

Hyperthermophilic *Thermotoga* Arginine Repressor Binding to Full-length Cognate and Heterologous Arginine Operators and to Half-site Targets

Amélie Morin¹, Nadine Huysveld², Frédérique Braun¹, Diliانا Dimova¹
Vehary Sakanyan^{1,3} and Daniel Charlier^{2*}

¹Laboratoire de Biotechnologie
FRE CNRS 2230 Unité
Biocatalyse, Faculté des
Sciences et Techniques
Université de Nantes, 2 rue de
la Houssinière, 44322 Nantes
France

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

³ProtNeteomix, Université de
Nantes, 2 rue de la Houssinière
44322 Nantes, France

The degree of sequence conservation of arginine repressor proteins (ArgR) and of the cognate operators (tandem pairs of 18 bp imperfect palindromes, ARG boxes) in evolutionarily distant bacteria is unusually high, and the global mechanism of ArgR-mediated regulation appears to be similar. However, here we demonstrate that the arginine repressor from the hyperthermophilic bacterium *Thermotoga neapolitana* (ArgR_{Tn}) exhibits characteristics that clearly distinguish this regulator from the well-studied homologues from *Escherichia coli*, *Bacillus subtilis* and *B. stearothermophilus*. A high-resolution contact map of ArgR_{Tn} binding to the operator of the biosynthetic *argGHCJBD* operon of *Thermotoga maritima* indicates that ArgR_{Tn} establishes all of its strong contacts with a single ARG box-like sequence of the operator only. Protein array and electrophoretic mobility-shift data demonstrate that ArgR_{Tn} has a remarkable capacity to bind to arginine operators from Gram-negative and Gram-positive bacteria, and to single ARG box-bearing targets. Moreover, the overall effect of L-arginine on the apparent K_d of ArgR_{Tn} binding to various cognate and heterologous operator fragments was minor with respect to that observed with diverse bacterial arginine repressors. We demonstrate that this unusual behaviour for an ArgR protein can, to a large extent, be ascribed to the presence of a serine residue at position 107 of ArgR_{Tn}, instead of the highly conserved glutamine that is involved in arginine binding in the *E. coli* repressor. Consistent with these results, ArR_{Tn} was found to behave as a superrepressor in *E. coli*, inhibiting growth in minimal medium, even supplemented with arginine, whereas similar constructs bearing the S107Q mutant allele did not inhibit growth. We assume that ArgR_{Tn}, owing to its broad target specificity and its ability to bind single ARG box sequences, might play a more general regulatory role in *Thermotoga*

© 2003 Elsevier Ltd. All rights reserved.

Keywords: transcriptional regulation; protein–DNA contacts; arginine metabolism; hyperthermophiles; *Thermotoga*

*Corresponding author

Introduction

Structural studies, sequence comparisons and functional analyses indicate that arginine repressors from very divergent bacteria, Gram-negatives (*Escherichia coli*, *Salmonella typhimurium*, *Moritella profunda*, *Thermotoga neapolitana* and *Thermotoga*

maritima) and Gram-positives (various *Bacillus* and *Streptomyces* species, *Enterococcus faecalis*) belong to the superfamily of wHTH transcriptional regulators, are surprisingly well conserved and bind to similar operators.^{1–25} The *Pseudomonas aeruginosa* repressor, however, is an exception, as it belongs to the AraC/XylS family.^{26,27} The three resolved ArgR structures, from *E. coli*, *B. stearothermophilus* and *B. subtilis*, show great similarity in their folding patterns, even though the enterobacterial and bacillar proteins share only 27% sequence identity.^{10,12,17,22} Sequence alignments and secondary

Abbreviations used: EMSA, electrophoretic mobility-shift assay; IRDyes, near-infrared fluorescence dyes.

E-mail address of the corresponding author:
dcharlie@vub.ac.be

structure predictions indicate that the repressor of *T. neapolitana* and *T. maritima* belongs to the same family.¹⁸ ArgR proteins can act as repressor of the arginine biosynthesis and as activator of the different catabolic pathways.^{28–33} Moreover, *E. coli* ArgR plays an essential role in the assemblage of the nucleoprotein complex that resolves multimeric forms of ColE1-type plasmids in the monomeric constituents.^{34,35}

Purified ArgR proteins have been isolated as homohexamers or lower oligomeric forms (mainly trimers) that can form hexamers at high concentrations of protein and/or in the presence of arginine, the co-repressor.^{1,5,11,18,24,25,36,37} The N-terminal domain of ArgR provides binding to DNA, whereas the C-terminal domain is responsible for oligomerization and arginine binding; these domains are connected by a short linker.^{10,12,17,22,38} Six arginine molecules bind at the trimer–trimer interface; they reinforce the interactions between subunits within one trimer and between trimers. Ligand binding has been shown to induce a rotation of one trimer with respect to the other in the *B. stearothermophilus* repressor. This conformational change is assumed to improve the affinity of the hexameric repressor for operator DNA.¹⁷

Biosynthetic arginine operators generally consist of a tandem pair of ARG boxes, 18 bp imperfect palindromes separated by a short spacer (2 or 3 bp) that exhibits a specific length and cognate base-pair composition; 3 bp spacers are composed exclusively of weak base-pairs, whereas 2 bp long spacers contain at least one strong pair.²⁴ Stoichiometry measurements of the *E. coli* ArgR–operator binding and *in vitro* contact probing analyses with the ArgR molecules of *E. coli* (ArgR_{Ec}), *S. typhimurium* (ArgR_{St}), *B. subtilis* (AhrC), *B. stearothermophilus* (ArgR_{Bst}), and *M. profunda* (ArgR_{Mp}) indicate that a single ArgR holohexamer binds to the tandem pair of boxes and establishes contacts with two major groove segments and the intervening minor groove of each box, all aligned on the same face of the helix.^{3–6,11,14,24,25} The ArgR protein from the hyperthermophilic *T. neapolitana* (ArgR_{Tn}), however, exhibits remarkable deviations from this general pattern. Mobility-shift assays, DNase I footprinting and high-resolution contact mapping of ArgR_{Tn} binding to its own operator (*argR_O*) indicate that arginine does not stimulate complex formation but instead exerts a slight negative effect, that structurally different complexes are formed in the absence and in the presence of the co-factor, and that the vast majority of the strongest contacts are clustered within a shorter region of the operator as compared to other bacterial ArgR proteins.^{18,24} Moreover, recently ArgR_{Tn} has been shown to bind to a single ARG box-bearing fragment of the *B. stearothermophilus argC* operator (*argC_O*) nearly as efficiently as to the intact operator, whereas ArgR_{Bst} does so only with a reduced affinity.^{39,40}

The unusual degree of structural and functional conservation of bacterial ArgR proteins and their

targets in general, and the distinct features detected recently for ArgR_{Tn} make the arginine regulatory system being an attractive test case for the study of evolution of microbial transcription regulation. In this study, we aimed at a better understanding of the atypical behaviour of the *Thermotoga* repressor and of the molecular strategies of adaptation occurring in the arginine regulatory mechanism. We have analyzed ArgR_{Tn} binding to various cognate and heterologous arginine operators, established a high-resolution contact map of the protein binding to the operator sequence involved in regulation of the biosynthetic *argGHCJBD* operon in *T. maritima*, studied complex formation with a mutant *T. neapolitana* ArgR protein affected in the arginine-binding domain (S107Q), and determined the capacity of ArgR_{Tn} to repress transcription *in vitro* and *in vivo*. The data indicate that ArgR_{Tn} establishes all of its important contacts with a single ARG box-like sequence of the *T. maritima argG_O* operator and is able to bind to a wide variety of arginine-specific operators composed of single and tandem pairs of ARG boxes in an arginine-independent fashion, an observation that suggests a more global regulatory role for ArgR_{Tn}, supported also by growth inhibition occasioned upon introduction of the *argR_{Tn}* gene in *E. coli* cells.

Results

High resolution *in vitro* contact probing of ArgR_{Tn} binding to the *argG_O* operator

Previously, we have shown that ArgR_{Tn} contacts an unusually short sequence stretch in its own control region (Figure 1(f))²⁴ and that the repressor binds a DNA fragment bearing the sole downstream box of the heterologous *argC* operator of *B. stearothermophilus* rather efficiently.⁴⁰ To determine whether this binding pattern constitutes a general characteristic of the *Thermotoga* repressor–operator interaction or whether it is restricted to particular operators, we analyzed the interaction of ArgR_{Tn} with a cognate biosynthetic operator. A sequence bearing a putative promoter–operator region has been identified upstream of the *argGHCJBD* operon¹⁸ in the *T. maritima* genome sequence.⁴¹ Attempts to clone the equivalent operon from *T. neapolitana* were unsuccessful, however, likely due to the presence of a strong promoter (A.M., F.B. and V.S., unpublished results) that causes high-level expression of the cytotoxic *argJ* encoded ornithine acetyltransferase.⁴² Therefore, we used the *argG* promoter–operator region from *T. maritima* to analyze *in vitro* the binding of ArgR_{Tn} to the biosynthetic *Thermotoga* operator.

Electrophoretic mobility-shift assays (EMSA) performed at 37 °C and in the presence of arginine indicate that ArgR_{Tn} binds specifically to the *T. maritima argG* control region (*argG_O*) with an apparent equilibrium dissociation constant (K_d) of

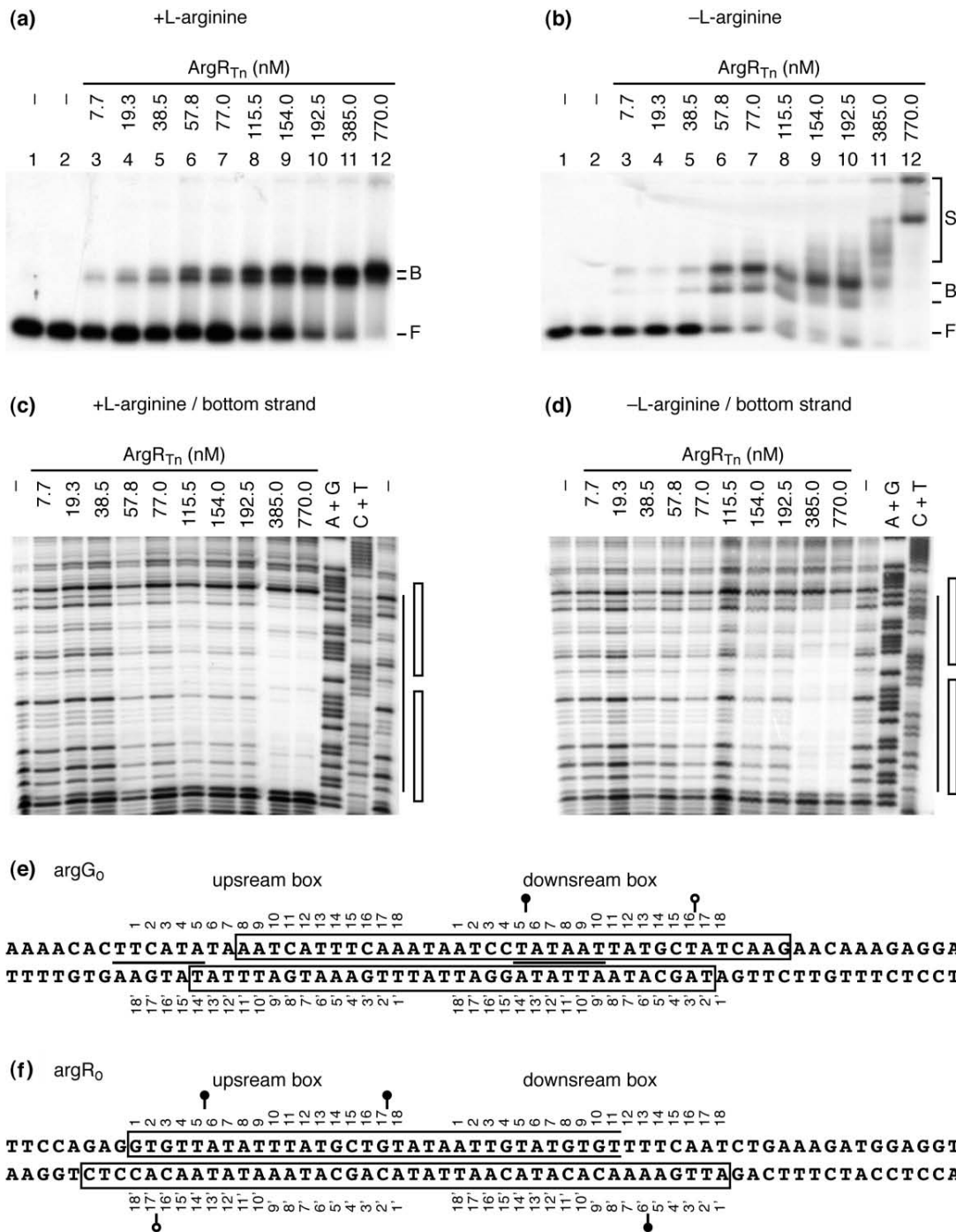


Figure 1. (a) and (b) EMSA and (c) and (d) DNase I footprinting of ArgR_{Tn} (as indicated, in nM monomer equivalents) binding to the (5'-³²P)-single-end-labelled *argG₀* operator DNA probe (bottom strand labelled) in the presence ((a) and (c)) and in the absence ((b) and (d)) of arginine. Labelled operator DNA was incubated at 37 °C and in the presence of an excess non-specific competitor with increasing concentrations of ArgR_{Tn} for 25 minutes in a total volume of 100 µl. Then, 10 µl aliquots were removed and immediately loaded onto pre-electrophoresed 5% polyacrylamide gels (with arginine in the gel and the running buffer for those assays performed in the presence of arginine, without arginine for the others). The positions of free DNA (F), of complexes (B) and of supershifted material (S) are indicated. The remaining 90 µl was used for DNase I footprinting and the reaction products analyzed by gel electrophoresis in denaturing conditions on 6% polyacrylamide gels. A + G and C + T are the corresponding Maxam–Gilbert sequencing ladders. Rectangles represent ARG box sequences, the global area of protection is indicated with a vertical bar. (e) and (f) Nucleotide sequence of the *argG* and *argR* (for comparison) control regions. Transcription proceeds in both cases to the right. ARG box sequences are numbered, rectangles represent regions protected against DNase I digestion by ArgR_{Tn} in the presence of arginine, vertical bars with a dot represent sites that become hyper-reactive to DNase I in the presence of the repressor (filled symbol for a strong effect, open symbol for a weaker effect). Potential promoter elements are underlined.

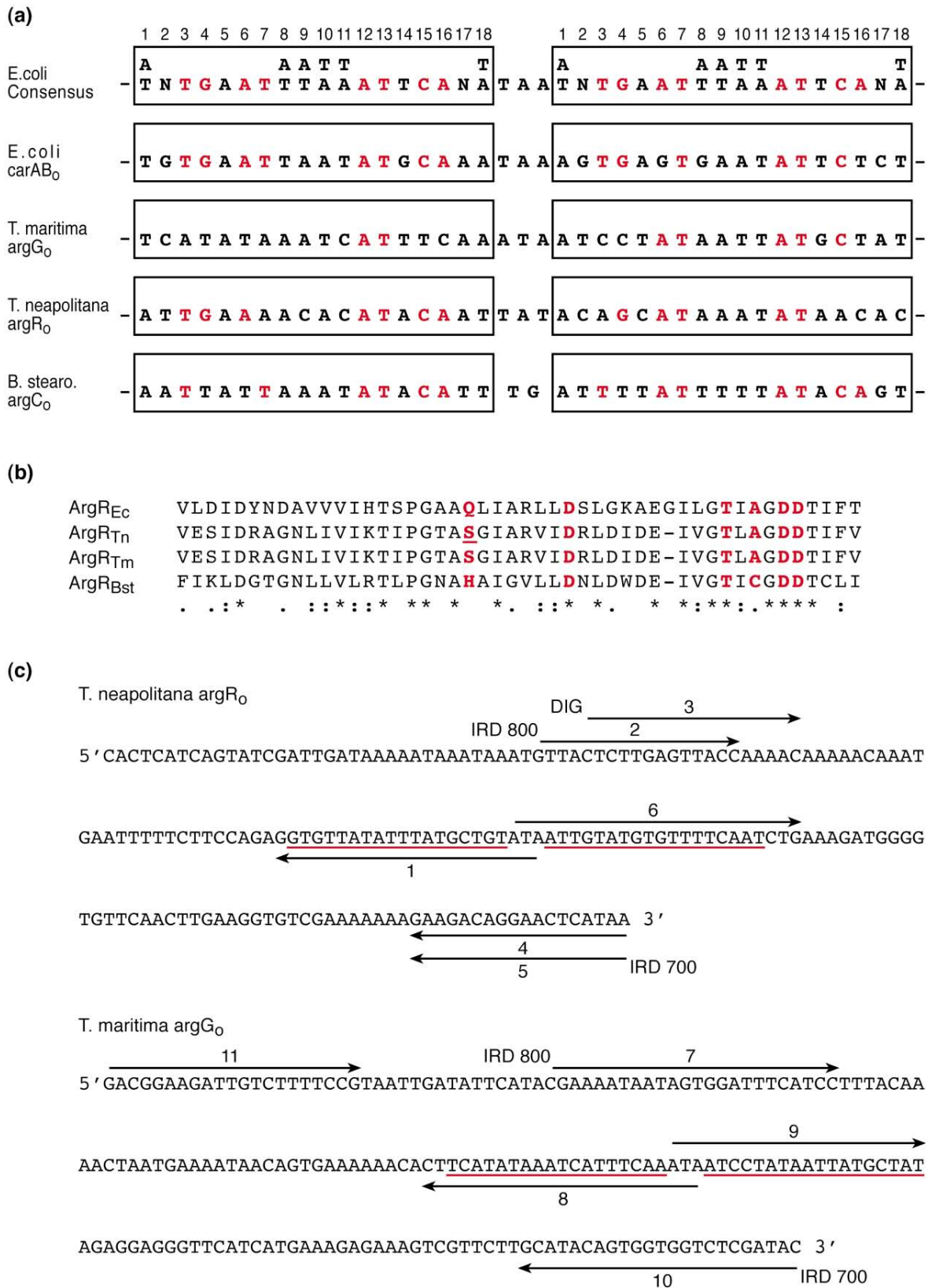


Figure 2. (a) Alignment of operator sequences and their dissection in ARG box equivalents for *E. coli* carAB₀ and consensus, *T. maritima* argG₀, *T. neapolitana* argR₀ and *B. stearothermophilus* argC₀. Bases that are highly conserved among the various *E. coli* operators studied are indicated in red, also when occurring in the operators from other organisms. Note that in this alignment the *argR* operator has been inverted with respect to the direction of transcription (to the left for *argR*, to the right for the others) to maximally align the sites of strong contacts established by ArgR_{Tn}

around 100 nM (Figure 1(a)). The retarded material consists of a doublet, but the differences between the two poorly resolved complexes are not clear. Both bands correspond to specific complexes, since their formations were equally resistant to increasing concentrations of non-specific competitor (not shown). No qualitative differences (extent of protected region, hyper-reactivity) could be observed in DNase I footprints at increasing concentrations of repressor, resulting in different ratios of the two complexes (Figure 1(a) and (c)), and no differences could be detected in the binding interference patterns of the separated upper and lower parts of the doublet in missing contact probing experiments (see below) or in the footprints of the DNA-protein complexes performed in the gel, using the 1,10-phenanthroline-copper ion (not shown). Therefore, it appears unlikely that this minor difference in migration velocities (not detected on agarose gels, see below) could be ascribed to binding of trimeric and hexameric repressor molecules. Possibly, the complexes correspond to the binding of liganded and unliganded hexameric ArgR_{Tn} molecules. The slight difference in net charge of the apo- and holorepressor, and small differences that might exist in the overall DNA bending induced by the two ArgR_{Tn} protein populations due to differences in the communication between subunits in the arginine-bound and unbound repressor molecules, could result in slightly different velocities. It is noteworthy that a similar kind of poorly resolved doublets has been observed in the presence of arginine upon binding of *E. coli* ArgR mutant proteins (S44F and S47L) affected in the DNA-binding domain and which still bind arginine normally, but no explanation was provided.⁸ Later, binding of the trimeric form of wild-type and mutant *E. coli* ArgR proteins has been proposed,^{9,43} but alternative explanations are not excluded and have been advanced since.³⁷

In the absence of arginine, ArgR_{Tn} binding was less specific and the complexes less stable, as indicated by the appearance of supershifts at higher concentrations of protein and an important smear (Figure 1(b)) but the effect on the apparent K_d was minor.

DNase I footprinting performed on the very same samples used for EMSAs in the presence and in the absence of arginine indicate that in both conditions ArgR_{Tn} protects against digestion 34–36 nt stretches that are offset by 3–5 nt towards the 3'-end on complementary strands of the *argG_O* operator (Figure 1(c)–(e)). This region, composed of between 21 nt and 59 nt upstream of the ATG

initiation codon of *argG*, overlaps the potential –10 and –35 promoter elements (Figure 1(e)). The sequence analysis of this stretch and comparisons with the *T. neapolitana argR_O* and other bacterial operator sequences, indicated that the *argG_O* operator can be dissected into two 18 bp long ARG box-like sequences separated by a 3 bp spacer (Figures 1(f) and 2(a)). The latter is composed exclusively of A-T and T-A base-pairs, as found in the *argR_{Tn}* operator, and in all 3 bp long spacers of arginine operators in general.^{24,25} ArgR_{Tn} binding induced the appearance of two sites hyper-reactive for DNase I, both located on the top strand of the downstream box, between residues T5 and A6 (strongest effect) and between T16 and A17 (weaker effect) (Figure 1(e)). The lack of qualitative differences in the footprints obtained in the presence and in the absence of arginine (Figure 1(c) and (d)) contrasts with our previous observations on binding of ArgR_{Tn} to its own control region (*argR_O*), where DNase I footprints of complexes formed in both conditions were significantly different, both in the extent of the protected regions and in the pattern of ArgR_{Tn}-induced hyper-reactivity (see also Discussion).²⁴

To identify the bases that play a primordial role in the ArgR_{Tn}–*argG_O* complex formation, we have performed missing contact probing experiments⁴⁴ applied to purines and pyrimidines. Sparingly depurinated or depyrimidated operator DNA molecules were incubated in the presence of arginine with ArgR_{Tn} at concentrations suited to obtain approximately 50% binding. ArgR_{Tn}–operator complexes (B1, B2) were separated from free DNA molecules (F1, F2) by mobility-shift electrophoresis (Figure 3(a) and (d)), both forms recovered from the gel, treated with piperidine to induce strand scission at abasic sites, and the reaction products analyzed by gel electrophoresis in denaturing conditions, alongside a similar amount of input material (I) representing all possible sites of modification (Figure 3(b), (c), (e) and (f)). In those conditions, a molecule modified at a site essential for ArgR_{Tn} recognition is supposed to be under-represented in the bound form and therefore to enrich the free DNA population with respect to the input material, whereas a molecule modified at a site irrelevant for ArgR_{Tn} binding is expected to be distributed evenly. Citrate-induced depurination binding interference experiments indicated that the removal of either one of nine purines of the downstream box (A6, A8, A9, A12, G14 of the top strand and A6', A8', A9', and A12' of the bottom strand) interfered strongly with repressor binding

with the *argR_O*²⁴ and *argG_O* operator fragments (this work, Figure 3). (b) Alignment of the C-terminal domain of the ArgR proteins from *E. coli*, *T. neapolitana*, *T. maritima* and *B. stearothermophilus*. The six highly conserved amino acids involved in binding of arginine are indicated in red bold. The serine residue at position 107 in ArgR_{Tn} is underlined. (c) Location of primers used for the synthesis of labelled probes containing full-length operators (tandem pairs of ARG boxes) or single ARG box derivatives of the *argR_O* and *argG_O* operators. ARG box sequences are underlined in red.

(Figure 3(b), (e) and (g)). Additional moderate effects were observed upon removal of A12 and A18 of the upstream box, and of A1, A3', G4', G15', G16' and A17' of the downstream box, whereas a weak effect was observed at A8, A9, A17, G3', A4', A5', G8' and A9' of the upstream box, A17, and A1' of the downstream box, and at three positions in the spacer separating the two boxes (Figure 3(g)). The same interference patterns were observed when, in a separate experiment, the upper and lower parts of the doublet of complexes were recovered and analyzed separately (not shown). Hydrazine-induced depyrimidation revealed that the removal of either one of seven thymine bases, all located in the downstream box (T7, T10, T11, T13, T10', T11' and T13') severely affected complex formation (Figure 3(c), (f) and (g)). Moderate effects were observed at positions T10, C11, and T13, of the upstream box and at C15, T2', C5', and T7' of the downstream box, whereas weak effects occurred at T1', T2', T11' and T12' of the upstream box, at T18' of the downstream box, and at three T residues in the spacer (Figure 3(g)). Clearly, ArgR_{Tn} establishes the vast majority of its crucial contacts with one ARG box-like sequence of the *argG_O* operator, as it does in *argR_O*.²⁴ It is noteworthy that the location of that box within the tandem pair of boxes constituting the full-length operator (see also binding to single boxes, below) is inverted with respect to the orientation of transcription in *argR_O* (upstream box) and *argG_O* (downstream box). A helical presentation of the data indicates that nearly all the critical contacts are restricted to two major-groove segments and the intervening minor groove, all aligned on one face of the helix (Figure 4).

ArgR_{Tn} binding to heterologous arginine operators

To further analyze the DNA-binding characteristics of ArgR_{Tn}, we performed a comparative study of its interaction with the operators *argR_O*, *argC_O* and *carAB_O* from *T. neapolitana*, *B. stearothermophilus* and *E. coli*, respectively. Complex formation at 37 °C in the presence and in the absence of arginine was monitored by EMSA on agarose gels. As observed for *argR_O*,¹⁸ binding of ArgR_{Tn} in the presence of arginine resulted in the formation of a single retarded species (Figure 5(a)) (which appears to be split into two poorly resolved bands on polyacrylamide gels), whereas binding in the absence of arginine resulted in two complexes exhibiting distinct migration velocities with all three operators (Figure 5(b)). In the absence of arginine, higher concentrations of repressor resulted in the formation of a more slowly migrating complex that might correspond to the apohexamer-operator complex, whereas the faster-migrating band might correspond to the apotrimer-operator complex, as suggested previously by some of us.¹⁸

The ability of ArgR_{Tn} to bind to the *argR_O*, *argC_O* and *carAB_O* operators in the absence and in the pre-

Table 1. Affinity of ArgR proteins for *E. coli carAB_O*, *B. stearothermophilus argC_O* and *T. neapolitana argR_O* operators in the absence and in the presence of arginine

ArgR protein		K _d (M, monomer)		
		<i>carAB_O</i>	<i>argC_O</i>	<i>argR_O</i>
ArgR _{Tn}	– Arg	1.4 × 10 ^{–9}	4.9 × 10 ^{–9}	1.3 × 10 ^{–9}
	+ Arg	1.4 × 10 ^{–9}	3.8 × 10 ^{–9}	1.4 × 10 ^{–9}
S107Q	– Arg	8.3 × 10 ^{–9}	2.0 × 10 ^{–8}	7.9 × 10 ^{–9}
	+ Arg	1.1 × 10 ^{–10}	4.8 × 10 ^{–9}	4.3 × 10 ^{–10}
ArgR _{Ec}	– Arg	–	–	–
	+ Arg	3.6 × 10 ^{–9}	6.6 × 10 ^{–6}	5.1 × 10 ^{–6}
ArgR _{Bst}	– Arg	1.1 × 10 ^{–8}	1.6 × 10 ^{–8}	1.2 × 10 ^{–8}
	+ Arg	4.2 × 10 ^{–9}	4.4 × 10 ^{–9}	1.1 × 10 ^{–8}

sence of arginine was further investigated by the surface plasmon resonance technique. The DNA binding affinity of ArgR_{Tn} for the three operators was found to be almost identical in the absence and in the presence of arginine (Table 1). Under similar conditions, arginine stimulated binding of ArgR_{Bst} to the *argC_O* and *carAB_O* operators nearly three- to fourfold,³⁹ whereas binding to the *argR_O* operator was similar in the absence and in the presence of arginine, but approximately tenfold weaker than binding of ArgR_{Tn} to the same target (Table 1). As expected from previous observations,^{24,36,39,45–48} a significant binding of ArgR_{Ec} was observed only to the *carAB_O* DNA and solely in the presence of arginine (Table 1).

These results indicate that ArgR_{Tn} binds to all, its own operator, the biosynthetic *Thermotoga argG_O* operator (see above) and the heterologous biosynthetic operators from *E. coli* and *B. stearothermophilus* with similar apparent affinities in the presence and in the absence of arginine. These features clearly distinguish ArgR_{Tn} from the *E. coli* and *M. profunda* repressors that are strongly arginine-dependent and hardly bind to arginine operators from heterologous origin,^{24,25,36,45–48} but are, to some extent, reminiscent of the *B. stearothermophilus* repressor, which is less arginine-dependent under similar conditions.³⁹ Thus, ArgR_{Bst} seemingly occupies a position in between ArgR_{Ec} and ArgR_{Tn}.

The S107Q substitution results in an enhanced arginine-dependence of ArgR_{Tn}–operator binding

An inspection of the C-terminal arginine-binding domain of ArgR_{Tn} reveals that five out of the six highly conserved amino acids involved in binding of arginine are present in the repressor from hyperthermophilic origin (see Figure 2(b)), but the glutamine residue at position 106 of the *E. coli* protein is a serine in ArgR_{Tn} (position 107) and in the *T. maritima* repressor protein.¹⁸ To verify whether this substitution might be responsible, in part, for the atypical arginine-independent behaviour of the *Thermotoga* repressor, we have constructed the S107Q derivative of ArgR_{Tn} and compared binding

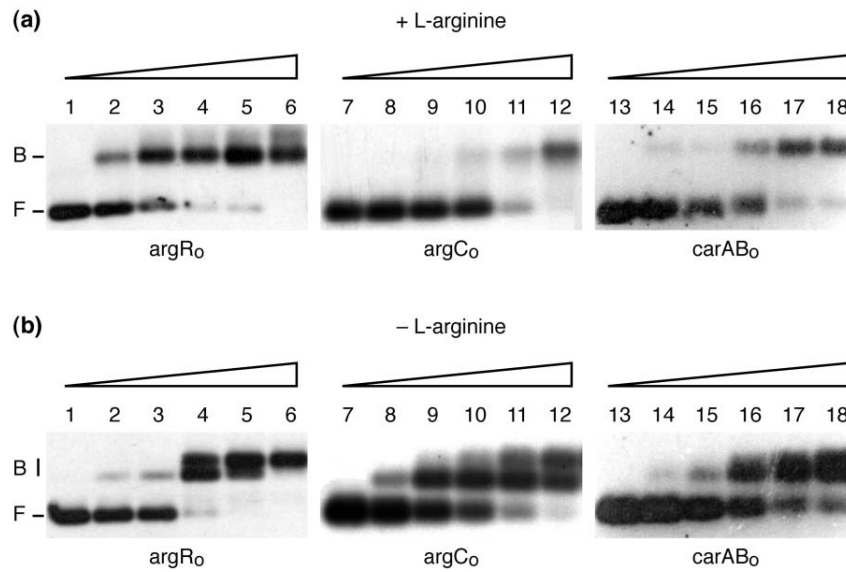


Figure 5. Gel retardation of ArgR_{Tn} binding to the *T. neapolitana argR_O*, *B. stearothermophilus argC_O* and *E. coli carAB_O* DNA probes in the presence (a) and in the absence of arginine (b). Labelled probes were obtained by PCR amplification using the oligonucleotides 3 and 4 for *argR_O*, 12 and 13 for *carAB_O* and 16 and 17 for *argC_O* (see Table 4). Binding to the *T. neapolitana argR_O* operator: lanes 1, without repressor, lanes 2–6 with 100, 300, 600, 1000 and 2000 nM ArgR_{Tn}. Binding to the *B. stearothermophilus argC_O* operator: lanes 7, without repressor, lanes 8–12 with 50, 100, 200, 400 and 800 nM ArgR_{Tn}. Binding to the *E. coli carAB_O* operator: lanes 13, without repressor, lanes 14–18 with 100, 200, 550, 1000 and 1600 nM ArgR_{Tn}. The position of free DNA (F) and of protein–DNA complexes (B) are indicated.

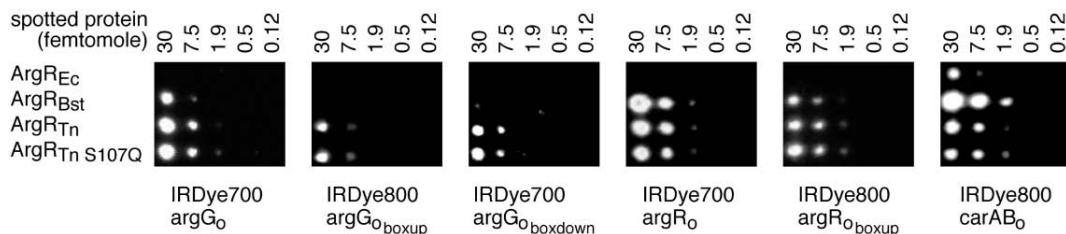


Figure 6. Evaluation of the ability of ArgR proteins to bind single and tandem pairs of ARG boxes by protein array. Purified ArgR proteins from *E. coli* (ArgR_{Ec}), *B. stearothermophilus* (ArgR_{Bst}), *T. neapolitana* (ArgR_{Tn}) and its S107Q derivative were serially fourfold diluted and spotted on NC membranes. The amount of repressor protein spotted is indicated on top of each panel. Binding reactions were performed with equal quantities of IRDye800 or IRDye700 labelled probes obtained by PCR amplification using the oligonucleotides 10 and 11 for *argG_Oboxup*, 7 and 8 for *argG_Oboxdown*, 9 and 10 for *argG_Oboxdown*, 2 and 4 for *argR_O*, 2 and 1 for *argR_Oboxup* and 14 and 15 for *carAB_O* target DNAs.

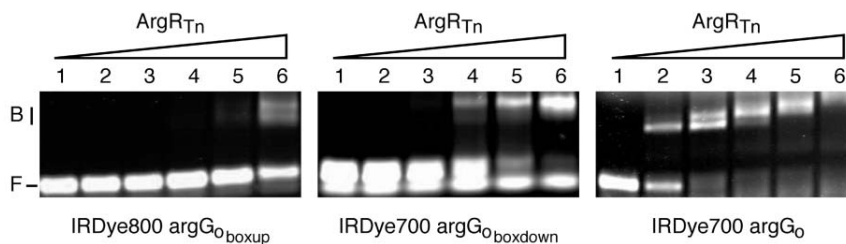


Figure 7. ArgR_{Tn} binding to the *argG_O*, *argG_Oboxup* and *argG_Oboxdown* DNA probes. IRDye labelled DNA fragments (3.0 ng) were incubated with increasing amounts of purified ArgR_{Tn} at 37 °C for 30 minutes in the presence of 10 mM L-arginine. Lanes 1, without repressor. Lanes 2–6, with 100, 200, 400, 800 and 1600 nM ArgR_{Tn}.

efficiently, respectively, than to the intact operator $argG_O$ (K_d nearly 100 nM) (Figure 7). ArgR_{Tn} binding to the full-length $argR_O$ probe proved to be only about twofold more efficient (K_d 120 nM) than to the $argR_{O_{boxup}}$ fragment (K_d 200 nM; not shown). Consistent with the *in vitro* contact probing results, ArgR_{Tn} bound the downstream ARG box more efficiently than the upstream one, though this quantitative difference was not detected in the protein array experiment (see above).

Combined, the results indicate that ArgR_{Tn} exhibits a high affinity for binding to various full-length operators, and it has a remarkable capacity to bind at least some single ARG box fragments.

ArgR_{Tn} represses transcription at the $argR_{Tn}$, $argG_{Tm}$ and $argC_{Bst}$ promoter–operators *in vitro*

To evaluate whether ArgR_{Tn} exerts repression or activation on transcription we analyzed reporter gene expression in a coupled transcription–translation system of *E. coli* in the absence and in the presence of purified ArgR_{Tn}. The potential regulatory effect of ArgR_{Tn} was tested on the expression of the *T. maritima argG* and *B. stearothermophilus argC* genes transcribed from the *T. maritima PargG* and *B. stearothermophilus PargC* promoter–operator regions, respectively. To avoid the autoregulatory interference of ArgR_{Tn} in the assay with the *T. neapolitana argR* template encoding the regulatory protein itself, the $argR_O$ promoter–operator region was fused to the *B. stearothermophilus argC* reporter gene. In parallel, the *B. stearothermophilus amaB* gene encoding L-carbamoylase fused to the phage T7 promoter region⁴⁹ was used as a negative control.

A significant decrease in protein synthesis was observed for all constructs bearing arginine-specific promoter–operator regions, proportional to the amount of ArgR_{Tn} added (Figure 8). At 110 pmol, the ArgR_{Tn} protein reduced the expression initiated from the *B. stearothermophilus*

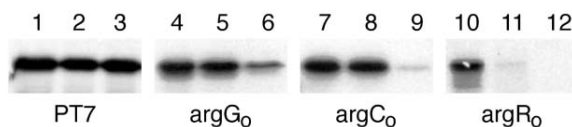


Figure 8. Autoradiogram of the *in vitro* protein expression from fused reporter genes in the presence of exogenous ArgR_{Tn}. Lanes 1–3, *B. stearothermophilus amaB* expressing from the phage T7 promoter (negative control); lanes 4–6, *T. maritima argG* expression from the $argG_O$ promoter–operator region; lanes 7–9, *B. stearothermophilus argC* expression from the $argC_O$ promoter–operator region; lanes 10–12, *B. stearothermophilus argC* expression from the *T. neapolitana argR_O* promoter–operator region. Lanes 1, 4, 7 and 10 without repressor, lanes 2, 5, 8 and 11 with 55 pmol of purified ArgR_{Tn}, and lanes 3, 6, 9 and 12 with 110 pmol of ArgR_{Tn}.

argC and the *T. maritima argG* promoter–operator regions by 11-fold and sixfold, respectively, and 55 pmol of the protein diminished the *argC* reporter gene expression initiated from the *T. neapolitana argR* control region by sixfold. Only a very slight negative effect (about 1.2-fold) was detected for the T7-*amaB* control. These data indicate that ArgR_{Tn} can act as a transcriptional regulator that represses both native and heterologous arginine promoters.

Wild-type and S107Q mutant ArgR_{Tn} proteins act as a repressor in *E. coli* cells

The capacity of recombinant ArgR proteins from *T. neapolitana*, *B. stearothermophilus* as well as *T. maritima*(ArgR_{Tn}) to repress gene transcription was further confirmed in *E. coli* cells bearing two compatible plasmids, one carrying a heterologous *argR* gene, the other bearing the *lacZ* reporter gene put under the control of the *B. stearothermophilus argC* promoter–operator.⁴⁶ As demonstrated previously,^{46–48,24} ArgR_{Ec} is unable to repress the *B. stearothermophilus argC* promoter activity. Hence, the β-galactosidase levels detected with the fusion construct in *E. coli* reflect the capacity of the heterologous repressor to block the *B. stearothermophilus argC* promoter activity. The results indicate that the expression of both *Thermotoga argR* genes had a weak negative effect on β-galactosidase production (1.5-fold), whereas the mutant $argR_{Tn}$ allele (S107Q) occasioned an almost fivefold drop in enzyme activity (Table 2). In similar conditions, the *B. stearothermophilus argR* gene occasioned an almost 14-fold reduction.

Moreover, we observed that *E. coli* transformants bearing the wild-type $argR_{Tn}$ gene on a multi-copy vector hardly grew in defined minimal medium, even when supplemented with arginine (Table 2). In contrast, recombinant cells carrying the plasmid bearing the S107Q mutant construct grew well, with and without arginine. Therefore, it appears that the S107Q substitution leads to the production of a mutant ArgR_{Tn} repressor protein with a higher specific binding and repression capacity in the presence of arginine (see above), but that has lost the growth-inhibitory effect when overexpressed in *E. coli* cells. As shown previously,⁴⁸ *E. coli* transformants overexpressing the *B. stearothermophilus argR* gene hardly grow on minimal medium supplemented or not with arginine.

Discussion

Arginine regulation in *Thermotoga* species is extremely poorly documented, but a small reduction (about twofold) in *argB* and *argH*-encoded enzyme activities has been observed upon addition of 5 mM arginine to *T. maritima* cells grown in a medium containing 0.05% yeast extract.⁵⁰ In this study, the data obtained both *in vitro* and *in vivo* indicate that the ArgR_{Tn} regulatory

Table 2. β -Galactosidase activity and repression levels mediated *in vivo* by the *Thermotoga* ArgR proteins

Plasmid	Source of the <i>argR</i> gene	β -Galactosidase ^a	Repression coefficient ^b	Growth in minimal medium supplemented	
				Arg +	Arg –
PCR2.1Topo	–	94.9 \pm 8.3	–	nd ^c	nd
pCRargR _{Tn}	<i>T. neapolitana</i>	63.3 \pm 6.1	1.5	+	+
pCRargR _{TnS107Q}	<i>T. neapolitana</i> Ser107Gln	18.0 \pm 2.0	5.3	+++	+
pCRargR _{Tm}	<i>T. maritima</i>	60.7 \pm 5.8	1.5	+	+
pCRargR _{Bst}	<i>B. stearothermophilus</i>	6.8 \pm 1.1	14.0	+	+

^a Expressed as nmol/minute per 1 mg of protein.

^b The repression level is calculated as a ratio of the enzymatic activity displayed in extracts of cells carrying the pCR2.1Topo vector without insert and cells carrying recombinant plasmids with a wild-type or mutant *argR* gene.

^c Not determined.

protein is able to exert transcriptional repression of reporter genes, as do other bacterial arginine repressors. However, several features clearly distinguish complex formation with the hyperthermophilic *Thermotoga* ArgR protein from similar interactions in evolutionarily distant bacteria, even though sequence comparisons indicate that ArgR_{Tn} belongs to the structurally and functionally well-conserved family of bacterial wHTH arginine repressors.

Here, we demonstrate that ArgR_{Tn} bound the operators from *E. coli* and *B. stearothermophilus* efficiently, whereas ArgR_{Ec} hardly binds the operators from *Thermotoga*, *B. subtilis* or *B. stearothermophilus*. Moreover, we found that ArgR_{Tn} has the remarkable property to establish all its strong contacts with a single ARG box of the *Thermotoga* *argG* operator. ArgR_{Tn} interacts primarily with two major-groove segments and the intervening minor groove, all aligned on one face of the helix in *argG_O* (Figure 4) and the *argR_O* operators.²⁴ The hyper-reactivity for DNase I induced between residues 5 and 6 of the ARG boxes in the *argG_O* and *argR_O* operators constitutes another signature of ArgR_{Tn} binding. It is noteworthy that ArgR_{Tn} interacts mainly with the downstream ARG box of the *argG_O* operator, whereas in the *argR_O* operator the upstream box constitutes the major recognition element. Therefore, this particular binding mode of ArgR_{Tn} is not restricted to autoregulation but instead constitutes a general property of the ArgR_{Tn}–operator complex. Consistent with this observation, ArgR_{Tn} was found to bind the single ARG box-bearing fragments *argG_Oboxdown* and *argR_Oboxup* rather well, as indicated by protein array binding experiments and mobility-shift assays. The ArgR_{Tn} protein presents an affinity only twofold and sixfold lower for the operators bearing the single *argG_Oboxdown* and *argR_Oboxup* than for the full-length operators containing a pair of ARG boxes. ArgR_{Tn} has been shown to bind to a DNA fragment bearing the single downstream ARG box of the *B. stearothermophilus* *argC* operator.⁴⁰ In contrast, in the presence of arginine, binding of the *E. coli* repressor shows a chelate effect that characterizes binding of a single homo-

hexameric repressor molecule to a tandem pair of correctly spaced ARG boxes.^{3,4} Binding of ArgR_{Ec} to a single box or to a pair of incorrectly spaced boxes is possible, but with a significantly reduced affinity.^{3,4,34,37,43}

It has been shown that ArgR_{Tn} is able to bind its own operator with a high degree of efficiency in the absence of arginine.^{18,24} In contrast, no binding of the ArgR_{Ec} repressor was observed in the absence of arginine and the affinity of the AhrC repressor of *B. subtilis* was found to be reduced by 100-fold.⁵¹ Our data indicate that the ligand has only a minor effect on the overall apparent *K_d* of ArgR_{Tn} binding to cognate and heterologous arginine operators, even though arginine affects the ratio of sequence-specific to non-specific DNA binding and the stability of the ArgR_{Tn}–operator complexes. Moreover, we have demonstrated that the arginine-independent binding of ArgR_{Tn} is related, at least in part, with the presence of a serine residue at position 107 of the repressors from *T. neapolitana* and *T. maritima*,¹⁸ instead of the highly conserved glutamine known to be involved in arginine binding in the *E. coli* and *B. stearothermophilus* repressors.^{10,17} Overexpression of the wild-type *argR_{Tn}* gene leads to growth inhibition of *E. coli* transformants, even in the presence of arginine, whereas the S107Q allele does not affect cell growth. These observations indicate that the substitution of serine 107 by glutamine might narrow the capacity of the *Thermotoga* repressor to bind potential (single) ARG-box like targets not belonging to the arginine regulon in *E. coli*.

A single amino acid substitution distinguishes the ArgR proteins from *E. coli* K-12 and *E. coli* B (L instead of P at position 70), which show different ratios of ArgR binding in the presence and in the absence of arginine, a property that alters the regulatory response significantly, the *E. coli* B repressor being a super-repressor that largely inhibits transcription even at limiting concentrations of arginine.⁷ In *E. coli*-K12, the concentration of ArgR molecules (about 600/cell in minimal medium, 300 when supplemented with arginine), the *K_d* of the liganded ArgR–operator binding (sub-nanomolar range), the affinity of the repressor for arginine

(K_d 10^{-4} M) and the large differences in binding of the liganded and unliganded repressor molecules are such that *in vivo* repression is dictated primarily by the concentration of activated ArgR molecules, and thus by the concentration of the corepressor, arginine.⁵² As *in vitro* the efficiency of ArgR_{Tn}-operator binding is influenced by arginine only marginally, we may expect a different situation in *Thermotoga*, though it is still possible that the increase in binding specificity observed in the presence of arginine is sufficiently important to be determinant for the *in vivo* response. Alternatively, arginine might exert its effect mainly by influencing the synthesis of ArgR, and/or operator-bound liganded and unliganded ArgR_{Tn} molecules might differently affect the *argG* promoter activity, possibly by direct contact with the polymerase or through subtle (since not detected in DNase I footprinting on linear molecules) alterations of the DNA structure. A comparison of the contact maps of ArgR_{Tn} binding to the *argR_O* and *argG_O* operators in conjunction with the differential effect of arginine on complex formation with the two *Thermotoga* operators allows us to propose that *Thermotoga* can use two different molecular strategies to modulate expression of the arginine biosynthetic operon and to autoregulate repressor synthesis. The property of ArgR_{Tn} to bind to native and heterologous arginine operators with similar efficiency in the presence and in the absence of arginine is likely related to its capacity to establish tight contacts with a single ARG box, a binding mode that does not require the arginine-induced transition that improves target recognition of the other bacterial arginine repressors to various extents. Therefore, it appears that ArgR_{Tn} can behave as a kind of global regulator in *Thermotoga* cells, as has been postulated by Maas⁵² for arginine regulators.

On the basis of the arginine-dependence and the target specificity, bacterial arginine repressors of the wHTH family can be grouped into three major classes. Class I repressors are highly arginine-dependent and very sequence-specific binders. The *E. coli*, *S. typhimurium* and *M. profunda* proteins belong to this class. These repressors exhibit a high degree of sequence conservation ($\geq 70\%$ amino acid sequence identity) and exist primarily as hexamers, even in the absence of arginine and at rather low concentrations of protein.^{1,5,25,37} Arginine strongly improves their binding to the cognate operators (about 200–1000-fold) and to operators from other class I organisms, but they hardly interact with operators from the two other classes (represented by *Bacillus* and *Thermotoga* systems). The arginine-mediated allosteric activation involves an alteration in the quaternary structure and requires communication among subunits of the ArgR hexamer, but subunit assembly appears to play no role in activation.³⁷ The targets of class I repressors consist of a tandem pair of ARG boxes, ideally separated by a 3 bp long spacer composed exclusively of A-T and T-A base-pairs. Class I repressors

induce DNA bending by approximately 70–90°^{4,9,37} and exert repression by competing with RNA polymerase for binding to partially overlapping sites,² but they are unable to block an elongating polymerase.^{53,54,25}

Class II comprises the repressor from various Gram-positive *Bacillus* species and the less well studied repressors from *Streptomyces* might belong to the same class as well. Class II and class I repressors share only 25–30% amino acid sequence identity but they adopt the same fold and show a similar organization,^{17,22} and both classes function as repressors and as co-activators for arginine catabolism. Purified *B. subtilis* AhrC and *B. stearothermophilus* ArgR are trimers, but they can form hexamers at high concentrations of protein and/or in the presence of arginine.^{11,36,39} Their binding to DNA is moderately arginine-dependent (between fourfold and 100-fold stimulation, dependent on the protein and the buffer composition). Their cognate targets consist of a tandem pair of ARG-box like sequences, separated by a 2 bp spacer that contains at least one G-C or C-G pair. Repressors of this class show a broad target specificity and can bind operators from all three classes rather efficiently, both *in vitro* and *in vivo*. The *B. stearothermophilus* repressor is able to bind a single ARG box but with a reduced affinity as compared to the full length operator.³⁹

The repressor from *T. neapolitana* and likely *T. maritima* (97% identity) constitute a third class of proteins that interact primarily with a single ARG box-like sequence of the cognate operators and exhibit a remarkable capacity to bind to arginine operators from heterologous origin and to single ARG box-bearing fragments. Their binding is influenced by arginine only marginally.

Thus, arginine repressors can be of different kinds; at one extreme there are the global kind of regulators exhibiting low repression efficiency and poor target specificity, as postulated for the *Thermotoga* ArgR. At the other extreme there are the highly specified arginine-dependent repressors present in the γ group of Proteobacteria, the *Bacillus* repressors being somewhere in between. Evidently, the degree of arginine-dependence and the potential to bind heterologous operators do not constitute sharp borderlines separating class II and class III repressors. It is even likely that the identification and characterization of an even larger variety of organisms would reveal the existence of a continuum in nature. The *Bacillus* and *Thermotoga* proteins differ, however, in the way they contact the targets; *Bacillus* repressors establish strong contacts with both boxes of the naturally occurring biosynthetic operators, whereas the *Thermotoga* repressor exhibit a strong bias in the distribution of the strong contacts to a single box, as if ArgR_{Tn} were not optimally suited to contact a pair of adjacent boxes simultaneously.

The existence of functionally different classes of ArgR molecules in evolutionarily distant bacteria might reflect the evolution of regulatory mechanisms

and networks, from a general regulator exhibiting poor target specificity and low efficiency into highly specialized and efficient molecules that may provide a selective advantage. Though the wHTH motif of the DNA-binding domain predominantly determines recognition/binding to specific operator sequences, other structures as the $\alpha 4$ helix in the oligomerization/arginine-binding domain can modulate the ARG box recognition *via* interdomain contacts in oligomeric repressor molecules.³⁹ Even point mutations in this oligomerization/arginine-binding domain affect the DNA-binding parameters of the repressors from *E. coli* and *B. stearothermophilus*.^{4,48} Reciprocal exchange of the linkers connecting the N and C-terminal domains of ArgR_{Ec} and ArgR_{Bst} weakly affect the DNA-binding characteristics³⁹ and, moreover, a single amino acid difference in the linker region (position 70) appears to be responsible for the different behaviour of the *E. coli* K-12 and *E. coli* B repressors (see above).⁷ Thus, in spite of a pronounced degree of structural conservation of bacterial ArgR molecules and of their wHTH motif in diverse evolutionarily distant organisms, the ArgR_{Tn} protein exhibits a rather high degree of plasticity to adapt to full-length operators and half-site targets. Consequently, it is possible that transcriptional arginine repressors that do not correspond by their DNA-binding pattern to the three classes of binders already identified might

be found among other representatives of the bacterial domain.

Materials and Methods

Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids are described in Table 3. *E. coli* strains were grown at 28 °C or 37 °C in Luria–Bertani media⁵⁵ supplemented with the appropriate antibiotics. Ampicillin was added at 100 µg/ml, kanamycin at 30 µg/ml and chloramphenicol at 25 µg/ml. Minimal medium M9⁵⁶ supplemented or not with arginine was used to assess arginine auxotrophy of *E. coli* Top10F' transformants carrying plasmid pCR-TOPO or derivatives thereof bearing an *argR* gene by cultivation at 37 °C for 48 hours, as described.⁴⁸

Construction, overexpression and purification of the mutant S107Q ArgR_{Tn} protein

The mutant *argR*_{Tn} allele was obtained by the overlap extension method⁵⁷ using Pfu DNA polymerase (Stratagene). DNA regions extending upstream and downstream of the site of interest were amplified by PCR, using as primers the pairs of oligonucleotides ArgRDeb (5'-AAGATGGAGGTGTTCAAC-3')-SerGlnDn (5'-GGA-ACTGCCCAAGGGATCGCCCGT-3') and ArgRFin (5'-CGAATCAGAGTATAGACG-3')-SerGlnUp (5'-GGC-GATCCCTTGGGCAGTTCTGG-3'). The two amplified fragments were subsequently combined in a PCR

Table 3. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant genotype or description	Reference or source
<i>E. coli</i> XL1-blue MRF'	$\Delta(mcrA)183 \Delta(mrr-hsdRMS-mcrBC)173 endA1 supE44 sth1-1 recA1 gyrA96 relA1 lac$ [F' <i>proAB lac</i> ^o Z Δ M15::Tn10(Tet ^r)]	Stratagene
<i>E. coli</i> BL21(DE3)	F ⁻ <i>ompT</i> [lon] <i>hsdS_B</i> (<i>r_B</i> ⁻ <i>m_B</i> ⁻) <i>gal dcm</i> (DE3)	Novagen
<i>E. coli</i> TOP10	F' <i>mcrA</i> $\Delta(mrr-hsdRMS-mcrBC)$ $\Phi 80 lacZ\Delta M15\Delta lacX74 deoR recA1 araD139 \Delta(ara-leu)7697 galU galK rpsL endA1 nupG$	Invitrogen
<i>B. stearothermophilus</i> NCIB8224	Wild-type	46
<i>T. maritima</i> MSB8	Wild-type	18
<i>T. neapolitana</i> DSM5068	Wild-Type	18
pET21d(+)	<i>bla</i> , T7 promoter expression vector	Novagen
pET21d-argR-His	<i>bla</i> , T7 promoter, the <i>T. neapolitana</i> DSM5068 wild-type <i>argR</i> gene fused to a C-terminal His-tag	18
pET21d-argRS107Q-His	<i>bla</i> , T7 promoter, carrying the <i>T. neapolitana</i> Ser107Gln <i>argR</i> mutant gene fused to a C-terminal His-tag	This work
pTN52	<i>bla</i> , <i>kan</i> , pCR2.1-TOPO carrying a 3 kb <i>argR</i> gene region from <i>T. neapolitana</i>	18
pHAS65	<i>cat</i> , <i>argC-lacZ</i> fusion in the pACYC184 vector under the control of the <i>B. stearothermophilus</i> <i>Parg</i> promoter	11
pETamaB	<i>bla</i> , pET3 T7 promoter, carrying the <i>amaB</i> gene from <i>B. stearothermophilus</i>	49
pHAV2	<i>bla</i> , pBTac2 carrying the <i>B. stearothermophilus</i> <i>PargC.argC</i> region	46
pCR2.1-TOPO	<i>bla</i> , <i>kan</i> , T7 and Plac promoters, PCR cloning vector	Invitrogen
pCR-argR _{Tm}	pCR2.1-TOPO, carrying the <i>T. maritima</i> MSB8 wild-type <i>argR</i> gene with its own Shine-Dalgarno and stop codon	This work
pCR-argR _{Tn}	pCR2.1-TOPO, carrying the <i>T. neapolitana</i> DSM5068 wild-type <i>argR</i> gene with its own Shine-Dalgarno and stop codon	This work
pCR-argR _{Bst}	pCR2.1-TOPO, carrying the <i>B. stearothermophilus</i> NCIB8224 wild-type <i>argR</i> gene with its own Shine-Dalgarno and stop codon	This work
pCR-argR _{S107Q}	pCR2.1-TOPO, carrying the <i>T. neapolitana</i> Ser107Gln <i>argR</i> mutant gene with its own Shine-Dalgarno and stop codon	This work

Table 4. Oligonucleotide primers used for amplification of operator DNA regions labeled with IRDye or Dig

Primer	Loc. ^a	Sequence 5'–3'
3' argRo Up	1	TATACAGCATAAATATAACACC
5' argRo 800	2	GTTACTCTTGAGTTACC
5' argRo Dig	3	CTCTTGAGTTACCAAAAC
3' argRo	4	TTATGAGTTCCTGTCTTC
3' argRo 700	5	TTATGAGTTCCTGTCTTC
5' argRo Down	6	ATAATTGTATGTGTTTTCAATCTG
5' argGo 800	7	CGAAAATAATATGTTGGATTTCATCC
3' argGoUp	8	TATTTGAAATGATTATATGAAG
5' argGo Down	9	ATAATGCTATAATTATGCTATC
3' argGo 700	10	GTATCGAGACCACCCTGTATGC
5' argGo	11	GACGGAGATTGTCTTTTCCG
3' carAB	12	CCGCCGAACCTG
5' carAB Dig	13	ACGTCATCATTGTGAATTAA
5' carAB	14	ACGTCATCATTGTGAATTAA
3' carAB 800	15	GCGCTGACTTAATCAAAACACCC
5' argCo Dig	16	GGCTGCCGGGACAAATCGG
3' argCo	17	CCCGTATGCCTCATGTAG

^a Location of primers, numbered from 1 to 11, is illustrated in Figure 2(c). DNA templates used for PCR amplifications are the plasmid pTN52 for primers 1–6, genomic DNAs from *T. maritima* for primers 7–11, from *E. coli* XL1 for primers 12–15 and from *B. stearothermophilus* for primers 16 and 17.

reaction using Goldstar Taq DNA polymerase (Eurogentec) with the flanking primers ArgRDeb and ArgRFin to amplify the full-length mutant *argR* gene, which was then inserted in the pCR2.1-TOPO vector (TOPO TA Cloning kit, Invitrogen). The presence of the mutation in the recombinant vector was verified by DNA sequencing.⁵⁸ To construct the expression vector, the coding part of the mutant *argR* allele was amplified using the TnR-NcoI (5'-TTCACCATGGAGGTGTCGAAAAAAG) and TnR-XhoI (5'-AACGAACTCGAGTATAGACG-AGAG) oligonucleotides as primers, and cloned into the NcoI and XhoI sites of pET21d + (Novagen) to obtain a C-terminal His₆-tagged protein. Expression was performed in *E. coli* BL21(DE3) induced with 0.1 mM IPTG at A_{600 nm} of 1.0 and further incubated for six to eight hours at 30 °C. The His₆-tagged repressor proteins were purified by affinity-chromatography on a Ni²⁺-NTA column as recommended (QIAexpressionist, Qiagen). Purified samples were analyzed by SDS-PAGE as described⁵⁹ and the concentration determined as described.⁶⁰

DNase I footprinting and missing contact probing

DNase I footprinting and missing contact probing experiments were performed with single (5'-³²P)-end labelled DNA fragments and non-tagged purified ArgR_{Tn} repressor, as described.^{14,24}

Electrophoretic mobility-shift assay (EMSA)

Mobility-shifty assays were performed either with single ³²P-end-labelled *argGo* fragments (see footprinting), DIG-labelled fragments⁶¹ or fluorescent IRDye-labelled DNA probes. The *T. neapolitana argRo*, *T. maritima argGo*, *B. stearothermophilus argCo* and *E. coli carABo* operator fragments were amplified by PCR using as templates the plasmid pTN52, and genomic DNA from *T. maritima* MSB8, *B. stearothermophilus* NCIB8224 and *E. coli* XL1, respectively, and a pair of primers, one of which was (5'-³²P)-labelled, 5'-end-DIG-labelled or labelled at the first

position with IRDye700 or 800 (Table 4). The PCR products were purified by polyacrylamide or agarose gel electrophoresis. Purified repressor proteins were diluted in the DNA-binding buffer (20 mM Tris-HCl (pH 7.9), 50 mM KCl, 50 mM NaCl, 0.1 mM DTT and 0.005% (v/v) surfactant P20 from Biacore). EMSA was performed as described,³⁹ except for the experiments containing IRDye-labelled fluorescent probes, when gels were screened immediately after migration with an Odyssey Imager (LI-COR, Inc.). The apparent equilibrium dissociation constants of the ArgR-DNA complexes (K_d , in monomer equivalents) correspond to the protein concentrations resulting in 50% binding.

Surface plasmon resonance

Biotin-labelled *argCo*, *argRo* and *carABo* operator fragments were prepared as described for the EMSA experiments, with a 5'-biotinylated oligonucleotide as primer. The fragments were injected at 5 ng/μl and a flow-rate of 5 μl/minute over the streptavidin-coated biosensor chip (Biacore BA). Binding assays were conducted in the DNA-binding buffer (see above) by injection of wild-type and mutant ArgR_{Tn} proteins at 10, 25, 50 and 100 nM (monomer equivalents) for eight minutes (association phase) followed by the injection of protein-free buffer for 15 minutes (dissociation phase) at a flow rate of 20 μl/minute, at room temperature, using a Biacore 2000 (Biacore AB) apparatus. When mentioned, L-arginine (at 10 mM) was included in the binding buffer. Data for protein-DNA interactions were evaluated from sensograms using the 1:1 binding model (BIA evaluation Software Handbook, 1999) and K_d values estimated for the ArgR monomer equivalents.

Protein arrays and fluorescent detection of DNA-protein interactions

His₆-tagged ArgR proteins were dialyzed against PBS, 0.3 M NaCl, 20% (v/v) glycerol and their concentration determined using the lab-on-chip protein 200 plus assay kit in the Bioanalyzer 2100 from Agilent technologies. Proteins at 30 μM were serially diluted fourfold and 1.0 nl was printed with a GSM 417 arrayer (Affymetrix) onto a BA83 nitrocellulose membrane (Schleicher & Schuell) attached manually to a glass slide.³⁹ Membranes were incubated in DNA-binding buffer (20 mM Tris-HCl (pH 7.9), 50 mM KCl, 50 mM NaCl, 0.1 mM DTT, 25 μg/ml of sonicated salmon sperm DNA) with 10 mM L-arginine. The labelled probes (at 1.0 ng/ml) were added to the pre-binding solution and incubated for 12 hours at room temperature, and rotated slowly. Membranes were washed with PBS containing 0.1% (v/v) Tween and fluorescent signals were detected at 700 nm or 800 nm with the Odyssey Imager (LI-COR, Inc.).

In vitro *argG*_{Tm}, *argR*_{Tn} and *argC*_{Bst} promoter-operator driven synthesis and ArgR_{Tn}-mediated repression

The structural *argC* gene with its control region was amplified from genomic DNA of *B. stearothermophilus* NCIB8224 with the primers upC_o (5'-CATAGACTTAGG-GAGGGG-3') and downargC (5'-CCGCTTGAAATCCT-TCCG-3'; located 95 nt downstream of the *argC* open reading frame), whereas the coding part was amplified with the primers upargC (5'-GATGGAGGTGTTCAACATGATGAACGTAGCCATTATCGGG-3') and downargC.

The *argG* gene and its corresponding control region was amplified from genomic DNA of *T. maritima* MSB8 with as primers the oligonucleotides up_{G_O} (5'-CCATACCC-CACGGCGAAAACG-3') and down_{argG} (5'-CCTGGAG-TGTACTATGGAGGCC-3'), whereas the *argR_{Tn}* promoter-operator region was amplified with pTNR52 plasmid DNA as template and as primers the oligonucleotides up_{R_O} (5'-GATCATCTTTGCACCCC-3') and down_{R_O} (5'-GTTGAACACCTCCATCTTTC-3'). To obtain the *argR_{Tn}* promoter-operator-*argC* fusion construct, the two fragments were mixed and amplified by overlap extension with the oligonucleotides up_{R_O} and down_{argG} as primers. The pT7_{amaB} template, used as a control, was amplified with the pET_{amaB} plasmid as template, and the universal primers PT7 and PT7_{ter}. *In vitro* transcription-translation assays were performed with S30 extracts prepared essentially as described⁶² but with minor modifications. Protein synthesis was performed with 60 ng of linear template DNA, at 37 °C for 90 minutes with 10 μCi of L-[³⁵S]-methionine (specific activity of 1000 Ci/mmol; 37 TBq/mmol; Amersham Pharmacia Biotech), when indicated in the presence of increasing amounts of Arg_{R_{Tn}} previously dialyzed against DNA-binding buffer. The synthesized products were separated by SDS-PAGE Gels were fixed on 3 mm paper and autoradiographs were made after treatment of the gels with an amplifier solution (Amersham Pharmacia Biotech). Quantification was performed with a Phosphor-Imager 6445 SI (Molecular Dynamics).

Detection of Arg_{R_{Tn}}-mediated repression in *E. coli* cells

It has been demonstrated that Arg_{R_{Ec}} is unable to repress the *B. stearothermophilus argC* promoter activity.^{46,24} The high copy number plasmid pCR-TOPO and derivatives carrying the wild-type or S107Q mutant *argR_{Tn}* transcribed from the *P_{lac}* promoter were transformed into *E. coli* Top10F' cells harboring the compatible low copy number plasmid pHAS65, bearing the *argC_{Bst}* :: *lacZ* reporter gene construct, in which *lacZ* transcription is driven by the *B. stearothermophilus argC* promoter-operator.⁴⁶ Hence, the β-galactosidase levels provided by the *lacZ* fusion construct in *E. coli* reflect the capacity of the heterologous ArgR protein to modulate the *B. stearothermophilus argC* promoter activity. A single colony of the transformants was taken from a fresh plate and used directly for growth to mid-exponential phase ($A_{600\text{ nm}}$ 0.4–0.6) in LB broth supplemented with chloramphenicol and kanamycin. To evaluate Arg_{R_{Tn}}-mediated repression, β-galactosidase assays were performed as described.⁵⁵

Acknowledgements

We greatly acknowledge J. -P. ten Have for the artwork and P. Vusio for the use of the Biacore 2000. This project has been supported by grants from the Région des Pays de la Loire (Contrat de Plan Etat-Région), the Ministère de L'Education et de la Recherche, CNRS (France) and the Fund for Scientific Research-Flanders (FWO Vlaanderen).

References

1. Lim, D., Oppenheim, J. D., Eckhardt, T. & Maas, W. K. (1987). Nucleotide sequence of the *argR* gene of *Escherichia coli* K12 and isolation of its product, the arginine repressor. *Proc. Natl Acad. Sci. USA*, **84**, 6697–6701.
2. Charlier, D., Weyens, G., Roovers, M., Piette, J., Bocquet, C., Piérard, A. & Glansdorff, N. (1988). Molecular interactions in the control region of the *carAB* operon encoding *Escherichia coli* carbamoylphosphate synthetase. *J. Mol. Biol.* **204**, 867–877.
3. Charlier, D., Roovers, M., Van Vliet, F., Boyen, A., Cunin, R., Nakamura, Y. *et al.* (1992). Arginine regulation of *Escherichia coli* K12. A study of repressor-operator interactions and of *in vitro* binding affinities versus *in vivo* repression. *J. Mol. Biol.* **226**, 367–386.
4. Tian, G., Lim, D., Carey, J. & Maas, W. K. (1992). How does the arginine repressor of *Escherichia coli* K12 bind to its operator sites. *J. Mol. Biol.* **226**, 387–397.
5. Lu, C.-D., Houghton, J. & Abdelal, A. T. (1992). Characterization of the arginine repressor from *Salmonella typhimurium* and its interactions with the *carAB* operator. *J. Mol. Biol.* **225**, 11–24.
6. Czaplewski, L. G., North, A. K., Smith, M. C. M., Baumberg, S. & Stockley, P. G. (1992). Purification and initial characterization of AhrC: the regulator of arginine metabolism genes in *Bacillus subtilis*. *Mol. Microbiol.* **6**, 267–275.
7. Tian, G., Lim, D., Oppenheim, J. D. & Maas, W. K. (1994). Explanation for different types of regulation of arginine biosynthesis in *Escherichia coli* B and *Escherichia coli* K-12 caused by a difference between their arginine repressors. *J. Mol. Biol.* **235**, 221–230.
8. Tian, G. & Maas, W. K. (1994). Mutational analysis of the arginine repressor of *Escherichia coli*. *Mol. Microbiol.* **13**, 599–608.
9. Burke, M., Merican, A. & Sherratt, D. (1994). Mutant *Escherichia coli* arginine repressor proteins that fail to bind L-arginine, yet retain the ability to bind their normal DNA-binding sites. *Mol. Microbiol.* **13**, 609–618.
10. Van Duyne, G. D., Ghosh, G., Maas, W. K. & Sigler, P. B. (1996). Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *J. Mol. Biol.* **256**, 377–391.
11. Dion, M., Charlier, D., Wang, H., Gigot, D., Savchenko, A., Hallet, J.-N. *et al.* (1997). The highly thermostable arginine repressor of *Bacillus stearothermophilus*: gene cloning and repressor-operator interactions. *Mol. Microbiol.* **25**, 385–398.
12. Sunnerhagen, M., Nilges, M., Otting, G. & Carey, J. (1997). Solution structure of the DNA binding domain and model for the complex multifunctional hexameric arginine repressor with DNA. *Nature Struct. Biol.* **4**, 819–826.
13. Rodriguez-Garcia, A., Ludovice, M., Martin, J. F. & Liras, P. (1997). Arginine boxes and the *argR* gene of *Streptomyces clavuligerus*: evidence for a clear regulation of the arginine pathway. *Mol. Microbiol.* **25**, 219–228.
14. Wang, H., Glansdorff, N. & Charlier, D. (1998). The arginine repressor from *Escherichia coli* K-12 makes direct contacts to minor and major groove determinants of the operators. *J. Mol. Biol.* **277**, 805–824.
15. Maghnouj, A. (1999). Le catabolisme anaérobie de l'arginine chez *Bacillus licheniformis* et sa régulation par les protéines ArgR et ArcR. PhD thesis, Université Libre de Bruxelles

16. Mironov, A. A., Koonin, E. V., Roytberg, M. A. & Gelfand, M. S. (1999). Computer analysis of transcription regulatory patterns in completely sequenced bacterial genomes. *Nucl. Acids Res.* **27**, 2981–2989.
17. Ni, J., Sakanyan, V., Charlier, D., Glansdorff, N. & Van Duyne, G. (1999). Structure of the arginine repressor from *Bacillus stearothermophilus*. *Nature Struct. Biol.* **6**, 427–432.
18. Dimova, D., Weigel, P., Takahashi, M., Marc, F., Van Duyne, G. D. & Sakanyan, V. (2000). Thermostability, oligomerization and DNA-binding properties of the regulatory protein ArgR from the hyperthermophilic bacterium *Thermotoga neapolitana*. *Mol. Gen. Genet.* **263**, 119–130.
19. Rodriguez-Garcia, A., de la Fuente, A., Pérez-Redondo, R., Matin, J. F. & Liras, P. (2000). Characterization and expression of the arginine biosynthetic gene cluster of *Streptomyces clavuligerus*. *J. Mol. Microbiol. Biotechnol.* **2**, 543–550.
20. Mironov, A. A., Koonin, E. V., Roytberg, M. A. & Gelfand, M. S. (1999). Computer analysis of transcription regulatory patterns in completely sequenced bacterial genomes. *Nucl. Acids Res.* **27**, 2981–2989.
21. Makarova, K. S., Mironov, A. A. & Gelfand, M. S. (2001). Conservation of the arginine repressor DNA-binding signal in all bacterial lineages. *Genome Biol.* **2** pp. research00131–research00138.
22. Dennis, C. A., Glykos, N. M., Rarsons, M. P. & Phillips, S. E. V. (2002). The structure of AhrC, the arginine repressor/activator protein from *Bacillus subtilis*. *Acta Crystallog. sect. D*, **58**, 421–430.
23. Barcelona-Andrés, B., Marina, A. & Rubio, V. (2002). Gene structure, organization, expression, and potential regulatory mechanisms of arginine catabolism in *Enterococcus faecalis*. *J. Bacteriol.* **184**, 6289–6300.
24. Song, H., Wang, H., Gigot, D., Dimova, D., Sakanyan, V., Glansdorff, N. & Charlier, D. (2002). Transcription regulation in thermophilic Bacteria: high resolution contact probing of *Bacillus stearothermophilus* and *Thermotoga neapolitana* arginine repressor–operator interactions. *J. Mol. Biol.* **315**, 255–274. doi: 10.1006/jmbi.2001.5236.
25. Xu, Y., Sun, Y., Huysveld, N., Gigot, D., Glansdorff, N. & Charlier, D. (2003). Regulation of arginine biosynthesis in the psychrophilic bacterium *Moritella profunda*: *in vivo* repressibility and *in vitro* repressor–operator contact probing. *J. Mol. Biol.* **326**, 353–369.
26. Park, S.-M., Lu, C.-D. & Abdelal, A. T. (1997). Cloning and characterization of *argR*, a gene that participates in regulation of arginine biosynthesis and catabolism in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **179**, 5300–5308.
27. Park, S.-M., Lu, C.-D. & Abdelal, A. T. (1997). Purification and characterization of an arginine regulatory protein, *argR*, from *Pseudomonas aeruginosa* and its interactions with the control regions for the *car*, *argF*, and *aru* operons. *J. Bacteriol.* **179**, 5309–5317.
28. Klingel, U., Moller, C. M., North, A. K., Stockley, P. G. & Baumberg, S. (1995). A binding site for activation by the *Bacillus subtilis* AhrC protein, a repressor/activator of arginine metabolism. *Mol. Gen. Genet.* **248**, 329–340.
29. Gardan, R., Rapoport, G. & Débarbouillé, M. (1995). Expression of the *rocDEF* operon involved in arginine catabolism in *Bacillus subtilis*. *J. Mol. Biol.* **249**, 843–856.
30. Miller, C. M., Baumberg, S. & Stockley, P. G. (1997). Operator interactions by the *Bacillus subtilis* arginine repressor/activator, AhrC: novel positioning and DNA-mediated assembly of transcriptional activator at the catabolic sites. *Mol. Microbiol.* **26**, 37–48.
31. Maghnoij, A., De Sousa Cabral, T. F., Stalon, V. & Vander Wauven, C. (1998). The *arcABDC* gene cluster, encoding the arginine deiminase pathway of *Bacillus licheniformis*, and its activation by the arginine repressor ArgR. *J. Bacteriol.* **180**, 6468–6475.
32. Lu, C.-D. & Abdelal, A. T. (1999). Role of ArgR in activation of the *ast* operon, encoding enzymes of the arginine succinyltransferase pathway in *Salmonella typhimurium*. *J. Bacteriol.* **181**, 1934–1938.
33. Kiupakis, A. & Reitzer, L. (2002). ArgR-independent induction and ArgR-dependent superinduction of the *astCADBE* operon in *Escherichia coli*. *J. Bacteriol.* **184**, 2940–2950.
34. Stirling, C. J., Szatmari, G., Stewart, G., Smith, M. C. M. & Sherratt, D. (1988). The arginine repressor is essential for plasmid-stabilizing site-specific recombination at the *ColE1 cer* locus. *EMBO J.* **7**, 4389–4395.
35. Alén, C., Sherratt, D. & Colloms, S. (1997). Direct interaction of aminopeptidase A with recombination site DNA in Xer site specific recombination. *EMBO J.* **16**, 5188–5197.
36. Holtham, C. A. M., Jumel, K., Miller, C. M., Harding, S. E., Baumberg, S. & Stockley, P. G. (1999). Probing activation of the prokaryotic arginine transcriptional regulator using chimeric proteins. *J. Mol. Biol.* **289**, 707–727.
37. Szwajkajzer, D., Dai, L., Fukayama, J. W., Abramczyk, B., Fairman, R. & Carey, J. (2001). Quantitative analysis of DNA binding by the *Escherichia coli* arginine repressor. *J. Mol. Biol.* **312**, 949–962.
38. Grandori, L., Lavoie, T. A., Pflumm, M., Tian, G., Niersbach, H., Maas, W. K. *et al.* (1995). The DNA binding domain of the hexameric arginine repressor. *J. Mol. Biol.* **254**, 150–162.
39. Gochikyan, A., Karaivanova, I., Lecocq, M., Vusio, P., Arnaud, M.-V., Snapyan, M. *et al.* (2002). Arginine operator binding by heterologous and chimeric ArgR repressors from *Escherichia coli* and *Bacillus stearothermophilus*. *J. Bacteriol.* **184**, 6602–6614.
40. Snapyan, M., Lecocq, M., Guevel, L., Arnaud, M.-C., Ghochikyan, A. & Sakanyan, V. (2003). Dissecting DNA–protein and protein–protein interactions involved in bacterial transcriptional regulation by a sensitive protein array method combining near-infrared fluorescence detection. *Proteomics*, **3**, 647–657.
41. Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H. *et al.* (1999). Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. *Nature*, **399**, 323–329.
42. Marc, F., Weigel, P., Legrain, C., Almeras, Y., Santrot, M., Glansdorff, N. & Sakanyan, V. (2000). Characterization and kinetic analysis of mono- and bifunctional ornithine acetyltransferases from thermophilic microorganisms. *Eur. J. Biochem.* **267**, 5217–5226.
43. Chen, S.-H., Merican, A. & Sherratt, D. (1997). DNA binding of *Escherichia coli* arginine repressor mutants altered in oligomeric state. *Mol. Microbiol.* **24**, 1143–1156.
44. Brunelle, A. & Schleif, R. F. (1987). Missing contact probing of DNA–protein interactions. *Proc. Natl Acad. Sci. USA*, **84**, 6673–6676.
45. Smiths, M. C. M., Czaplowsky, L., North, A.,

- Baumberg, S. & Stockley, P. (1989). Sequences required for regulation of arginine biosynthesis promoters are conserved between *Bacillus subtilis* and *Escherichia coli*. *Mol. Microbiol.* **3**, 23–28.
46. Savchenko, A., Charlier, D., Dion, M., Weigel, P., Hallet, J.-N., Holtham, C. *et al.* (1996). The arginine operon of *Bacillus stearothermophilus*: characterization of the control region and its interaction with the heterologous *B. subtilis* arginine repressor. *Mol. Gen. Genet.* **252**, 69–78.
47. Savchenko, A., Weigel, P., Dimova, D., Lecocq, M. & Sakanyan, V. (1998). The *Bacillus stearothermophilus* *argCJBD* operon harbours a strong promoter as evaluated in *Escherichia coli*. *Gene*, **212**, 167–177.
48. Miltcheva Karaivanova, Y., Weigel, P., Takahashi, M., Fort, C., Versavaud, A., Van Duyne, G. *et al.* (1999). Mutational analysis of the thermostable arginine repressor from *Bacillus stearothermophilus*: dissecting residues involved in DNA binding properties. *J. Mol. Biol.* **291**, 843–855.
49. Batisse, N., Weigel, P., Lecocq, M. & Sakanyan, V. (1997). Two amino acid amidohydrolase genes encoding L-stereospecific carbamoylase and amidohydrolase are organized in a common operon in *Bacillus stearothermophilus*. *Appl. Environ. Microbiol.* **63**, 763–766.
50. Van de Castele, M., Demarez, M., Legrain, C., Glansdorff, N. & Piérard, A. (1990). Pathways of arginine biosynthesis in extreme thermophilic archaeo- and eubacteria. *J. Gen. Microbiol.* **136**, 1177–1183.
51. Stockley, P. G., Baron, A. J., Wild, C. M., Parsons, I. D., Miller, C. M., Holtham, C. A. M. & Baumberg, S. (1998). Dissecting the molecular details of prokaryotic transcriptional control: the methionine and arginine repressor proteins. *Biosensors. Bioelect.* **13**, 637–650.
52. Maas, W. K. (1994). The arginine repressor from *Escherichia coli*. *Microbiol. Rev.* **58**, 631–640.
53. Piette, J., Cunin, R., Boyen, A., Charlier, D., Crabeel, M., Van Vliet, F. *et al.* (1982). The regulatory region of the divergent *argECBH* operon in *Escherichia coli* K-12. *Nucl. Acids Res.* **10**, 8031–8048.
54. Piette, J., Nyunoya, H., Lusty, C. J., Cunin, R., Weyens, G., Crabeel, M. *et al.* (1984). DNA sequence of the *carA* gene and the control region of *carAB*: tandem promoters, respectively controlled by arginine and the pyrimidines, regulate the synthesis of carbamoyl-phosphate synthetase in *Escherichia coli* K-12. *Proc. Natl Acad. Sci. USA*, **81**, 4134–4138.
55. Miller, J. H. (1992). *A Short Course in Bacterial Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
56. Sakanyan, V., Kochikyan, A., Mett, I., Legrain, C., Charlier, D., Piérard, A. & Glansdorff, N. (1992). A re-examination of the pathway for ornithine biosynthesis in a thermophilic and two mesophilic *Bacillus* species. *J. Gen. Microbiol.* **138**, 125–130.
57. Ho, N. S., Hunt, D. H., Horton, M. R., Pullen, K. J. & Pease, R. L. (1989). Site directed mutagenesis by overlap extension with the polymerase chain reaction. *Gene*, **77**, 51–59.
58. Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463–5469.
59. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D., Seidman, J. G., Smith, J. A. & Struhl, K. (1993). *Current Protocols in Molecular Biology*, Wiley, New York.
60. Bradford, M. M. (1976). A rapid sensitive method for the quantitation of microgram quantities of protein utilising the principle of protein–dye binding. *Anal. Biochem.* **72**, 248–254.
61. Hölte, H. J., Ankenbauer, W., Muhlegger, K., Rein, R., Gagner, G., Seibl, R. & Walter, T. (1995). The digoxigenin (DIG) system for non-radioactive labeling and detection of nucleic-acids. An overview. *Cell Mol. Biol.* **41**, 883–905.
62. Zubay, G. (1973). *In vitro* synthesis of protein microbial systems. *Annu. Rev. Genet.* **7**, 267–287.

Edited by M. Yaniv

(Received 16 April 2003; received in revised form 18 July 2003; accepted 18 July 2003)