Arginine regulation in *Thermotoga neapolitana* and *Thermotoga maritima*

D. Charlier

Erfelijkheidsleer en Microbiologie, Vrije Universiteit Brussel, Pleinlaan, 2, B-1050 Brussels, Belgium

**Abstract**

Experimental data and *in silico* analyses of sequenced bacterial genomes indicate that arginine repressor (ArgR) proteins and their respective target sites are surprisingly well conserved in very diverse bacteria. Arginine regulation therefore constitutes an interesting model system from the study of evolutionary aspects of bacterial regulation. Moreover, arginine repressor molecules are multifunctional; they repress the arginine biosynthetic genes and are involved in the activation of the various arginine catabolic pathways. Studies on the arginine repressor from the hyperthermophiles *Thermotoga neapolitana* and *Thermotoga maritima* have reinforced the uniform view of the bacterial ArgR–operator interaction, but have also revealed that the *Thermotoga* repressor exhibits unique features. Thus, its DNA-binding activity is nearly arginine-independent and exhibits poor sequence specificity. ArgR<sub>in</sub> has a remarkable capacity to bind heterologous arginine operators and half-site targets.

**Introduction**

The study of arginine regulation and the discovery of the arginine repressor in *Escherichia coli* dates back to the 1950s with the studies of Henry Vogel, Werner Maas and Luigi Gorini. In conjunction with the pioneering work on induction of the lactose genes and repression of the tryptophan biosynthesis, their work made a significant contribution to the understanding of gene regulation in general (for an historical and anecdotal review see [1]). Since then, numerous *in vitro* and *in vivo* studies performed by our own and several other groups on distantly related organisms, including Gram-negative and Gram-positive bacteria, indicate that the global mode of transcriptional regulation of arginine biosynthetic genes and operons is surprisingly well conserved throughout the bacterial domain [2–14]. This is unusual since most other regulatory systems are quite different in these diverse organisms. Consequently, the arginine regulatory system constitutes an interesting model system for the study of the origin and evolution of regulatory molecules and their implication in regulatory networks. With respect to the latter aspect it is particularly interesting to note that arginine repressor molecules (ArgR, called AhrC for arginine hydroxamate repressor in *Bacillus subtilis*) are multifunctional proteins; they repress the biosynthetic pathway and are involved in the activation of various arginine catabolic pathways (arginine succinyltransferase, arginine deiminase, arginase) in both Gram-negative and Gram-positive bacteria [15–22]. Moreover, the arginine repressor of *Streptomyces clavuligerus* participates in regulation of clavulanic acid production [23] and *E. coli* ArgR is an essential component of a site-specific DNA recombination mechanism that resolves multimeric forms of the ColE1 plasmid into the monomeric constituents [24]. The studies on arginine regulation, with special emphasis on arginine repressor–operator interactions, were further extended to extremophiles, the psychrophilic *Moritella profunda* [25], the moderate thermophile *Bacillus steatorrhophilus* [26,27] and the hyperthermophiles *Thermotoga neapolitana* and *Thermotoga maritima*, which grow up to 90°C [27–29]. No ArgR homologue could however be detected in the completed genome sequence of another hyperthermophilic bacterium, *Aquifex aeolicus* [30], nor in any archaeal genome.

Our studies on the thermophilic and hyperthermophilic repressors conducted in collaboration with Vehary Sakanyan (University of Nantes, Nantes, France) have reinforced the uniform view of the arginine regulatory system in bacteria, but also have revealed that the *Thermotoga* repressor–operator interaction exhibits unique features that clearly distinguish this system from the previously studied homologues [26–29]. Based on these and previous data, we propose a classification of bacterial arginine repressors of the wHTH (winged helix-turn-helix) family into three major types.

**Bacterial ArgR molecules of the wHTH class**

The first indication that ArgR molecules from divergent bacteria might be structurally and functionally similar came from the observation that an *E. coli* argR mutant could be complemented by introduction of the *B. subtilis* homologue *ahrC* for both the regulatory function and in the plasmid-resolution reaction [31]. This idea is now further corroborated by the three-dimensional determination of three arginine-repressor structures, from *E. coli* [32,33], *B. steatorrhophilus* [34] and *B. subtilis* [35], that show a similar organization and fold, even though the *E. coli* and the *Bacillus* proteins show only limited sequence.

**Key words:** arginine biosynthesis, arginine repressor, bacteria, protein-DNA interaction, thermophile, transcription.

**Abbreviations used:** ArgR, arginine repressor; wHTH, winged helix-turn-helix.

©2004 Biochemical Society
The transcriptional regulation of the arginine biosynthetic pathway in *T. neapolitana* and *T. maritima* is extremely poorly documented but Van de Casteele et al. [40] have observed an approx. 2-fold reduction of the *N*-acetylglutamate 5-phosphotransferase (*argB*) and argininosuccinate synthetase (*argG*) enzyme specific activities upon addition of 5 mM arginine to a *T. maritima* culture grown in the presence of a limited amount of yeast extract (corresponding to approx. 0.05 mM arginine), indicating that an ArgR/arginine-mediated repression of promoter activity might possibly be operative in hyperthermophiles. Besides, *in vitro* studies on the ArgR protein of *B. stearothermophilus* had indicated that the purified repressor remains active at temperatures far above the maximum growth temperature of the organism [26]. Therefore it became plausible that a similar regulatory mechanism might be present in organisms adapted to temperatures of 80°C or above.

Using a pair of degenerated oligonucleotides corresponding to conserved regions in the DNA-binding domain of arginine repressors, Vehary Sakanyan and his collaborators succeeded in the amplification of part of the *argR* gene of *T. neapolitana* and *T. maritima* [28]. These fragments were subsequently used to screen genomic libraries to isolate the intact repressor genes. The *Thermotoga* *argR* genes were found to code for approx. 17 kDa proteins, close to the molecular mass of arginine repressors of mesophilic bacteria. Sequence determination and alignments indicated that the two hyperthermophilic repressors exhibit 97% amino acid sequence identity and share identical or chemically similar amino acids in the functionally important secondary structural elements with other ArgR proteins and likely adopt the same fold, even though the overall sequences show rather low similarity (around 25% with *E. coli* ArgR, 36% with the *B. stearothermophilus* homologue) [28]. With the determination of the genome sequence of *T. maritima*, the biosynthetic *argGHCJBD* and *purBA-argF* gene clusters could also be identified and the *argR* sequence confirmed [41].

***Figure 1*** | Diagram representing binding of a single hexameric ArgR protein to a biosynthetic operator constituting a pair of ARG boxes (open rectangles)

ARG boxes (open rectangles) are ideally separated by 3 bp in class I organisms and by a 2 bp spacer in class II organisms. Each ARG box constitutes an 18 bp-long imperfect palindrome (arrows). In all cases investigated, the operator overlaps the promoter elements (grey bars) at least partially, and repression by *E. coli* ArgR was shown to be by steric exclusion of RNA polymerase binding.
Figure 2 | High-resolution contact mapping of ArgRTn binding to its own control region (argR) and to the operator of the biosynthetic argGHJBD operon of T. maritima (argG)

ARG box sequences are numbered. Rectangles indicate zones of protection against DNase I cleavage, vertical bars with a dot indicate positions that become hyper-reactive for DNase I (filled symbol for a strong effect, open symbol for weak effects). The vertical grey lines indicate the limits of protection in the absence of arginine. Squares and triangles represent purines and pyrimidines, respectively, the removal of which interferes strongly (dark filled symbols), moderately (grey-filled symbols) or weakly (open symbols) with repressor binding.

The recombinant T. neapolitana ArgR repressor (ArgRTn) was overexpressed in E. coli and purified as a homotrimeric protein that assembles into hexamers at higher protein concentrations and/or in the presence of arginine [27,28]. ArgRTn proved to be intrinsically thermostable, complex formation was more efficient at temperatures up to 80–90°C than at 37°C, and repressor binding increased the melting temperature of the DNA fragments by approx. 15°C [28].

The interaction of ArgRTn with its own control region and with the biosynthetic argG operator of T. maritima was analysed in vitro by mobility-shift analysis, DNase I and hydroxyl radical footprinting, missing contact probing and various base-specific premodification binding-interference techniques (Figure 2) [27–29]. Binding of ArgRTn to its own control region proved to be about 2-fold more efficient in the absence than in the presence of arginine [27,28]. Moreover, the DNase I footprinting indicated that the protected areas of complexes formed in the presence of arginine are slightly longer and show an extra hyper-reactive site, two observations indicative of the formation of structurally different complexes in the two conditions [27].

A close inspection of the arginine-binding motifs of arginine repressors indicates that ArgRTn and its homologue from T. maritima have a serine instead of a glutamine at position 107 of ArgRTn. Substitution of Ser-107 by glutamine results in a mutant protein that exhibits a marked need for arginine as co-factor (about 18-fold stimulation on binding to its own operator) and an enhanced target specificity that is also reflected in the observation that, unlike the wild-type repressor, the mutant protein does not give rise to growth inhibition of E. coli transformants [29]. Therefore, the
unusual behaviour of ArgRTn can at least in part be ascribed to the presence of serine at position 107.

**Classification of ArgR proteins**

The analyses of the hyperthermophilic arginine repressors, combined with previous data and detailed studies of operator sequences, allowed us to propose a grouping of bacterial arginine repressors of the wHTH family in three major classes [29]. Class I repressors are very sequence-specific binders (they hardly bind operators from class II and class III organisms) that are highly arginine-dependent (200–1000-fold stimulation). These repressors are typically found in the γ-group of proteobacteria (*E. coli, S. typhimurium* and *M. profunda*). Their high-affinity targets in the biosynthetic operators consist of a tandem pair of ARG boxes ideally separated by a 2 bp-long spacer that is invariably folded [34,35]. Their native targets consist of a tandem pair of same class. Class I and class II repressors share only 25–30% amino acid sequence identity but they adopt the same fold [34,35]. Their native targets consist of a tandem pair of ARG boxes ideally separated by a 2 bp spacer that invariably contains at least one strong base pair [27]. The repressors from *B. subtilis* and *B. stearothermophilus* were isolated as trimeric proteins but they exist in a concentration-dependent equilibrium with the hexameric form, which is stabilized in the presence of arginine [26,43]. Finally, the repressors from the hyperthermophiles *T. maritima* and *T. neapolitana* constitute a third class. Their binding is only marginally influenced by arginine, they primarily interact with a single ARG-box-like sequence of the operator and exhibit a very broad target specificity. ArgRTn was purified as a trimeric protein but as the class II repressors it can form hexamers.

The occurrence of various types of arginine repressors in evolutionary distant bacteria may very well reflect the evolution of ArgR proteins from a general regulator exhibiting poor target specificity and low regulatory efficiency (class III proteins) to highly specialized and remarkably efficient molecules (class I) that might provide a selective advantage.

**References**


Received 19 September 2003