L.C., Pillay V. and Danckwerts M.P.** (2006) Tuberculosis chemotherapy: current drug delivery approaches**.** Respiratory Research; **7**:118

25- **Martinez, J. L. and Baquero, F.** (2000) Mutation frequencies and antibiotic resistance. Antimicrob. Agents Chemother.; **44**:1771-1777

26- **World Health Organization** (2009) Global tuberculosis control: epidemiology.

27- **Katti, M. K.** (2001) Plasmids of mycobacteria. J. Med. Microbiol.; **50**:575-576

28- **Filliol, I., Motiwala, A. S., Cavatore, M., Qi, W., Hazbon, M. H., Bobadilla del Valle, M., Fyfe, J., Garcıa-Garcıa, L., Rastogi, N. and Sola, C.** (2006) Global phylogeny of Mycobacterium tuberculosis based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA

fingerprinting systems, and recommendations for a minimal standard SNP set. J. Bacteriol.; **188**:759-772

29- **Cubillos-Ruiz, A., Morales, J. and Zambrano, M. M.** (2008) Analysis of the genetic variation in Mycobacterium tuberculosis strains by multiple genome alignments. BMC Res. Notes; **1**:110

30- **Loudon, R. G., Bumgarner, L. R., Lacy, J. and Coffman, G. K.** (1969) Aerial transmission of mycobacteria. Am. Rev. Resp. Dis. **100**:165-171

31- **Peccia, J. L., Werth, H. M. and Hernandez, M. T.** (2000) Effects of relative humidity on the UV-induced inactivation of bacterial bioaerosols. J. Aerosol. Sci.; **31**:959-960

32- **Adams, L. B., Dinauer, M. C., Morgenstern, D. E. and Krahenbuhl, J. L.** (1997) Comparison of the roles of reactive oxygen and nitrogen intermediates in the host response to Mycobacterium tuberculosis using transgenic mice. Tuber. Lung Dis.; **78**:237-246

33- **Burney, S., Caulfield, J. L. C., Niles, C. J., Wishnok, J. S. and Tannenbaum, S. R.** (1999) The chemistry of DNA damage from nitric oxide and peroxynitrite. Mutat. Res.; **424**:37-49

34- **Moller, P., Loft, S., Lundby, C. and Olsen, N. V.** (2001) Acute hypoxia and hypoxic exercise induce DNA strand breaks and oxidative DNA damage in humans. FASEB J. **15**:1181-1186

35- **Grishko, V., Solomon, M., Breit, J. F., Killilea, D. W., LeDoux, S. P., Wilson, G. L. and Gillespie, M. N.** (2001) Hypoxia promotes oxidative base modifications in the pulmonary artery endothelial cell VEGF gene. FASEB J.; **15**:1267-1269

36- **Mizrahi V. and Andersen S. J.** (1998) DNA repair in Mycobacterium tuberculosis. What have we learnt from the genome sequence? Molecular biology; **29**:1331-1339

37- **Dos Vultos T., Mestre O., Tonjum T. and Gicquel B. (**2009) DNA repair in Mycobacterium tuberculosis revisited. FEMS Microbiol Rev.; **33**:471-87

38- **Kang J. and Blaser M.J.** (2006) Bacterial populationas perfect gases: genomic integrity and diversification tension in Helicobacter pylori. Nat Rev Microbiol; **4**:826-836

39- **Dos Vultos T., Mestre O., Rauzier J., Golec M., Rastogi N., Rasolofo V., Tonjum T., Sola C., Matic I., Gicquel B.** (2008). Evolution and diversity of clonal bacteria: the paradigm of Mycobacterium tuberculosis. PLoS One; **3**:e1538.

40- **Dalhus B, Laerdahl JK, Backe PH, Bjoras M.** (2009). DNA base repair, recognition and initiation of catalysis. FEMS Microbiol. Rev.; **33**:1044-1078.

41- **Shrivastav N., Li D., Essigmann J.M.** (2010). Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. Carcinogenesis; **3**:59 –70.

42- **Sedgwick B. and Lindahl T.** (2002) Recent progress on the Ada response for inducible repair of DNA alkylation damage. Oncogene; **21**:8886-94

43- **Saget B.M., Walker G.C.** (1994)The Ada protein acts as both a positive and a negative modulator of Escherichia coli's response to methylating agents. Proc. Natl. Acad. Sci. U.S.A.; **91**:9730–9734

44- **Yang M., Aamodt R.M., Dalhus B., Balasingham S., Helle I., Andersen P., Tønjum T., Alseth I., Rognes T., Bjørås M.** (2011) The ada operon of Mycobacterium tuberculosis encodes two DNA methyltransferases for inducible repair of DNA alkylation damage. DNA Repair (Amst); **10**:595-602.

45- **Durbach S.I., Springer B., Machowski E.E., North R.J., Papavinasasundaram K.G., Colston M.J., Bottger E.C., Mizrahi V.** (2003) DNA alkylation damage as a sensor of nitrosative stress in *Mycobacterium tuberculosis*. Infect. Immun.; **71**:997–1000

46- **Ebrahimi-Rad M., Bifani P., Martin C., Kremer K., Samper S., Rauzier J., Kreiswirth B., Blazquez J., Jouan M., van S.D., Gicquel B**. (2003) Mutations in putative mutator genes of Mycobacterium tuberculosis strains of the W-Beijing family. Emerg. Infect. Dis.; **9**:838–845

47- **Nouvel L.X., Dos V.T., Kassa-Kelembho E., Rauzier J., Gicquel B.** (2007) A non-sense mutation in the putative anti-mutator gene ada/alkA of Mycobacterium tuberculosis and M. bovis isolates suggests convergent evolution. BMC Microbiol.; **7**:39

48- **Marginson G.P. Santibanez Koref, M.F.** (2002) O6-alkylguianine-DNA alkyltransferase: role in carcinogenesis and chemoterapy. Bioessay; **24**:255-266

49- **Pegg A.E.** (2011). Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools. Chem. Res. Toxicol.; **24**:618–639.

50- **Kanugula S., Goodtzova K., Pegg A.E.** (1998). Probing of conformational changes in human O6-alkylguanine-DNA alkyl transferase protein in its alkylated and DNA-bound states by limited proteolysis. Biochem. J.; **329**:545–550.

51- **Xu-Welliver M., Pegg A.E.** (2002). Degradation of the alkylated form of the DNA repair protein, O(6)-alkylguanine-DNA alkyltransferase. Carcinogenesis; **23**:823– 830.

52- **Moore M.H., Gulbis J.M., Dodson E.J., Demple B., Moody P.C.** (1994) Crystal structure of a suicidal DNA repair protein: the Ada O6-

methylguanine-DNA methyltransferase from E. coli. EMBO J.; **13**:1495-1501

53- **Daniels D.S., Mol C.D., Arvai A.S., Kanugula S., Pegg A.E., Tainer J.A.** (2000) Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical binding. EMBO J.; **19**:1719–1730.

54- **Daniels D.S., Woo T.T., Luu K.X., Noll D.M., Clarke N.D., Pegg A.E., Tainer J.A.** (2004) DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nat. Struct. Mol. Biol.; **11**:714-720.

55- **Duguid E.M., Rice P.A., He C.** (2005) The structure of the human AGT protein bound to DNA and its implications for damage detection. J. Mol. Biol.; **350**:657-666

56- **Hashimoto H., Inoue T., Nishioka M., Fujiwara S., Takagi M., Imanaka T., Kai Y.** (1999) Hyperthermostable protein structure maintained by intra- and inter-helix ion pairs in archael O6-methylguanine-DNA methyltransferase. J. Mol. Biol.; **292**:707–716

57- **Miggiano R., Casazza V., Garavaglia S., Ciaramella M., Perugino G., Rizzi M., Rossi F.** (2013) Biochemical and structural studies of the Mycobacterium tuberculosis O6-methylguanine methyltransferase and mutated variants. J Bacteriol.; **195**:2728-2736

58- **Tubbs J.L., Pegg A.E., Tainer J.A**. (2007) DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. DNA Repair (Amst.); **6**:1100–1115

59- **Adams C.A., Melikishvili M., Rodgers D.W., Rasimas J.J, Pegg A.E, Fried M.G.** (2009) Topologies of complexes containing O6-alkylguanine-DNA alkyltransferase and DNA. J. Mol. Biol.; **389**:248–263

60- **Goodtzova K., Kanugula S., Edara S., Pauly G.T., Moschel R.C., Pegg A.E.** (1997) Repair of O6-benzylguanine by the Escherichia coli Ada and Ogt and the human O6-alkylguanine-DNA alkyltransferase J. Biol. Chem.; **272**:8332–8339

61- **Pegg A.E. (**2000) Repair of O6-alkylguanine by alkyltransferases**.** Mutation Research/Reviews in Mutation Research; **462**:83-100

Outline of the thesis

Tuberculosis is a global health emergency with nearly 1.5 million deaths that occur annually and more than two billion people currently infected with TB bacilli. As a facultative intracellular pathogen, *Mycobacterium tuberculosis* (MTB), the etiologic agent of TB, survives and replicates inside the human macrophages. Accordingly, during its entire life cycle MTB is exposed to a variety of DNA-damaging stresses including alkylating agents that may compromise both persistence in the host's infected macrophages as well as the reactivation of the bacillus from the dormant state.

Similarly to what observed in other organisms MTB attains alkylated-base repair, maintaining its genome integrity and a low level of genetic variability, either by using multi-enzymatic repair systems or by direct removing the alkyl group from the damaged base sacrificing one molecule of a DNA-protein alkyltransferase such as O^6-methylguanine methyltransferase (OGT). Notably, a number of geographically widely distributed MTB W-Bejing strains and multi-drug resistant (MDR) isolates show nonsynonymous SNPs in their OGT encoding sequence resulting in amino acid substitutions at the poorly functionally characterized N-terminal domain of the enzyme (T15S and R37L). MDR-TB is reaching epidemic proportions and DNA repair systems is supposed to play a main role in MDR strains occurrence as well as in mycobacterial pathogenesis. For this reason biochemical and structural

investigations on O^6-methylguanine DNA methyltransferase and mutated versions mimicking the ones occurring in clinical isolates could provide new insights on hypermutator phenotype genesis *in vivo* as a consequence of catalytic defects in alkylated DNA repair.

Moreover, the improved understanding of MTB DNA repair mechanisms may allow to reveal biochemical details of the mycobacterial life cycle that could be integrated in the complex framework describing the MTB systems biology.

In the chapter three of the present PhD dissertation, I report on the kinetic analysis and structural characterization of MTB OGT and point-mutated variants designed to mimic the mentioned SNPs, in order to explore the functional effects of these mutations. Biochemical studies reveal that these amino acid substitutions affect, to different extents, alkyl-guanine repair activity suggesting a new role of the N-terminal domain of the protein in the coordination of the catalytic cycle and/or in mediating protein cooperative assembly at damaged DNA sites. As described in chapter four we extended this analysis to other mutants in order to better describe the structural and functional elements of novelty displayed by *Mt*OGT when compared to other ligand-free or DNA-bound OGT models. Overall, the results discussed in this thesis provide experimental evidence of a possible contribution of a defective alkylated DNA repair to the “mutator” phenotype proposed for several MDR MTB strains.

Biochemical and Structural Studies of the *Mycobacterium tuberculosis* O^6-Methylguanine Methyltransferase and Mutated Variants

Riccardo Miggiano[a*], Valentina Casazza[a*], Silvia Garavaglia[a], Maria Ciaramella[b], Giuseppe Perugino[b], Menico Rizzi[a], Franca Rossi[a].

[a]Dipartimento di Scienze del Farmaco, University of Piemonte Orientale A.Avogadro,Novara,Italy; [b]Istituto di Biochimica delle Proteine, IBP-CNR

Published in Journal of Bacteriology, June 2013; 195 (12): 2728–2736. Epub 5 April 2013

Summary

Mycobacterium tuberculosis displays remarkable genetic stability despite continuous exposure to the hostile environment represented by the host's infected macrophages. Similarly to other organisms, *M. tuberculosis* possesses multiple systems to counteract the harmful potential of DNA alkylation. In particular, the suicidal enzyme O^6-methylguanine-DNA methyltransferase (OGT) is responsible for the direct repair of O^6-alkylguanine in double-stranded DNA and is therefore supposed to play a central role in protecting the mycobacterial genome from the risk of G·C-to-A·T transition mutations. Notably, a

number of geographically widely distributed *M. tuberculosis* strains shows nonsynonymous single-nucleotide polymorphisms in their OGT-encoding gene, leading to amino acid substitutions at position 15 (T15S) or position 37 (R37L) of the N-terminal domain of the corresponding protein. However, the role of these mutations in *M. tuberculosis* pathogenesis is unknown. We describe here the in vitro characterization of *M. tuberculosis* OGT (*Mt*OGT) and of two point-mutated versions of the protein mimicking the naturally occurring ones, revealing that both mutated proteins are impaired in their activity as a consequence of their lower affinity for alkylated DNA than the wild-type protein. The analysis of the crystal structures of *Mt*OGT and *Mt*OGT-R37L confirms the high level of structural conservation of members of this protein family and provides clues to an understanding of the molecular bases for the reduced affinity for the natural substrate displayed by mutated *Mt*OGT. Our in vitro results could contribute to validate the inferred participation of mutated OGTs in *M. tuberculosis* phylogeny and biology.

Introduction

During its entire life, *Mycobacterium tuberculosis* is exposed to a variety of potential physical and chemical DNA-damaging stresses that could compromise the settling, containment, and reactivation of the infection (1). In order to maintain high genome stability, signaled by a remarkably low level of genetic diversity among isolates (2), it is therefore

mandatory for *M. tuberculosis* to possess efficient systems to counteract the effects of such environmental and host-generated DNA-endangering assaults (3–6). In particular, during its long-term persistence inside infected macrophages, *M. tuberculosis* must deal with endogenous DNA-alkylating chemical species originated by the action of highly reactive oxidative and nitrosative radicals (7). The majority of living organisms attain alkylated base repair by deploying different strategies, depending upon the chemical nature of the alkyl group, the entity of the damage, and the physiological condition of the cell: (i) the multistep excision of a short, lesion-containing strand followed by new DNA synthesis; (ii) the substitution of the modified base in toto; or (iii) the direct surgical removal of the alkyl-substituting group from the base by sacrificing one molecule of a DNA-protein alkyltransferase (8, 9), such as the O^6-methylguanine-DNA methyltransferase (OGT) (EC 2.1.1.63). All OGT members studied until now invariably act through a suicidal mechanism (10), by performing the stoichiometric transfer of the O^6-alkyl group from the modified guanine to a strictly conserved cysteine residue in the protein active site, which is hosted in the C-terminal domain of the protein. This covalent modification leaves OGT permanently inactivated and possibly more prone to degradation (11, 12) (Fig. 1).

Paralleling observations of other bacteria (13), the mycobacterial OGT-encoding sequence (*adaB*; Rv1316c) is part of an adaptive response operon, which was recently functionally characterized (14). Gene

inactivation experiments demonstrated that *adaB* (here referred to as OGT, in analogy to orthologs in other species) is not essential for infectivity and survival either in vitro or in the mouse model of *M. tuberculosis* infection (15, 16). However, several observations support the importance of OGT activity in protecting the mycobacterial GC-rich DNA from the promutagenic potential of guanine O^6-methylation. First, the OGT gene expression profile changes during infection and in response to alkylating compounds (14, 17, 18), signaling a requirement for fine-tuned OGT modulation under stressful conditions; moreover, the heterologous expression of *M. tuberculosis* OGT (*Mt*OGT) in the *Escherichia coli* KT233 *ada-ogt*-defective strain suppresses its MNNG (*N*-methyl-*N*'-nitro-*N*-nitrosoguanidine) sensitivity (14) and rescues the hypermutator phenotype.

Interestingly, two nonsynonymous single-nucleotide polymorphisms (nsSNPs) have been identified in the OGT genes of both *M. tuberculosis* Beijing strains and multidrug-resistant isolates (19–21), leading several authors to postulate that corresponding mutated OGTs may contribute to the success of these strains in terms of worldwide distribution; indeed, it has been suggested that a defective OGT might result in an increased mutation frequency and, thus, in a better capability of the bacterium to rapidly adapt to the host (14, 19). Notably, both these nsSNPs result in amino acid substitutions at the poorly structurally conserved and functionally characterized N-terminal domain of OGT, thus limiting the

possibility of predicting their effect on protein activity exclusively based upon preexisting information.

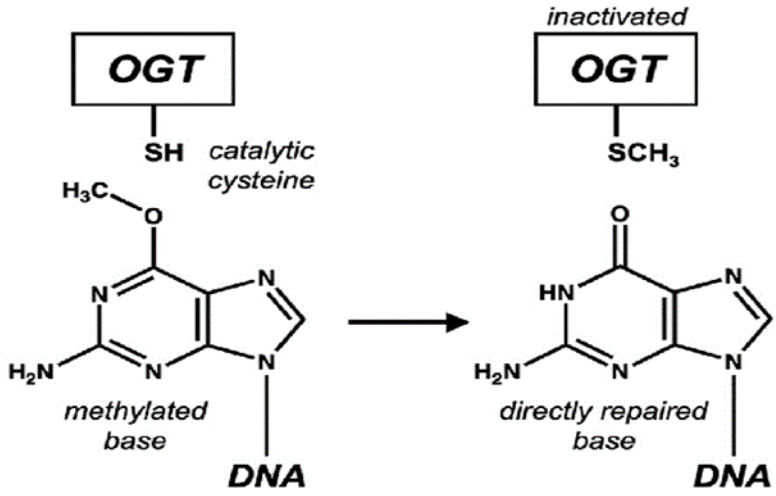


Figure 1 Schematic representation of transfer of O^6-alkyl group from modified guanine to strictly conserved cysteine in the protein active site.

Many years of intense investigations rendered a detailed picture of the catalytic mechanism used by these proteins as well as of the structural elements of the C-terminal domain playing essential roles in DNA binding and alkyl group removal (8, 10); in contrast, a more limited number of studies put the focus onto the N-terminal domain, which could play a role in the coordination of the catalytic cycle (22) and/or in mediating protein assembly at the site of damage upon DNA binding (23–25).

In the present work, we exploited a recently described fluorescence-based approach (26) to analyze the kinetics of the reaction performed by wild-type *Mt*OGT and by two point-mutated variants of the protein

mimicking the ones occurring in *M. tuberculosis* clinical isolates (i.e., *Mt*OGT-R37L and *Mt*OGT-T15S). The method relies on the use of a fluorescent derivative of O^6-benzylguanine, an inhibitor of human O^6-alkylguanine-DNA alkyltransferase (AGT), and allows the dissection of the OGT reaction in its DNA binding and alkylguanine transfer steps, thereby leading to the determination of the DNA binding affinity and the alkyltransferase reaction rate.

The results of our biochemical studies reveal that although neither mutation affects the intrinsic alkyl-guanine-transferase reaction rate, both variants display reduced DNA binding affinity although to different extents, with *Mt*OGT-R37L showing a 10-fold-lower affinity for alkylated double-stranded DNA (dsDNA) than the wild-type protein. Moreover, the crystal structures of *Mt*OGT and its R37L mutant reported here give a contribution to the overall description of this structurally conserved class of proteins and provide a framework to investigate the possible molecular bases of the observed DNA binding defect.

Materials and method

Chemicals. All reagents were obtained from Sigma-Aldrich unless otherwise specified.

Construction of expression vectors. The open reading frame coding for *M. tuberculosis* O^6-methylguanine methyltransferase (Rv1316c) was isolated by PCR using the clone MTCY130 (Institute Pasteur, Paris, France) as the template, the Hot-Star PCR system (Qiagen), and primers *Mt*OGTfwd and *Mt*OGTrev (see Table S1 in the supplemental material). The NcoI-BamHI double-digested PCR product was then ligated into the similarly digested pET16b vector (Novagen) by standard techniques (27), resulting in the pET-*Mt*OGT construct. Plasmids encoding the T15S and R37L point mutants of *Mt*OGT (pET-*Mt*OGT-T15S and pET-*Mt*OGTR37L) were obtained by using the pET-*Mt*OGT construct as the DNA template, the two primers pairs T15Sfwd/T15Srev and R37Lfwd/R37Lrev (see Table S1 in the supplemental material), and the QuikChange II site-directed mutagenesis kit (Stratagene). In each construct, the region encoding the wild-type protein or its point-mutated variant was verified by sequencing (Eurofins MWG Operon).

Expression and purification of wild-type *Mt*OGT and point mutants. The following procedure was invariably used to express and purify the wild-type protein and the relative point-mutated versions used in biochemical analyses. *E. coli* strain BL21(DE3) bacteria (Novagen), freshly transformed with one of the three expression constructs, were

spread onto LB agar plates with 50 µg/ml ampicillin and grown at 37°C overnight. The next day, colonies were scraped and inoculated in 1 liter of selective ZYP- 5052 medium (28) to reach a starting optical density at 600nm (OD_{600}) of 0.1. This culture was further grown at 37°C for 3 h and then brought to 17°C for 16 h with vigorous shaking to autoinduce the expression of the recombinant protein. A bacterial pellet was obtained by centrifugation (11,000 ×*g* for 15 min at 4°C), washed once in phosphate-buffered saline (PBS), dissolved in 50 ml of buffer A (20 mM Tris-HCl, pH 7.8), and disrupted by ultrasonication. Upon the addition of a protease inhibitor cocktail, the insoluble material of the lysate was removed by centrifugation (14,000 ×*g* for 45 min at 4°C), and the recombinant protein was purified from the supernatant by fast protein liquid chromatography (FPLC) (Akta Basic Instrument; GE Healthcare), using, in sequence, HiTrapQ, MonoQ, HiTrap heparin, and Superdex 200 prepacked chromatography media (GE Healthcare), as further detailed in the supplemental material (see Fig. S1 in the supplemental material). During the entire procedure, the recombinant protein was monitored by standard SDS-PAGE analysis (29), and the protein concentration was determined by the Bradford assay (30), using bovine serum albumin as the standard.

Activity assays. To measure the alkyltransferase activity of wild-type *Mt*OGT and relative mutants, a fluorescence assay using SNAP-Vista Green reagent (New England BioLabs) (here referred to as VG for brevity) was used, as previously described (26) (see Fig. S2A in the

supplemental material). Protein bands from SDS-PAGE were visualized by direct gel imaging (VersaDoc 4000; Bio-Rad), and their fluorescence intensities (FIs), corresponding to the reacted protein, were corrected for the actual amount of protein loaded into each lane by measuring the intensity of bands after Coomassie brilliant blue R250 staining. Activity assays at different pH values were performed with 1× Fluo reaction buffer containing 50.0 mM different buffering species, allowing the determination of the pH optimum of the reaction used in all subsequent experiments (data not shown).

Enzyme kinetics analysis. Assuming a 1:1 substrate-enzyme stoichiometry and irreversible binding, incubation of a fixed amount of protein (5.0 µM) with VG (in the range of 0.1 to 20.0 µM) for 1 h at 25°C gave corrected FI data, which were fitted by a linear equation whose slope gave a direct reference value of FI/µM protein, which was then used to estimate the amount of covalently modified protein (expressed in picomoles) in time course experiments. The latter assays were performed at 25°C using different protein/VG molar ratios and taking 10.0-µl aliquots at time intervals, as detailed in Fig. S2B in the supplemental material. Plots of the picomoles of reacted protein versus time were fitted by exponential equations to determine the apparent rates for covalent modification (k_{obs}).

These values were plotted versus the VG doses ([VG]) by using the following hyperbolic equation:

$$k_{obs} = \frac{k}{1+K_{VG}/[VG]} \quad (1)$$

Table 1 Kinetic constants of the reaction catalysed by *Mt*OGT and mutated proteins.

	ds-DNA^{met} (µM)	**Mean *k* (s^{-1}) ± SD**	**Mean K_{VG} (µM) ± SD**	**Mean K_{DNA} (µM) ± SD**
***Mt*OGT**	0	0.12 ± 0.02	1.82 ± 0.4	**0.24± 0.11**
	0.63	0.14 ± 0.03	3.53 ± 0.5	
	1.00	0.10 ± 0.03	5.29 ± 0.9	
	1.25	0.14 ± 0.02	8.68 ± 0.8	
***Mt*OGT-T15S**	0	0.10 ± 0.01	1.73 ± 0.5	**0.48 ± 0.26**
	0.63	0.11 ± 0.01	2.29 ± 0.4	
	1.00	0.13 ± 0.02	3.46 ± 0.5	
	1.25	0.12 ± 0.03	5.40 ± 0.7	
***Mt*OGT-R37L**	0	0.07 ± 0.02	3.17± 0.5	**2.16 ± 0.23**
	0.63	0.09 ± 0.02	4.23± 0.7	
	1.00	0.07 ± 0.01	4.55± 0.6	
	1.25	0.08 ± 0.02	5.11± 0.5	

where *k* and K_{VG} are the rate of covalent linkage and the dissociation constant for the free enzyme and free VG reagent during the first collision step (before covalent modification), respectively (see Fig. S2C in the supplemental material). In order to calculate the apparent K_{VG} values (K_{VG}^{app}) of proteins for VG as a function of the double-stranded methylated DNA ($dsDNA^{met}$) concentration, time course experiments such as those described above were performed in the presence of

increasing fixed doses of $dsDNA^{met}$ (see Table S1 and Fig. S3 in the supplemental material).

Data were fitted according to the following linear equation:

$$K_{VG}{}^{app} = \frac{K_{VG}}{K_{DNA}} \left[dsDNA^{met}\right] + K_{VG} \qquad (2)$$

where K_{DNA} is the dissociation constant of the protein for the $dsDNA^{met}$ substrate. The results of these enzyme kinetics analyses are summarized in Table 1.

Electrophoretic mobility shift assay (EMSA). A 40-bp 6-carboxytetramethylrhodamine (TAMRA)-labeled double-stranded DNA probe (0.1 µM) was incubated at 25°C for 10 min with different amounts of proteins in the range of 0.0 to 120.0 µM in a total volume of 10.0 µl in 1× binding buffer (20.0 mM Tris-HCl [pH 7.5], 50.0 mM KCl, 0.1 mM dithiothreitol [DTT], and 10% glycerol). Samples were loaded onto an 8% polyacrylamide-bisacrylamide native gel in 1× Tris-borate-EDTA (TBE). Signals were visualized by direct gel imaging using a green light-emitting diode (LED)/605-nm-band-pass filter as excitation/emission parameters, respectively.

Crystallization. Initial conditions for *Mt*OGT crystallization were identified by means of a robot-assisted (Oryx4; Douglas Instruments), sitting-drop-based spare-matrix strategy using screen kits from Hampton Research. The best crystals were obtained by mixing 2 µl of a protein solution at 5 mg/ml with an equal volume of a reservoir solution

containing 0.1 M HEPES (pH 7.5), 6% (wt/vol) polyethylene glycol 8000 (PEG 8000), and 6% (wt/vol) ethylene glycol and equilibrating the drop against 800 µl of the reservoir solution at 4°C in sitting-drop format. Crystals used in X-ray diffraction experiments grew to maximum dimensions of 0.2 mm in about 2 weeks. Similarly, crystallization of the MtOGT-R37L mutant enzyme was obtained by mixing 2 µl of the protein solution at 4 mg/ml with an equal volume of crystallization buffer (0.1 M HEPES [pH 7.5], 12% [wt/vol] PEG 8000, 4% ethylene glycol) and equilibrating the drop against 800 µl of the reservoir at 4°C in sitting-drop format, yielding 0.2-mm crystals in about 2 weeks.

Structure determination. For X-ray data collection, crystals of both wild-type *Mt*OGT and *Mt*OGT-R37L were taken directly from the crystallization droplet, rapidly equilibrated in a solution containing the crystallization buffer and 15% glycerol as a cryoprotectant, and flash-frozen at 100 K under liquid nitrogen.

Table 2 Data collection, phasing and refinement statistics

	***Mt*OGT**	***Mt*OGT-R37L**
Data collection		
Space Group	$P2_12_12$	$P2_12_12$
Wavelength (Å)	0.933	0.99
Resolution (Å)	1.8	2.8
Total reflections	70150	16209
Unique reflections	17690	4572
Mean(I)/sd(I)	15.8 (4.4)	10.2 (3.6)
Completeness (%)	99.5 (100)	96.2 (100)
Multiplicity	4.0 (4.0)	3.5 (3.6)
R_{merge} (%)	4.8 (27.2)	8.2 (31.6)
R_{meas} (%)	5.5 (31.3)	9.4 (36.4)
R_{pim} (%)	2.7 (15.3)	4.5 (17.5)
Refinement		
R_{factor}/R_{free} (%)	21.8/26.4	18.6/26
Protein Atoms	1251	1258
Ligand atoms	19	-
Water molecules	141	48
R.m.s.d. bond (Å)	0.007	0.009
R.m.s.d. angles (°)	1.016	1.151
Average B ($Å^2$)		
Protein	34.3	63.4
Solvent	40.5	50.2
Residues in Ramachandran plot areas (%)		
Preferred	96.3	98
Allowed	3.7	2

A complete data set of the wild-type *Mt*OGT crystal was collected at 100 K to a 1.8-Å resolution using synchrotron radiation (λ= 0.933 Å) at the ID14-EH1 beam line (European Synchrotron Radiation Facility [ESRF], Grenoble, France), equipped with an ADSC detector. Analysis of the collected diffraction data set allowed us to assign the crystal to the orthorhombic space group $P2_12_12$, with the cell dimensions a = 59.13 Å, b = 81.75 Å, and c = 38.16 Å, containing 1 molecule per asymmetric unit, with a corresponding solvent content of 52%. For all data collections, diffraction intensities were evaluated, integrated, and scaled by using the CCP4 suite of programs (31). The structure of *Mt*OGT was solved by molecular replacement using the program Phaser (32) and the structure of *E. coli* O^6-methylguanine-DNA methyltransferase AdaC protein as the search model (Protein Data Bank [PDB] accession number 1SFE) (33). The resulting electron density map was of high quality and allowed automatic tracing by the program ARP/wARP (34), which was also used for adding solvent molecules. The program Coot (35) was used for manual rebuilding, and the program PHENIX (36) was used for crystallographic refinement.

A complete data set of the *Mt*OGT-R37L crystal was collected at 100 K to a 2.8-Å resolution using synchrotron radiation (λ= 0.99 Å) at the ESRF ID23 beam line equipped with an ADSC detector. Since the analysis of the diffraction data set showed the *Mt*OGT-R37L crystal to be isomorphic with those of the wild-type protein, we used the refined set of atomic coordinates of the *Mt*OGT protein (in which water/ligand

molecules and Arg37 side-chain atoms were omitted) for *Mt*OGT-R37L structure determination; model building and crystallographic refinement were performed following the same procedure described above for the wild-type protein. Data collection, phasing, and refinement statistics are given in Table 2. Structural superpositions were performed with the Superpose program of the CCP4 suite (31). Figures 3 through 6 were generated with PyMol (37).

Protein structure accession numbers. The atomic coordinates and structure factors of *Mt*OGT and *Mt*OGT-R37L have been deposited in the Protein Data Bank (http://www.rcsb.org/) under accession numbers 4BHB and 4BHC, respectively.

Result and discussion

Expression and purification of wild-type and point-mutated *Mt*OGT variants. In order to perform a molecular characterization of the *M. tuberculosis* OGT protein, the corresponding open reading frame was subcloned into the pET16b vector within the NcoI/BamHI sites of the polylinker to eliminate the standard polyhistidine coding region. Once transformed into *E. coli* strain BL21(DE3), the resulting pET-*Mt*OGT expression construct drives the synthesis of an untagged version of the protein (*Mt*OGT) (predicted molecular mass of 17,844 Da; 165 amino acid residues; Ile in position 2 replaced by Val as a consequence of the adopted subcloning strategy) that could be reproducibly purified at a

high yield by FPLC-based standard techniques. pET-*Mt*OGT was also used as the template for PCR-based site-directed mutagenesis to generate the expression vectors for the production in *E. coli* of two point-mutated variants of the enzyme (*Mt*OGTT15S and *Mt*OGT-R37L), which mimic those naturally occurring in *M. tuberculosis* strains bearing nsSNPs at codon 15 (ACC mutated to AGC) or codon 37 (CGC mutated to CTA) of their OGT coding sequence (19–21). We obtained pure and highly homogeneous preparations of both *Mt*OGT-R37L and *Mt*OGT-T15S, and we observed that, similar to other OGTs, our recombinant proteins behave as monomers in solution (see Fig. S1 in the supplemental material). These observations indicate that the R37L and T15S amino acid substitutions do not affect the overall stability of the corresponding mutated proteins and do not alter their quaternary structure in the absence of ligands.

Biochemical characterization of *Mt*OGT. From a biochemical point of view, OGT proteins are at the same time the catalyst and one of the stoichiometric reagents of the alkyl group transfer from the modified base to the reactive cysteine (Cys126 in *Mt*OGT) in the active site, resulting in permanent inactivation of the protein (10). We exploited this feature to characterize the kinetics of the reaction performed by *Mt*OGT and its variants by adopting a recently developed procedure (26) that makes use of derivatives of O^6-benzylguanine (a well-known inhibitor of AGTs), which acts by covalent transfer of the benzylic group to the active-site cysteine (see Fig. S2A in the supplemental material). The

method allows the determination of the *Mt*OGT catalytic activity from the direct measurement of the covalent protein-inhibitor fluorescent complex. Specifically, the fluorescein-labeled O^6-benzylguanine derivative SNAP-Vista Green (VG) was used to measure the activities of *Mt*OGT, *Mt*OGT-T15S, and *Mt*OGTR37L by quantifying the intensity of the signal emitted by the VG-modified protein analyzed by SDS-PAGE. In all three cases, the fluorescence intensity of the band showed a linear dependence on the VG reagent concentration, reaching a plateau at a 1:1 protein/VG ratio, thus confirming the predicted stoichiometry and the suicide nature of the reaction catalyzed by *Mt*OGT, regardless of the presence of either the T15S or R37L mutation (data not shown).

Notably, this analysis also revealed that the *Mt*OGT active site is in principle capable of housing bulky base adducts, as shown previously for the *Sulfolobus solfataricus* protein (26); in addition, it confirmed the robustness of the fluorescent VG-based assay and its applicability for the careful characterization of OGT proteins in general as a reliable alternative to classical and laborious assays for DNA-alkyltransferase activity.

Time course experiments, performed with different protein/VG ratios, allowed us to determine the kinetic constants for the covalent modification of *Mt*OGT, *Mt*OGT-T15S, and *Mt*OGT-R37L. The values obtained in the presence of the synthetic substrate VG were similar for

all the three proteins (Table 1), suggesting that neither mutation affects the alkyltransferase reaction rate.

Previous work with the *S. solfataricus* OGT protein showed that the assay can also be applied to determine the activity of OGT proteins toward their natural substrate (i.e., alkylated DNA) in competition experiments (26). We thus repeated the time course experiments described above in the presence of increasing amounts of a double-stranded DNA fragment bearing an internal O^6-methylated guanine ($dsDNA^{met}$), which is expected to compete with the VG reagent in the alkyltransferase reaction. According to equation 2, a linear plot of the K_{VG} values as a function of the $dsDNA^{met}$ concentration (see Fig. S3 in the supplemental material) allowed the calculation of the dissociation constant (K_{DNA}) of *Mt*OGT and of the two point mutants for this substrate. Our data indicate that both mutations affect the protein affinity for methylated DNA albeit to different extents: whereas the *Mt*OGT-R37L point mutant exhibited a 10-fold-lower affinity for $dsDNA^{met}$ (K_{DNA}= 2.16 ± 0.23 µM) than the wild-type protein (K_{DNA}= 0.24 ± 0.11 µM), this difference appeared less pronounced for *Mt*OGT-T15S (Table 1). Thus, although neither mutation results in impaired intrinsic alkyltransferase activity, both *Mt*OGT variants, and *Mt*OGT-R37L in particular, exhibit reduced affinity for an alkylated DNA molecule.

To investigate the reasons for the reduced activity of the mutated *Mt*OGTs toward their physiological substrate, we studied direct protein-DNA associations using a fluorescent dsDNA probe (see Table S1 in the

supplemental material) in an EMSA based analysis. Figure 2 shows that wild-type *Mt*OGT binds ds-DNA in a cooperative manner with a plateau at a DNA/protein molar ratio of 1:150 and a dissociation constant value (*K*) of approximately 7 μM, a value comparable to that determined for human AGT by using the same method (38). In contrast, *Mt*OGTT15S and *Mt*OGT-R37L showed a significant reduction in dsDNA binding, with a plateau at DNA/protein molar ratios of 1:300 and 1:600 (*K* values of approximately 14 μM and 26 μM), respectively.

Taken together, our data indicate that although neither mutation affects specifically the kinetics of the alkyl group transfer from the synthetic VG substrate to the protein, both mutations significantly decrease the protein's capability of associating with DNA molecules and, consequently, the protein's ability to efficiently perform alkylated-DNA repair. To assess whether this catalytic defect translates in vivo into a "hypermutator" phenotype, possibly beneficial to the fitness of the bacillus under circumstances of selective pressure, will require further microbiological studies.

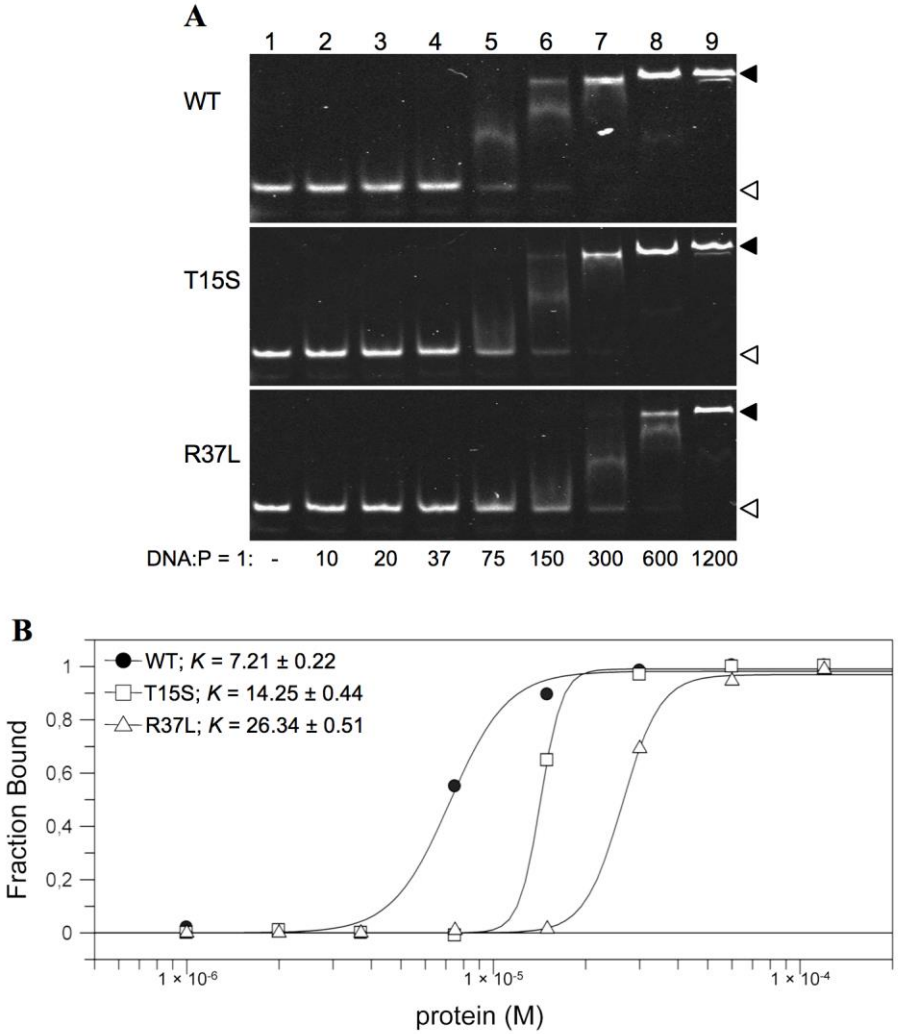


Figure 2 DNA binding activity of wild-type *Mt*OGT and point-mutated proteins. (A) Band shift analysis of *Mt*OGT (wild type [WT]), *Mt*OGT-T15S (T15S), and *Mt*OGT-R37L (R37L) proteins. Lane 1, 1 pmol of fluorescent double-stranded DNA probe (DNA); lanes 2 to 9, increasing amounts of protein (P) incubated in the presence of the probe at the indicated DNA/protein molar ratios. The white and black arrowheads point to the free and protein-bound probes, respectively. (B) Plot of the DNA-bound protein fractions at the DNA/protein molar ratios indicated in panel A. *K*, dissociation constant (μM).

Structural analysis of *Mt*OGT. The N-terminal domain of the DNA-protein alkyltransferase family members features a low degree of both sequence and length conservation between different species, thus precluding an *in silico* mapping of the residues affected by the nsSNPs under study based exclusively upon information currently available in the Protein Data Bank (see Fig. S4 in the supplemental material). Therefore, to shed light on the molecular basis of the functional defect characterizing the *Mt*OGT variants, we determined the crystal structures of the wild-type protein and of the *Mt*OGT-R37L mutant at 1.8-Å and 2.8-Å resolutions, respectively, revealing in both cases one monomer in the asymmetric unit.

An excellent electron density was visible for the entire *Mt*OGT protein chain (residues 2 to 165) and allowed the modeling of two molecules of glycerol used as the cryoprotectant. Since the crystals of *Mt*OGT-R37L were isomorphic with those of *Mt*OGT, we used the model of the wild-type protein to solve the structure of the mutant. The stereochemistry of the refined *Mt*OGT and *Mt*OGT-R37L structures was assessed with the program PROCHECK (39), revealing 96% and 98% of residues in the preferred region of the Ramachandran plot, respectively (Table 2).

***Mt*OGT overall structure**. Similar to structures of other archaeal (40, 41) and eubacterial (33) OGTs as well as of human AGT in its ligand-free (42) and alkylated (43) states or in complex with modified dsDNA (25, 44), both *Mt*OGT and *Mt*OGT-R37L fold into a roughly globular molecular architecture built up by two domains connected by a long loop

and ending in a 10-residue-long tail. In both structures, the tail adopts a fully extended conformation, pointing straight outwards from the bulk body of the protein (Fig. 3) and contacting the C-terminal domain of the adjacent symmetry mate within the crystal lattice (see Fig. S5 in the supplemental material).

The N-terminal domain consists of an antiparallel three-stranded β-sheet and connecting loops (residues 2 to 27) sandwiched between a mainly randomly coiled region (residues 28 to 46) containing a single helical turn at its middle (residues 35 to 39) (h1 in Fig. 3) on one side and a structurally conserved -helix on the opposite one (residues 47 to 59) (H2 in Fig. 3). The C-terminal domain (residues 75 to 155) adopts the prototypical all-α-fold and houses the highly conserved functional elements that previous biochemical and mutational studies (45, 46) demonstrated to be required to perform efficient catalysis. Referring to the *Mt*OGT primary sequence, these motifs are (i) the helix-turn-helix (HTH) motif, which is responsible for DNA binding at its minor groove and bears Arg109, whose structural equivalent in human AGT (*i.e.*, Arg128) has been demonstrated to act as a temporary substitute for the modified base upon its flipping out from the regular base stacking (25, 44); (ii) the "Asn hinge" building up one wall of the deep ligand binding cavity that accepts the modified base for repairing; (iii) the strictly conserved PCHR consensus motif surrounding the catalytic cysteine (Cys126); (iv) the active-site loop that participates in the correct positioning of the alkylated base inside the ligand binding pocket; and

(v) the structurally conserved H6 helix that, by shielding the ligand binding cavity on the opposite side of the Asn hinge, contributes essential residues for completing both the catalytic triad (His127 and Glu153) and the modified-base-bonding network (Lys146) (Fig. 3, left) (25, 33, 40–44).

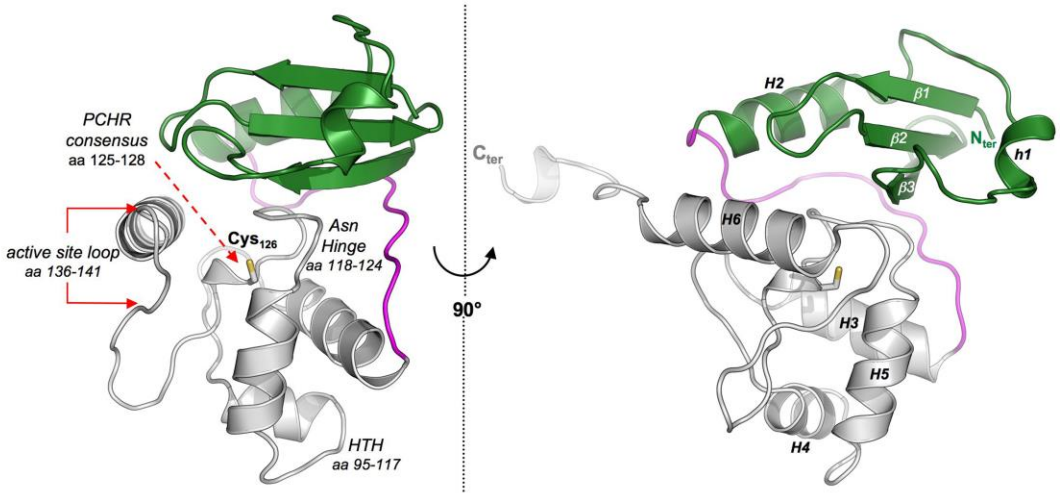


Figure 3 Crystal structure of *Mt*OGT. Shown are cartoon representations of the *Mt*OGT structure, as observed from two different points of view. In both images, the N-terminal domain, the C-terminal domain, and the connecting loop are colored in green, gray, and magenta, respectively. The catalytic cysteine residue (Cys126) is invariably drawn as sticks. Functional element labeling and secondary structure element numbering appear in the image at the left and at the right, respectively. aa, amino acids.

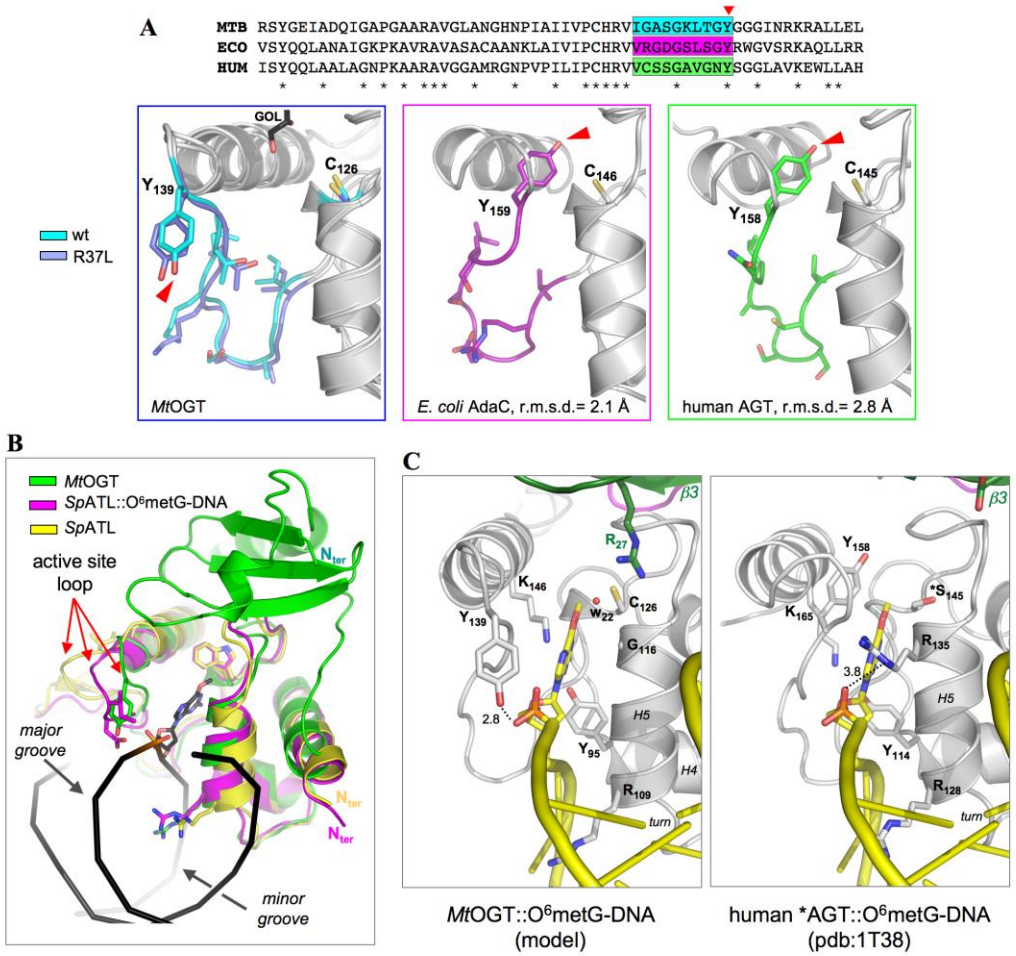


Figure 4 The active-site loop of *Mt*OGT is observed in a solvent-exposed conformation. (A) Close-up views of equivalent regions of *Mt*OGT (MTB) (both the wild-type protein and the R37L variant are shown at the left), *E. coli* AdaC (ECO) (PDB accession number 1SFE), and human AGT in the absence of ligand (HUM) (PDB accession number 1EH6), upon optimal structural superposition of their C-terminal domains. In each panel, residues of the region highlighted in the sequence alignment (top) and the catalytic cysteine are shown as sticks. The red arrow points to the OH group of the structurally equivalent Tyr residue. GOL, glycerol. (B)

Cartoon representation of optimally superimposed *Mt*OGT (green), the *Schizosaccharomyces pombe* alkyltransferase-like protein (*Sp*ATL) in the absence of ligands (yellow) (PDB accession number 3GVA; average RMSD 1.8 Å), and *Sp*ATL in complex with O^6-methylguanine dsDNA (magenta) (PDB accession number 3GX4; average RMSD 1.5 Å) (the DNA strands appear as black ribbons). (C) Possible mode of binding of alkylated dsDNA to *Mt*OGT (left), resulting upon optimal superposition of *Mt*OGT on the structure of a Cys145Ser mutant of human AGT (*AGT) in complex withO^6-methylguanine (O^6-metG) dsDNA (right) (PDB accession number 1T38; average RMSD 2.4 Å). Cys126 of *Mt*OGT is observed at 3.8 Å from the methyl group of the O^6-methylguanine. In panels B and C, residues discussed in the text appear as sticks, and the water molecule in panel C (W22) is rendered as a sphere.

The active-site loop of *Mt*OGTs is observed in a solvent-exposed conformation. Our crystallographic analysis confirms the high degree of structural conservation displayed by the catalytic C-terminal domain among members of the protein-DNA alkyltransferase protein family. However, different from what was observed in all the crystal structures of OGTs, in both the *Mt*OGT and *Mt*OGT-R37L models the active-site loop is oriented toward the exterior of the protein body (Fig. 4A), adopting a conformation that closely resembles the one displayed by the equivalent region in the structures of alkyltransferase-like proteins (ATLs) (47) in complex with alkylated DNA (48, 49) (Fig. 4B).

In principle, such a solvent-exposed conformation of the *Mt*OGT active-site loop still allows the HTH motif to peculiarly bind the substrate DNA

at its minor groove, as disclosed by structural studies of AGT and ATLs (25, 44, 48, 49), and is fully compatible with the housing of the modified base inside the active site (Fig. 4C, left). In fact, by overlaying our structures upon that of an inactive variant of human AGT (catalytic Cys145 mutated to Ser) in complex with O^6-guanine-methylated dsDNA (Fig. 4C, right) (25), we noted that the ligand molecule perfectly fits into the *Mt*OGT active site, which appears fully accessible to the flipped-out base without requiring any side-chain repositioning. In this unprecedented conformation, the hydroxyl group of the highly conserved Tyr139 residue is correctly oriented to establish a hydrogen bond with the DNA sugar phosphate backbone at the 5' side of the modified base (Fig. 4C, left), apparently compensating for the absence at position 116 of the recognition helix H5 of a residue functionally equivalent to Arg135 of human AGT (Gly116 in *Mt*OGT) (Fig. 4C). Although in this conformation, Tyr139 is no longer available for stacking against the alkylated base on the opposite side with respect to the recognition helix H5, the presence in the active site of Arg27 could contribute to anchoring the base by contacting its O^6 position through a water-mediated hydrogen bond (Fig. 4C).

At present, we cannot univocally discern whether the conformation adopted by the active-site loop in our *Mt*OGT structures represents a unique, distinctive trait of the mycobacterial proteins or is a consequence of crystal packing (see Fig. S5 in the supplemental material); nonetheless, the reported data provide further independent support for

the proposed conformational plasticity of this region and its central role in the control of alkylated-DNA binding (25, 40–44, 48, 49).

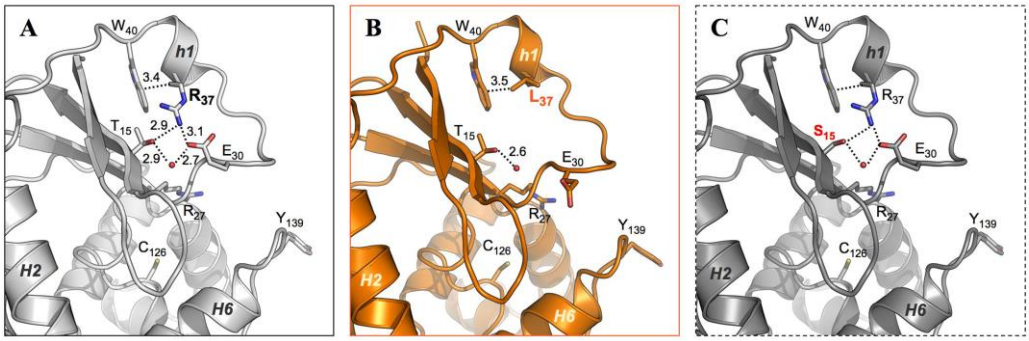


Figure 5 Possible structural role of Arg37. (A and B) Close-up views of the superposed N-terminal domain and part of the active site of wild-type *Mt*OGT (A) and *Mt*OGT-R37L (B), as observed in their respective crystal structures. (C) Equivalent regions of a structure-based model of the *Mt*OGT-T15S variant. In all images, residues cited in the text appear as sticks, dotted lines indicate close interactions (distances appear in angstroms), and water molecules are rendered as spheres.

***Mt*OGT structures provide clues to explain the DNA binding impairment displayed by *Mt*OGT-R37L**. Optimal superposition of the available OGT three-dimensional (3D) models from different species confirmed the limited structural conservation of their N-terminal domain (see Fig. S4 in the supplemental material). In this respect, the *Mt*OGT structure allowed us to directly observe the intradomain contacts involving the mutated residues under study. In particular, Thr15 and Arg37 seem to play a central role in coordinating the network of contacts

established between a small number of residues of strands 1 and 2 and the region bearing the h1 helix of the facing random coil. In fact, in the structure of *Mt*OGT (Fig. 5A), the Thr15 OH group is observed within hydrogen- bonding distance of the Arg37 NH2 atom, which is engaged in an ion pair with Glu30; in turn, the latter residue coordinates a water molecule that is also in hydrogen-bonding contact with the Thr15 OH group. The observed hydrogen bond arrangement could provide a likely explanation for why the *Mt*OGT-R37L protein displays a more reduced affinity toward alkylated DNA than does *Mt*OGT-T15S. Indeed, in the structure of the *Mt*OGT-R37L variant (Fig. 5B), the interactions involving Arg37 and the chemical identity of the local environment are lost, whereas the analysis of a model of the *Mt*OGT-T15S mutant (Fig. 5C) suggests that these biochemical features remain unaltered, due to the presence in Ser15 of a hydroxyl group structurally equivalent to the one present in the substituted threonine.

How might a nonconservative substitution of *Mt*OGT Arg37, which maps far away from both the Cys126 reactive center and the currently identified DNA binding motifs, produce an impact on protein function, as revealed by our in vitro biochemical characterization? In principle, there are three possible - and not mutually exclusive - scenarios: the R37L mutation could (i) induce conformational changes of the N-terminal random-coiled region that can affect catalysis, (ii) alter the assembly of protein complexes at the damaged site during cooperative

DNA binding, and (iii) impair an unexpected participation of the Arg37-bearing random coil to direct DNA binding (Fig. 6).

For the first scenario, it must be noted that by comparing the structures of wild-type *Mt*OGT and the *Mt*OGT-R37L variant (Fig. 5), we did not observe any dramatic conformational change affecting the protein backbone (average root mean square deviation [RMSD] 0.3 Å). Moreover, even taking into account the more limited resolution of the *Mt*OGT-R37L model, the only appreciable differences at the level of the N-terminal domain are represented by Glu30 and Arg27, whose side chains appear largely undefined in the structure of the mutated protein (Fig. 5B). Similarly, the low increase of the thermal factor values of the N-terminal random-coil residues is comparable for both the *Mt*OGT and *Mt*OGT-R37L models (not shown). Thus, our data do not indicate an increased conformational flexibility of this region in the mutated protein, possibly directly interfering with the catalysis.

Regarding the second scenario, seminal studies on human AGT revealed a participation of discrete regions of the N-terminal domain in the cooperative behavior displayed by the protein during DNA recognition (10, 23, 24), and our biochemical characterization shows a severe undermining of *Mt*OGT DNA binding activity upon R37L mutation (Table 1 and Fig. 2); these observations prompted us to speculate that the lack of Arg37 might alter the possibility of establishing optimal protein-protein contacts required for a tight assembly of OGT monomers at the alkylated site.

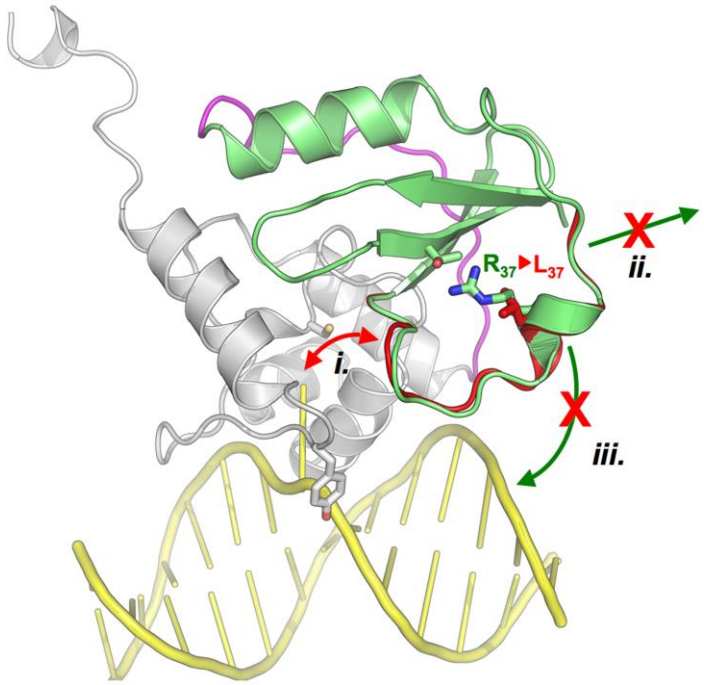


Figure 6 Schematic representation of the possible functional consequences of the replacement of Arg37 with Leu in *Mt*OGT. (i) Direct interference in the dealkylation reaction; (ii) reduced capability of establishing protein-protein interactions and tight monomer packing during cooperative DNA binding; (iii) direct participation of Arg37 in DNA binding. The N-terminal domain, the domain-connecting loop, and the C-terminal domain of the wild-type protein are colored green, magenta, and white, respectively. The *Mt*OGT-R37L structure is uniformly rendered in red; the modeled dsDNA appears in yellow.

Consequently, although we did not observe any significant structural difference between the wild-type and mutated *Mt*OGT, we cannot rule out that a repositioning of the N-terminal random coil could become

evident only in the presence of DNA, therefore affecting the cooperativity of DNA binding. The third hypothesis mentioned above, implying a direct involvement of the Arg37-bearing protein moiety in DNA binding, is strongly suggested by the results of our biochemical analysis but needs further site-directed mutagenesis analysis and structural investigations to be definitively assessed. Overall, our results, besides revealing a high level of structural conservation among DNA-protein alkyltransferases, could provide a framework to orient future studies aimed at an understanding of the molecular mechanism leading to the observed functional defect shown by the analyzed *Mt*OGT variants. In conclusion, we showed that while the catalytic activity of the mutated *Mt*OGT variants toward the synthetic substrate used is not diminished relative to the wild-type protein, the same mutations result in decreased affinity toward the physiological alkylated-DNA substrate. Consequently, the mutants under analysis are less efficient than the wild-type protein in carrying out the complete catalytic cycle required for alkylated-DNA direct repair in vitro. However, due to the critical contribution of the physiological context in determining whether or not a specific mutation in a protein impacts its function, definitive proof of a role of these *Mt*OGT nsSNPs in altering alkylation damage sensitivity and mutation frequency of the corresponding *M. tuberculosis* strain will come only from future *in vivo* studies.

Acknowledgments

This work was supported by the European Community (project SysteMTb HEALTH-F4-2010-241587).

We acknowledge the European Synchrotron Radiation Facility (Grenoble, France) for provision of synchrotron radiation at beam lines ID14 and ID23.

SUPPLEMENTAL MATERIAL

BIOCHEMICAL AND STRUCTURAL STUDIES ON THE *M. tuberculosis* O^6-METHYLGUANINE METHYLTRANSFERASE AND MUTATED VARIANTS

Riccardo Miggiano[a], Valentina Casazza[a], Silvia Garavaglia[a], Maria Ciaramella[b], Giuseppe Perugino[b], Menico Rizzi[a] and Franca Rossi[a].

[a]Dipartimento di Scienze del Farmaco, University of Piemonte Orientale A.Avogadro,Novara,Italy; [b]Istituto di Biochimica delle Proteine, IBP-CNR

Published in Journal of Bacteriology, June 2013; 195 (12): 2728–2736. Epub 5 April 2013

SUPPLEMENTAL METHODS

Expression and purification of wild-type MtOGT and point mutants

The three recombinant proteins *Mt*OGT, *Mt*OGT-T15S and *Mt*OGT-R37L were purified by FPLC (Akta Basic Instrument GE Healthcare)

from 1 L of the corresponding auto-induced bacterial culture adopting the following workflow. 50 mL of clarified lysate in buffer A [Tris-HCl 20mM pH 7.8 and a commercial protease inhibitor cocktail (Sigma)] were loaded onto a HiTrapQ pre-packed anion exchange column. The retained proteins were then eluted using a linear 0-0.5 M NaCl gradient. Recombinant protein-containing fractions were dialysed against buffer A, applied onto a MonoQ pre-packed column (GE Healthcare) and protein elution was obtained as above.

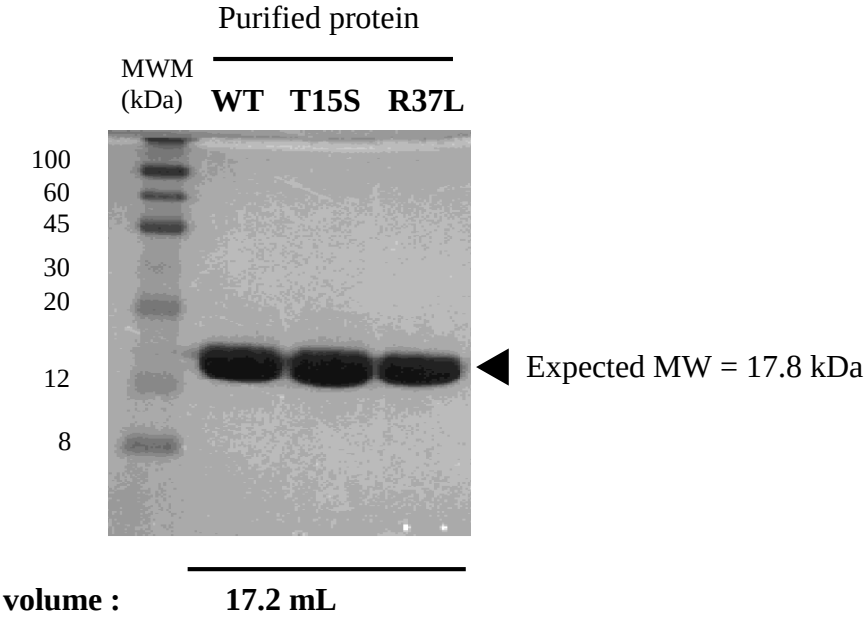
Recombinant protein-containing fractions from the last anion-exchange chromatography step were pooled, desalted by dialysis against buffer A and loaded onto a HiTrap Heparin disposable column (GE Healthcare). Protein elution was obtained by a linear 0-1.0 M NaCl gradient and the eluted fractions containing the recombinant protein were pooled, concentrated to 1 mg/mL by using 10 kDa NMWCO ultrafiltration device (Vivaspin, Vivascience) and loaded onto a Superdex 200 10/300 size exclusion chromatography column (GE Healthcare) using buffer B [20 mM Tris-HCl pH 7.8, 150 mM NaCl] as the mobile phase. Fractions corresponding to the main peak were concentrated to 5 mg/mL by using a 10 kDa NMWCO Vivaspin device. During the entire procedure the recombinant protein was monitored by standard SDS⁄PAGE analyses (example given in Supplementary Figure 1) and protein concentration was determined by the Bradford assay (Sigma). If not immediately used, the purified proteins were stored in aliquots at -80 °C.

Table S1 ***Oligonucleotides used in the present work***

***Mt*OGTfwd**[a]	5'-ATATCCATGGTTCACTACCGCACCATCGATAGC-3' (*Nco* I)
***Mt*OGTrev**[a]	5'-AAAGGATCCTCAGTCGAAGAGCGTCAAGTCTG-3' (*Bam* HI)
T15S fwd[b]	5'-CCCATCGGGCCATTA A**G**C CTGGCCGGGCATGGC-3' ACC
T15S rev	5'-GCCATGCCCGGCCAGGCTTAATGGCCCGATGGG-3'
R37L fwd[c]	5'-CGTATGAGCCAAGC C**T**C ACACACTGGACACCC-3 CGC
R37L rev	5'-GGGTGTCCAGTGTGTGAGGCTTGGCTCATACG-3'
UP^{met} [d]	5'-GGACACTGTACGTTAAGGC**G***ATCGAATTAGGATTAA-3'
DOWN	5'-GGTTAATCCTAATTCGATCGCCTTAACGTACAGTGT-3'
A^{+} oligo[e]	5'-**TMR**-GCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTA-3'
D^{-} oligo	5'-TAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGC-3'

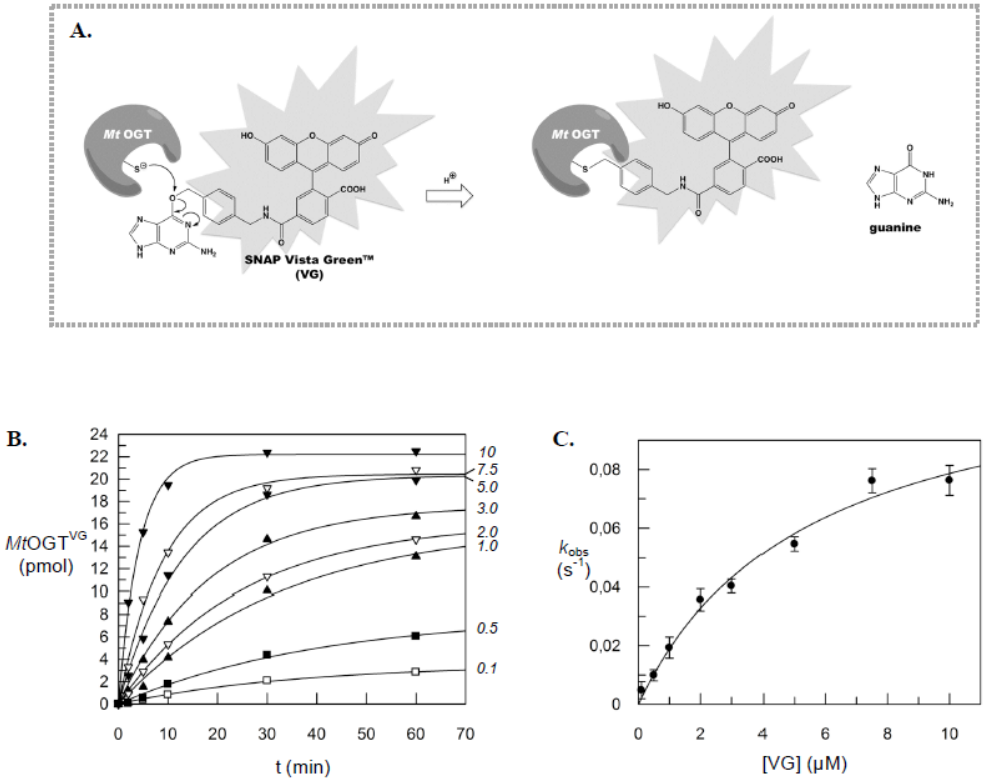
[a] Restriction sites for the endonucleases indicated in brackets are underlined; [b] [c] point-mutation in each oligonucleotide appears in bold; [d] the methylguanine residue in the UP^{met} oligonucleotide is labeled by an asterisk; the ds-DNAmet probe used in enzyme kinetics analysis has been prepared by annealing **UP^{met}** and **DOWN** oligonucleotides [e] "**TMR**" indicates the TAMRA^{TM} fluorescent group; the dsDNA probe used in EMSA analysis has been prepared by annealing **A+** and **D-** oligonucleotides.

Fig. S1



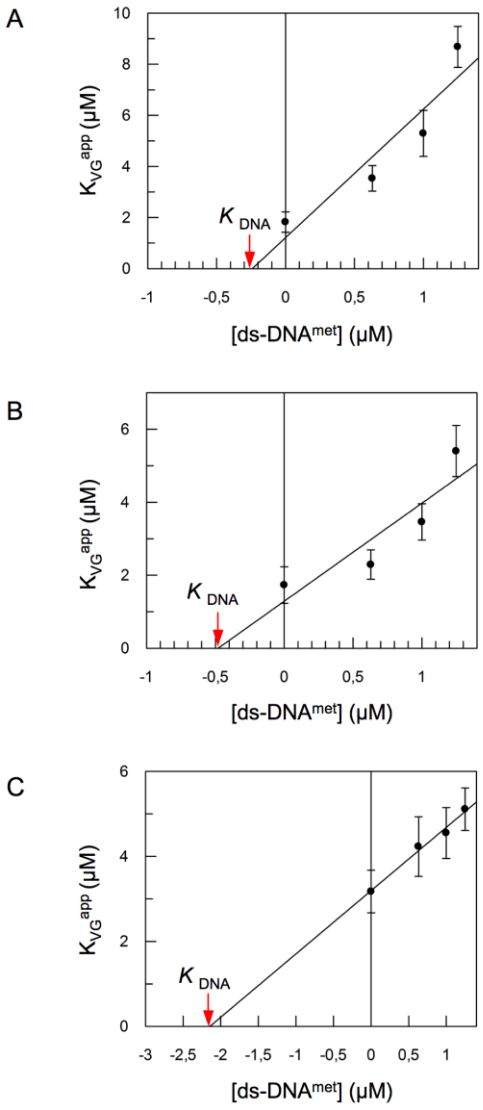
Purification of MtOGT, MtOGT-T15S and MtOGT-R37L. 4 μL of each purified protein (5 mg/mL) were separated using a 15% acrylamide-bisacrylamide gel in SDS-PAGE and visualized by Coomassie brilliant blue staining. The three proteins, independently loaded onto a Superdex 200 10/300 size exclusion chromatography column (GE Healthcare), invariably eluted at Ve=17.2 mL, which corresponds to a MW of 18 kDa, as calculated on the basis of the calibration curve of the column.

Fig. S2



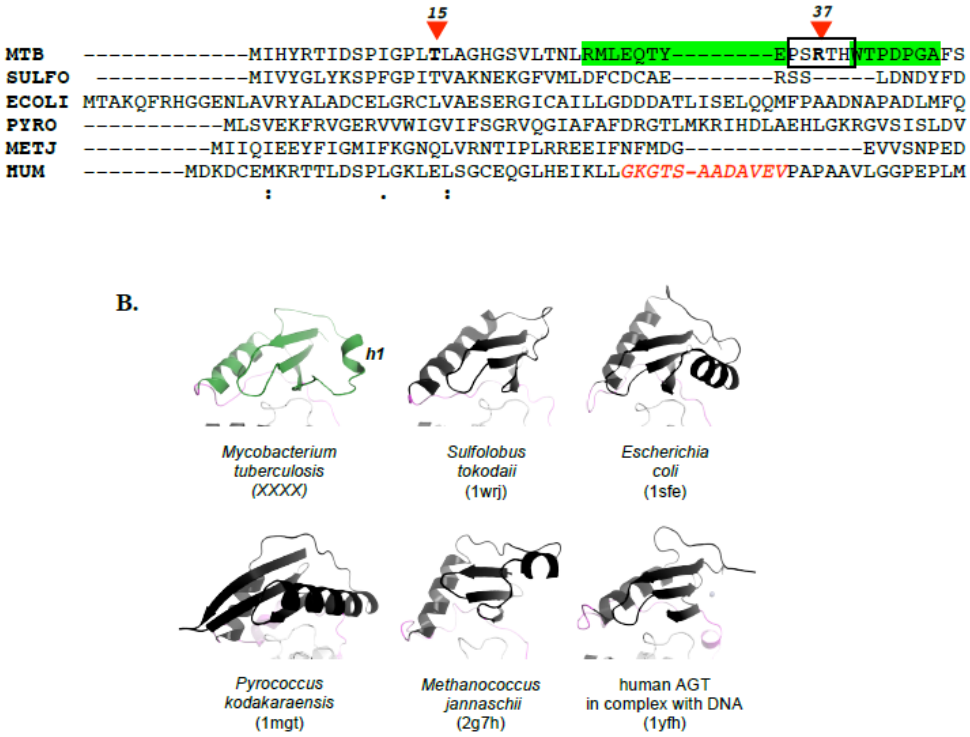
Example of the determination of the kinetic constants of the reaction catalyzed by the recombinant proteins under study. **A.** Scheme of the assay with the fluorescent substrate/inhibitor SNAP-Vista Green™ (VG); **B.** Time-course experiments of covalent modification of 5 μM *Mt*OGT incubated with VG in the range of 0.1-10 μM (VG concentration appears at the right of each corresponding curve) and taking aliquots at 0, 2, 5, 10, 30, 60 min. Similar sets of measurements were carried out for both the *Mt*OGT-T15S and *Mt*OGT-R37L mutated proteins (not shown); **C.** Plot of the first-order rate constants for covalent modification of *Mt*OGT (*k*obs) as a function of SNAP-Vista Green™ (VG) concentration. Rate values obtained in the time course experiments were fitted according to equation **1** (see the "*Materials and Methods*" section of the main text).

Fig. S3



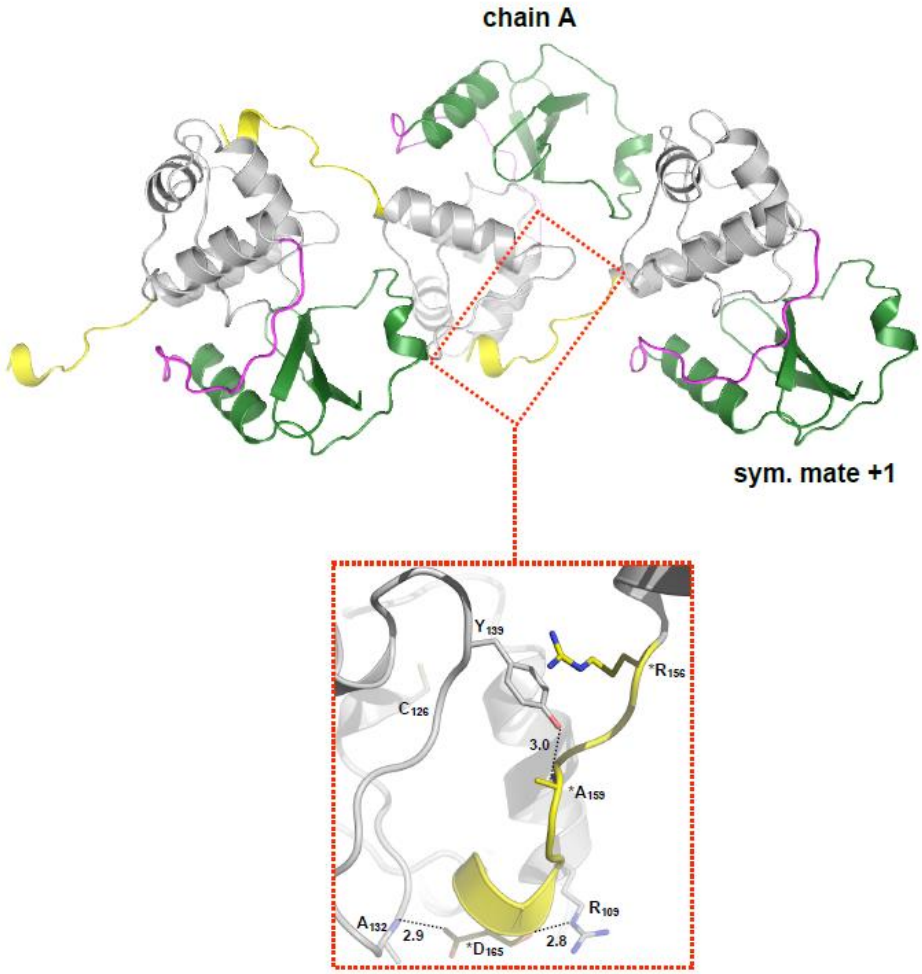
Determination of the DNA dissociation constants of the reaction catalyzed by MtOGT, MtOGT-T15S and MtOGT-R37L. Plots of the variation of *K*VG of *Mt*OGT (A), *Mt*OGT-T15S (B) and *Mt*OGT-R37L (C) as a function of the dsDNA^{met} concentration. Data were fitted according to equation **2** (see the "*Materials and Methods*" section of the main text) that allowed the calculation of the dissociation constant of each enzyme for the dsDNA^{met} ligand (K_{DNA}), indicated by a red arrow in each panel.

Fig. S4



Structural comparison of MtOGT with OGTs from different species. **A.** The figure shows the primary ClustalW-generated sequence alignment of equivalent moieties of the N-terminal domains of OGTs whose structures have been deposited with the protein data bank. **MTB**: *Mycobacterium tuberculosis*; **SULFO**: *Sulfolobus tokodaii*; **ECOLI**: *Escherichia coli*; **PYRO**: *Pyrococcus kodakaraensis*; **METJ**: *Methanococcus jannaschii*; **HUM**: human AGT. The random-coiled region in *M. tuberculosis* OGT is highlighted in green and residues belonging to the h1 helical turn are boxed. The red arrowheads label the residues mutated in *Mt*OGT-T15S and *Mt*OGT-R37L, respectively. Residues rendered in red in the human sequence are not visible in the human AGT structures. **B.** Cartoons of the indicated structures after optimal superposition to the *Mt*OGT model; each panel shows only the N-terminal domain (depicted in green in *Mt*OGT and in black in all other structures) and part of the inter-domain connecting loop (invariably colored in magenta).

Fig. S5



Crystal packing of MtOGT. Close up view of the interacting regions of the C-terminal domain of one chain (gray) and the tail of its symmetry mate (yellow, residues labeled by an asterisk) as observed in the *Mt*OGT crystal lattice (upper image); only the residues involved in the closest interactions are labeled and shown as sticks; bond distances (Å) are indicated as dotted lines.

Bibliography

1**. Gorna AE, Bowater RP, Dziadek J.** 2010. DNA repair systems and the pathogenesis of Mycobacterium tuberculosis: varying activities at different stages of infection. Clin. Sci. (Lond.) **119**:187-202.

2. **Dos Vultos T, Mestre O, Rauzier J, Golec M, Rastogi N, Rasolofo V, Tonjum T, Sola C, Matic I, Gicquel B.** 2008. Evolution and diversity of clonal bacteria: the paradigm of Mycobacterium tuberculosis. PLoS One **3**:e1538. doi:10.1371/journal.pone.0001538.

3. **Mizrahi V, Andersen SJ.** 1998. DNA repair in Mycobacterium tuberculosis. What have we learnt from the genome sequence? Mol. Microbiol. **29**:1331-1339.

4. **Warner DF, Mizrahi V.** 2007. The survival kit of Mycobacterium tuberculosis. Nat. Med. **13**:282-284.

5. **Dos Vultos T, Mestre O, Tonjum T, Gicquel B.** 2009. DNA repair in Mycobacterium tuberculosis revisited. FEMS Microbiol. Rev. **33**:471-487.

6. **Kurthkoti K, Varshney U.** 2012. Distinct mechanisms of DNA repair in mycobacteria and their implications in attenuation of the pathogen growth. Mech. Ageing Dev. **133**:138-146.

7. **Voskuil MI, Bartek IL, Visconti K, Schoolnik GK.** 2011. The response of Mycobacterium tuberculosis to reactive oxygen and nitrogen species. Front. Microbiol. **2**:105. doi:10.3389/fmicb.2011.00105.

8. **Dalhus B, Laerdahl JK, Backe PH, Bjoras M.** 2009. DNA base repair - recognition and initiation of catalysis. FEMS Microbiol. Rev. **33**:1044-1078.

9. **Shrivastav N, Li D, Essigmann JM.** 2010. Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. Carcinogenesis **31**:59-70.

10. **Pegg AE.** 2011. Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools. Chem. Res. Toxicol. **24**:618-639.

11. **Kanugula S, Goodtzova K, Pegg AE.** 1998. Probing of conformational changes in human O6-alkylguanine-DNA alkyltransferase protein in its alkylated and DNA-bound states by limited proteolysis. Biochem. J. **329**: 545-550.

12. **Xu-Welliver M, Pegg AE.** 2002. Degradation of the alkylated form of the DNA repair protein, O(6)-alkylguanine-DNA alkyltransferase. Carcinogenesis **23**:823-830.

13. **Samson L, Cairns J.** 1977. A new pathway for DNA repair in Escherichia coli. Nature **267**:281–283.

14. **Yang M, Aamodt RM, Dalhus B, Balasingham S, Helle I, Andersen P, Tonjum T, Alseth I, Rognes T, Bjoras M.** 2011. The ada operon of Mycobacterium tuberculosis encodes two DNA methyltransferases for inducible repair of DNA alkylation damage. DNA Repair (Amst.) **10**:595-602.

15. **Sassetti CM, Boyd DH, Rubin EJ.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol. Microbiol. **48**:77-84.

16. **Durbach SI, Springer B, Machowski EE, North RJ, Papavinasasundaram KG, Colston MJ, Bottger EC, Mizrahi V.** 2003. DNA alkylation damage as a sensor of nitrosative stress in Mycobacterium tuberculosis. Infect. Immun. **71**:997-1000.

17. **Boshoff HI, Myers TG, Copp BR, McNeil MR, Wilson MA, Barry CE, III.** 2004. The transcriptional responses of Mycobacterium tuberculosis to inhibitors of metabolism: novel insights into drug mechanisms of action. J. Biol. Chem. **279**:40174–40184.

18. **Schnappinger D, Ehrt S, Voskuil MI, Liu Y, Mangan JA, Monahan IM, Dolganov G, Efron B, Butcher PD, Nathan C, Schoolnik GK.** 2003. Transcriptional adaptation of Mycobacterium tuberculosis within macrophages: insights into the phagosomal environment. J. Exp. Med. **198**:693-704.

19. **Ebrahimi-Rad M, Bifani P, Martin C, Kremer K, Samper S, Rauzier J, Kreiswirth B, Blazquez J, Jouan M, van Soolingen D, Gicquel B.** 2003. Mutations in putative mutator genes of Mycobacterium tuberculosis strains of the W-Beijing family. Emerg. Infect. Dis. **9**:838-845.

20. **Olano J, Lopez B, Reyes A, Lemos MP, Correa N, Del Portillo P, Barrera L, Robledo J, Ritacco V, Zambrano MM.** 2007. Mutations in DNA repair genes are associated with the Haarlem lineage of Mycobacterium tuberculosis independently of their antibiotic resistance. Tuberculosis (Edinb.) **87**:502–508.

21. **Mestre O, Luo T, Dos Vultos T, Kremer K, Murray A, Namouchi A, Jackson C, Rauzier J, Bifani P, Warren R, Rasolofo V, Mei J, Gao Q, Gicquel B.** 2011. Phylogeny of Mycobacterium tuberculosis Beijing strains constructed from polymorphisms in genes involved in DNA replication, recombination and repair. PLoS One **6**:e16020. doi:10.1371/journal.pone.0016020.

22. **Fang Q, Kanugula S, Pegg AE.** 2005. Function of domains of human O6-alkylguanine-DNA alkyltransferase. Biochemistry **44**:15396-15405.

23. **Adams CA, Melikishvili M, Rodgers DW, Rasimas JJ, Pegg AE, Fried MG.** 2009. Topologies of complexes containing O6-alkylguanine-DNA alkyltransferase and DNA. J. Mol. Biol. **389**:248-263.

24. **Adams CA, Fried MG.** 2011. Mutations that probe the cooperative assembly of O6-alkylguanine-DNA alkyltransferase complexes. Biochemistry **50**:1590-1598.

25. **Daniels DS, Woo TT, Luu KX, Noll DM, Clarke ND, Pegg AE, Tainer JA.** 2004. DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nat. Struct. Mol. Biol. **11**:714-720.

26. **Perugino G, Vettone A, Illiano G, Valenti A, Ferrara MC, Rossi M, Ciaramella M.** 2012. Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. J. Biol. Chem. **287**:4222-4231.

27. **Sambrook J, Fritsh EF, Maniatis T.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

28. **Studier FW.** 2005. Protein production by auto-induction in high-density shaking cultures. Protein Expr. Purif. **41**:207-234.

29. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685.

30. **Bradford MM.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248-254.

31. **Collaborative Computational Project Number 4.** 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50**:760-763.

32. **McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ.** 2007. Phaser crystallographic software. J. Appl. Crystallogr. **40**: 658-674.

33. **Moore MH, Gulbis JM, Dodson EJ, Demple B, Moody PC.** 1994. Crystal structure of a suicidal DNA repair protein: the Ada O6-methylguanine-DNA methyltransferase from E. coli. EMBO J. **13**:1495-1501.

34. **Perrakis A, Morris R, Lamzin VS.** 1999. Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. **6**:458-463.

35. **Emsley P, Cowtan K.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60**:2126-2132.

36. **Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH.** 2010. PHENIX: a comprehensive Pythonbased system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. **66**:213-221.

37. **DeLano WL.** 2002. The PyMOL molecular graphics system. DeLano Scientific, Palo Alto, CA. http://www.pymol.org/.

38. **Spratt TE, Wu JD, Levy DE, Kanugula S, Pegg AE.** 1999. Reaction and binding of oligodeoxynucleotides containing analogues of O6- methylguanine with wild-type and mutant human O6-alkylguanine-DNA alkyltransferase. Biochemistry **38**:6801-6806.

39. **Laskowski RA, MacArthur MW, Moss DS, Thornton JM.** 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**:283-291.

40. **Hashimoto H, Inoue T, Nishioka M, Fujiwara S, Takagi M, Imanaka T, Kai Y.** 1999. Hyperthermostable protein structure maintained by intra and inter-helix ion-pairs in archaeal O6-methylguanine-DNA methyltransferase. J. Mol. Biol. **292**:707-716.

41. **Roberts A, Pelton JG, Wemmer DE.** 2006. Structural studies of MJ1529, an O6-methylguanine-DNA methyltransferase. Magn. Reson. Chem. **44**(Spec No):S71–S82. doi:10.1002/mrc.1823.

42. **Wibley JE, Pegg AE, Moody PC.** 2000. Crystal structure of the human O(6)-alkylguanine-DNA alkyltransferase. Nucleic Acids Res. **28**:393- 401.

43. **Daniels DS, Mol CD, Arvai AS, Kanugula S, Pegg AE, Tainer JA.** 2000. Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding. EMBO J. **19**:1719-1730.

44. **Duguid EM, Rice PA, He C.** 2005. The structure of the human AGT protein bound to DNA and its implications for damage detection. J. Mol. Biol. **350**:657-666.

45. **Crone TM, Goodtzova K, Pegg AE.** 1996. Amino acid residues affecting the activity and stability of human O6-alkylguanine-DNA alkyltransferase. Mutat. Res. **363**:15-25.

46. **Xu-Welliver M, Pegg AE.** 2000. Point mutations at multiple sites including highly conserved amino acids maintain activity, but render O6- alkylguanine-DNA alkyltransferase insensitive to O6-benzylguanine. Biochem. J. **347**:519-526.

47. **Margison GP, Butt A, Pearson SJ, Wharton S, Watson AJ, Marriott A, Caetano CM, Hollins JJ, Rukazenkova N, Begum G, Santibáñez-Koref MF.** 2007. Alkyltransferase-like proteins. DNA Repair **6**:1222-1228.

48. **Tubbs JL, Latypov V, Kanugula S, Butt A, Melikishvili M, Kraehenbuehl R, Fleck O, Marriott A, Watson AJ, Verbeek B, McGown G, Thorncroft M, Santibanez-Koref MF, Millington C, Arvai AS, Kroeger MD, Peterson LA, Williams DM, Fried MG, Margison GP, Pegg AE, Tainer JA.** 2009. Flipping of alkylated DNA

damage bridges base and nucleotide excision repair. Nature **459**:808-813.

49. **Wilkinson OJ, Latypov V, Tubbs JL, Millington CL, Morita R, Blackburn H, Marriott A, McGown G, Thorncroft M, Watson AJ, Connolly BA, Grasby JA, Masui R, Hunter CA, Tainer JA, Margison GP, Williams DM.** 2012. Alkyltransferase-like protein (Atl1) distinguishes alkylated guanines for DNA repair using cation-pi interactions. Proc. Natl. Acad. Sci. U. S. A. **109**:18755-18760.

Site-directed mutagenesis analysis of *Mycobacterium tuberculosis* O^6-methylguanine methyltransferase outlines the functional role of catalysis-coordinating residues.

Riccardo Miggiano[a], Maria Ciaramella[b], Giuseppe Perugino[b], Menico Rizzi[a] and Franca Rossi[a].

[a]Dipartimento di Scienze del Farmaco, University of Piemonte Orientale "A. Avogadro", Novara, Italy; [b]Institute of Biosciences and BioResources, IBBR-CNR

Unpublished results

Summary

In *Mycobacterium tuberculosis* the O^6-methylguanine methyltransferase protein (OGT/AdaB) plays a main role in protecting the bacterial GC-rich genome from the pro-mutagenic potential of O^6-methylated guanine. Interestingly, several multi-drug resistant *M. tuberculosis* strains show nonsynonymous single-nucleotide polymorphisms (nsSNP) in their OGT-encoding sequence (T15S and R37L). Although the functional consequences of such point-mutated OGT variants on the biology of the corresponding *M. tuberculosis* strains remain undetermined, we recently demonstrated *in vitro* that the recombinant mutated proteins (*i.e. Mt*OGT-T15S and *Mt*OGT-R37L) display an

impaired activity, with *Mt*OGT-R37L showing a 10-fold lower affinity for double stranded DNA containing O^6-methylguanine with respect to the wild-type protein. Our concomitant structural investigations provided us with possible scenarios explaining the molecular bases of such a catalytic defect, and added elements of novelty to currently available structural description of this class of protein, namely the conformation adopted in *Mt*OGT by the active site loop.

We describe here the biochemical characterization of two structure-based designed point-mutated variants of *Mt*OGT (*i.e.* *Mt*OGT-R37K and *Mt*OGT-R37E) aimed at better understanding the functional role of Arginine 37 during the catalysis. Similar analyses on the *Mt*OGT-Y139F - a *Mt*OGT version mutated at the level of a highly conserved tyrosine of the protein active site loop-, confirms the fundamental role of this residue in catalysis, although disclosing an unprecedented mode of action.

Introduction

Mycobacterium tuberculosis (MTB), the causative agent of tuberculosis (TB), displays a remarkable genetic stability (1) despite the continuous exposure to pro-mutagenic and genotoxic stresses that could compromise both the persistence in the host as well as the reactivation of the bacillus from the dormant state (2). In particular, during its long-term stay inside the infected macrophages, MTB must deal with potent

endogenous DNA-alkylating chemical species originating by the action of highly reactive oxidative and nitrosative radicals (3). Similarly to what observed in other organisms, *M. tuberculosis* attains alkylated-base repair either using multi-enzymatic systems or by directly removing the alkyl adduct from the modified base by sacrificing a molecule of a DNA-protein alkyltransferase (4, 5). O^6-methylguanine methyltransferase (OGT) (E.C. 2.1.1.63) acts through a suicidal mechanism (6) by performing the irreversible transfer of the O^6-alkyl group from the modified guanine to a strictly conserved cysteine residue buried in its active site, which is hosted in the C-terminal domain of the protein. Ground-breaking X-rays crystallography studies on the human ortholog O^6-alkylguanine DNA alkyltransferase (AGT) (6, 7) in complex with double stranded olignucleotides, containing either the physiological substrate O^6-methylguanine or substrate analogues (8, 9), revealed the molecular details of the protein association with the alkylated double stranded DNA (dsDNA). These results showed that human AGT - as well as alkyl-transferase like proteins (10) - binds the substrate DNA through a novel mechanism, by establishing contacts between a canonical helix-turn-helix motif in the C-terminal domain of the protein and the DNA minor groove. The alkylated guanine residue appears flipped out from the regular base stacking and housed into the active site pocket, thus resulting correctly positioned with respect to the catalytic cysteine (7, 8, 9). The protein-substrate complex is further stabilized by residues of the active site loop, a crucial structural element of the C-

terminal domain that has been invariably observed in a "pointing inwards" conformation in all other ligand-free or DNA-bound OGT crystal structures.

As in most bacterial species, mycobacterial *ogt* gene (*adaB*; Rv1316c) is part of an adaptive response operon that has been recently fully functionally characterized, confirming its role in protecting mycobacterial genome against the pro-mutagenic potential of DNA alkylation (11). Several lines of experimental evidences support this observation: i) the OGT gene expression profile changes during infection and in response to pro-alkylating compounds (11, 12, 13); ii) the hyper-mutator phenotype displayed by a MNNG (N-methyl-N'-nitro-N-nitrosoguanidine)-treated MTB strain lacking the *ogt* operon could be rescued by gene trans-complementation (14); and iii) the trans-specific knock-in of an Ada/OGT-defective *Escherichia coli* strain with the *M. tuberculosis ogt* gene alone, suppresses its alkylation damage sensitivity and restores mutation frequencies to wild-type levels (11). Notably, a number of geographically widely distributed MTB W-Beijing strains and multi-drug resistant isolates show nonsynonymous SNPs in their OGT-encoding sequence, resulting in amino acid substitutions at the poorly functionally characterized N-terminal domain of the enzyme (15, 16, 17).

In order to reveal a possible direct impact of these mutations (*i.e.* threonine 15 and arginine 37 replaced by serine and leucine, respectively) on *Mt*OGT protein function, we recently performed the

biochemical and structural characterization of the *Mt*OGT-T15S and *Mt*OGT-R37L recombinant proteins that mimic the ones occurring in clinical isolates (18). Interestingly, the *Mt*OGT-R37L mutant displays a severely impaired catalytic activity as a consequence of its ten-fold lower affinity for methylated double-stranded DNA with respect to the wild-type protein. The high-resolution crystal structures of both *Mt*OGT and *Mt*OGT-R37L provided clues to understand the molecular bases of the observed catalytic defect showed by the *Mt*OGT-R37L variant, that was suggested to be possibly due to an unexpected participation of Arg37 to direct DNA binding and/or to its role in the cooperative protein assembly at the alkylated-DNA site during substrate recognition. Our analysis also revealed that, notwithstanding the high degree of overall structural conservation at the level of the catalytic C-terminal domain of members of the alkyltransferase protein family, in both the wild-type and mutated *Mt*OGT crystal structures the active site loop invariantly adopts a "pointing outwards" conformation, thus exposing a highly conserved tyrosine residue (Y139 in *Mt*OGT) to the bulk solvent (18).

By adopting an experimental approach that combines biochemical and X-rays crystallography studies, in the present work we carried out the determination of the enzyme kinetics and the structural characterization of a new panel of recombinant *Mt*OGT point-mutated variants that were specifically designed to further confirm our previous conclusions and clarify the issues left unanswered. In particular, our current results underline the requirement of a positive charged residue at position 37 of

the N-terminal domain of the protein, corroborating the hypothesis that sees Arg37 directly involved in protein assembling and/or direct DNA binding during catalysis. Moreover, the crystal structure of *Mt*OGT-Y139F confirms the peculiar conformation of the active site loop in *Mt*OGT and allow us to propose a specific role for Y139 in *Mt*OGT-performed alkylated-DNA repair.

Materials and methods

Chemicals. All reagents were obtained from Sigma-Aldrich unless otherwise specified.

Construction of pET-*Mt*OGT-R37K, pET-*Mt*OGT-R37E and pET-*Mt*OGT-Y139F expression vectors - The pET-*Mt*OGT construct coding for the wild-type *M. tuberculosis* O^6-methylguanine methyltransferase (*orf*: Rv1316c) (19) was invariably used as the DNA template in PCR-based site-directed mutagenesis experiments, using the QuikChange II site-directed mutagenesis kit reagents (Stratagene) and the three primer pairs R37Kfwd/R37Krev, R37Efwd/R37Erev and Y139Ffwd/Y139rev (Table 1). The region encoding the corresponding point-mutated OGT variant in each resulting expression construct (namely pET-*Mt*OGT-R37K, pET-*Mt*OGT-R37E and pET-*Mt*OGT-Y139F) was verified by sequencing (Eurofins MWG Operon).

Table 1- ***Oligonucleotides used in the present work.***

R37K fwd[(a)]	5'- GACGTATGAGCCAAGC**AAG**ACACACTGGACACCCGACC -3'
R37K rev	5'- GGTCGGGTGTCCAGTGTGTCTTGCTTGGCTCATACGTC -3'
R37E fwd[(a)]	5'- GACGTATGAGCCAAGC**GAG**ACACACTGGACACCCGACC-3
R37E rev	5'- GGTCGGGTGTCCAGTGTGTCTCGCTTGGCTCATACGTC-3'
Y139F fwd[(a)]	5'- CGGAAAGCTCACCGGGT**T**CGGCGGTGGAATCAAC-3'
Y139Frev	5'- GGTTGATTCCACCGCCGAACCCGGTGAGCTTTCCG-3'
UP^{met} [(b)]	5'- GGACACTGTACGTTAAGGC**G***ATCGAATTAGGATTAA-3'
DOWN	5'- GGTTAATCCTAATTCGATCGCCTTAACGTACAGTGT-3'
A^{+} oligo[(c)]	5'-**TMR**-GCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTA-3'
D^{-} oligo	5'-TAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGC-3'

(a) point-mutation in the "sense" oligonucleotide appears in bold; (b) the methylguanine residue is labeled by an asterisk; (c) "TMR" indicates the TAMRA™ fluorescent group.

Expression and purification of point-mutated *Mt*OGT variants -

The procedure adopted to express and purify the *Mt*OGT-R37K, *Mt*OGT-R37E and *Mt*OGT-Y139F point-mutated versions of *Mt*OGT used in biochemical and structural analyses was previously reported (18).

Activity assays and enzyme kinetics analysis - To measure the alkyltransferase activity of wild-type *Mt*OGT and its point-mutated variants *Mt*OGT-R37K, *Mt*OGT-R37E and *Mt*OGT-Y139F, a fluorescence assay using SNAP-Vista Green reagent (New England BioLabs) (here referred to as VG for brevity) was used, as previously described (20). Protein bands from SDS-PAGE were visualized by direct gel imaging (VersaDoc 4000; Bio-Rad), and their fluorescence

intensities (FIs), corresponding to the reacted protein, were corrected for the actual amount of protein loaded into each lane by measuring the intensity of bands after Coomassie brilliant blue R250 staining. Activity assays at different pH values were performed with 1× Fluo reaction buffer containing 50.0mM different buffering species, allowing the determination of the pH optimum of the reaction used in all subsequent experiments (data not shown). Assuming a 1:1 substrate-enzyme stoichiometry and irreversible binding, incubation of a fixed amount of protein (5.0 μM) with VG (in the range of 0.1 to 20.0 μM) for 1 h at 25°C gave corrected FI data, which were fitted by a linear equation whose slope gave a direct reference value of FI/μM protein, which was then used to estimate the amount of covalently modified protein (expressed in picomoles) in time course experiments. The latter assays were performed at 25°C using different protein/VG molar ratios and taking 10 μl aliquots at different time intervals. Plots of the picomoles of reacted protein versus time were fitted by exponential equations to determine the apparent rates for covalent modification (k_{obs}). These values were plotted versus the VG doses ([VG]) by using the following hyperbolic equation:

$$k_{\mathrm{obs}} = \frac{k}{1+\mathrm{K_{VG}}/[\mathrm{VG}]} \quad (1)$$

where k and K_{VG} are the rate of covalent linkage and the dissociation constant for the free enzyme and free VG reagent during the first collision step (before covalent modification), respectively. In order to

calculate the apparent K_{VG} values (K_{VG}^{app}) of proteins for VG as a function of the double-stranded methylated DNA ($dsDNA^{met}$) concentration, time course experiments such as those described above were performed in the presence of increasing fixed doses of $dsDNA^{met}$. Data were fitted according to the following linear equation:

$$K_{VG}^{app} = \frac{K_{VG}}{K_{DNA}} [dsDNA^{met}] + K_{VG} \quad (2)$$

where K_{DNA} is the dissociation constant of the protein for the $dsDNA^{met}$ substrate. The results of these enzyme kinetics analyses are summarized in Table 2.

Electrophoretic mobility shift assay (EMSA) - A 40-bp 6-carboxytetramethylrhodamine (TAMRA)-labeled double-stranded DNA probe (0.1 µM) was incubated at 25°C for 10 min with different amounts of proteins in the range of 0.0 to 120.0 µM in a total volume of 10.0 µl in 1× binding buffer (20.0 mM Tris-HCl [pH 7.5], 50.0 mM KCl, 0.1 mM dithiothreitol [DTT], and 10% glycerol). Samples were loaded onto an 8% polyacrylamide-bisacrylamide native gel in 1× Tris-borate-EDTA (TBE). Signals were visualized by direct gel imaging using a green light-emitting diode (LED)/605-nm-band-pass filter as excitation/emission parameters, respectively.

Crystallization - Crystals of either *Mt*OGT-R37K or *Mt*OGT-Y139F were grown by the hanging drop vapor diffusion method mixing 2 µl of the corresponding protein solution at 5 mg/ml with an equal volume of a reservoir solution containing 0.1 M HEPES (pH 7.5), 4% (wt/vol) polyethylene glycol 8000 (PEG 8000), and 4% (wt/vol) ethylene glycol for MtOGT-R37K or 0.1 M HEPES (pH 7.5), 4% (wt/vol) polyethylene glycol 8000 (PEG 8000), and 8% (wt/vol) ethylene glycol for *Mt*OGT-Y139F. The drop was equilibrated against 800 µl of the reservoir solution at 4°. Crystals used in X-ray diffraction experiments grew to maximum dimensions of 0.2 mm in about 2 weeks.

Structure determination - For X-ray data collection, crystals of both *Mt*OGT-R37K and *Mt*OGT-Y139F were taken directly from the crystallization droplet, rapidly equilibrated in a solution containing the crystallization buffer and 15% glycerol as a cryoprotectant, and flash-frozen at 100 K under liquid nitrogen. A complete data set of the *Mt*OGT-R37K crystal was collected at 100 K to a 2.3-Å resolution using synchrotron radiation (λ= 0.979 Å) at the ID14-EH4 beam line (European Synchrotron Radiation Facility [ESRF], Grenoble, France), equipped with an ADSC detector. Analysis of the collected diffraction data set allowed us to assign the crystal to the orthorhombic space group $P2_12_12$, with the cell dimensions a = 59.10 Å, b = 81.88 Å, and c = 38.27 Å, containing one molecule per asymmetric unit, with a corresponding solvent content of 52%. For all data collections, diffraction intensities

were evaluated, integrated, and scaled by using the CCP4 suite of programs (21). The structure of *Mt*OGT-R37K was solved by molecular replacement using the program Phaser (22) and the structure of *M. tuberculosis* O^6-methylguanine-DNA methyltransferase as the search model (18) (Protein Data Bank [PDB] accession number 4BHB) omitting water/ligand molecules and Arg37 side-chain. The resulting electron density map was of high quality and allowed automatic tracing by the program ARP/wARP (23), which was also used for adding solvent molecules. The program Coot (24) was used for manual rebuilding, and the program PHENIX (25) was used for crystallographic refinement.

A complete data set of the *Mt*OGT-Y139F crystal was collected at 100 K to a 2.6-Å resolution using synchrotron radiation (λ= 0.99 Å) at the ESRF ID14-EH4 beam line equipped with an ADSC detector.

The structure of *Mt*OGT-Y139F was solved by molecular replacement using the program Phaser (22) and the structure of *M. tuberculosis* O_6-methylguanine-DNA methyltransferase as the search model (Protein Data Bank [PDB] accession number 4BHB) (in which water/ligand molecules and Tyr139 side-chain atoms were omitted).

Model building and crystallographic refinement were performed following the same procedure described above for *Mt*OGT-R37K. Data collection, phasing, and refinement statistics are given in Table 2. Structural superpositions were performed with the Superpose program of the CCP4 suite (21). Figures illustrating the structural analysis were generated with PyMol (26).

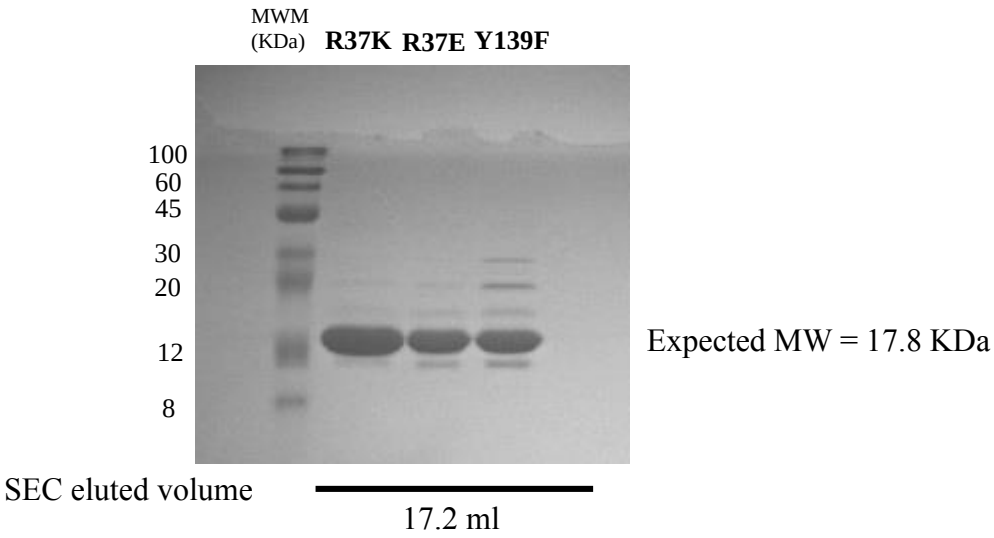
Results and Discussion

Expression and purification of *Mt*OGT-R37K, *Mt*OGT-R37E and *Mt*OGT-Y139F - As described in the "*Introduction*" section, previous results from the *in vitro* characterization of *Mt*OGT and its mutated variants mimicking the ones occurring in clinically relevant MTB strains (18) highlighted the requirement of the R37 residue for optimal catalysis. This result was quite unexpected, since R37 belongs to the OGT N-terminal domain that has never been reported to assist the removal of the alkyl group from the modified base or to directly participate in DNA binding. The second element of novelty disclosed from those studies consisted in the observation of the solvent-exposed conformation adopted by the Y139 residue of the active site loop that has never been observed before in the crystal structure of ligand free or DNA-bound OGTs of other species (8, 9, 27, 28, 29).

Therefore, to specifically investigate the molecular role of R37 and Y139 in *Mt*OGT, we constructed the pET-*Mt*OGT-R37K, the pET-*Mt*OGT-R37E and pET-*Mt*OGT-Y139F expression plasmids, which drive the synthesis in *E. coli* of the three (untagged) OGT variants that could be reproducibly purified at high yield by FPLC-based standard techniques. We obtain pure and highly homogeneous preparations of the three mutated proteins that, as already demonstrated for the wild-type *Mt*OGT (18), behave as monomers in solution (Figure 1); this observation confirms that the substitution of R37 by a lysine or a glutamate residue

(in *Mt*OGT-R37K and *Mt*OGT-R37E, respectively) and the replacement of Y139 by a phenylalanine (in *Mt*OGT-Y139F) do not affect the overall stability of the corresponding mutated protein and do not alter its quaternary structure in the absence of ligands.

Figure 1



Purification of MtOGT-R37K, MtOGT-R37E and MtOGT-Y139F. 5 µL of each purified protein (5 mg/mL) were separated using a 15% acrylamide-bisacrylamide gel in SDS-PAGE and visualized by Coomassie brilliant blue staining. The three proteins, independently loaded onto a Superdex 200 10/300 size exclusion chromatography column (GE Healthcare), invariably eluted at elution volume of 17.2 mL, which corresponds to a MW of 18 kDa, as calculated on the basis of the calibration curve of the column.

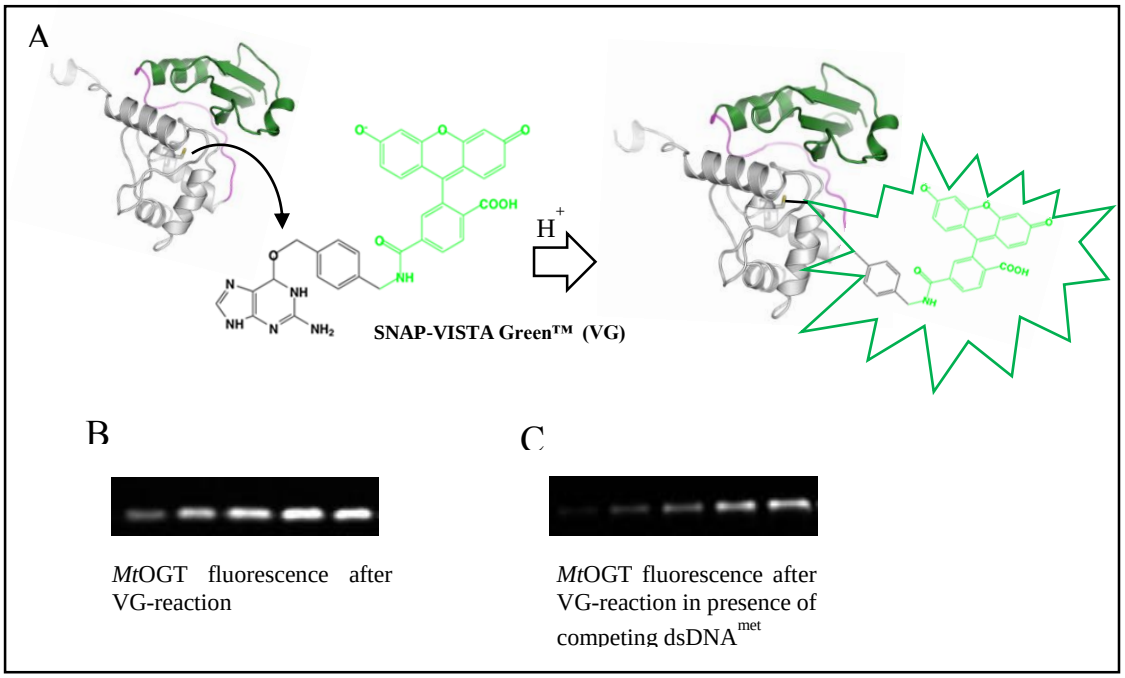
Enyme kinetics study of *Mt*OGT-R37K, *Mt*OGT-R37E, and *Mt*OGT-Y139F.

By adopting a biochemical assay based on the SNAP-VISTA Green™ molecule (VG), a fluorescent derivative of O^6-benzylguanine (18, 20), we determined the kinetic constants for the covalent modification of the three point-mutated variants of *Mt*OGT in the presence of increasing amounts of a double-stranded DNA fragment containing an internal O^6-methylated guanine ($dsDNA^{met}$), which represents the physiologic substrate of OGT and consequently competes with the VG reagent during the trans-alkylation reaction (*i.e.* from the modified base to the catalytic cysteine of the enzyme, Figure 2). According to equation 2 in the "*Material and Methods*" section, a linear plot of the K_{VG} values as a function of the $dsDNA^{met}$ concentration allowed the calculation of the dissociation constant (K_{DNA}) of *Mt*OGT-R37K, *Mt*OGT-R37E and *Mt*OGT-Y139F for $dsDNA^{met}$ (Table 2).

Our data show that the *Mt*OGT-R37E point mutant exhibits a five-fold lower affinity for $dsDNA^{met}$ (K_{DNA}= 1.14 ± 0.15 μM) with respect to *Mt*OGT (K_{DNA}= 0.24 ± 0.11 μM), while the replacement of Arg37 by a lysine residue does not produce such a significant effect; indeed the $dsDNA^{met}$ binding affinity of *Mt*OGT-R37K is in the range of the one displayed by the wild-type protein (K_{DNA}= 0.38 ± 0.2 μM).

These kinetics data corroborate our working hypothesis that the functional role of Arg37 could be to providing a positively charged

Figure 2



Fluorescence-based assay used for the characterization of MtOGT enzyme kinetics. (A) Scheme of the transalkylation reaction of *Mt*OGT with the fluorescent substrate/inhibitor SNAP-VISTA Green™ (VG); (B) SDS-PAGE analysis of increasing amount of *Mt*OGT covalently bound to the VG. (C) SDS-PAGE analysis of *Mt*OGT incubated in the presence of both VG and dsDNA^{met} that competes with VG for alkyltransferase reaction.

group at this specific position of the *Mt*OGT N-terminal domain, that would be required for establishing optimal protein-protein and/or protein-DNA contacts during substrate recognition.

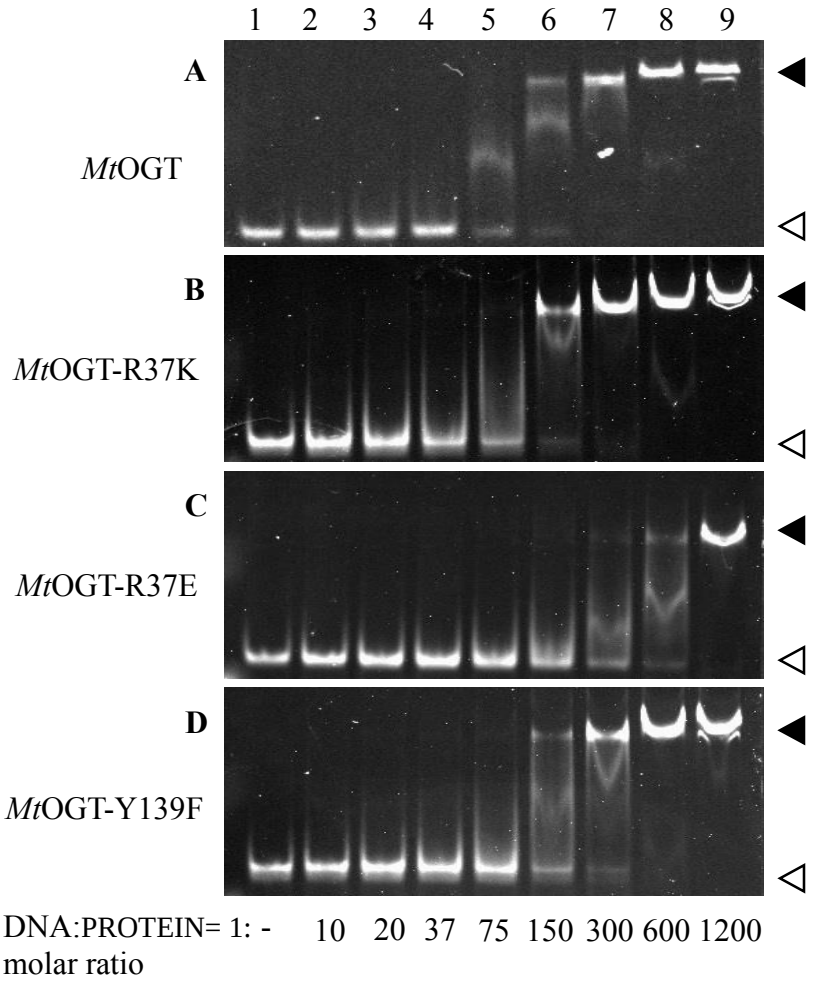
Considering the *Mt*OGT-Y139F mutant, its ten-fold lower affinity for dsDNA^{met} (K_{DNA}= 2.19 ± 0.5 µM) than the wild type enzyme confirms the crucial role of Tyr139 of the active site loop in *Mt*OGT-mediated

repairing of methylated guanine in dsDNA, as already fully established for the equivalent residue in other OGTs. (30, 31, 32, 33).

Table 2- Kinetic constants of the reaction catalysed by *Mt*OGT (18) and corresponding mutants.

	ds-DNA^{met} (μM)	Mean k (s^{-1}) ± SD	Mean K_{VG} (μM) ± SD	Mean K_{DNA} (μM) ± SD
***Mt*OGT**	0	0.12 ± 0.02	1.82 ± 0.4	**0.24 ± 0.11**
	0.63	0.14 ± 0.02	3.53 ± 0.5	
	1	0.10 ± 0.02	5.29 ± 0.9	
	1.25	0.14 ± 0.02	8.68 ± 0.8	
***Mt*OGT-R37K**	0	0.07 ± 0.01	0.42 ± 0.1	**0.38 ± 0.2**
	0.63	0.07 ± 0.01	0.88 ± 0.2	
	1	0.07 ± 0.01	0.96 ± 0.5	
	1.25	0.06 ± 0.01	1.43 ± 0.5	
***Mt*OGT-R37E**	0	0.03 ± 0.01	1.72 ± 0.4	**1.14 ± 0.15**
	0.63	0.04 ± 0.01	2.81 ± 0.6	
	1	0.05 ± 0.01	3.47 ± 0.2	
	1.25	0.04 ± 0.02	3.61 ± 0.4	
***Mt*OGT-Y139F**	0	0.02 ± 0.02	0.69 ± 0.3	**2.19 ± 0.5**
	0.63	0.03 ± 0.02	0.73 ± 0.4	
	1	0.02 ± 0.01	0.93 ± 0.4	
	1.25	0.02 ± 0.02	1.05 ± 0.5	

Figure 3



Band shift analysis of MtOGT (wild type) (18), MtOGT-R37K, MtOGT-R37E and MtOGT-Y139F proteins. Lane 1, 1 pmol of fluorescent double-stranded DNA probe (DNA); lanes 2 to 9, increasing amounts of protein (P) incubated in the presence of the probe at the indicated DNA/protein molar ratio.

Direct DNA binding analysis of *Mt*OGT-R37K, *Mt*OGT-R37E, and *Mt*OGT-Y139F.

The enzyme kinetics characterization reported in the previous section shows that a non-conservative substitution of Arg37 - as observed in the recombinant *Mt*OGT-R37L protein mimicking the one characterizing a number of frequently isolated MTB strains as well as in the *Mt*OGT-R37E variant described in the present study- results in a catalytically inactive *Mt*OGT (Table 2) (18). To understand whether this defect could be attributable to a reduced capability of the protein to bind the substrate, we performed an EMSA-based analysis (Figure 3). Interestingly, the *Mt*OGT-R37E mutant showed a significant reduction in dsDNA band shift activity, with a plateau at a DNA/protein molar ratio of 1:600 (Figure 3C), whereas the wild type *Mt*OGT and the *Mt*OGT-R37K variant induce a complete shift at a DNA/protein molar ratio of 1:150 (Figure 3, panels A and B respectively). Taken together our data indicate that the presence of a positively charged group at position 37 of the N-terminal domain of *Mt*OGT is required for efficiently recognizing and binding dsDNA. Since an arginine residue at this position does not appear as a conservative trait in OGTs, it emerges as a peculiar feature of the mycobacterial enzyme.

An equivalent analysis using the *Mt*OGT-Y139F protein reveals that this point-mutated variant does not show any impairment in unmodified dsDNA cooperative binding (Figure 3D); indeed, *Mt*OGT-Y139F

behaves as the wild-type enzyme in EMSA-based experiment, reaching a plateau in band shift activity at a DNA/protein molar ratio of 1:150. Notably the *Mt*OGT-Y139F DNA binding defect is detectable only in presence of a dsDNA^{met} (see Table 2), suggesting an involvement of Tyr139 in specifically stabilizing the modified base upon DNA binding.

*Mt*OGT-R37K structure analysis

Previous X-ray crystallography-based structural investigations on MTB OGT (18), allowed us to speculate about the possible mechanism/s by which the Arg37 residue -mapping at the poorly structurally conserved and functionally characterized N-terminal domain of the protein- could take part in the de-alkylation reaction and/or in mediating protein recruiting at the alkylated-site on DNA, two mandatory steps along catalysis previously attributed solely to elements/motifs of the C-terminal domain of the protein. We postulated that the non-conservative substitution of *Mt*OGT Arg37 by leucine -as observed in the OGT-R37L mutant- could affect protein function either by altering the local conformation of the N-terminal domain, or by interfering with the assembling of protein complexes at the site of damage or by impeding a direct participation of Arg37 to direct DNA binding.

To discriminate between these hypotheses, we expanded our comparative structural analysis of MTB OGT to include the 2.3-Å crystal structure of *Mt*OGT-R37K (Table 3) and a model of the *Mt*OGT-

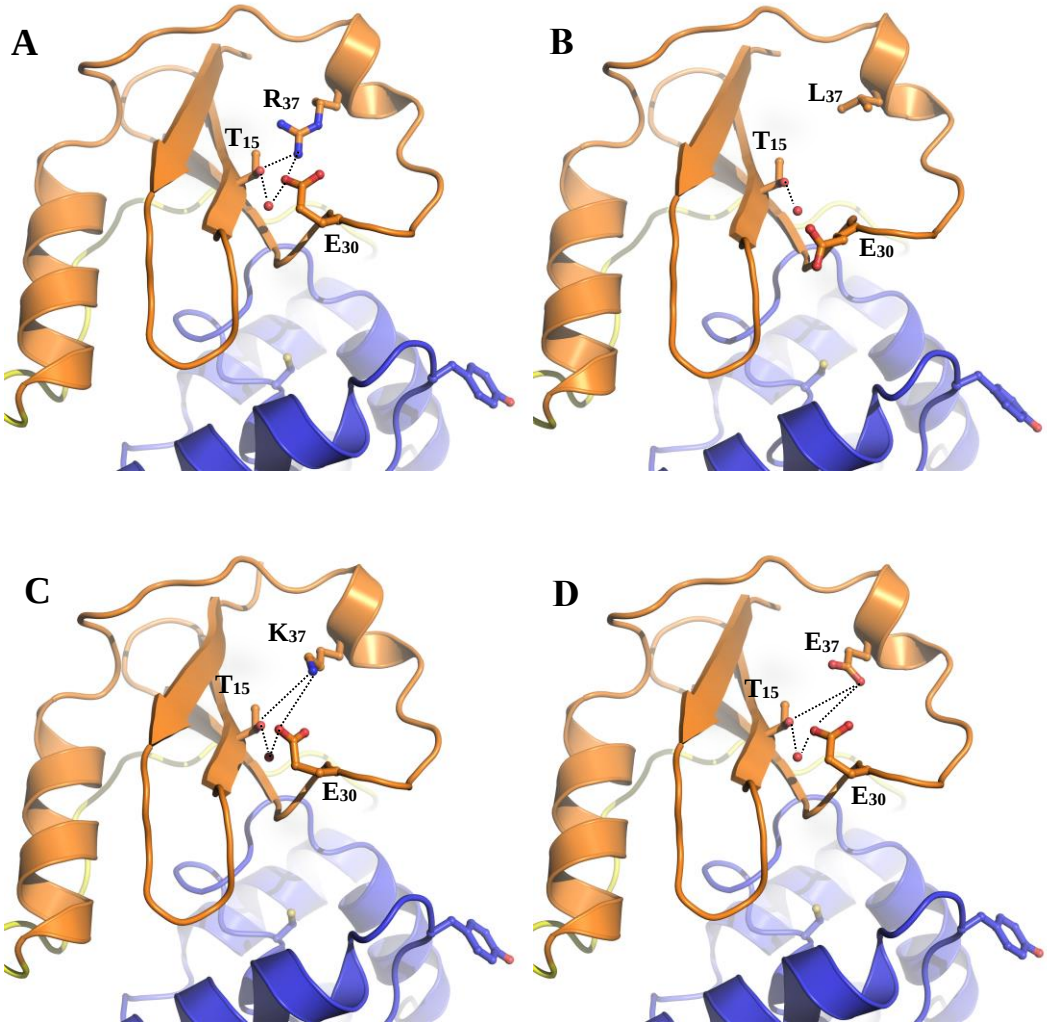
R37E, since in this latter case we did not obtain single crystals suitable for X-ray diffraction experiments.

Figure 4A shows that in wild type *Mt*OGT Arg37 is involved in a hydrogen bonding network established between residues of strand β1 and of the facing random coil. As already reported, the amino acid substitution of Arg37 by leucine (*Mt*OGT-R37L, Figure 4B) indeed destroys this network of contacts, with no evident effects on the local fold of the protein. Notably, in both the *Mt*OGT-R37K crystal structure (Figure 4C) and the structure-based model of the *Mt*OGT-R37E mutant (Figure 4D) the hydrogen bonding network present in the wild-type enzyme appears fully rebuilt; however, our biochemical studies using the VG-based assay (Table 2) reveal that these two latter *Mt*OGT variants profoundly differ in their ability of sustaining catalysis, with *Mt*OGT-R37E displaying a five-fold lower affinity for $dsDNA^{met}$ with respect to *Mt*OGT-R37K.

Therefore, we conclude that the role of Arg37 in *Mt*OGT is to provide a positive charge at that specific position, and not to simply stabilize the overall fold of the N-terminal domain by coordinating the local bonding scheme. We still need to assess whether this positive charged residue is directly involved in DNA binding or if it mediates protein-protein interactions. To this end we performed co-crystallization experiments using the wild type protein and a modified double stranded DNA oligonuclotide, which can be covalently bound by the protein. These

trials yielded single crystals diffracting at low resolution, and the optimization of crystallization conditions have been undertaken.

Figure 4



Close-up views of the N-terminal domain of MtOGT and mutated variants. (A) wild-type *Mt*OGT (pdb: 4BHB); (B) *Mt*OGT-R37L (pdb: 4BHC); (C) *Mt*OGT-R37K (present work) upon optimal superposition of the corresponding crystal structures. (D) structure-based model of the *Mt*OGT-R37E variant.

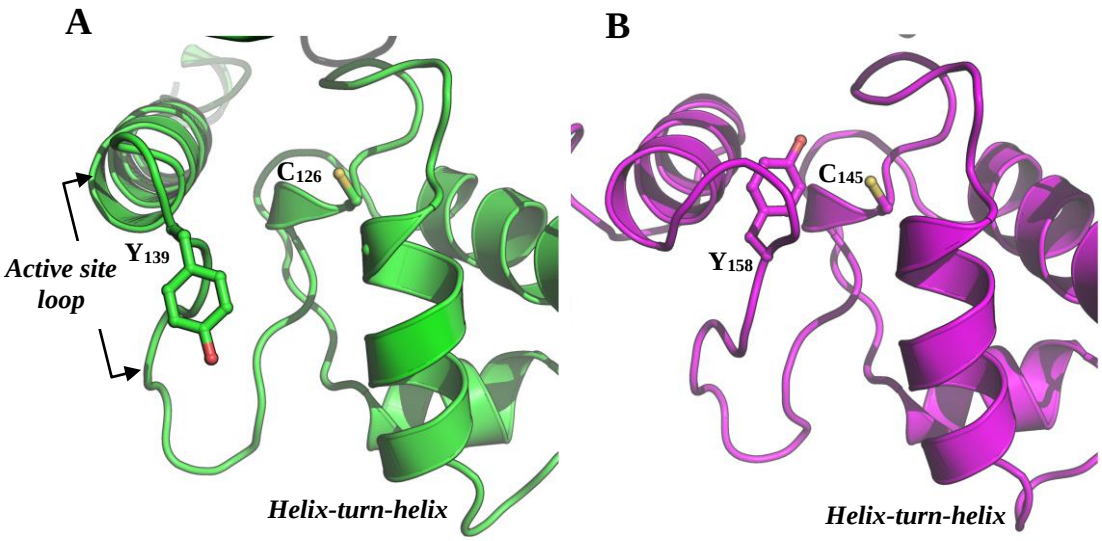
Table 3- Data collection, phasing and refinement statistics.

	***Mt*OGT-R37K**	***Mt*OGT-Y139F**
Data collection		
Space Group	$P2_12_12$	$P2_12_12$
Wavelength (Å)	0.979	0.99
Resolution (Å)	2.3	2.6
Total reflections	50946	17965
Unique reflections	8668	5706
Mean(I)/sd(I)	26.7 (8.5)	10.8 (2.4)
Completeness (%)	99.5 (100)	95.9 (100)
Multiplicity	5.9 (6.0)	3.1 (2.9)
R_{merge} (%)	3.9 (17.4)	6.1 (51.9)
R_{meas} (%)	4.3 (19.1)	7.3 (62.7)
R_{pim} (%)	1.8 (7.8)	3.9 (34.7)
Refinement		
R_{factor}/R_{free} (%)	18.68/22.98	20.1/27.2
Protein Atoms	1257	1258
Ligand atoms	12	-
Water molecules	66	23
R.m.s.d. bonds (Å)	0.008	0.009
R.m.s.d. angles (°)	1.053	1.30
Average B (Å^2)		
Protein	41.5	57.8
Solvent	42	53.5

*Mt*OGT-Y139F structure analysis.

As mentioned in the "*Introduction*" section, the main exception to the otherwise high degree of structural conservation displayed by the catalytic C-terminal domain of *Mt*OGT is represented by the conformation adopted by the active site loop, (residues from 136 to 141), (Figure 5).

Figure 5

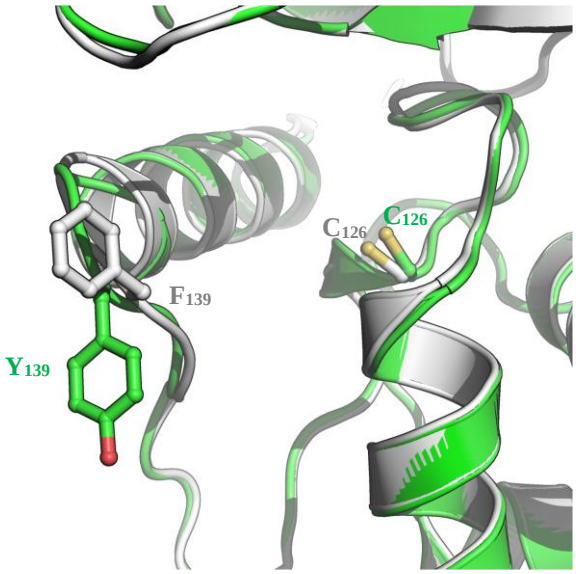


The active site loop of MtOGT is observed in a new conformation. Close-up view of equivalent regions of *Mt*OGT (in green; pdb:4BHB) (18) and human AGT in the absence of ligands (in magenta; pdb:1EH6).

In fact, in all crystal structures of OGT/AGT of other species reported until now, this element is observed in a "pointing inwards" conformation that appears functional to narrowing the catalytic pocket. In particular, the highly evolutionary conserved tyrosine residue of the active site loop (Tyr158 in human AGT) has been suggested to provide an aromatic environment for locking the alkyl adduct of the flipped-out modified base inside the active site. Of outmost interest, a point-mutated variant of human AGT, in which Tyr158 is replaced by a phenylalanine, is only slightly less effective than the wild-type AGT in protecting the alkyltransferases-lacking *E. coli* strain GWR109 from MNNG killing (34). In sharp contrast, our biochemical analysis revealed that an equivalent *Mt*OGT-Y139F variant displays a ten-fold lower affinity for the $dsDNA^{met}$ with respect to the wild-type *Mt*OGT (Table 2), while the ability of binding unmodified DNA appears unaltered (Figure 3, panel D).

The analysis of the 2.6-Å resolution crystal structure of *Mt*OGT-Y139F (Table 3) does not reveal any structural rearrangement at the level of the active site loop Cα-backbone with respect to the wild type *Mt*OGT structure (Figure 6). Consequently the observed catalytic defect could not be imputable to a mere destabilization of local conformation at the level of catalytic C-terminal domain.

Figure 6



MtOGT and MtOGT-Y139F active-site loop conformation. Close-up view of MtOGT (in green;pdb:4BHB)(18) and *Mt*OGT-Y139F (in grey; present work) C-terminal domain upon, optimal superposition of the corresponding crystal structures. The mutated 139 residues and the catalytic cysteine are rendered as ball and sticks.

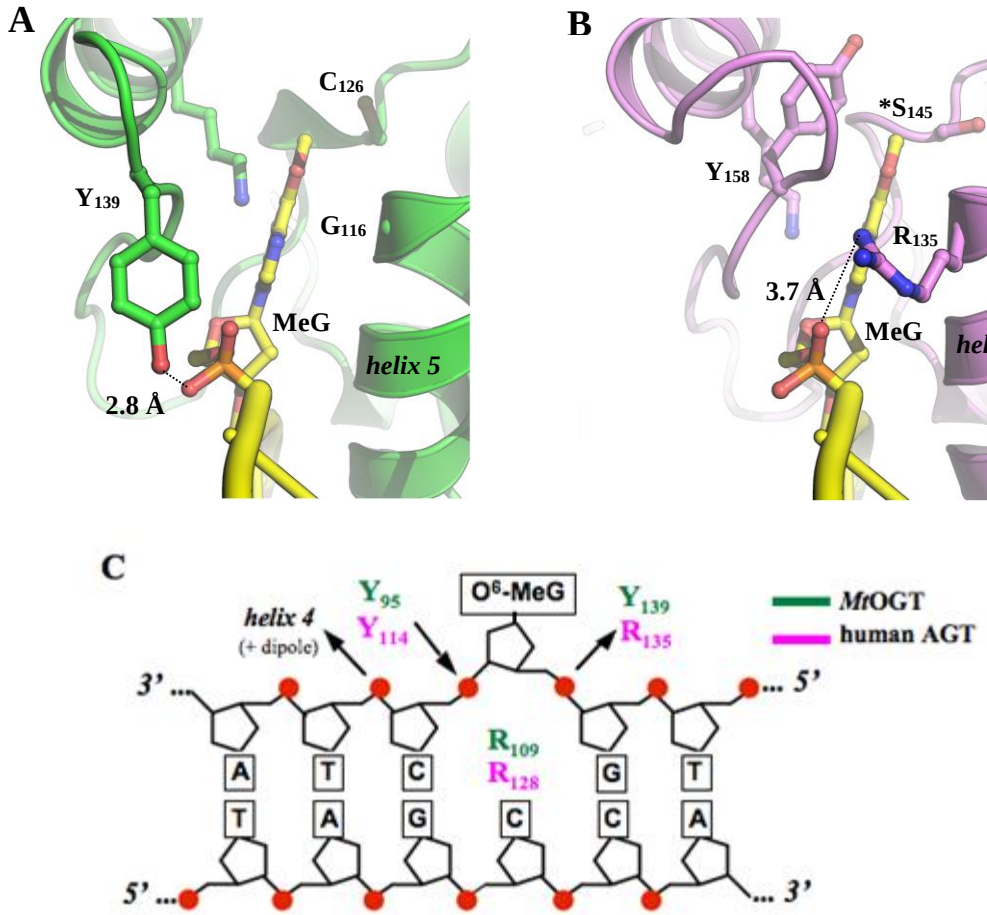
On the basis of structural comparison analyses to the human AGT in complex with methylated dsDNA, we previously showed that Tyr139 in *Mt*OGT is correctly oriented to establish a hydrogen bond with the phosphodiester group at the 5' side of the flipped-out modified base (Figure 7A) (18). Interestingly, by adopting this unprecedented conformation, *Mt*OGT Y139 seems to functionally compensate for the absence of an arginine residue at position 116 of the recognition helix 5,

which in human AGT (*i.e.* Arg135) is engaged in an equivalent contact with the DNA sugar-phosphate backbone (Figure 7B).

Figure 8 illustrates structure-based models of the possible mode of binding of unmodified (black) and methylated (magenta) dsDNA to *Mt*OGT, showing that the flipping-out of the modified base brings its 5' phosphate group closer to the Tyr139 hydroxyl-group, in order to establish a hydrogen bond.

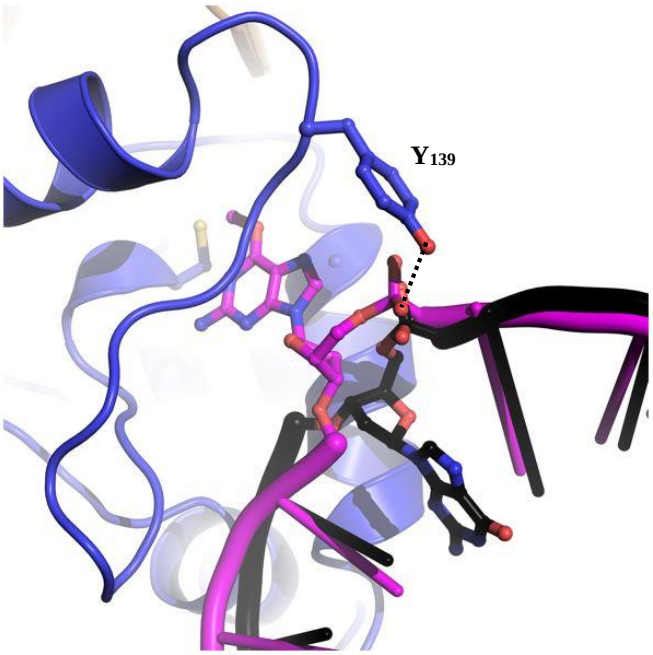
Taken as a whole, our biochemical and structural characterization of the *Mt*OGT-Y139F variant pinpoints Tyr139 as a molecular determinant discriminating between unmodified and alkylated DNA substrates through a novel mechanism, whose detailed description should await the resolution of the crystal structure of *Mt*OGT in complex with DNA.

Figure 7



Solvent exposed conformation of the invariant Tyr139 would not hamper DNA binding. (A) Possible mode of binding of methylated guanine bearing dsDNA to *Mt*OGT resulting upon optimal superposition of *Mt*OGT on the structure of a Cys145Ser human AGT mutated version in complex with O^6-metgylguanine dsDNA (pdb:1T38). (B) Schematic of DNA contacts of *Mt*OGT and human AGT. Phosphates are labeled in red. Arrows point to contacts between indicated residues side chain and DNA sugar-phosphate backbone.

Figure 8



Structure based models of MtOGT in complex with DNA substrates. Cartoon representation of the possible mode of binding of unmodified DNA (black) and O^6-methylguanine bearing DNA (magenta), upon optimal superposition of *Mt*OGT structure (pdb:4BHB) to that of human AGT in complex with O^6-MeG-dsDNA (pdb:1T38)

Bibliography

1- **Dos Vultos T, Mestre O, Rauzier J, Golec M, Rastogi N, Rasolofo V, Tonjum T, Sola C, Matic I, Gicquel B.** (2008). Evolution and diversity of clonal bacteria: the paradigm of Mycobacterium tuberculosis. PLoS One **3**:e1538. doi:10.1371/journal.pone.0001538

2- **Gorna AE, Bowater RP, Dziadek J.** (2010). DNA repair systems and the pathogenesis of Mycobacterium tuberculosis: varying activities at different stages of infection. Clin. Sci. (Lond.) **119**:187-202.

3- **Voskuil MI, Bartek IL, Visconti K, Schoolnik GK.** (2011). The response of Mycobacterium tuberculosis to reactive oxygen and nitrogen species. Front. Microbiol. **2**:105. doi:10.3389/fmicb.2011.00105.

4- **Dalhus B, Laerdahl JK, Backe PH, Bjoras M.** (2009). DNA base repair - recognition and initiation of catalysis. FEMS Microbiol. Rev. **33**:1044-1078.

5- **Shrivastav N, Li D, Essigmann JM.** (2010). Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. Carcinogenesis **31**:59-70.

6- **Pegg AE.** (2011). Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools. Chem. Res. Toxicol. **24**:618-639

7- **Tubbs JL, Pegg AE, Tainer JA**. (2007) DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair* (Amst). **6**:1100-15.

8- **Duguid EM, Rice PA, He C.** (2005). The structure of the human AGT protein bound to DNA and its implications for damage detection. J. Mol. Biol. **350**:657-666.

9- **Daniels DS, Woo TT, Luu KX, Noll DM, Clarke ND, Pegg AE, Tainer JA.** (2004). DNA binding and nucleotide flipping by the human DNA repair protein AGT. Nat. Struct. Mol. Biol. **11**:714-720.

10- **Tubbs JL, Latypov V, Kanugula S, Butt A, Melikishvili M, Kraehenbuehl R, Fleck O, Marriott A, Watson AJ, Verbeek B, McGown G, Thorncroft M, Santibanez-Koref MF, Millington C, Arvai AS, Kroeger MD, Peterson LA, Williams DM, Fried MG, Margison GP, Pegg AE, Tainer JA.** (2009) Flipping of alkylated DNA damage bridges base and nucleotide excision repair. *Nature*. **459**:808-13

11- **Yang M, Aamodt RM, Dalhus B, Balasingham S, Helle I, Andersen P, Tonjum T, Alseth I, Rognes T, Bjoras M.** (2011). The ada operon of Mycobacterium tuberculosis encodes two DNA methyltransferases for inducible repair of DNA alkylation damage. DNA Repair (Amst.) **10**:595-602.

12- **Boshoff HI, Myers TG, Copp BR, McNeil MR, Wilson MA, Barry CE, III.** (2004). The transcriptional responses of Mycobacterium tuberculosis to inhibitors of metabolism: novel insights into drug mechanisms of action. J. Biol. Chem. **279**:40174–40184.

13- **Schnappinger D, Ehrt S, Voskuil MI, Liu Y, Mangan JA, Monahan IM, Dolganov G, Efron B, Butcher PD, Nathan C, Schoolnik GK.** (2003). Transcriptional adaptation of Mycobacterium tuberculosis within macrophages: insights into the phagosomal environment. J. Exp. Med. **198**:693-704.

14- **Durbach SI, Springer B, Machowski EE, North RJ, Papavinasasundaram KG, Colston MJ, Böttger EC, Mizrahi V.** (2003) DNA alkylation damage as a sensor of nitrosative stress in Mycobacterium tuberculosis. Infect Immun.; **71**:997-1000.

15- **Ebrahimi-Rad M, Bifani P, Martin C, Kremer K, Samper S, Rauzier J, Kreiswirth B, Blazquez J, Jouan M, van Soolingen D, Gicquel B.** (2003). Mutations in putative mutator genes of Mycobacterium tuberculosis strains of the W-Beijing family. Emerg. Infect. Dis. **9**:838-845.

16- **Olano J, Lopez B, Reyes A, Lemos MP, Correa N, Del Portillo P, Barrera L, Robledo J, Ritacco V, Zambrano MM.** (2007). Mutations in DNA repair genes are associated with the Haarlem lineage of Mycobacterium tuberculosis independently of their antibiotic resistance. Tuberculosis (Edinb.) **87**:502–508.

17- **Mestre O, Luo T, Dos Vultos T, Kremer K, Murray A, Namouchi A, Jackson C, Rauzier J, Bifani P, Warren R, Rasolofo V, Mei J, Gao Q, Gicquel B.** (2011). Phylogeny of Mycobacterium tuberculosis Beijing strains constructed from polymorphisms in genes involved in DNA replication, recombination and repair. PLoS One **6**:e16020. doi:10.1371/journal.pone.0016020.

18- **Miggiano R., Casazza V., Garavaglia S., Ciaramella M., Perugino G., Rizzi M., Rossi F.** (2013) Biochemical and structural studies of the Mycobacterium tuberculosis O6-methylguanine methyltransferase and mutated variants. J Bacteriol.; **195**:2728-2736

19- **Sambrook J, Fritsh EF, Maniatis T.** (1989). Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

20- **Perugino G, Vettone A, Illiano G, Valenti A, Ferrara MC, Rossi M, Ciaramella M.** (2012). Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. J. Biol. Chem. **287**:4222-4231.

21- **Collaborative Computational Project Number 4.** (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50**:760-763.

22- **McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ.** (2007). Phaser crystallographic software. J. Appl. Crystallogr. **40**: 658-674.

23- **Perrakis A, Morris R, Lamzin VS.** (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. **6**:458-463.

24- **Emsley P, Cowtan K.** (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60**:2126-2132.

25- **Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH.** (2010). PHENIX: a comprehensive Pythonbased system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. **66**:213-221.

26- **DeLano WL.** (2002). The PyMOL molecular graphics system. DeLano Scientific, Palo Alto, CA. http://www.pymol.org/.

27- **Moore M.H., Gulbis J.M., Dodson E.J., Demple B., Moody P.C.** (1994) Crystal structure of a suicidal DNA repair protein: the Ada O6-methylguanine-DNA methyltransferase from E. coli. *EMBO J.*; **13**: 1495-1501

28- **Daniels D.S., Mol C.D., Arvai A.S., Kanugula S., Pegg A.E., Tainer J.A.** (2000) Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical binding. *EMBO J.*; **19**:1719–1730.

29- **Hashimoto H., Inoue T., Nishioka M., Fujiwara S., Takagi M., Imanaka T., Kai Y.** (1999) Hyperthermostable protein structure

maintained by intra- and inter-helix ion pairs in archael O6-methylguanine-DNA methyltransferase. *J. Mol. Biol.*; **292**:707–716

30- **Edara S, Goodtzova K and Pegg AE** (1995) The role of tyrosine-158 in O6-alkylguanine-DNA alkyltransferase activity. Carcinogenesis; **16**:1637-1642.

31- **Xu-Welliver M, Leitão J, Kanugula S, and Pegg AE** (1999) Alteration of the Conserved Residue Tyrosine-158 to Histidine Renders Human O6-Alkylguanine-DNA Alkyltransferase Insensitive to the Inhibitor O6-Benzylguanine. Cancer Res **59**:1514-1519

32- **Griffin RJ, Arris CE, Bleasdale C, Boyle FT , Calvert AH, Curtin NJ, Dalby C, Kanugula S, Lembicz NK, Newell DR, Pegg AE, and Golding BT.** (2000) Resistance-Modifying Agents. 8.1 Inhibition of O6-Alkylguanine-DNA Alkyltransferase by O6-Alkenyl-, O6-Cycloalkenyl-, and O6-(2-Oxoalkyl)guanines and Potentiation of Temozolomide Cytotoxicity in Vitro by O6-(1-Cyclopentenylmethyl)guanine. J. Med. Chem., **43**:4071–4083

33- **Luu KX, Kanugula S, Pegg AE, Pauly GT and Moschel RC** (2002) Repair of oligodeoxyribonucleotides by O(6)-alkylguanine-DNA alkyltransferase. Biochemistry; **41**:8689-8697.

34- **Crone TM, Goodtzova K, Pegg AE.** (1996) Amino acid residues affecting the activity and stability of human O6-alkylguanine-DNA alkyltransferase. Mutat Res.; **363**:15-25

Conclusions

Mycobacterium tuberculosis, the etiologic agent of TB, is an extremely well adapted human pathogen surviving and replicating inside the human macrophages. During the infectious process, *M. tuberculosis* is exposed to a variety of potentially DNA-damaging assaults including alkylating agents that may compromise both the persistence in the host's infected macrophages and the reactivation of the bacillus from the dormant state. Accordingly, DNA repair pathways are crucial in maintaining genome integrity as well as a low level of genetic variability. *M. tuberculosis* counteracts the effect of DNA alkylation either by using multi-step repair systems or by direct removal of the alkyl group from the damaged base performed by DNA-protein alkyltransferases such as the "suicidal" O^6-methylguanine DNA methyltransferase (OGT). Since a number of broadly geographically distributed *M. tuberculosis* strains show nonsynonymous single-nucleotide polymorphisms (nsSNPs) in their *ogt* gene, it was supposed a possible contribution of corresponding mutated OGT to the success of this strains in terms of adaptation to the human host. Detected nsSNPs map at the region of the gene encoding the poorly functionally characterized N-terminal domain of the protein, and result in threonine 15 and arginine 37 substituted by serine and leucine, respectively.

In the present work we report the structural and biochemical characterization of wild type *M. tuberculosis* OGT (*Mt*OGT) and of the

*Mt*OGT-T15S and *Mt*OGT-R37L mutated proteins, revealing that the R37L variant is characterized by a ten-fold lower affinity for alkylated double stranded DNA than the wild type enzyme. These results provided experimental evidence to a possible contribution of this SNP to the "mutator" phenotype proposed for the corresponding widely distributed strains.

*Mt*OGT crystal structure analysis revealed an element of novelty when compared to other ligand-free or DNA-bound OGT structures: The C-terminal active site loop adopts an unprecedented conformation exposing the invariant Tyrosine 139 to the bulk solvent.

Further biochemical investigations on structure-based designed *Mt*OGT point-mutated variants (*i.e.* *Mt*OGT-R37K, *Mt*OGT-R37E and *Mt*OGT-Y139F) provided new insights in the molecular mechanisms adopted by the analysed residues in coordinating the catalytic cycle. In particular, we showed that a positively charged residue at position 37 is required for an optimal DNA binding activity and that the conserved tyrosine residue at position 139 can possibly establish a hydrogen bond with alkylated-DNA through its hydroxyl group, a feature never observed for OGTs from other species. Confirmation of such a proposal, together with a complete mechanistic overview of the *Mt*OGT damaged-DNA recognition and repair, will be provided by our ongoing investigations on the crystal structure of *Mt*OGT in complex with alkylated double stranded DNA.

Overall, our *in vitro* biochemical and structural studies on mycobacterial OGT contributed to better understand DNA-direct repair in *Mycobacterium tuberculosis.*

The future activity for this research project will focus on the study of protein-protein and protein-DNA interactions involving *Mt*OGT and other key enzymes of DNA-repair pathways, to possibly disclose cross-talks between different DNA repair systems in *M. tuberculosis*.

List of publications

1- Miggiano R, Casazza V, Garavaglia S, Ciaramella M, Perugino G, Rizzi M, Rossi F. (2013) Biochemical and structural studies of the *Mycobacterium tuberculosis* O^6-methylguanine methyltransferase and mutated variants. *J Bacteriol.*; 195: 2728-2736.

2- Spyrakis F, Felici P, Bayden AS, Salsi E, Miggiano R, Kellogg GE, Cozzini P, Cook PF, Mozzarelli A, Campanin B. (2013) Fine tuning of the active site modulates specificity in the interaction of O-acetylserine sulfhydrylase isozymes with serine acetyltransferase. *Biochim. Biophys. Acta*; 1834: 169-81.

Acknowledgments

This work has been partially supported by the program SysteMTb funded by the European Community FP7 program (SysteMTb HEALTH-F4-2010-241587)

Ringrazio il professor Menico Rizzi per avermi dato l'opportunità di intraprendere il percorso di dottorato raccontato in queste pagine e la possibilità di viverlo come un'esperienza fantastica in ogni suo dettaglio.
Un grazie immenso alla dottoressa Franca Rossi, Maestra onnipresente, mai stanca di accompagnarmi in questo viaggio. Grazie…
Grazie alla dottoressa Silvia Garavaglia per gli insegnamenti su tutto ciò che brilla (luce, cristalli e nuvole elettroniche).
Grazie di cuore a tutti i compagni di laboratorio…Davide, Stefano, Valentina, Andrea (Cava), Serena, Andrea (Moretti), Laura e Samarpita.
Grazie alla dottoressa Maria Ciaramella, al dottor Giuseppe Perugino e a tutti i compagni di Napoli per tutto quello che è stato nelle due trasferte napoletane.
Grazie al dottor Carlo Carolis e a tutto il laboratorio del CRG di Barcellona per l'indimenticabile esperienza.
Il grazie più grande va alla mia famiglia, per esserci sempre donandomi giorno dopo giorno la libertà di essere (in particolare grazie al fratello/cugino Antonio che continua a seguirmi sempre ed ovunque).
Infine grazie a tutti gli amici per i sorrisi, i momenti di gioia e il supporto che mi avete sempre dato…grazie Federico, Luigi, Federica, Martina, Antonio, Gerry, Chiara, Irene, Riccardo, Irhu, Giuseppe, Peppino, Francesco, Luca, Valeria, Livio, Rosario…