Tuberculosis 108 (2018) 169–177

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: www.elsevier.com/locate/tube

Review

Mycobacterial nucleoid associated proteins: An added dimension in gene regulation

Nastassja L. Kriel*, James Gallant, Niël van Wyk, Paul van Helden, Samantha L. Sampson, Robin M. Warren, Monique J. Williams

DST/NRF Centre of Excellence for Biomedical Tuberculosis Research, SAMRC Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, South Africa

ARTICLE INFO

Keywords:
Mycobacterium tuberculosis
Nucleoid associated proteins
Chromosome structure
Transcription
Gene regulation

ABSTRACT

Nucleoid associated proteins (NAPs) are known organisers of chromosomal structure and regulators of transcriptional expression. The number of proposed NAPs in mycobacteria are significantly lower than the number identified in other organisms. An interesting feature of mycobacterial NAPs is their low sequence similarity with those in other species, a property that has hindered their identification. In this review, we discuss the current evidence for the proposed classification of six mycobacterial proteins, Lsr2, EspR, mIHF, HupB, MDP2 and NapM, as NAPs in mycobacterial species with an emphasis on their roles in modulating chromosome structure and transcriptional regulation. In addition, we highlight the technical difficulties associated with investigating and providing evidence for the classification of proteins as NAPs in mycobacteria. We also address the role of mycobacterial NAPs as mediators of stress responses and highlight the recent developments aimed at targeting NAP-DNA interactions for the development of novel anti-TB drugs.

1. Introduction

The ability of *Mycobacterium tuberculosis* to adapt to the adverse conditions encountered during infection is crucial to its survival within the host, and involves appropriate changes in gene expression in response to the changing environment. Numerous studies have investigated the transcriptional response of *M. tuberculosis* following exposure to host-derived stresses such as hypoxia [1–4], nutrient limitation [5], oxidative and nitrosative stress [6], and extensive work has been done to characterise the transcription factors involved in mediating these responses. More recently, the shift to a 'systems biology' approach has sought to understand the complex interactions between various components of these regulatory networks [7–9]. In bacteria, genomic DNA is compacted to a structure called the nucleoid, which is roughly 10^4 times smaller than the volume of linear DNA of equivalent length [10]. The organisation and plasticity of this structure is crucial for gene expression, since it must allow proteins involved in transcription to access the DNA. Nucleoid associated proteins (NAPs), also known as histone-like proteins, are a group of small, highly abundant DNA binding proteins implicated in maintaining chromosome structure [11–13], and they therefore play an important role in regulating gene expression, both globally and locally [14,15]. NAPs are often positively charged, low molecular weight, dimeric proteins which modulate the bacterial chromosome through bending, wrapping and bridging of DNA [11,13]. These proteins are distinct from homo-tetrameric single stranded DNA binding proteins (SSBs), which are required for chromosome replication in eubacteria [16]. Unlike NAPs, these helix destabilising proteins are known to possess an oligonucleotide binding fold (OB-fold) on the N-terminal domain which is essential for oligomerization and DNA binding [17]. Furthermore, the primary function of SSBs is to prevent the formation of secondary DNA structures through the occupation of ~35, ~56 or ~65 nucleotides [18–20]. Like NAPs, SSBs interact with DNA in a sequence independent manner and physically protect DNA from chemical attacks and nucleases [16,21]. In *Escherichia coli* the differential expression of NAPs is proposed to drive growth-phase dependent changes in chromosome structure [22]. The identification of NAPs in mycobacteria has been hindered by low sequence conservation with other well-characterised NAPs. This review aims to summarize our current knowledge of proteins proposed to function as NAPs in mycobacteria (Table 1).

2. Lsr2 is a novel DNA-bridging protein in mycobacteria

Lsr2 is a small (12kDa) basic protein which is highly conserved in

* Corresponding author. DST/NRF Centre of Excellence for Biomedical Tuberculosis Research, SAMRC Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, P.O. Box 241, Cape Town, 8000, South Africa.
E-mail address: nastassja@sun.ac.za (N.L. Kriel).

https://doi.org/10.1016/j.tube.2017.12.004
Received 17 August 2017; Received in revised form 8 December 2017; Accepted 11 December 2017

Table 1
Nucleoid-associated proteins of the genus *Mycobacterium*.

NAP	*M. tb* Rv number	*M. smegmatis* MSMEG number	*M. bovis* Mb number	*M. marinum* MMAR number	*M. leprae* ML number	% identity to *E. coli* homologue[a]	Essential in *M. tb*[b]	Molecular Mass (kDa)	Isoelectric point	Native Protomer	Binding Groove	DNA binding properties
Lsr2	Rv3597c	MSMEG_6092	Mb3628c	MMAR_5101	ML0234	–	Yes	12.09	10.69	dimer	Minor Groove [28]	Bridging [29–31]
EspR	Rv3849	MSMEG_6431	Mb3879	MMAR_5399	ML0069	30%	No[d]	14.7	9.04	dimer	Major Groove [41]	Bending, Bridging [38,41,89]
HupB (Hlp)	Rv2986c	MSMEG_2389	Mb3010c	MMAR_1728	ML1683	43%	Yes	22.18	12.48	dimer	Minor Groove [46]	Bending [46]
mIHF	Rv1388	MSMEG_3050	Mb1423	MMAR_2201	ML0540	–	Yes	20.81 (12.10)[c]	11.02 (10.06)	dimer	Unknown	Bending, Bridging, Wrapping [63,64]
NapM	Rv0047c	MSMEG_6903	Mb0048c	MMAR_0066	ML2691	–	No	20.41	10.39	–	Major Groove	Bridging [65]
MDP2 (H-NS)	Rv3852	No homologue	Mb3882	MMAR_5402	ML0067	–	No	13.82	11.34	dimer	Unknown	Unknown

[a] Obtained from BLASTp search using *M. tuberculosis* H37Rv protein sequence.
[b] Based on (Sassetti et al. [23]; Griffin et al. [24]).
[c] Alternative starting site (Mishra et al. [63]).
[d] Slow growth mutant.

mycobacteria and related actinomycetes. Forward genetic screens predicted *lsr2* to be essential for survival of *M. tuberculosis in vitro* [23,24], nevertheless, a slow growth *Δlsr2* deletion mutant has been generated [25]. Characterization of the *Δlsr2* deletion mutant revealed that Lsr2 was required by *M. tuberculosis* for growth under normoxic and hyperoxic conditions, as well as for the adaptation to anaerobiosis [26]. Initial evidence identifying Lsr2 in *M. tuberculosis* as a novel histone-like protein included the ability to: (a) form large multimeric complexes with DNA, (b) bind preferentially to AT-rich sequences (commonly found within promoter regions and foreign acquired DNA), (c) induce modest supercoiling in relaxed plasmid DNA, and (d) inhibit transcription and topoisomerase I *in vitro* [27]. In ChIP-chip studies Lsr2 co-precipitates with 21% and 13% of the *M. tuberculosis* and *Mycobacterium smegmatis* genomes respectively, showing a preference for binding to regions with low GC content, including those regions acquired by horizontal gene transfer [28]. Whole genome expression data demonstrated that the majority of genes identified in ChIP-chip studies were upregulated in the *Δlsr2* deletion mutant, supporting the proposed role of Lsr2 as a transcriptional repressor [26,28].

Atomic force microscopy (AFM) revealed that Lsr2 has DNA-bridging properties [29], and complementation studies suggest that it functions analogously to the *E. coli* DNA-bridging NAP H-NS [30]. Similar to the *E. coli* NAP H-NS, Lsr2 contains an N-terminal dimerization domain (residues 1–65) and a C-terminal DNA-binding domain (residues 51–112), although the structure of the C-terminal domain is distinct from that of H-NS [28]. Despite their unique C-terminal domain structures, H-NS and Lsr2 share a common binding mechanism to the minor groove of DNA, and in both cases, their increased affinity for AT-rich sequences is determined by the width of the minor groove [31]. The narrower minor groove produced by AT-rich DNA sequences is proposed to help H-NS and Lsr2 bind the minor groove in an “AT-Hook-Like” grip [31–33]. Following binding to the minor groove, Lsr2 and H-NS are able to bridge distant DNA fragments into hairpins and loops [29]. *In vitro*, Lsr2 binds co-operatively along extended DNA to form a rigid Lsr2-DNA nucleoprotein filament, while at low DNA tension it causes DNA aggregation [34], suggesting that the variation in DNA tension across the chromosome may act to regulate the formation of Lsr2-DNA structures. Structural studies of the Lsr2 N-terminal domain demonstrated that removal of the first three amino acid residues by trypsin resulted in a shift from dimerization to oligomerization, leading authors to speculate that proteolytic processing of Lsr2 may be a mechanism of regulating its binding *in vivo* [35]. Interestingly, some mycobacteriophages appear to have acquired *lsr2* from their host, although the significance of this is unclear [36].

3. EspR and HupB, transcriptional regulators and NAPs?

The primary role of NAPs is to maintain chromosome structure, and in most instances this is facilitated through binding to numerous sites in the chromosome in a sequence-independent manner. However, some NAPs also bind high-affinity binding sites, making it difficult to distinguish these proteins from transcriptional regulators, particularly if their binding has an impact on gene expression.

EspR, encoded by *Rv3849* in *M. tuberculosis*, is a small (14.7 kDa) protein that is highly conserved in mycobacteria and related actinomycetes. Initial interest in EspR was due to its role in regulating ESX-1, a virulence-associated type VII secretion system in *M. tuberculosis*, mediated via induction of the *espACD* (Rv3616c-Rv3614c) operon [37]. Like other ESX-1 secretion mutants, a transposon mutant producing low levels of EspR induced high levels of IL-12 in macrophages and was defective for survival in mice [37]. EspR was originally thought to be a secreted substrate of ESX-1, however more recent work demonstrated that it is mainly cytosolic, and its intracellular concentration increases throughout the cell cycle, reaching approximately 100 000 molecules per cell in stationary phase [38]. Recently the two-component systems PhoP-PhoR and MprA-MprB were shown to directly regulate EspR,