## Rho-dependent attenuation of sulfur metabolism in mycobacteria

V.A. Lyubetsky, S.A. Korolev, A.V. Seliverstov, O.A. Zverkov

*Institute for Information Transmission Problems RAS (Kharkevich Institute), lyubetsk@iitp.ru*

This work further develops the pioneer studies by C. Yanofsky of Rho-dependent attenuation regulation of tryptophanase *tna* in *Escherichia coli*. The authors predicted attenuation of transcription of gene *cysK* mediated by the Rho factor and cysteine concentration (Seliverstov et al., 2005). This regulation is characterized by an open reading frame that encodes the leader peptide and contains a short run of cysteine codons preceding the stop codon. We present a study of Rho-dependent expression attenuation of genes involved in sulfur metabolism and transport of sulfur-containing substances for all annotated genes of mycobacteria contained in GenBank. The regulation of metabolism and transport differs considerably in different species. The mRNA 5'-regions in all annotated mycobacteria was searched for a regulation based on transcription-translation coupling (both Rho-dependent and classic sense Yanofsky). We predicted the presence or lack of such regulation, described the structure of corresponding regulons, and assigned putative roles to proteins with unknown function. Thus, all proteins with the DUF4395 domain (Pfam nomenclature), ROK domain, and some others, were predicted to be involved in sulfur metabolism. We also showed affinity of some transporters to taurine. The obtained results considerably extend earlier large-scale studies of classic and non-classic attenuations that depend on amino acids other than cysteine (Lopatovskaya et al., 2010). In particular, we show that the cysteinyl-tRNA synthetase gene is not regulated by attenuation, although this is a known regulation mechanism for genes of some other aminoacyl-tRNA synthetases. The Rho transcription factor was known to be present in all mycobacteria, with the exception of *M. africanum* (GenBank:NC_015758.1). Protein alignment of the Rho sequence from *M. bovis* and its predicted homolog from *M. africanum* shows strong overlaps at the C-terminus (residues 217–602), and N-terminus (residues 1–114). This observation suggests that the Rho protein in *M. africanum* modified closer to the N-terminus may still retain function, while its gene is considered a pseudogene in current annotation. Alternatively, the Rho protein in *M. africanum* may consist of two

different subunits translated independently in two different frames. In *M. tuberculosis*, *M. bovis*, and *M. canettii* the Rho sequences are very similar, albeit being diverged from those in other mycobacteria. *M. leprae*, *M. ulcerans* and *M. marinum* also express diverged patterns. The thiosulfate sulfurtransferase gene is either regulated by classic attenuation or does not have the leader peptide gene. If a DUF4395-domain protein is genome-encoded, one of its paralogs is regulated by attenuation. This is the only exception from the general rule that we found. Our predictions of cysteine-independent expression attenuation allow to identify proteins with poorly characterized domains ROK and DUF4395 involved in sulfur metabolism. For example, presence of the leader peptide with four consecutive cysteine residues associated with an ABC-transporter transmembrane co-factor with unknown function in *Amycolicicoccus subflavus* GenBank:YP_004491913.1) suggests that this is a subunit of an ABC-transporter for taurine or other sulfur-containing compounds. A diverse regulon with cysteine-dependent attenuation that includes genes encoding DUF4395-domain proteins was predicted in *Mycobacterium* sp. JLS, *Mycobacterium* sp. KMS, *Mycobacterium* sp. MCS, *M. abscessus*, *M. massiliense*, *M. rhodesiae*, and *M. smegmatis*. In some species it encodes subunits of the taurine transporter. However, presence of transposases in the regulons of *M. gilvum* and *M. ulcerans* may suggest a later acquisition of attenuation in the result of horizontal gene transfers between mycobacteria, and, in particular, the gain of the new gene associated with the leader peptide in *M. marinum* and *M. ulcerans*. Among studied mycobacteria, *M. massiliense*, *M. smegmatis* and some others are species that were found to possess classic cysteine-dependent attenuation. This observation conforms well with the earlier evidence that classic attenuation is more common in actinomycetes of the genera *Streptomyces* and *Corynebacterium* than in mycobacteria. The Rho-dependent nature of attenuation in mycobacteria is also advocated by presence of a pyrimidine-rich region in the neighborhood of the cysteine codon run in the leader peptide gene sequence. These regions may be viewed as binding sites. However, the hypothesis of Rho involvement is also based on the parsimony principle as not requiring other factors. Our predictions were verified with modeling Rho-dependent attenuation of transcription using principles of the previously developed original model of classic attenuation.