DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
60

STUDIES OF THE TOL PLASMID TRANSCRIPTION FACTOR XylS

NIILO KALDALU

TARTU 2000

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
60

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
60

STUDIES OF THE TOL PLASMID TRANSCRIPTION FACTOR XylS

NIILO KALDALU

TARTU UNIVERSITY
PRESS

Department of Microbiology and Virology, Institute of Molecular and Cell Biology, Tartu University, Estonia

Dissertation is accepted for the commencement of the degree of Doctor of Philosophy (in Molecular Biology) on April 3rd, 2000 by the Council of the Institute of Molecular and Cell Biology, Tartu University

Opponent: Prof. Dr. Paul Broda (University of Manchester, UK)

Commancement: May 15th, 2000

The publication of this dissertation is granted by the University of Tartu

Tartu Ülikooli Kirjastuse trükikoda
Tiigi 78, Tartu 50410
Tellimus nr. 231

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

I. **Kaldalu, N.,** Mandel, T. and Ustav, M. (1996) TOL plasmid transcription factor XylS binds specifically to the Pm operator sequence. Mol Microbiol 20(3), 569–579.

II. **Kaldalu, N.,** Lepik, D., Kristjuhan, A. and Ustav, M. (2000) Monitoring and purification of proteins using bovine papillomavirus E2 epitope tags. Biotechniques 28(3), 456–462.

III. **Kaldalu, N.,** Toots, U., de Lorenzo, V. and Ustav, M. (2000) Functional domains of the TOL plasmid transcription factor XylS. J Bacteriol 182(4), 1118–1126.

ABBREVIATIONS

aa	amino acid(s)
bp	base pair(s)
C terminus	carboxyterminus
CTD	carboxyterminal domain
DNase I	deoxyribonuclease I
EHEC	enterohemorrhagic *Escerichia coli*
EIEC	enteroinvasive *Escerichia coli*
EPEC	enteropathogenic *Escerichia coli*
GFP	green fluorescent protein
HU	bacterial histone-like protein HU
H-NS	bacterial histone-like protein H-NS
HTH	helix-turn-helix
IHF	integration host factor
kb	kilobase pair(s)
kD	kilodalton(s)
MAb	monoclonal antibody
N terminus	aminoterminus
NTR	aminoterminal region
N-XylS	aminoterminally tagged XylS
NC-XylS	amino- and carboxyterminally tagged XylS
1E2-XylS	XylS, tagged with the epitope recognized by 1E2 Mab
3F12-XylS	XylS, tagged with the epitope recognized by 3F12 Mab
Om	*xylXYZLTEGFJQKIH* operator
P_{araBAD}	*araBAD* promoter
P_{araC}	*araC* promoter
P_{araFGH}	*araFGH* promoter
Pm	*xylXYZLTEGFJQKIH* promoter
*Pr*1 and *Pr*2	*xylR* promoters
*Ps*1 and *Ps*2	*xylS* promoters
*Ptet**	*tet* promoter modified by *lac* operator
Pu	*xylUWCMABN*
RNAP	RNA polymerase
UP element	upstream element
wt	wild type

1. INTRODUCTION

The level of expression of every gene in a genome is modulated by physiological conditions. A complex regulatory network mediates signals from inner and outer environment of the organism, thereby switching genes on and off or tuning them up and down. The physiological response to different stimuli is executed through regulatory proteins, either by changing their levels of expression or modifying their activities. Regulators are modified either covalently (e.g. by phosphorylation) or noncovalently, via binding of some other proteins or low molecular-weight allosteric effectors. In their classical works, Francois Jacob and Jacque Monod inferred that regulatory proteins must undergo allosteric conformational changes upon interactions with their ligands or effectors.

In recent history, human activities have challenged adaptiveness of living creatures by polluting the environment with a large variety of xenobiotics, i.e. man-made compounds. Some microorganisms, soil bacteria and fungi, have taken the challenge and have acquired an ability to use these compounds for their own advantage. Thus, diverse xenobiotics do not accumulate in nature but are turned into microbial biomass. Degradation of xenobiotics by soil bacteria requires new genes, enzymes, and pathways which must be properly regulated and integrated into global cell metabolism. These novel traits have been acquired and adapted by microorganisms in a relatively short period of time and are often plasmid-encoded or transposon-encoded. TOL plasmid pWWO of *Pseudomonas putida* mt2 specifies degradation of toluene and xylenes. The genes, which encode enzymes for catabolism of these hydrocarbons, are co-located within two nested transposons, are grouped into two operons and are positively regulated by two regulatory genes, *xyl*R and *xyl*S. One catabolic operon specifies oxidation of aromatic hydrocarbons to benzoates and the second catabolic operon specifies further oxidation of benzoates.

Activation of transcription initiation is the main checkpoint of TOL plasmid gene regulation and both regulators function at this level in substrate-dependent manner. Enzymes required for the catabolism of aromatics are produced only when respective substrates are present in the growth medium. XylR induces transcription of the first operon and *xylS* gene in the presence of toluene. XylS activates transcription of the second catabolic operon in the presence of benzoates and when overproduced. Thus, both regulators are controlled by allosteric effectors. However, XylR and XylS belong to different families of bacterial transcription activators and use different mechanisms for transcriptional activation.

The objective of the present work has been to characterize the DNA binding and modular structure of the TOL plasmid transcription factor XylS.

2. REVIEW OF LITERATURE

2.1. Metabolism of toluene and its regulation

2.1.1. TOL plasmid pWWO: an extrachromosomal element for toluene/xylene degradation

TOL plasmids specify degradation of toluene and xylenes. Such plasmids have been isolated from soil bacteria in different locations all over the world (Assinder and Williams, 1990). The TOL plasmids of *Pseudomonas putida* encode similar catabolic pathways. However, they belong to various incompatibility groups, differ in size and genetic organisation.

The prototype TOL plasmid pWWO was isolated from *P. putida* strain mt2 (Williams and Murray, 1974). pWWO is about 117 kb in size, belongs to the IncP9 incompatibility group and is self-transmissible. Conjugal transfer between pseudomonads occurs at high frequencies (10^{-1} to 1 transconjugants per recipient cell on agar plates) and transfer to *Enterobacteriaceae* has been shown as well (Benson and Shapiro, 1978; Ramos-Gonzalez *et al.*, 1991). *Pseudomonas putida* containing pWWO can grow on toluene, *meta*- and *para*-xylene, and 1,2,4-trimethylbenzene (Worsey and Williams, 1975). The genes, which encode enzymes for catabolism of these hydrocarbons, are grouped into two operons. The upper pathway operon specifies oxidation of toluene to benzoate and xylenes to alkylbenzoates. The *meta*-pathway operon specifies further oxidation of these compounds to Krebs cycle intermediates. That pathway of benzoic acid degradation is called *meta*-pathway because the aromatic ring in catechols, the pathway intermediates, is cleaved in *meta*-fission. Two regulatory proteins, XylR and XylS, positively regulate the catabolic operons (Ramos *et al.*, 1997). In the presence of upper pathway substrates, XylR activates the *Pu* promoter of the upper pathway operon and the *Ps1* promoter of the *xylS* gene. Subsequently, overproduced XylS protein activates the *Pm* promoter of the *meta*-cleavage operon (Inouye *et al.*, 1987a; Ramos *et al.*, 1987). Furthermore, XylS protein is constitutively expressed at a low level from the weak *Ps2* promoter (Gallegos *et al.*, 1996a) and in the presence of benzoates, *i. e.* the degradation products of the upper pathway and substrates for the *meta* pathway, it activates the *Pm* promoter at low protein concentrations. Both TOL catabolic operons are poorly induced in cells growing at the early-exponential-growth phase and strongly induced in cells at late-exponential-growth phase (Cases *et al.*, 1996; Hugouvieux-Cotte-Pattat *et al.*, 1990; Kessler *et al.*, 1992; Marques *et al.*, 1995; Marques *et al.*, 1994; Miura *et al.*, 1998). Several observations have shown that *Pu* and *Ps1* are subjected to catabolite repression (Duetz *et al.*, 1994; Duetz *et al.*, 1996; Holtel *et al.*, 1994; Marques *et al.*, 1994). Therefore, the XylR-dependent expression of *meta*-pathway is also under catabolite repression control. However, XylR expression and the XylR-inde-

pendent activation of *Pm* in the presence of XylS effector molecules are not subjected to catabolite repression. The catabolite repression of the TOL operons is not regulated by levels of cAMP and its regulatory mechanisms are unknown (Holtel *et al.*, 1994).

The catabolic genes of TOL plasmid pWWO are located in nested transposons (Figure 1) and, within these transposons, can be integrated into host chromosome. The toluene transposons, Tn*4651* (56 kb in size) and Tn*4653* (70 kb in size), are class II transposons and transpose by cointegrate formation (Tsuda and Iino, 1987; Tsuda and Iino, 1988; Tsuda *et al.*, 1989). They belong to the Tn3 family and Tn*4653* is related to the Tn*1721* subgroup of the family. Both transposons encode transposases (*tnpA*) of their own but share *res* site and common genes for site-specific resolution (*tnpS* and *tnpT*). The catabolic genes are deleted at high frequencies due to homologous recombination between two identical 1275 bp direct repeats (formerly described as 1.4-kb direct repeats), located at the boundaries of the 39 kb catabolic region (Bayley *et al.*, 1977; Meulien *et al.*, 1981; Reddy *et al.*, 1994). These sequences contain two open reading frames of unknown function, have terminal inverted repeats of 12 bp and were designated IS1246 for significant homology to a number of bacterial insertion sequences (Reddy *et al.*, 1994).

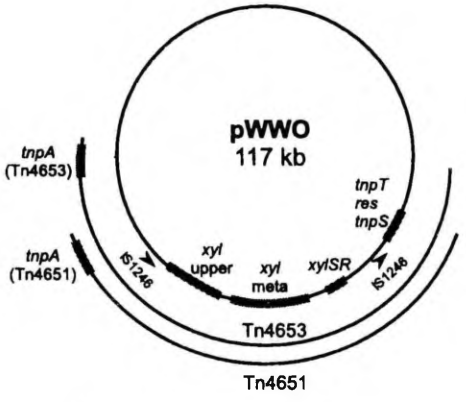

Figure 1. Functional map of the TOL plasmid pWWO (Tsuda and Iino, 1988).

2.1.2. The catabolic pathway for toluene/xylene degradation

Reactions involved in the catabolism of toluene and its methylated derivatives to Krebs cycle intermediates are depicted on Figure 2. First, the methyl group at carbon 1 of the aromatic ring is oxidised to yield benzyl alcohol. Thereafter, alcohol is sequentially oxidised to benzaldehyde and benzoate. These reactions constitute so-called upper pathway. The enzymes that carry out these reactions

are encoded by the genes of the upper operon (Assinder and Williams, 1990; Ramos *et al.*, 1987).

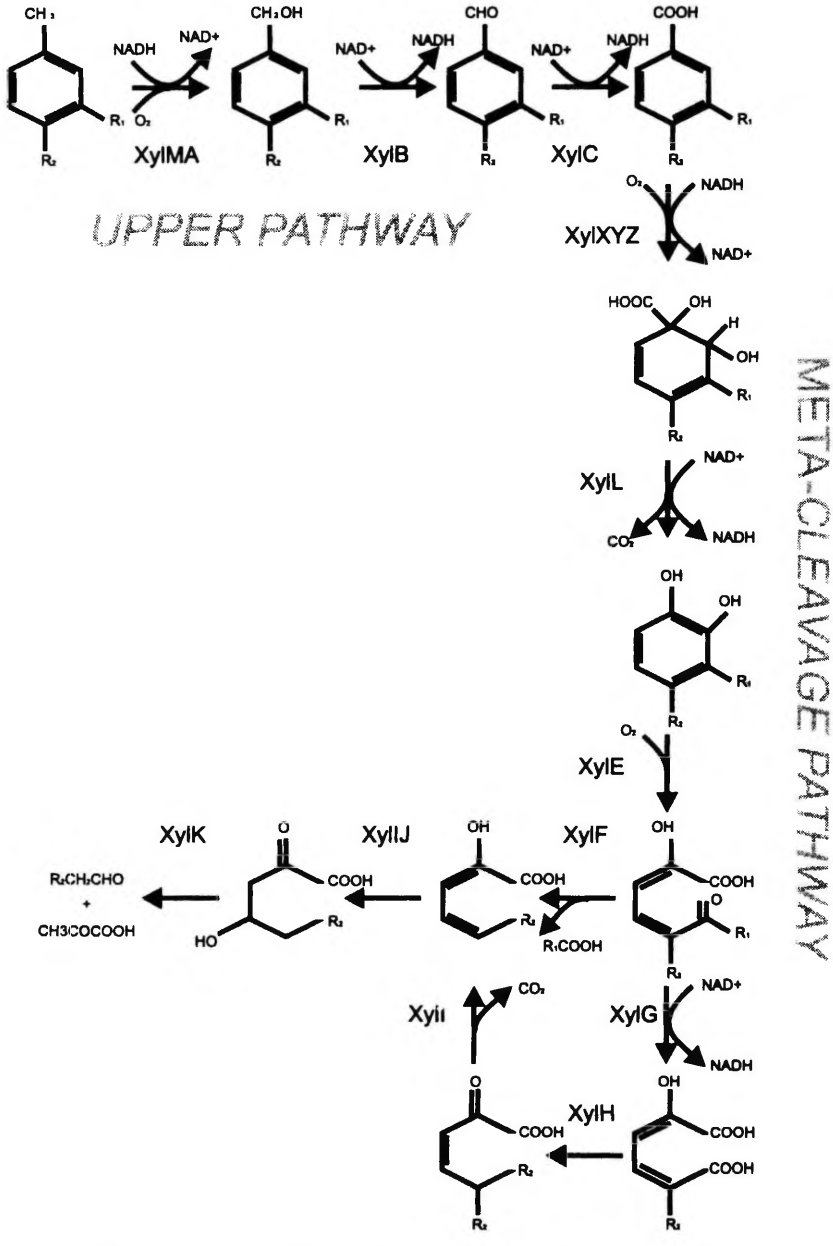

Figure 2. Degradation of toluene (R_1=R_2=H) and xylenes (R_1 and/or R_2=CH_3) as specified by the TOL plasmid pWWO. The Xyl proteins involved in the pathway are indicated next to each step (Ramos *et al.*, 1997).

The aromatic carboxylic acids are further metabolized through the *meta*-cleavage pathway. Benzoate, which may be methylated at positions *meta*, *para*, or both, is first oxidized and then decarboxylated to produce the corresponding catechol. The aromatic ring of catechol is cleaved in *meta*-fission to yield 2-hydroxymuconic acid semialdehyde or its alkyl derivatives (Assinder and Williams, 1990). Metabolism of semialdehydes continues through a branched pathway further to 2-oxopent-4-enoate, a common intermediate. When semialdehyde is produced from *meta*-toluate, it is hydrolized, whereas semialdehyde produced from benzoate or *para*-methylbenzoate is metabolised through the oxalocrotonate branch, which involves several enzymatic steps (Harayama *et al.*, 1987). 2-Oxopent-4-enoate is oxidized to 2-oxo-4-hydroxypentonate which finally is converted to Krebs cycle intermediates (Ramos *et al.*, 1987).

2.1.3 Catabolic operons of the TOL plasmid

The TOL upper operon *xylUWCMABN* encodes three enzymes: toluene oxidase (XylMA), benzyl alcohol dehydrogenase (XylB), and benzaldehyde dehydrogenase (XylC) (Harayama *et al.*, 1989). The function of the *xylU*, *xylW*, and *xylN* gene products is unknown and these genes are not required for growth on toluene/xylenes (Williams *et al.*, 1997). The *Pu* promoter of the upper operon is inducible by XylR, σ^{54}-dependent and fully induced in the late-exponential- and stationary-growth phases (Ramos *et al.*, 1987).

The *meta*-pathway operon, *xylXYZLTEGFJQKIH*, contains 13 genes and extends over 11kb in size (Harayama and Rekik, 1990). The first four genes, *xylXYZL*, are involved in the oxidation of benzoate and its derivatives to catechols. *xylXYZ* encode toluate 1,2 dioxygenase and *xylL* encodes the next enzyme in the pathway that produces catechol. The *xylEGFJQKIH* genes are required for the transformation of catechols into Krebs Cycle derivatives. The *xylT* product is a ferrodoxine that is required for the regeneration of catechol 2,3-dioxygenase XylE when cells grow on *para*-toluate. Isofunctional and homologous genes to *xylTEGFJQKIH* have been found in other catabolic operons that encode enzymes for catechol degradation. The *Pm* promoter of the *meta*-pathway operon does not show similarity to the –10/–35 consensus sequence for binding of σ^{70} RNA polymerase, but matches better with the consensus sequences of σ^{32} and σ^{38} RNA polymerases in the –10 region (Marques *et al.*, 1999). The XylS-dependent transcription from the *Pm* promoter is fully induced in cells at late exponential and stationary-growth phase (Hugouvieux-Cotte-Pattat *et al.*, 1990). In the stationary phase, the *Pm* promoter is dependent on stationary phase-specific sigma factor σ^{38} (σ^{S}) both in *P. putida* and *E. coli* (Marques *et al.*, 1995; Miura *et al.*, 1998). In the exponential growth phase *E. coli*, transcription from *Pm* is dependent on heat-shock sigma factor σ^{32} (Marques *et al.*, 1999). However, transcription from *Pm* by σ^{32} and σ^{38} RNA

polymerases has not been shown *in vitro*. Since dependence on these alternative sigma factors has been demonstrated by *in vivo* experiments in the σ-deficient *E. coli* of strains, these results may also indicate indirect effects.

2.1.4. Regulatory genes and gene regulation in the TOL plasmid

Regulatory cirquits controlling expression of the *xyl* operons are shown on Figure 3. Organization of the *xyl* operon promoters and location of the binding sites of involved regulatory proteins is depicted on Figure 4. TOL plasmid contains two regulatory genes, *xylR* and *xylS*. In the presence of the upper pathway substrates and several other aromatics, XylR activates the *Pu* promoter of the upper pathway operon and the *Ps1* promoter of the *xylS* gene. In addition, XylR controls its own synthesis (Ramos *et al.*, 1987). XylR recognises symmetric binding sites with the consensus sequence 5'-TTGATCAATTGATCAA-3'. It occupies two recognition sites in the upstream region of both promoters, spanning –172 to –157 and –142 to –127 in *Pu*, and located at –184 to –169 and –154 to –139 in *Ps1* (Abril *et al.*, 1991; de Lorenzo *et al.*, 1991; Perez-Martin and de Lorenzo, 1996c). Both binding sites are necessary for the full activation of these promoters and are simultaneously and co-operatively occupied by XylR. The activator is always bound to its recognition sites; the presence of an effector is not required for the DNA binding (Abril *et al.*, 1991). The XylR-binding sites can be moved about 1kb upstream, without loss of their function (Gomada *et al.*, 1992; Inouye *et al.*, 1990).

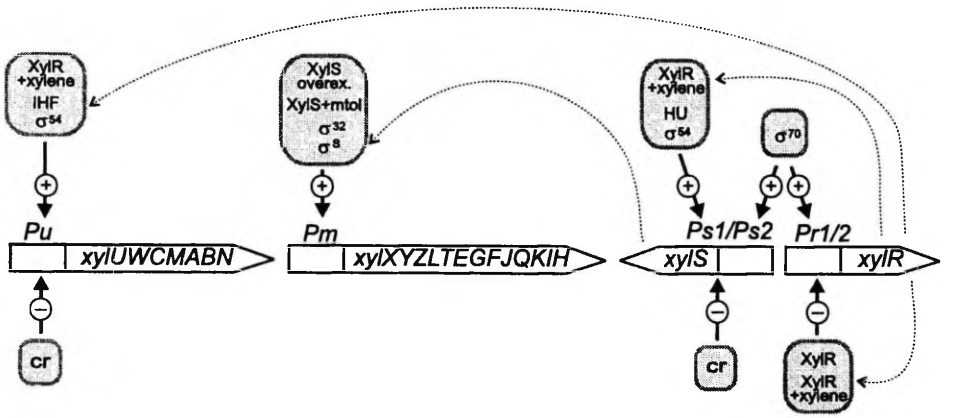

Figure 3. Regulatory cirquits controlling expression of *xyl* operons. +: stimulation of transcription; –: inhibition of transcription (Ramos *et al.* 1997).

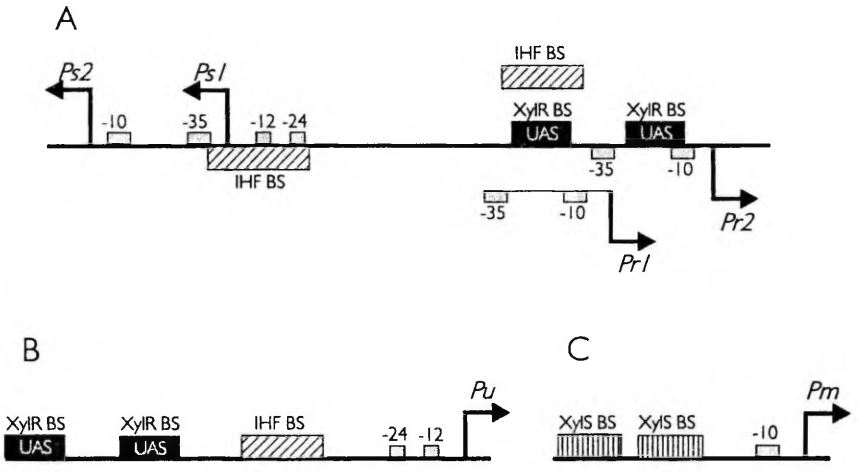

Figure 4. Organization of the *xyl* operon promoters. Ligth-gray boxes indicate conserved sequense-motifs of the RNAP binding sites. XylR BS: XylR binding site; XylS BS: XylS binding site; IHF BS: IHF binding site (Ramos *et al.* 1997)

XylR belongs to the NtrC/NifA family of transcription activators (Inouye *et al.*, 1988). Proteins of this family exhibit four domains, three of which, B, C, and D, are highly conserved among members of the family. The D domain is a DNA-binding domain, the central C domain contacts σ^{54} and has ATPase activity, whereas the B domain serves probably as a linker between C and A domains (Kustu *et al.*, 1989). The nonhomologous, regulatory A domain is required for signal reception (Morett and Segovia, 1993). In XylR, it binds toluene and the other specific effector molecules (Delgado and Ramos, 1994; Delgado *et al.*, 1995). The A domain of XylR interacts with the C domain and inhibits activity of the regulator. Deletion of the inhibitory domain converts XylR into a constitutive activator. Effector binding relieves the intramolecular repression and permits productive interaction with the promoter-bound σ^{54} RNA polymerase (Perez-Martin and de Lorenzo, 1995). Using the N-terminally truncated XylR which lacks the A domain, Pérez-Martín and de Lorenzo have shown that the co-operative binding and oligomerization of XylR is dependent on ATP binding and subsequent conformational change (Perez-Martin and de Lorenzo, 1996a; Perez-Martin and de Lorenzo, 1996b). Therefore, after binding of a ligand, XylR binds ATP, oligomerizes and undergoes several conformational changes that lead to enhanced ATPase activity and transcriptional activation.

Furthermore, nucleoid-associated DNA-bending proteins IHF and HU stimulate transcription from the σ^{54} dependent TOL promoters. The integration host factor (IHF) positively regulates *Pu* (Abril *et al.*, 1991; de Lorenzo *et al.*, 1991) and HU stimulates transcription from *Ps1* (Perez-Martin and de Lorenzo, 1995). At the *Pu* promoter, binding of IHF stimulates the recruitment of the enzyme to

the –12/–24 sequence, i.e. increases efficiency of closed complex formation. For such a recruitment, which is fully independent of XylR, the C-terminal domain of the RNAP α subunit must interact with specific, UP element-like DNA sequences upstream of the IHF binding site (Bertoni *et al.*, 1998a; Carmona *et al.*, 1999).

As mentioned above, in addition to stimulation of *Pu* and *Ps1*, the XylR protein also downregulates its own level by repressing the σ^{70}-dependent tandem promoters *Pr1* and *Pr2* when liganded. The upstream activating sequences of *Ps1* overlap with the –10/–35 hexamers of these divergent *Pr* promoters (Gomada *et al.*, 1992; Inouye *et al.*, 1985). As deduced from experiments with the N-terminally truncated XylR, the downregulation is apparently caused by the ATP-dependent multimerization and strong DNA binding of XylR (Bertoni *et al.*, 1998b).

Briefly, XylR may serve as an example of a transcription activator which contains a distinct autorepressive effector-binding domain. The functional activity of the regulator is controlled through intramolecular interdomain interactions and ligand binding to the regulatory domain causes repression release. However, the effector has only a small effect on DNA-binding by the protein and host factors play an important role in regulation of the XylR-dependent promoters.

XylS, the second activator of the TOL plasmid, is precisely described below, in a special section. The *xylS* gene is transcribed from two independent promoters: from the weak, constitutive, σ^{70}-dependent *Ps2* promoter and, in the presence of the upper pathway substrates, from the XylR- and σ^{54}-dependent *Ps1* promoter (Gallegos *et al.*, 1996a).

2.2. AraC/XylS family of transcription activators

2.2.1 Functions, distribution and evolution

XylS belongs to the AraC/XylS family of bacterial transcription regulators. Proteins of this family are characterized by significant homology over a 100-residue stretch, a region that is proposed to be necessary for DNA binding and stimulation of transcription. The AraC/XylS family includes numerous bacterial transcription activators which regulate metabolism, stress-response, and virulence (Gallegos *et al.*, 1997). Some members of the family may function also as repressors (e.g. AraC at P_{araBAD} and P_{araC}, CelD from *E. coli*) or may have additional activities (Ada). To several proteins of the family no specific regulatory function has yet been assigned and sequencing of bacterial genomes gradually reveals new open reading frames which encode the putative AraC/XylS family regulators of unknown function. Members of the family are widely distributed in diverse eubacterial genera. Until now, most of the genes

encoding for these regulators have been found from the γ subdivision of proteobacteria (purple bacteria) (Gallegos *et al.*, 1997). However, the observed distribution may be nonrepresentative because many prokaryotic genera are genetically and genomically less studied. Proteins of the family vary in size. The most small among them contain about 100 a.a. (AarP, MarA, PqrA, RamA, SoxS), several mammalian virulence regulators are about 200 a.a. in size (e.g. InvF from *Salmonella typhimurium*, MxiE from *Shigella spp.*), the majority of the AraC/XylS family proteins contain about 300 a.a., whereas the regulators of plant pathogen hypersensitive response and Ada from *Mycobacterium tuberculosis* contain about 500 a.a. (Gallegos *et al.*, 1997).

Regulators of carbon metabolism of the family control the degradation of sugars (e.g. arabinose — AraC, melibiose — MelR and MsmR, rhamnose — RhaR and RhaS), amino acids (valine — MmsR, arginine — AdiY and ArgR, ornithine — OruR) and some other compounds, mostly xenobiotics (e.g. methylbenzoate — XylS, herbicide S-dipropylthiocarbamate — TchR). Several members of the family are involved in the response to stressors, e.g. alkylating agents (Ada from diverse bacteria), oxidative stress (SoxS), antibiotics and organic solvents (MarA and Rob from *E. coli*). Virulence regulators of the family control expression of virulence factors in plant pathogens (e.g. HrpB, regulator of the hypersensitive response from *Burkholderia solancearum*) and mammalian pathogens. The latter include regulators of adhesine and fimbria production (e.g. PerA from EHEC and EPEC — intimin production, AggR from enteroaggregative *E. coli*, Rns from enterotoxigenic *E. coli*), invasine production (e.g. InvF from *Salmonella typhimurium*, MxiE from *Shigella spp.*) and other virulence traits (e.g. VirF and LcrF from *Yersinia* spp — activators of *yop* and *ysc* genes; UreR from uropathogenic enterobacteria — urease production; ExsA from *Pseudomonas aeruginosa* — exoenzyme production and secretion) (Gallegos *et al.*, 1997). Virulence regulators can be plasmid-encoded or cromosomally encoded. Often, the genes of the chromosomally encoded virulence regulators are located within pathogenicity islands (e.g. InvF and SprA in SPA-1 of *Salmonella typhimurium*) (Eichelberg *et al.*, 1999). The AraC/XylS family virulence regulators respond to environmental factors such as temperature, osmolarity of the growth medium, and concentration of Ca^{++}. Thermoregulation of bacterial pathogenicity by several AraC/XylS family proteins is physically based on a temperature-dependent structural transition of DNA. Temperature-dependent overproduction of these regulators is modulated by supercoiling of the promoter region or nucleoid-associated host factors (e.g. H-NS dependent overproduction of VirF in *Shigella* and EIEC) (Falconi *et al.*, 1998).

The characterized binding sites of the AraC/XylS family proteins show certain similarities: they are asymmetric, approximately 17 bp in length, include two major-groove regions and an intervening minor-groove region on the protein-contacting surface, and are often arranged as either direct or inverted repeats (Bhende and Egan, 1999; Carra and Schleif, 1993; Egan and Schleif, 1994; Lu *et al.*, 1992; Tobin and Schleif, 1990). The large genetic distances

between bacteria encoding AraC/XylS family regulators and big differences in G+C content of their genes (from 28% to 67%) suggest that a progenitor arose early in eubacterial evolution (Gallegos *et al.*, 1997).

2.2.2 Domain organization

The AraC/XylS family proteins are grouped on the basis of significant homology over a 100-residue stretch (Gallegos *et al.*, 1993; Gallegos *et al.*, 1997; Ramos *et al.*, 1990b) which has been shown to enable DNA binding and stimulation of transcription for several family members (Bustos and Schleif, 1993; Lauble *et al.*, 1989; Michan *et al.*, 1995). This DNA-binding region has been proposed to contain two helix-turn-helix (HTH) motifs (Brunelle and Schleif, 1989; Francklyn and Lee, 1988; Mermod *et al.*, 1987; Pabo and Sauer, 1984) which base-specifically interact with two adjacent major groove regions of an asymmetric target-site, as indicated by footprinting and missing-contact probing (Brunelle and Schleif, 1989; Hendrickson and Schleif, 1985; Lu *et al.*, 1992; Park *et al.*, 1997). A crystal structure of a family member MarA, in complex with its binding site, confirmed this unique fold of the DNA-binding region (Rhee *et al.*, 1998). Several small monomeric activators in the family (e.g. MarA and SoxS from *E. coli*) match to the conserved sequence and do not contain additional domains (Amabile-Cuevas and Demple, 1991; Cohen *et al.*, 1993). To exert their regulatory roles, these AraC/XylS family proteins are overproduced in response to environmental stimuli. However, majority of the AraC/XylS family proteins contain also a nonconserved region. In these proteins, the conserved part is usually located C-terminally, but it may reside also in the N terminus (e.g. CafR and Rob from *E. coli*) or constitute the central part of the protein (Ybbb from *B. subtilis*, Ada) (Gallegos *et al.*, 1997). In diverse AraC/XylS family proteins, the nonconserved region has been shown or proposed to be necessary for dimerization, ligand binding and ligand-regulated control of transcription activation through interaction with the DNA-binding domain (Bustos and Schleif, 1993; Lauble *et al.*, 1989; Michan *et al.*, 1992a; Michan *et al.*, 1992b; Saviola *et al.*, 1998a). The regulators of metabolism in the family are signal receptors and stimulate transcription from their cognate promoters in the presence of specific low molecular weight effectors. In addition, some other regulators often transcriptionally control levels of these proteins. However, most of the virulence regulators also contain the nonconserved region but, except for UreR, have not been shown to respond to any allosteric effector. Since the proteins which contain the nonconserved part bind to recognition sites which are organized as direct or inverted repeats, dimerization could be a general function of the N-terminal region (Gallegos *et al.*, 1997).

2.2.3. Characteristics of individual family members

2.2.3.1. XylS

The role of XylS in TOL plasmid gene regulation has been exhaustively described above. The protein contains 321 a.a. Activation of the *Pm* promoter by the XylS protein requires the operator sequence *Om*. Two direct repeats 5'-TGCAAPuAAPuPyGGnTA-3', are essential for activation of transcription by XylS (Kessler *et al.*, 1993). Thus, the operator of *Pm* partially overlaps with the –35 region of promoter, similarly with the operator sequences recognized by VirF of *Shigella flexneri* (Tobe *et al.*, 1993) and *Yersinia enterocolica* (Wattiau and Cornelis, 1994), Rns from *E. coli* ETEC (Munson and Scott, 1999), RhaS (Egan and Schleif, 1994), Mar at several marboxes, and AraC of *E. coli* at P_{araBAD} (Lee *et al.*, 1981) and P_{araFGH} (Hendrickson *et al.*, 1992). Proper phasing and integrity of the repeats have proved to be necessary for the functioning of XylS. Point mutations of nucleotides which are identical between the the half-sites resulted in a dramatic decrease in *Pm* induction, whereas substitutions in the variable positions or inter-repeat region had minor effect. In addition, the exact 6 bp spacing between the repeats, which enables the spacing of 21 bp, i.e. two helical turns, between the equivalent bases of the half-sites, appeared to be necessary for promoter activity (Kessler *et al.*, 1993). By using the tagged and matrix-bound XylS protein we have shown that the protein really binds to *Om* sequence (reference I in this thesis). An alternative operator has been suggested in a work from J. L. Ramos lab on the basis of mutagenesis of the promoter region (Gallegos *et al.*, 1996b). However, further genetic analysis carried out at that lab has shown that the former definition of the XylS binding site was correct (Gonzalez-Perez *et al.*, 1999). Results of this study completely coincide with those obtained at our lab.

Transcriptional activation by XylS is stimulated by alkylbenzoates and modulated by the intracellular level of the protein (Inouye *et al.*, 1987a). Diverse benzoate derivatives have different ability to induce XylS-dependent transcription. Benzoates with different substitutions (-CH_3, -C_2H_5, -OCH_3, -F, -Cl, -Br, and -I) can be XylS effectors, although not all positions in the aromatic ring are equivalent. Position 3 is highly permissible, whereas positions 2 and 4 pose more restrictions to substituents. Some disubstitutions involving positions 2 and 3 or positions 3 and 4 are also permissible (Ramos *et al.*, 1986). Expression of XylS from strong promoters has shown that overproduction of the protein, naturally mediated by XylR, is sufficient for activation of *Pm* in the absence of benzoate effectors (Inouye *et al.*, 1987b; Mermod *et al.*, 1987; Spooner *et al.*, 1987). On that basis, Mermod and co-workers have suggested a hypothesis about a dynamic equilibrium between an inactive and an active, DNA-binding, conformations of the protein in the cell. The putative role of effector would be to shift the equilibrium towards the active conformation (Mermod *et al.*, 1987).

To identify which parts of the protein mediate the effect of benzoates, extensive mutagenesis of the *xylS* gene has been carried out. Several amino-acid substitutions have been found that altered effector specificity of XylS (Michan *et al.*, 1992b; Ramos *et al.*, 1990a; Ramos *et al.*, 1986), or produced semiconstitutive phenotype. Mutations of both types were scattered all over the *xylS* gene, however, many of them were clustered in a small glycine-rich N-terminal region P37-R45. The mutants XylSR41C, XylSP37G, XylSS229I, XylSD274V, XylSD274E had increased uninduced, basal level of transcriptional activation which was still inducible by benzoate (Michan *et al.*, 1992b; Zhou *et al.*, 1990), whereas semiconstitutive mutants XylSG44S, XylSP37L, XylSG44S exhibited increased stability (Michan *et al.*, 1992b). Arg-41 seems to be a critical residue, since changes at this position resulted in multiple different phenotypes. It has been shown that some mutations in the C terminus are intraallelically dominant over substitutions in the N terminus and, in contrary, a mutation in the N terminus (R45T) can restore the effector control that has been lost due to these C-terminal mutations. All these data suggest that N and C termini of XylS may interact and benzoate effectors regulate the activity of the protein by modulation of that interaction (Michan *et al.*, 1992a).

2.2.3.2. AraC

The AraC protein of *E. coli*, the most well characterised member of the protein family, regulates metabolism of L-arabinose. In the absence of arabinose, AraC represses transcription of the *araBAD* and *araC* genes from respective promoters (P_{araBAD} and P_{araC}) (Huo *et al.*, 1988; Lee *et al.*, 1992; Lee *et al.*, 1981). In the presence of arabinose, AraC activates transcription of the catabolic operon *araBAD* and the arabinose transport genes *araE* and *araFGH* (Hendrickson *et al.*, 1992; Hendrickson *et al.*, 1990; Johnson and Schleif, 1995). Expression of all these *ara* operons is also regulated by cAMP and the CAP protein (Hahn *et al.*, 1984). AraC is dimeric at cellular concentrations (Wilcox and Meuris, 1976) and binds to a sequence $AGCN_7TCCATA$, which appears as a tandem repeat in the *ara* operator regions (Brunelle and Schleif, 1989; Hendrickson *et al.*, 1992; Hendrickson and Schleif, 1985; Hendrickson *et al.*, 1990; Lu *et al.*, 1992). Binding of AraC bends DNA about 90° (Saviola *et al.*, 1998b).

Regulation of the P_{araBAD} and P_{araC} promoters has been extensively studied. In the absence of arabinose, one monomer of the AraC dimer occupies the I_1 binding site while the other occupies the O_2 half-site approximately 200 bp away. This way, the dimer bound to target sites generates a DNA loop and prevents transcription from P_{araBAD} and P_{araC} (Carra and Schleif, 1993; Lobell and Schleif, 1990; Lobell and Schleif, 1991). When arabinose is added, the regulator undergoes a conformational change and shifts to occupy the adjacent half-sites I_1 and I_2 (Lobell and Schleif, 1990; Lobell and Schleif, 1991). As a result, P_{araBAD} is induced. Therefore, the main consequence of arabinose binding is to

change the affinity of the protein for different spatial arrangements of half-sites. A work from the lab of B. Müller-Hill has shown that arabinose differently modulates the affinity of AraC towards the two ends of a half-site, designated A and B boxes, i.e. interaction of the two HTH motifs with respective DNA major-groove surfaces (Niland *et al.*, 1996).

Transcription activation parameters at P_{araBAD} have been measured (Zhang *et al.*, 1996) and it has been demonstrated that formation of the transcriptionally competent AraC-RNAP-promoter open complex depends of the assembly order of the three components (Zhang and Schleif, 1996). When AraC was added first, open complexes formed rapidly, but when RNA polymerase was added first the kinetics of open complex formation was much slower. Since cells do not appear to possess a mechanism for controlling assembly order, interpretation of this result remains unclear.

The I_2 half-site of P_{araBAD} overlaps with the –35 region of the promoter. As mentioned above, such an arrangement of binding sites is characteristic to several proteins in the AraC/XylS family. Using a series of alanine substitutions, it has been shown that the residues 590 to 603 in region 4 of the *E. coli* RNAP σ^{70}subunit are involved in transcription activation by AraC. These residues of *E. coli* σ^{70} are probably located in a long recognition helix that interacts with the –35 hexamer of promoters and are ideally positioned to permit interaction with activators. Interaction with the same region of the σ^{70} subunit is involved in transcription activation by a mutationally altered CAP derivative at type II promoters (Lonetto *et al.*, 1998). It has been shown, that the methylated Ada protein, another member of the family, also activates transcription by contacting σ^{70} region 4 (Landini *et al.*, 1998; Landini and Busby, 1999). Transcription activation by AraC depends also on the C-terminal domain RNAP α subunit, since *E. coli* rpoA341 mutation impairs AraC-dependent, as wall as MelR-dependent but not XylS-regulated transcription (Kessler *et al.*, 1994a; Thomas and Glass, 1991). In addition, DNA bending by AraC has apparently some effect on transcriptional activation, because an AraC mutant that could normally bind DNA but caused less bending was defective in stimulation of transcription (Saviola *et al.*, 1998b).

For the AraC proteins of *E. coli* and *Salmonella typhimurium*, the nonconserved N terminus and the conserved C terminus have been shown to form separable functional domains (Bustos and Schleif, 1993; Lauble *et al.*, 1989). The modular structure of the *Salmonella typhimurium* regulator has been characterized by limited proteolysis and cross-linking experiments, while the domains of the AraC protein from *E. coli* were mapped using protein chimeras. AraC consists of two functional domains: the conserved C-terminal domain carries sequence-specific DNA-binding capability while the nonconserved N-terminal domain mediates effector responsiveness and carries dimerization capability (Bustos and Schleif, 1993; Lauble *et al.*, 1989). These domains are connected with a flexible linker region (residues 167 to 177) (Eustance *et al.*,

1994). It has been shown that individual amino acids of this region may be altered without substantial effect on the ability of transcription activation and effector response of the protein (Eustance and Schleif, 1996). The DNA-binding domain most probably contains all the determinants which are necessary for transcriptional activation, since the separately expressed C-terminal domain of AraC has residual ability to activate transcription without the arabinose effector (Menon and Lee, 1990). The crystal structure of the N-terminal domain of AraC from *E. coli* revealed the parts of the molecule which are necessary for dimerization and sugar binding. The christallographic data show that two monomers of AraC associate by an antiparallel coiled-coil formed between the terminal α helices of the regulatory domain. Arabinose binds into a β-barrel, formed by the N-terminus of the protein, and the short NH_2-terminal arm covers the sugar-binding pocket as a lid. In the absence of arabinose, this N-terminal arm is disordered (Soisson *et al.*, 1997). Both deletion of the N-terminal arm and certain substitutions in it converted AraC into a strong, constitutive activator and could suppress the loss of ability of transcription activation caused by mutations in the DNA-binding domain. Likely, the extreme N-terminal arm of AraC binds to the C-terminal domain as an intramolecular repressor of transcription activation, whereas binding of arabinose causes rearrangement of the arm and repression release (Saviola *et al.*, 1998a). Hemiplegic mutations in AraC protein that specifically block either induction or repression at the P_{araBAD} promoter were isolated (Reed and Schleif, 1999). The induction-deficient mutations located in the N-terminal arm or arabinose-binding pocket, whereas the repression-deficient mutations resided in the DNA-binding domain, in the HTH region that contacts the B-box of the O_2 half-site (Niland *et al.*, 1996).

3. AIMS OF THE STUDY

1. To develop a suitable tagging system for detection and affinity purification of XylS and other proteins which are intrinsically prone to aggregation.
2. To purify physiologically active XylS protein.
3. To map the binding site of XylS.
4. To characterize the mode of DNA binding of the XylS protein.
5. To identify which regions of XylS are required for the DNA binding, transcriptional activation and effector responsiveness.

4. RESULTS AND DISCUSSION

4.1. Expression and purification of XylS (Reference I)

The DNA sequences, which confer XylS responsiveness, had been mapped within the *Pm* promoter/operator region. Although these two direct repeats could probably constitute the binding site of XylS, DNA-binding properties of XylS have not been studied *in vitro*, because it has been difficult to purify the soluble protein (de Lorenzo *et al.*, 1993). We fused the influenza virus haemagglutinin (HA) epitope (Field *et al.*, 1988) to the N terminus of XylS for monitoring and purification of the protein. The resulting tagged N-XylS protein was identified and immunopurified using the monoclonal antibody 12CA5 (Field *et al.*, 1988; Zhou *et al.*, 1992). The *xylS* sequence, amplified by the polymerase chain reaction (PCR) and verified by sequencing, was inserted into several expression vectors. Both wt XylS and epitope-tagged N-XylS were expressed from the phage T7 promoter of the pET11 vector in *E. coli* strain BL21(DE3) under variable conditions of cultivation and induction. The overexpressed protein was completely insoluble (Figure1 a and b, in reference I). We tried several methods of solubilisation and renaturation of the aggregated protein but all of the procedures were unsuccessful.

It has been reported that the expression of XylS from the *tet*-promoter of pBR322 facilitated activation of the responsive promoter *in vivo* (Mermod *et al.*, 1987), indicating that a substantial amount of correctly folded and functionally active XylS is produced. Therefore, N-*xylS* was cloned into pBR322 under the *tet* promoter and transformed into *E. coli* strain DH5α. We saw that a considerable fraction of the expressed N-XylS protein remained soluble and was readily detectable on immunoblots (Figure 1 c and d, in reference I).

The epitope-tagged N-XylS was purified by single-step, bachwise immunoaffinity binding to the 12CA5 monoclonal antibody, coupled to Protein A-Sepharose or Protein A-Fractogel (TSK) beads. Purity and intactness of the N-XylS preparations were estimated by silver staining of SDS-PAGE and immunoblotting (Figure 1 c and d, in reference I). Two contaminating bands could be seen on SDS-PAGE: one corresponding to the light chain of the antibody and another to a copurified protein whose molecular mass was approximately 65 kD.

Like XylS, many proteins aggregate during overexpression and produce insoluble inclusion bodies. That has been considered to be the result of unfolding and increased solvent exposure of hydrophobic amino acids. Several proteins of the AraC/XylS family turned out to be completely insoluble when overexpressed at high levels (de Lorenzo *et al.*, 1993; Egan and Schleif, 1994), while some pose solubility problems but retain a soluble fraction sufficient for biochemical analysis (Caswell *et al.*, 1992; Schleif and Favreau, 1982; Tobin and Schleif, 1990). We show that reduction of the level of expression along

with the use of a tag enables the affinity purification of such a protein. More, we found that batch-wise affinity binding with rotation was necessary to avoid aggregation of N-XylS, whereas column was clogged with the protein.

4.2. N-XylS specifically binds to *Om* (Reference I)

The ability of N-XylS to bind specifically to the putative recognition site within the *Pm* region was tested using a DNA immunoprecipitation assay, both in the presence and absence of *meta*-toluate. The plasmid pUPM190, containing the cloned *Pm* region, was digested with *Hpa*II and *Hin*fI restriction enzymes and generated fragments were end-labelled using the Klenow fragment of *E. coli* DNA polymerase I. Plasmid pUC19 was treated in the same way and served as a control for the specificity of DNA immunoprecipitation. The protein-loaded beads were incubated with a mixture of end-labeled restriction fragments. The experiment was carried out at oversaturating levels of the DNA probe. After several washes, the bound DNA was released and identified by gel electrophoresis, using the input fragments as a marker. We demonstrated that N-XylS-loaded beads bind a single 115 bp *Hpa*II (–97) *Hin*fI (+19) *Om*-containing fragment of the pUPM190 digest, whereas binding of any other fragment of this plasmid or any fragment of the pUC19 digest could not be observed (Figure 2, lanes 1 and 2, in reference I). Beads, which were incubated with the lysate of bacteria expressing no N-XylS, were unable to bind DNA (data not shown). These data indicate that N-XylS binds to a specific site within the *Pm* region and completely conform with the results of transcriptional activation *in vivo* (Kessler *et al.*, 1993; Kessler *et al.*, 1994b)

Furthermore, we studied the influence of *meta*-toluate on the specificity and affinity of the interaction. We saw that *meta*-toluate facilitated the specific DNA binding by N-XylS (Figure 2, lanes 1 and 3, in reference I), but had no effect on the non-specific DNA binding. Further, when we worked with the tagged XylS, which was purified from high-salt extracts, we saw that *meta*-toluate strongly enhanced specific DNA binding by XylS (Figure 5A, in reference III). Therefore, stimulation of DNA binding is a major, if not the single mechanism by which ligand regulates the XylS activity.

4.3. DNase I footprinting of the N-XylS-*Om* complexes (Reference I)

In order to map the binding site of XylS in more detail, we analyzed the interaction of N-XylS with the *Om* region by modified procedure of DNase I footprinting (Ustav *et al.*, 1991). The *Om* containing DNA fragment was end-

labelled with the Klenow fragment of *E. coli* DNA polymerase I and the N-XylS-*Om* complex was formed on the beads at oversaturating levels of the DNA probe. Non-bound DNA was removed by washing of the beads and retained fragment was treated with DNase I. After cleavage, the bound DNA was extracted from the beads, purified, and analyzed on sequencing gels. The DNase I footprinting data are presented in Figure 3, reference I, and summarized in Figure 5, reference I. Both strands of *Om* show a 44 bp area of protection, centered at the two direct repeats. The footprint extends from position –72 to –28 on the upper strand (Figure 3a, in reference I) and from –74 to –30 on the lower strand (Figure 3b, in reference I). Two DNase hypersensitive sites occur within the footprint at nucleotides –58/–56 and –37 on the upper strand while four hypersensitive sites, separated by one helical turn, appear at positions –69, –59/–58, –48, and –38 on the lower strand. The DNase I hypersensitive sites occur at equivalent positions within both direct repeats, but the cleavage is far more effective in the left repeat than the promoter-proximal repeat. The presence of *meta*-toluate in the binding buffer enhanced cleavage at position –75 on the lower strand (Figure 3b, in reference I), whereas other differences were not found. Consequently, ligand does not have a major effect on the mode of interaction between N-XylS and its target DNA.

4.4. Hydroxyl radical footprinting of the N-XylS-*Om* complexes (Reference I)

To study phosphoribose-backbone contacts of *Om* with N-XylS, we performed hydroxyl radical footprinting. In order to prevent hydroxyl radical scavenging by the polysaccharide matrix of Sepharose these experiments were done with N-XylS-Fractogel (TSK) beads. Results of the hydroxyl radical footprinting of the non-coding strand are presented in Fig. 3b, reference I, and summarized in Figure 5, reference I. Four protected 5 bp regions with equal 5 bp spacing, centered at positions –64, –53, –43 and –33occur. Less clear protection emerges around position –72. Both half-sites are protected equally , but protections in the promoter-distal ends of the half-sites, near –64 and –43, are less obvious. A three-dimensional representation of the hydroxyl radical footprinting data is presented in Fig. 6, reference I. Four protected regions of a half-helical turn extent are separated by a half turn and, consequently, are aligned on the same face of the DNA helix. Therefore, N-XylS contacts the sugar-phosphate back-bone on one face of the helix in four adjacent helical turns. We did not see any influence of *meta*-toluate on the hydroxyl radical footprints.

4.5. Base-specific interactions in the N-XylS-Om complexes (Reference I)

We used methylation interference and methylation protection experiments to find out which nucleotides in *Om* are critical for binding of XylS. Dimethyl sulfate (DMS) methylates guanines at the N^7 position in the major groove and helps to identify sequence-specific interactions or close contacts of major groove-contacting proteins. For methylation interference, the DNA probe was labeled as for DNase I footprinting, methylated with DMS and incubated with the N-XylS-loaded beads. The use of the matrix-bound N-XylS allowed us to omit the gel-retardation step for separation of the protein-DNA complex from the unbound DNA. After washing, the retained DNA was extracted from the beads, cleaved with piperidine and analyzed on sequencing gels (Figure 4, a and b in reference I). The relative intensity of the bands was measured on Phosphoimager (Figure 4, c and d in reference I). Methylation of guanines at positions –68, –59, –58, –49, –47, –38 and –37 on the top strand and at –67 and –46 on the bottom strand interfered with the binding of N-XylS. All these guanines except G-49 are identical nucleotides between the *Om* half-sites. At the same time, methylation of the three guanines in variable positions of the left half-site (Om_L) and those in the inter-repeat region does not influence DNA binding by N-XylS. The spacing between the interfering guanines is approximately 10 bp on the upper strand and 20 bp on the lower strand, i.e., one or two helical turns respectively (Figures 5 and 6 in reference I). From these results we conclude that N-XylS binds to one face of DNA, covering four helical turns, and makes base-specific contacts in four adjacent major-groove regions. The pattern of methylation interference is repeated over one major-groove region, reflecting the direct repeat structure of *Om*.

To confirm that indicated bases are engaged in the interaction with XylS, we performed methylation protection experiments of the N-XylS-*Om* complexes. All guanines in the recognition sequence, methylation of which lead to reduction of the specific affinity to N-XylS, were readily protected by the bound N-XylS (data not shown).

The results of methylation interference are in accordance with the prediction that XylS binds DNA by two putative helix-turn-helix (HTH) (Mermod *et al.*, 1987; Ramos *et al.*, 1990b). If XylS binds to its target site as a dimer, a XylS monomer should contain two separate DNA-binding units for the base-specific interactions in two adjacent major-groove regions in a half-site. Further, resolved structure of an AraC/XylS family member MarA in complex with its DNA binding site confirmed the presence of two HTH motifs (Rhee *et al.*, 1998) and, most probably, that is true for all other proteins in the family as well.

4.6. Construction of vectors for the expression of proteins tagged with Bovine Papillomavirus E2 epitopes (Reference II)

Like some other proteins of the AraC/XylS family, XylS is prone to aggregation both inside the cell as well as in the course of purification (de Lorenzo *et al.*, 1993; Egan and Schleif, 1994). Therefore, we expressed the protein at a low level and purified through a tag, primarily the 12CA5 specific HA epitope. To use an antibody available at the lab for that purpose and develop a new tagging system we chose two newly mapped Bovine Papillomavirus type 1 (BPV-1) E2 protein epitopes. Interaction of the monoclonal antibodies with these epitopes was specific and had high affinity in a variety of conditions. The 3F12 antibody — epitope interaction tolerates high salt concentrations (up to 2M), which allows to immunoprecipitate and immunopurify the tagged proteins in the presence of high salt concentrations to avoid the coimmunoprecipitation of contaminating proteins.

We constructed vector plasmids for expression of single or double-tagged proteins. The vectors for the moderate level expression of the tagged proteins were based on pBR322 and proteins were expressed from the *tet* promoter of that plasmid. The vectors pBR-3F12 and pBR-1E2 were constructed for the expression of proteins with N-terminally fused peptides GVSSTSSDFRDR and TTGHYSVRD, recognized by anti-BPV E2 monoclonal antibodies 3F12 and 1E2, respectively (Kurg *et al.*, 1999). Both vectors contain *Xba*I and *Bam*HI sites for cloning of the recombinant sequence. We also constructed the vector pBR-NC for expression of proteins with different immunotags in both N- and C-termini. The N-terminal tag was TTGHYSVRD as in pBR-1E2 and the C-terminal tag was TSSDFRDR, a shorter version of the epitope recognized by 3F12 MAb. The cloning sites for a coding sequence in pBR-NC are *Xba*I and *Kpn*I (Table 1 in reference II).

We transformed *E. coli* DH5α with plasmids expressing the tagged versions of XylS and analyzed expression of the tagged XylS proteins by Western blotting. 3F12 antibody recognized both 3F12-XylS and NC-XylS proteins, whereas 1E2 antibody recognized 1E2-XylS and NC-XylS proteins as single bands on the Western blot (Figure 1A in reference II, lanes 1 to 6). No cross-reaction with cellular proteins was observed. However, when detected with 3F12 MAb, NC-XylS gave much lighter band than with 1E2 MAb or the 3F12-XylS protein with 3F12 MAb (Figure 1A in reference II, lanes 6, 4, and 3, respectively). That could be explained by the use of the shorter version of the 3F12-specific epitope in the double-tagged protein. For the assay of XylS activity, expression plasmids containing tagged *xylS* variants were transformed also into *E. coli* strain CC118*Pm-lacZ* with a chromosomal copy of the XylS responsive *Pm* promoter fused to the *lacZ* gene (Kessler *et al.*, 1994b). The tags had no effect on transcriptional activation by XylS (data not shown).

To test whether the BPV E2-derived epitope does not interfere with the site-specific DNA binding of XylS *in vitro*, we used immunobound 3F12-XylS protein that was attached to the TSK beads through the N-terminal tag and 3F12 antibody. Western blot analysis showed that high salt concentrations, up to 2 M NaCl, do not hinder interaction of 3F12 MAb with the specific epitope (data not shown). Therefore, to avoid aggregation of XylS and coimmunoprecipitation of contaminating proteins, the crude lysate was prepared in a high-salt lysis buffer, containing 1.5 M NaCl, and 3F12-XylS was bound to the TSK-coupled 3F12 MAb by a single-step, batchwise procedure. The matrix-bound protein was used for DNA precipitation and DNase I footprinting (Figure 2 in reference II). Since we obtained the same results using the HA-epitope tagged N-XylS and 3F12-XylS, the BPV E2-derived epitope has no effect on the site-specific DNA binding.

4.7. Stimulation of Pm by truncated XylS proteins (Reference III)

To examine whether the N- and C-terminus of XylS constitute separable functional domains, we constructed two sets of progressing terminal deletions, from both ends of the coding sequence (Fig. 11 in reference III). We truncated the protein in putative loop regions, indicated by a prediction of the secondary structure (not shown). The deletion mutants of *xylS* were generated by PCR and verified by sequencing. The BPV E2-derived tag, recognised by the 3F12 antibody, was fused to the N-terminus of the proteins for monitoring and purification. The tagged full-size XylS was applied as a wild-type control (wt N-XylS). Since expression of several truncated XylS proteins (ΔN209 and ΔC310 in particular) from the *tet* promoter, which we had used previously for production of the tagged XylS protein, was apparently toxic to *E. coli* and caused plasmid instability, we inserted the *lac* operator sequence downstream of the promoter to reduce expression.

For assay of *Pm* activation, plasmids containing these *xylS* variants were transformed into *E. coli* strain CC118*Pm-lacZ* containing a chromosomal copy of the *Pm* promoter fused to the *lacZ* gene (Kessler *et al.*, 1994b). Figure 3 in reference III shows β-galactosidase (β-Gal) levels, mediated by wt N-XylS and various deletion mutants, in the presence and absence of *meta*-toluate. Note, that the level of expression of wt N-XylS from the modified *tet* promoter (labelled as *Ptet** below) mimics the XylR mediated overexpression of XylS in *P. putida* — it produces full activation of *Pm* without effector and addition of the ligand has no further effect on the promoter activity. We found that deletion mutant ΔN209, which corresponds to the putative DNA-binding domain, stimulates transcription from *Pm* as efficiently as wt N-XylS. When expressed

from *Ptet**, both mediated β-Gal levels close to 10^4 Miller units (Fig. 3 in reference III, lines 2 and 15).

To probe whether ΔN209 is inducible by benzoates, ΔN209 and wt N-XylS were expressed from the weak *Ps2* promoter (Fig. 4 in reference III). The level of expression from *Ps2* was so low that both proteins remained undetectable at the Western blot of crude lysates even by enhanced chemoluminescence (ECL) detection. However, in the presence of *meta*-toluate, such a small amount of wt N-XylS was enough to produce the same β-galactosidase level as wt N-XylS overexpressed from *Ptet**. We found that ΔN209 is not inducible by effector and provides the phenotype of constitutive activator. When activator proteins were expressed from *Ps2*, the stimulation of *Pm* caused by ΔN209 was almost five-fold higher than that produced by wt N-XylS without effector. However, it reached only 4% of that produced by wild-type activator in the presence of *meta*-toluate (Fig. 4 in reference III).

More, we found that the other deletion mutants which retained the putative DNA-binding domain: ΔN8, ΔN30, ΔN39, ΔN105, and ΔN134, were able to stimulate transcription from *Pm* to some extent, when overproduced from *Ptet** (Fig. 3, lines 10–14 in reference III). These N-terminally truncated proteins mediated 4–16 fold higher β-galactosidase levels than the uninduced basal level of the strain. However, it makes only 0,6–2,6 % of the β-Gal level produced by wt N-XylS. The XylS variants with longer N-terminal deletions than ΔN209 were unable to stimulate *Pm* and could not be detected by immunoblotting, obviously due to instability (data not shown). All deletions from the C-terminus produced proteins which could not activate the *Pm* promoter (Fig 3, lines 3–7 in reference III).

In conclusion, these results suggest that C-terminus of XylS indeed forms a DNA-binding domain and contains all the elements necessary for activation of transcription. The fact that presumably monomeric ΔN209 mediated almost five-fold higher *Pm* activation than wt N-XylS without effector, when these proteins were expressed from *Ps2*, suggests that the N terminus of XylS works as an intramolecular repressor.

Previously, Kessler *et al.* (Kessler *et al.*, 1994b) have characterized several deletion mutants of XylS *in vivo* and found that all of them were unable to activate *Pm* or modulate the activity of wild-type XylS. These mutants were expressed as a result of the readthrough transcription at a very low level, much lower than those produced from the *Ptet** or *Ps2* promoter as we can deduce from the β-Gal activities mediated by wild type XylS. Therefore, stimulation of *Pm* by the deletion mutants presumably remained undetectable due to the low levels of expression.

4.8. DNA binding by XylS CTD (Reference III)

All deletion mutants which activate transcription from *Pm in vivo*, should specifically bind to *Om*. In order to confirm the site-specific DNA binding of the C-terminal domain (CTD) of XylS *in vitro*, we purified both epitope-tagged ΔN209 and wt N-XylS proteins by single-step, batchwise immunoaffinity binding to the 3F12 monoclonal antibody, coupled to TSK beads. The DNA binding activity was studied with the immobilized protein, because purified XylS tends to aggregate in solution making it impossible to use the regular gel-shift assays for this purpose. The cell-free extracts were prepared in a high-salt lysis buffer, containing 1.5 M NaCl. The protein preparations were examined by silver staining of SDS/PAGE and immunoblotting (Fig. 2 B and C in reference III). We succeeded in isolating both wt N-XylS and ΔN209 that were functionally active in DNA-binding, however, the yield of wt N-XylS was lower and preparations contained some degradation products or contaminating proteins.

We used the matrix-bound N-XylS and ΔN209 in the specific DNA binding and DNase I footprinting assays. The mixture of end-labeled restriction fragments of the *Om*-containing plasmid pUPM190 (reference I) was incubated with the protein-loaded TSK beads in the presence and absence of *meta*-toluate. After the removal of free probe, only DNA that was bound to the immunopurified protein was retained on the beads. Since estimation of the amounts of proteins used in the assay was complicated, the experiment was carried out at oversaturating levels of the DNA probe, so that less than 2% of the input labelled probe retained to the beads. Figure 5A in reference III shows that both ΔN209 and wt N-XylS bind specifically the *Om*-containing 115 bp fragment of the pUPM190 *Hpa*II/*Hin*fI digest, whereas binding of any other fragment could not be observed. DNA binding by wt N-XylS was strongly (up to 100 times) induced by *meta*-toluate. Such a strong effect of *meta*-toluate could be observed only when wt N-XylS was purified from the high-salt lysate (1.5 M NaCl) and was not observed earlier when we used low-salt buffer conditions (220 mM KCl). We did not detect any effect of *meta*-toluate on DNA binding by ΔN209. The 3F12 beads lacking XylS did not bind any DNA (data not shown).

Further, we analyzed the interaction of both protein variants with *Om* by DNase I footprinting. The *Om*-containing DNA fragment was end-labelled and the DNA-protein complex was formed on the beads. With wt N-XylS the binding was done in the presence of *meta*-toluate. Again, the experiments were carried out at oversaturating levels of the DNA probe. After removal of unbound DNA, the complex was subjected to DNase I cleavage. The treated DNA was extracted from the beads and analyzed on the sequencing gels. We found that both ΔN209 and wt N-XylS protect a 44bp area, extending from position –74 to –30 on the lower strand (Fig. 5B). The DNase I footprints of these two protein variants were almost identical, indicating that both have a

similar mode of interaction with the binding site. The only difference is hypersensitivity to DNase I at nucleotide –75, observed in the complex with wt N-XylS. Enhanced cleavage at that position was detected, when the N-XylS-*Om* complex was formed in the presence of *meta*-toluate. That slight effect has been previously overlooked by us (in reference I).

The results of DNA-immunoprecipitation and DNase I footprinting confirm that, firstly, the DNA binding domain of XylS is located within the C-terminal 112 residues of the protein and, secondly, *meta*-toluate strongly facilitates DNA binding by matrix-attached N-XylS but does not affect DNA binding by XylS CTD *in vitro* (Fig. 4A).

Therefore, stimulation of DNA binding must be at least one, if not the single, major effect of ligand. However, additional, co-operative effects should not be excluded. Stimulation of dimer formation by effector is possible, but this should not enhance DNA binding in our assay since we used immobilized wt N-XylS which was attached to the beads through the N terminus and presumably had not enough freedom to change its multimeric state.

4.9. The complete N-terminal domain is required for effector-responsiveness of XylS (Reference III)

The N-terminal portion of XylS is believed to be necessary for effector binding and ligand-dependent regulation of activity of the protein. We demonstrated that the truncated proteins ΔN8, ΔN30, ΔN39, ΔN105, and ΔN134, which contain a part of the putative regulatory domain in addition to the complete DNA-binding domain, are non-inducible by effector, and when expressed from *Ptet**, produce much lower β-Gal activity than wt N-XylS or ΔN209 (Fig. 3 in reference III, lanes 2 and 10–15). When placed under the control of *Ps2* promoter, these deletion mutants were unable to stimulate the *Pm* promoter (data not shown). The deletion mutants ΔN8, ΔN30, ΔN39, ΔN105 and ΔN134 are expressed at the different intracellular concentrations, probably due to different stability of the truncated proteins (Fig. 2A in reference III). However, the activity of the N-terminal deletion mutants does not merely correlate with the protein levels seen in Figure 2A. As the levels of these proteins are readily detectable, they must be far more abundant than wt N-XylS and ΔN209, expressed from the *Ps2* promoter. As we have mentioned above, wt N-XylS and ΔN209 produce substantial *Pm* activation even at the levels of expression, which are undetectable by Western blot analysis (Fig. 4 in reference III). Thus, *in vivo* transcriptional activation data reflect, at least partially, the intrinsic properties of the truncated proteins, and not only different level of expression. Consequently, N-terminal

deletions in the putative regulatory domain cause the loss of the regulatory function, and reduce the activity of the C-terminal domain of XylS.

For further characterization of the regulatory portion of XylS and the requirements for its proper functioning, we constructed additional deletion and insertion mutants and tested their ability of Pm stimulation (Fig. 3 in reference III, lines 16–23). Most of the below-described mutants were expressed from *Ptet** at the level, comparable with that of wt N-XylS (Fig. 2A in reference III). Only GA/XylS and E2h/Δ140–209 were apparently very unstable. The latter could be detected at the Western blot only by ECL and is not shown on the figure.

We deleted two different portions of the central part of XylS, from the other end of the putative regulatory domain. The deletion mutants Δ140–209 and Δ174–209 were more active than N-terminal truncations but also had lost their effector responsiveness and showed reduced ability of *Pm* stimulation (Fig. 3 in reference III, lines 16 and 17; Fig. 4 in reference III). To ascertain whether that effect could be caused by the loss of hinge flexibility which does not allow the domains to interact properly, we replaced the deleted residues 140–209 with long unstructured regions: the 80-residue hinge region of BPV-1 E2 protein (Giri and Yaniv, 1988) and a 69-residues region, consisting mainly of glycine/alanine, which has been shown to substitute effectively the BPV1 E2 hinge (D. Örd and M. Ustav, unpublished). As we found that the region containing residues 140–209 is responsible for dimerization of XylS (N. Kaldalu and M. Ustav, unpublished), it was also substituted with the dimerization domain of λ CI repressor. Resultant proteins E2h/Δ140–209, GA/Δ140–209 and λCI/Δ140–209 remained unresponsive to *meta*-toluate. E2h/Δ140–209 and GA/Δ140–209 were much less active than wt N-XylS, whereas λCI/Δ140–209, when overexpressed, mediated about one half of the β-Gal level produced by wt N-XylS (Fig. 3 in reference III, lines 18, 20 and 22). Expression from the *Ps2* promoter shows, however, that λCI/Δ140–209 is a much weaker transcription activator than wt N-XylS and ΔN209 (Fig 4 in reference III). When the same heterologous protein portions were inserted into the putative hinge region of XylS, retaining the N-terminal region intact, we found that these hybrid XylS activators were responsive to the effector. λCI/XylS, E2h/XylS, and GA/XylS, which carry the interdomain insertions, were weaker transcription activators than wt N-XylS but were clearly inducible by *meta*-toluate (Fig3, lines 19, 21 and 23; Fig. 4). Since XylS variants with the inserted λ CI dimerization domain were relatively more active than the others, we substituted the entire N-terminal region with the λ CI domain and expressed the resultant fusion protein λCI/ΔN209 from both *Ptet** and*Ps2* promoters. β-galactosidase levels mediated by λCI/ΔN209 remained several times lower, than those produced by the C-terminal domain ΔN209 itself (Fig3, line 24; Fig. 4 in reference III). These data suggest that heterologous dimerization domain as well as the incomplete N-terminal regions of XylS interfered with the tran-

scriptional activation by XylS CTD. Therefore, we do not have evidence whether the λCI domain-containing proteins were more active due to their dimeric state or due to the lower level of interdomain interference.

In AraC protein, the N-terminal arm of the regulatory domain has been shown to interact with the C-terminal domain as an intramolecular repressor of binding to the adjacent binding-sites, which is necessary for transcription activation. Leucines in the N-terminal arm were crucial for that activity (Saviola *et al.*, 1998a). Using mutational analysis, we tried to specify a subdomain of a similar function within the N-terminal domain of XylS. Several mutations that caused the semiconstitutive behaviour of XylS are clustered in a small glycine-rich N-terminal region P37-R45 (Michan *et al.*, 1992b; Zhou *et al.*, 1990)(Fig1 in reference III), which was expected to form a loop by a prediction of the secondary structure. To remove the putative loop, we produced a deletion mutant Δ39–47. We also constructed a double mutant XylS(L5R,L6K), to test whether the leucines in the extreme N-terminus have a role in regulation of the XylS activity. Both Δ39–47 and XylS(L5R,L6K), as well as the N-terminal deletion mutant ΔN8, were unresponsive to effector. However, these mutants were very weak and not constitutive activators (Fig 3 in reference III, lines 8–10). Therefore, we did not find any region in the XylS NTR, which could be deleted to produce constitutive phenotype, like it has been shown in the case of AraC.

Thus, the data of the mutational analysis indicate that 210 N-terminal residues of XylS provide the ligand responsiveness to the protein, and the entire region is necessary for that activity. Since it has not been demonstrated that this part of XylS folds independently or functions (binds benzoate) independently of CTD, it would be correct to use the term the N-terminal region (NTR) of XylS. All the examined deletions in the XylS NTR reduced the activity of the XylS CTD. Since the level of the Pm activation by the mutant proteins positively correlates with the strength of the promoter used for the expression of these proteins, the reduced activation must be an intrinsic characteristic of the mutants and not caused by aggregation due to overexpression. It is possible that mutations in the N terminus, which convert XylS non-inducible by effector and strongly reduce its ability to activate *Pm*, disrupt the native structure of the N-terminus of the protein. Fusion of several aggregation-prone proteins to green fluorescent protein (GFP) has shown, that folding of the N-terminal portion of a protein affects the activity and, consequently, folding of the downstream portion (Waldo *et al.*, 1999). However, the mechanism of that effect has not been discussed. We suggest that the DNA-binding domain may contain a surface or surfaces that interact with the regulatory part of XylS. That is in agreement with the knowledge that mutations in both, N and C termini of XylS can yield the semiconstitutive phenotype (Ramos *et al.*, 1990a) and mutations in one domain can be suppressed by mutations in the other (Michan *et al.*, 1992a). It is possible that the misfolded N-terminus provides a new site or sites for irreversible

interaction with the CTD and, that way, inhibits DNA binding and/or transcriptional activation.

Binding of the effector to the intact NTR may cause a conformational change in it and release the intramolecular repression by similar mechanism which has been validated for AraC (Saviola *et al.*, 1998a). In the presence of effector, XylS NTR may either passively release the inhibition or actively facilitate the function of CTD, e.g. by assistance in local folding of the DNA-binding regions.

5. CONCLUSIONS

1. The BPV E2 protein-derived peptides GVSSTSSDFRDR and TTGHYSVRD, recognized by anti-monoclonal antibodies 3F12 and 1E2, respectively, are suitable for tagging and immunoaffinity purification of physiologically active XylS.
2. XylS binds specifically to the operator sequence *Om* of the *Pm* promoter consisting of two 5'-TGCAAPuAAPuPyGGnTA-3'direct repeats.
3. XylS binds along one side of the DNA helix covering four helical turns and has base-specific contacts in four adjacent major groove regions on the same helical face. The pattern of contacts made by XylS with *Om* is in accordance with the presence of two helix-turn-helix motifs within a single XylS molecule.
4. Methylbenzoate strongly facilitates specific DNA binding by XylS. Therefore, stimulation of occupation of the target site is a major effect of ligand.
5. The DNA binding and transcription activation domain of XylS is located within the C-terminal 112 residues of the protein. The separately expressed C-terminal domain has an ability to activate transcription and bind DNA independently of the benzoate effector.
6. The entire region of 210 N-terminal residues is required for ligand responsiveness of XylS.

REFERENCES

Abril, M. A., Buck, M. and Ramos, J. L. (1991) Activation of the *Pseudomonas* TOL plasmid upper pathway operon. Identification of binding sites for the positive regulator XylR and for integration host factor protein. *J Biol Chem*, **266**, 15832–8.

Amabile-Cuevas, C. F. and Demple, B. (1991) Molecular characterization of the *soxRS* genes of *Escherichia coli*: two genes control a superoxide stress regulon. *Nucleic Acids Res*, **19**, 4479–84.

Assinder, S. J. and Williams, P. A. (1990) The TOL plasmids: determinants of the catabolism of toluene and the xylenes. *Adv Microb Physiol*, **31**, 1–69.

Bayley, S. A., Duggleby, C. J., Worsey, M. J., Williams, P. A., Hardy, K. G. and Broda, P. (1977) Two modes of loss of the Tol function from *Pseudomonas putida* mt-2. *Mol Gen Genet*, **154**, 203–4.

Benson, S. and Shapiro, J. (1978) TOL is a broad-host-range plasmid. *J Bacteriol*, **135**, 278–80.

Bertoni, G., Fujita, N., Ishihama, A. and de Lorenzo, V. (1998a) Active recruitment of σ^{54}-RNA polymerase to the Pu promoter of *Pseudomonas putida*: role of IHF and αCTD. *EMBO J*, **17**, 5120–8.

Bertoni, G., Marques, S. and de Lorenzo, V. (1998b) Activation of the toluene-responsive regulator XylR causes a transcriptional switch between σ^{54} and σ^{70} promoters at the divergent Pr/Ps region of the TOL plasmid. *Mol Microbiol*, **27**, 651–9.

Bhende, P. M. and Egan, S. M. (1999) Amino acid-DNA contacts by RhaS: an AraC family transcription activator. *J Bacteriol*, **181**, 5185–92.

Brunelle, A. and Schleif, R. (1989) Determining residue-base interactions between AraC protein and *araI* DNA. *J Mol Biol*, **209**, 607–22.

Bustos, S. A. and Schleif, R. F. (1993) Functional domains of the AraC protein. *Proc Natl Acad Sci U S A*, **90**, 5638–42.

Carmona, M., de Lorenzo, V. and Bertoni, G. (1999) Recruitment of RNA polymerase is a rate-limiting step for the activation of the σ^{54} promoter Pu of *Pseudomonas putida*. *J Biol Chem*, **274**, 33790–4.

Carra, J. H. and Schleif, R. F. (1993) Variation of half-site organization and DNA looping by AraC protein. *EMBO J*, **12**, 35–44.

Cases, I., de Lorenzo, V. and Perez-Martin, J. (1996) Involvement of σ^{54} in exponential silencing of the *Pseudomonas putida* TOL plasmid Pu promoter. *Mol Microbiol*, **19**, 7–17.

Caswell, R., Williams, J., Lyddiatt, A. and Busby, S. (1992) Overexpression, purification and characterization of the *Escherichia coli* MelR transcription activator protein. *Biochem J*, **287**, 493–9.

Cohen, S. P., Hachler, H. and Levy, S. B. (1993) Genetic and functional analysis of the multiple antibiotic resistance (mar) locus in *Escherichia coli*. *J Bacteriol*, **175**, 1484–92.

de Lorenzo, V., Eltis, L., Kessler, B. and Timmis, K.N. (1993) Analysis of *Pseudomonas* gene products using lacIq/Ptrp-lac plasmids and transposons that confer conditional phenotypes. *Gene*, **123**, 17–24.

de Lorenzo, V., Herrero, M., Metzke, M. and Timmis, K. N. (1991) An upstream XylR- and IHF-induced nucleoprotein complex regulates the σ^{54}-dependent Pu promoter of TOL plasmid. *EMBO J*, **10**, 1159–67.

Delgado, A. and Ramos, J. L. (1994) Genetic evidence for activation of the positive transcriptional regulator Xy1R, a member of the NtrC family of regulators, by effector binding. *J Biol Chem*, **269**, 8059–62.

Delgado, A., Salto, R., Marques, S. and Ramos, J. L. (1995) Single amino acids changes in the signal receptor domain of XylR resulted in mutants that stimulate transcription in the absence of effectors. *J Biol Chem*, **270**, 5144–50.

Duetz, W. A., Marques, S., de Jong, C., Ramos, J. L. and van Andel, J. G. (1994) Inducibility of the TOL catabolic pathway in *Pseudomonas putida* (pWW0) growing on succinate in continuous culture: evidence of carbon catabolite repression control. *J Bacteriol*, **176**, 2354–61.

Duetz, W. A., Marques, S., Wind, B., Ramos, J. L. and van Andel, J. G. (1996) Catabolite repression of the toluene degradation pathway in *Pseudomonas putida* harboring pWW0 under various conditions of nutrient limitation in chemostat culture. *Appl Environ Microbiol*, **62**, 601–6.

Egan, S. M. and Schleif, R.F. (1994) DNA-dependent renaturation of an insoluble DNA binding protein. Identification of the RhaS binding site at *rhaBAD*. *J Mol Biol*, **243**, 821–9.

Eichelberg, K., Hardt, W. D. and Galan, J. E. (1999) Characterization of SprA, an AraC-like transcriptional regulator encoded within the *Salmonella typhimurium* pathogenicity island 1. *Mol Microbiol*, **33**, 139–52.

Eustance, R. J., Bustos, S. A. and Schleif, R. F. (1994) Reaching out. Locating and lengthening the interdomain linker in AraC protein. *J Mol Biol*, **242**, 330–8.

Eustance, R. J. and Schleif, R. F. (1996) The linker region of AraC protein. *J Bacteriol*, **178**, 7025–30.

Falconi, M., Colonna, B., Prosseda, G., Micheli, G. and Gualerzi, C. O. (1998) Thermoregulation of Shigella and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of *virF* promoter to transcriptional repressor H-NS. *Embo J*, **17**, 7033–43.

Field, J., Nikawa, J., Broek, D., MacDonald, B., Rodgers, L., Wilson, I. A., Lerner, R. A. and Wigler, M. (1988) Purification of a RAS-responsive adenylyl cyclase complex from Saccharomyces cerevisiae by use of an epitope addition method. *Mol Cell Biol*, **8**, 2159–65.

Francklyn, C. S. and Lee, N. (1988) AraC proteins with altered DNA sequence specificity which activate a mutant promoter in *Escherichia coli*. *J Biol Chem*, **263**, 4400–7.

Gallegos, M. T., Marques, S. and Ramos, J. L. (1996a) Expression of the TOL plasmid *xylS* gene in *Pseudomonas putida* occurs from a σ^{70}-dependent promoter or from σ^{70}- and σ^{54}-dependent tandem promoters according to the compound used for growth. *J Bacteriol*, **178**, 2356–61.

Gallegos, M. T., Marques, S. and Ramos, J. L. (1996b) The $TACAN_4TGCA$ motif upstream from the –35 region in the σ^{70}-σ^{S}-dependent Pm promoter of the TOL plasmid is the minimum DNA segment required for transcription stimulation by XylS regulators. *J Bacteriol*, **178**, 6427–34.

Gallegos, M. T., Michan, C. and Ramos, J. L. (1993) The XylS/AraC family of regulators. *Nucleic Acids Res*, **21**, 807–10.

Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K. and Ramos, J. L. (1997) Arac/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev*, **61**, 393–410.

Giri, I. and Yaniv, M. (1988) Structural and mutational analysis of E2 trans-activating proteins of papillomaviruses reveals three distinct functional domains. *EMBO J*, **7**, 2823–9.

Gomada, M., Inouye, S., Imaishi, H., Nakazawa, A. and Nakazawa, T. (1992) Analysis of an upstream regulatory sequence required for activation of the regulatory gene *xylS* in xylene metabolism directed by the TOL plasmid of *Pseudomonas putida*. *Mol Gen Genet*, **233**, 419–26.

Gonzalez-Perez, M. M., Ramos, J. L., Gallegos, M. T. and Marques, S. (1999) Critical nucleotides in the upstream region of the XylS-dependent TOL meta-cleavage pathway operon promoter as deduced from analysis of mutants. *J Biol Chem*, **274**, 2286–90.

Hahn, S., Dunn, T. and Schleif, R. (1984) Upstream repression and CRP stimulation of the *Escherichia coli* L- arabinose operon. *J Mol Biol*, **180**, 61–72.

Harayama, S., Mermod, N., Rekik, M., Lehrbach, P. R. and Timmis, K. N. (1987) Roles of the divergent branches of the *meta*-cleavage pathway in the degradation of benzoate and substituted benzoates. *J Bacteriol*, **169**, 558–64.

Harayama, S. and Rekik, M. (1990) The *meta* cleavage operon of TOL degradative plasmid pWW0 comprises 13 genes. *Mol Gen Genet*, **221**, 113–20.

Harayama, S., Rekik, M., Wubbolts, M., Rose, K., Leppik, R. A. and Timmis, K.N. (1989) Characterization of five genes in the upper-pathway operon of TOL plasmid pWW0 from *Pseudomonas putida* and identification of the gene products. *J Bacteriol*, **171**, 5048–55.

Hendrickson, W., Flaherty, C. and Molz, L. (1992) Sequence elements in the *Escherichia coli araFGH* promoter. *J Bacteriol*, **174**, 6862–71.

Hendrickson, W. and Schleif, R. (1985) A dimer of AraC protein contacts three adjacent major groove regions of the araI DNA site. *Proc Natl Acad Sci U S A*, **82**, 3129–33.

Hendrickson, W., Stoner, C. and Schleif, R. (1990) Characterization of the *Escherichia coli araFGH* and *araJ* promoters. *J Mol Biol*, **215**, 497–510.

Holtel, A., Marques, S., Mohler, I., Jakubzik, U. and Timmis, K.N. (1994) Carbon source-dependent inhibition of *xyl* operon expression of the *Pseudomonas putida* TOL plasmid. *J Bacteriol*, **176**, 1773–6.

Hugouvieux-Cotte-Pattat, N., Kohler, T., Rekik, M. and Harayama, S. (1990) Growth-phase-dependent expression of the *Pseudomonas putida* TOL plasmid pWW0 catabolic genes. *J Bacteriol*, **172**, 6651–60.

Huo, L., Martin, K. J. and Schleif, R. (1988) Alternative DNA loops regulate the arabinose operon in *Escherichia coli*. *Proc Natl Acad Sci U S A*, **85**, 5444–8.

Inouye, S., Gomada, M., Sangodkar, U. M., Nakazawa, A. and Nakazawa, T. (1990) Upstream regulatory sequence for transcriptional activator XylR in the first operon of xylene metabolism on the TOL plasmid. *J Mol Biol*, **216**, 251–60.

Inouye, S., Nakazawa, A. and Nakazawa, T. (1985) Determination of the transcription initiation site and identification of the protein product of the regulatory gene *xylR* for *xyl* operons on the TOL plasmid. *J Bacteriol*, **163**, 863–9.

Inouye, S., Nakazawa, A. and Nakazawa, T. (1987a) Expression of the regulatory gene xylS on the TOL plasmid is positively controlled by the *xylR* gene product. *Proc Natl Acad Sci U S A*, **84**, 5182–6.

Inouye, S., Nakazawa, A. and Nakazawa, T. (1987b) Overproduction of the *xylS* gene product and activation of the *xylDLEGF* operon on the TOL plasmid. *J Bacteriol*, **169**, 3587–92.

Inouye, S., Nakazawa, A. and Nakazawa, T. (1988) Nucleotide sequence of the regulatory gene *xylR* of the TOL plasmid from *Pseudomonas putida*. *Gene*, **66**, 301–6.

Johnson, C. M. and Schleif, R. F. (1995) *In vivo* induction kinetics of the arabinose promoters in *Escherichia coli*. *J Bacteriol*, **177**, 3438–42.

Kessler, B., de Lorenzo, V. and Timmis, K. N. (1992) A general system to integrate lacZ fusions into the chromosomes of gram-negative eubacteria: regulation of the Pm promoter of the TOL plasmid studied with all controlling elements in monocopy. *Mol Gen Genet*, **233**, 293–301.

Kessler, B., de Lorenzo, V. and Timmis, K. N. (1993) Identification of a cis-acting sequence within the Pm promoter of the TOL plasmid which confers XylS-mediated responsiveness to substituted benzoates. *J Mol Biol*, **230**, 699–703.

Kessler, B., Herrero, M., Timmis, K. N. and de Lorenzo, V. (1994a) Genetic evidence that the XylS regulator of the *Pseudomonas* TOL meta operon controls the Pm promoter through weak DNA-protein interactions. *J Bacteriol*, **176**, 3171–6.

Kessler, B., Timmis, K. N. and de Lorenzo, V. (1994b) The organization of the Pm promoter of the TOL plasmid reflects the structure of its cognate activator protein XylS. *Mol Gen Genet*, **244**, 596–605.

Kurg, R., Parik, J., Juronen, E., Sedman, T., Abroi, A., Liiv, I., Langel, U. and Ustav, M. (1999) Effect of bovine papillomavirus E2 protein-specific monoclonal antibodies on papillomavirus DNA replication. *J Virol*, **73**, 4670–7.

Kustu, S., Santero, E., Keener, J., Popham, D. and Weiss, D. (1989) Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism. *Microbiol Rev*, **53**, 367–76.

Landini, P., Bown, J. A., Volkert, M. R. and Busby, S. J. (1998) Ada protein-RNA polymerase sigma subunit interaction and alpha subunit- promoter DNA interaction are necessary at different steps in transcription initiation at the *Escherichia coli* Ada and *aidB* promoters. *J Biol Chem*, **273**, 13307–12.

Landini, P. and Busby, S. J. (1999) The *Escherichia coli* Ada protein can interact with two distinct determinants in the σ^{70} subunit of RNA polymerase according to promoter architecture: identification of the target of Ada activation at the *alkA* promoter. *J Bacteriol*, **181**, 1524–9.

Lauble, H., Georgalis, Y. and Heinemann, U. (1989) Studies on the domain structure of the Salmonella typhimurium AraC protein. *Eur J Biochem*, **185**, 319–25.

Lee, D. H., Huo, L. and Schleif, R. (1992) Repression of the araBAD promoter from araO1. *J Mol Biol*, **224**, 335–41.

Lee, N. L., Gielow, W. O. and Wallace, R. G. (1981) Mechanism of araC autoregulation and the domains of two overlapping promoters, Pc and PBAD, in the L-arabinose regulatory region of *Escherichia coli*. *Proc Natl Acad Sci U S A*, **78**, 752–6.

Lobell, R. B. and Schleif, R. F. (1990) DNA looping and unlooping by AraC protein. *Science*, **250**, 528–32.

Lobell, R. B. and Schleif, R. F. (1991) AraC-DNA looping: orientation and distance-dependent loop breaking by the cyclic AMP receptor protein. *J Mol Biol*, **218**, 45–54.

Lonetto, M. A., Rhodius, V., Lamberg, K., Kiley, P., Busby, S. and Gross, C. (1998) Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase σ^{70} subunit. *J Mol Biol*, **284**, 1353–65.

Lu, Y., Flaherty, C. and Hendrickson, W. (1992) AraC protein contacts asymmetric sites in the *Escherichia coli araFGH* promoter. *J Biol Chem*, **267**, 24848–57.

Marques, S., Gallegos, M. T. and Ramos, J. L. (1995) Role of σ^S in transcription from the positively controlled Pm promoter of the TOL plasmid of *Pseudomonas putida*. *Mol Microbiol*, **18**, 851–7.

Marques, S., Holtel, A., Timmis, K. N. and Ramos, J. L. (1994) Transcriptional induction kinetics from the promoters of the catabolic pathways of TOL plasmid pWW0 of *Pseudomonas putida* for metabolism of aromatics. *J Bacteriol*, **176**, 2517–24.

Marques, S., Manzanera, M., Gonzalez-Perez, M. M., Gallegos, M. T. and Ramos, J. L. (1999) The XylS-dependent Pm promoter is transcribed in vivo by RNA polymerase with σ^{32} or σ^{38} depending on the growth phase. *Mol Microbiol*, **31**, 1105–13.

Menon, K. P. and Lee, N. L. (1990) Activation of ara operons by a truncated AraC protein does not require inducer. *Proc Natl Acad Sci U S A*, **87**, 3708–12.

Mermod, N., Ramos, J. L., Bairoch, A. and Timmis, K. N. (1987) The *xylS* gene positive regulator of TOL plasmid pWWO: identification, sequence analysis and overproduction leading to constitutive expression of meta cleavage operon. *Mol Gen Genet*, **207**, 349–54.

Meulien, P., Downing, R. G. and Broda, P. (1981) Excision of the 40kb segment of the TOL plasmid from *Pseudomonas putida* mt-2 involves direct repeats. *Mol Gen Genet*, **184**, 97–101.

Michan, C., Kessler, B., de Lorenzo, V., Timmis, K. N. and Ramos, J. L. (1992a) XylS domain interactions can be deduced from intraallelic dominance in double mutants of *Pseudomonas putida*. *Mol Gen Genet*, **235**, 406–12.

Michan, C., Zhou, L., Gallegos, M. T., Timmis, K. N. and Ramos, J. L. (1992b) Identification of critical amino-terminal regions of XylS. The positive regulator encoded by the TOL plasmid. *J Biol Chem*, **267**, 22897–901.

Michan, C. M., Busby, S. J. and Hyde, E. I. (1995) The *Escherichia coli* MelR transcription activator: production of a stable fragment containing the DNA-binding domain. *Nucleic Acids Res*, **23**, 1518–23.

Miura, K., Inouye, S. and Nakazawa, A. (1998) The *rpoS* gene regulates OP2, an operon for the lower pathway of xylene catabolism on the TOL plasmid, and the stress response in *Pseudomonas putida* mt-2. *Mol Gen Genet*, **259**, 72–8.

Morett, E. and Segovia, L. (1993) The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J Bacteriol*, **175**, 6067–74.

Munson, G. P. and Scott, J. R. (1999) Binding site recognition by Rns, a virulence regulator in the AraC family. *J Bacteriol*, **181**, 2110–7.

Niland, P., Huhne, R. and Muller-Hill, B. (1996) How AraC interacts specifically with its target DNAs. *J Mol Biol*, **264**, 667–74.

Pabo, C. O. and Sauer, R. T. (1984) Protein-DNA recognition. *Annu Rev Biochem*, **53**, 293–321.

Park, S. M., Lu, C. D. and 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**, 5309–17.

Perez-Martin, J. and De Lorenzo, V. (1995) The amino-terminal domain of the prokaryotic enhancer-binding protein XylR is a specific intramolecular repressor. *Proc Natl Acad Sci U S A*, **92**, 9392–6.

Perez-Martin, J. and de Lorenzo, V. (1995) The σ^{54}-dependent promoter Ps of the TOL plasmid of *Pseudomonas putida* requires HU for transcriptional activation in vivo by XylR. *J Bacteriol*, **177**, 3758–63.

Perez-Martin, J. and de Lorenzo, V. (1996a) ATP binding to the σ^{54}-dependent activator XylR triggers a protein multimerization cycle catalyzed by UAS DNA. *Cell*, **86**, 331–9.

Perez-Martin, J. and de Lorenzo, V. (1996b) In vitro activities of an N-terminal truncated form of XylR, a σ^{54}-dependent transcriptional activator of *Pseudomonas putida*. *J Mol Biol*, **258**, 575–87.

Perez-Martin, J. and de Lorenzo, V. (1996c) Physical and functional analysis of the prokaryotic enhancer of the σ^{54}-promoters of the TOL plasmid of *Pseudomonas putida*. *J Mol Biol*, **258**, 562–74.

Ramos, J. L., Marques, S. and Timmis, K. N. (1997) Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid — encoded regulators. *Annu Rev Microbiol*, **51**, 341–73.

Ramos, J. L., Mermod, N. and Timmis, K. N. (1987) Regulatory circuits controlling transcription of TOL plasmid operon encoding *meta*-cleavage pathway for degradation of alkylbenzoates by *Pseudomonas*. *Mol Microbiol*, **1**, 293–300.

Ramos, J. L., Michan, C., Rojo, F., Dwyer, D. and Timmis, K. (1990a) Signal-regulator interactions. Genetic analysis of the effector binding site of XylS, the benzoate-activated positive regulator of *Pseudomonas* TOL plasmid *meta*-cleavage pathway operon. *J Mol Biol*, **211**, 373–82.

Ramos, J. L., Rojo, F., Zhou, L. and Timmis, K. N. (1990b) A family of positive regulators related to the *Pseudomonas putida* TOL plasmid XylS and the *Escherichia coli* AraC activators. *Nucleic Acids Res*, **18**, 2149–52.

Ramos, J. L., Stolz, A., Reineke, W. and Timmis, K. N. (1986) Altered effector specificities in regulators of gene expression: TOL plasmid *xylS* mutants and their use to engineer expansion of the range of aromatics degraded by bacteria. *Proc Natl Acad Sci U S A*, **83**, 8467–71.

Ramos-Gonzalez, M. I., Duque, E. and Ramos, J. L. (1991) Conjugational transfer of recombinant DNA in cultures and in soils: host range of *Pseudomonas putida* TOL plasmids. *Appl Environ Microbiol*, **57**, 3020–7.

Reddy, B. R., Shaw, L. E., Sayers, J. R. and Williams, P. A. (1994) Two identical copies of IS1246, a 1275 base pair sequence related to other bacterial insertion sequences, enclose the *xyl* genes on TOL plasmid pWW0. *Microbiology*, **140**, 2305–7.

Reed, W. L. and Schleif, R. F. (1999) Hemiplegic mutations in AraC protein. *J Mol Biol*, **294**, 417–25.

Rhee, S., Martin, R. G., Rosner, J. L. and Davies, D. R. (1998) A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. *Proc Natl Acad Sci U S A*, **95**, 10413–8.

Saviola, B., Seabold, R. and Schleif, R. F. (1998a) Arm-domain interactions in AraC. *J Mol Biol*, **278**, 539–48.

Saviola, B., Seabold, R. R. and Schleif, R. F. (1998b) DNA bending by AraC: a negative mutant. *J Bacteriol*, **180**, 4227–32.

Schleif, R. F. and Favreau, M. A. (1982) Hyperproduction of AraC protein from *Escherichia coli*. *Biochemistry*, **21**, 778–82.

Soisson, S. M., MacDougall-Shackleton, B., Schleif, R. and Wolberger, C. (1997) Structural basis for ligand-regulated oligomerization of AraC. *Science*, **276**, 421–5.

Spooner, R. A., Bagdasarian, M. and Franklin, F. C. (1987) Activation of the *xylDLEGF* promoter of the TOL toluene-xylene degradation pathway by overproduction of the *xylS* regulatory gene product. *J Bacteriol*, **169**, 3581–6.

Thomas, M. S. and Glass, R. E. (1991) *Escherichia coli rpoA* mutation which impairs transcription of positively regulated systems. *Mol Microbiol*, **5**, 2719–25.

Tobe, T., Yoshikawa, M., Mizuno, T. and Sasakawa, C. (1993) Transcriptional control of the invasion regulatory gene *virB* of *Shigella flexneri*: activation by *virF* and repression by H-NS. *J Bacteriol*, **175**, 6142–9.

Tobin, J. F. and Schleif, R. F. (1990) Purification and properties of RhaR, the positive regulator of the L-rhamnose operons of *Escherichia coli*. *J Mol Biol*, **211**, 75–89.

Tsuda, M. and Iino, T. (1987) Genetic analysis of a transposon carrying toluene degrading genes on a TOL plasmid pWW0. *Mol Gen Genet*, **210**, 270–6.

Tsuda, M. and Iino, T. (1988) Identification and characterization of Tn4653, a transposon covering the toluene transposon Tn4651 on TOL plasmid pWW0. *Mol Gen Genet*, **213**, 72–7.

Tsuda, M., Minegishi, K. and Iino, T. (1989) Toluene transposons Tn4651 and Tn4653 are class II transposons. *J Bacteriol*, **171**, 1386–93.

Ustav, M., Ustav, E., Szymanski, P. and Stenlund, A. (1991) Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. *EMBO J*, **10**, 4321–9.

Waldo, G. S., Standish, B. M., Berendzen, J. and Terwilliger, T. C. (1999) Rapid protein-folding assay using green fluorescent protein [see comments]. *Nat Biotechnol*, **17**, 691–5.

Wattiau, P. and Cornelis, G. R. (1994) Identification of DNA sequences recognized by VirF, the transcriptional activator of the *Yersinia yop* regulon. *J Bacteriol*, **176**, 3878–84.

Wilcox, G. and Meuris, P. (1976) Stabilization and size of AraC protein. *Mol Gen Genet*, **145**, 97–100.

Williams, P. A. and Murray, K. (1974) Metabolism of benzoate and the methylbenzoates by *Pseudomonas putida* (arvilla) mt-2: evidence for the existence of a TOL plasmid. *J Bacteriol*, **120**, 416–23.

Williams, P. A., Shaw, L. M., Pitt, C. W. and Vrecl, M. (1997) *xylUW*, two genes at the start of the upper pathway operon of TOL plasmid pWW0, appear to play no essential part in determining its catabolic phenotype. *Microbiology*, **143**, 101–7.

Worsey, M. J. and Williams, P. A. (1975) Metabolism of toluene and xylenes by *Pseudomonas putida* (arvilla) mt-2: evidence for a new function of the TOL plasmid. *J Bacteriol*, **124**, 7–13.

Zhang, X., Reeder, T. and Schleif, R. (1996) Transcription activation parameters at *ara* pBAD. *J Mol Biol*, **258**, 14–24.

Zhang, X. and Schleif, R. (1996) Obligatory activator-polymerase addition order at promoters. *Nucleic Acids Res*, **24**, 2575–7.

Zhou, L. M., Timmis, K. N. and Ramos, J. L. (1990) Mutations leading to constitutive expression from the TOL plasmid *meta*-cleavage pathway operon are located at the C-terminal end of the positive regulator protein XylS. *J Bacteriol*, **172**, 3707–10.

Zhou, Q., Lieberman, P. M., Boyer, T. G. and Berk, A. J. (1992) Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes Dev*, **6**, 1964–74.

SUMMARY

STUDIES OF THE TOL PLASMID TRANSCRIPTION FACTOR XylS

XylS, an AraC family transcription factor, positively regulates the alkylbenzoate degradation genes of *Pseudomonas putida* TOL plasmid from the *Pm* promoter. A tandem of 15 bp homologous direct repeats, separated by 6bp and overlapping with the –35 region of the promoter, is required for the activation of *Pm* by XylS. To characterize interaction of XylS with these operator sequences the protein was expressed with an epitope tag in the N terminus, immunopurified, and was shown to specifically bind to the *Pm* operator region *Om*. Binding of XylS protects 44 bp in the *Om* region on both strands from DNaseI digestion and generates hypersensitive sites within the protected area, which lie on the same face of the DNA helix. Results of hydroxyl radical footprinting and methylation interference assays indicate that XylS binds along one side of the DNA and covers four helical turns. The protein has base-specific contacts in four adjacent major-groove regions on the same helical face. Benzoate effector facilitated binding of XylS to its recognition site but had no major effect on the mode of interaction between the protein and DNA. The contact pattern are in accord with the presence of two HTH DNA-binding regions in a XylS molecule which are involved in base-specific contacts with two adjacent major-groove regions of an operator half-site.

To identify which parts of XylS are responsible for the DNA binding, transcriptional activation, and benzoate inducibility we used deletion mutants and hybrid proteins. We found that a 112-residue C-terminal fragment of XylS binds specifically to *Om,* protects the operator from DNase I digestion identically to the wild-type protein, and activates the *Pm* promoter *in vivo*. All the truncations, which incorporated these 112 C-terminal residues, were able to activate transcription at least to some extent when overproduced. Intactness of the 210-residue N-terminal portion was found to be necessary for benzoate responsiveness of XylS. Deletions in the N-terminal and central regions seriously reduced the activity of XylS and caused the loss of effector control, whereas insertions into the putative interdomain region did not change the basic features of the XylS protein. Consequently, XylS consists of two parts, which probably interact with each other. The C-terminal domain carries DNA-binding and transcriptional activation abilities, while the N-terminal region carries effector-binding and regulatory functions.

KOKKUVÕTE

UURIMUS XylS-VALGUST, TOL PLASMIIDI TRANSKRIPTSIOONIFAKTORIST

XylS on AraC perekonda kuuluv transkriptsiooni aktivaator, mis reguleerib alküülbensoaadi lagundamise geenide transkriptsiooni *Pseudomonas putida* TOL plasmiidis. TOL plasmiidi *Pm* promootori aktiveerimiseks XylS-i poolt on vajalikud kaks 15 aluspaarilist (ap.) järjestust, mis asuvad üksteisest 6 ap. kaugusel ja kattuvad *Pm* promootori –35 piirkonnaga. Et uurida XylS-valgu seostumist sellele operaatoralale, lisati valgu aminoterminusse teisest valgust pärit epitoop, puhastati valk immuunoafiinsusmeetodil ja näidati, et see seostub järjestusspetsiifiliselt *Pm* promootori operaatoralaga (*Om*). XylS-i seostumine DNA-ga kaitseb DNaasI eest 44 ap. pikkust ala mõlemal DNA ahelal, samas põhjustades DNaasI suhtes ülitundlike kohtade tekke kaitstud alas, DNA heeliksi ühel ja samal küljel. DNA ja valgu kompleksi uurimine hüdroksüülradikaale ja DNA eelnevat metüleerimist kasutades näitas, et XylS seostub DNA-ga, kattes neli järjestikust heeliksikeeret selle ühel küljel, ja kontakteerub spetsiifiliselt lämmastikalustega neljas kõrvuti asuvas suure vagumuse regioonis. Bensoaadid soodustavad DNA sidumist valgu poolt, kuid ei muuda oluliselt valgu ja DNA omavahelisi kontakte. Saadud tulemused on kooskõlas kahe heeliks-pööre-heeliks-tüüpi DNA-d siduva regiooni olemasoluga XylS-i molekulis, mille abil valk kontakteerub alus-spetsiifiliselt kahe kõrvuti asetseva suure vagumuse regiooniga operaatoris.

Et kindlaks teha, millised XylS osad on vajalikud DNA sidumiseks, transkriptsiooni aktivatsiooniks ja efektormolekulide mõju vahendamiseks, kasutati deletsioonidega ja hübriidseid XylS-valke. Selgus, et XylS-i karboksüterminaalne osa (112 aminohapet) seostub järjestusspetsiifiliselt *Om*-ga, kaitseb operaatorala DNaasI eest sarnaselt terve XylS-valguga ja on võimeline aktiveerima transkriptsiooni *Pm* promootorilt. Kõik deletsioonidega valgud, mis sisaldasid neid 112 aminohapet, olid suutelised aktiveerima transkriptsiooni. Valgu 210 aminohappeline N-terminaalne osa osutus vajalikuks bensoaatide toime vahendamiseks. Deletsioonid valgu aminoterminuses ja keskosas kahandasid oluliselt XylS-i aktiivsust, samuti kadus valgu võime vastata efektoritele. Samas ei mõjutanud insertsioonid oletatavate domeenide vahele XylS-valgu põhilisi omadusi. Järelikult koosneb XylS kahest, tõenäoliselt teineteisega interakteeruvast osast. Valgu C-terminaalsel domeenil on DNA-ga seostumise ja transkriptsiooni aktiveerimise võime, N-terminaalne osa täidab efektori sidumise ja regulatsiooni ülesannet.

ACKNOWLEDGEMENTS

This work was carried out in the Department of Virology and Microbiology at Tartu University, during the years 1994–1999.

Most of all I want to thank my parents, wife and children for their love and everything they have done.

I have great pleasure to thank my supervisor Prof. Mart Ustav for his guidance, support and patience. He has been my teacher during all these years and a good adviser from our first meeting. His advice — "do not quit!" — has encouraged me in the worst trouble and certainly he was right.

I wish to thank Urve Toots, my nearest co-worker during last two years and the best co-worker I ever had an opportunity to work with, for her great impact into our joint work. I enjoyed this teamwork very much.

I am grateful to all my former and present colleges and friends in the lab for help, advice and discussions, for creative and friendly atmosphere. Especially I would like to thank:

Dr. Maia Kivisaar and the other people from the lab "101" for giving me first lessons in molecular biology and for many fruitful discussions.

Drs. Jaanus Remme and Juhan Sedman for their help and their efforts to teach me science.

Dr. Jüri Parik for his expertise and advice.

Drs. Tanel Tenson, Reet Kurg, Marko Piirsoo and Mrs. Kadrin Wilfong for reading my manuscripts.

Dr. Tago Sarapuu who generously provided me reagents for hydroxyl radical footprinting from his own supplies.

Tiiu Mandel with whom we have shared the bitter setbacks and the first success in the beginning of the XylS project.

The guys from the lab "414" for their continuous support and tolerance.

Dr. Aksel Soosaar for supporting young people and to Merike Petuhov for her help in paper-work.

The people from the lab of Dr. Maia Kivisaar for many fruitful discussions.

Prof. Ain Heinaru and Dr. Einari Aavik who introduced me into molecular genetics and taught independence.

I am deeply thankful to Dr. Victor de Lorenzo who has taught and helped me a lot during three short months in his lab. He has kindly provided plasmids and bacterial strains for the XylS work and supported it with his expertise both in bacterial genetics and tapas.

I wish to thank all the people from VdL lab, especially Drs. Juan-Manuel Sanchez, Giovanni Bertoni and Manuel Carmona, for their help and friendship.

FEMS for the fellowship.
Estonian Society for Microbiology and Dr. Tiina Alamäe for support.

I would like to thank Father Christofer Wrembek, S.J. for his guidance during spiritual exercises. He has taught me the importance of explicitness and analytical thinking in all things, both spiritual and material.

This work was partly supported by grants from Estonian Science Foundation, The Independent Estonian Cancer Research Fund and Howard Hughes Medical Institute to MU.

PUBLICATIONS

I

Kaldalu, N., Mandel, T. and Ustav, M. (1996)
TOL plasmid transcription factor XylS binds specifically to the Pm operator sequence.
Mol Microbiol 20(3), 569–579.

Molecular Microbiology (1996) 20(3), 569–579

TOL plasmid transcription factor XylS binds specifically to the *Pm* operator sequence

Niilo Kaldalu, Tiiu Mandel and Mart Ustav*
Department of Microbiology and Virology, Institute of Molecular and Cell Biology, Tartu University and Estonian Biocentre, 23 Riia Street, EE2400 Tartu, Estonia.

Summary

XylS, an AraC family transcription factor, positively regulates transcription of *Pseudomonas putida* TOL plasmid *meta* operon from the *Pm* promoter. A tandem of 15 bp homologous direct repeats, separated by 6 bp and overlapping with the −35 hexamer of the promoter, is required for the activation of *Pm* by XylS *in vivo*. In this study we have characterized specific binding of XylS to the *Pm* operator *Om*. XylS was overexpressed with an epitope tag in its N-terminus. Tagged XylS (N-XylS) was immunopurified and was shown to specifically bind to *Om*. We have used matrix-bound N-XylS in DNA footprinting and methylation interference experiments. Binding of N-XylS protects 44 bp in the *Om* region on both strands from DNase I digestion and generates hypersensitive sites (within the protected area) which lie on the same face of the DNA helix. Results of hydroxyl radical footprinting and methylation interference assays indicate that XylS binds along one side of the DNA and covers four helical turns. The protein has base-specific contacts in four adjacent major groove regions on the same helical face. Our data are in accord with the prediction of the presence of two separate DNA-binding units in an XylS molecule which are involved in base-specific contacts in two adjacent major-groove regions of a half-site. The direct repeat arrangement of the binding site and the mode of DNA binding of XylS are similar to the arrangement of recognition sites and the DNA contact pattern of AraC protein from *Escherichia coli*.

Introduction

Among the transcription factors several protein families have been identified, based on conserved motifs in primary structure, common domain composition, and similar modes of target-site DNA binding. About 30 proteins, showing a high degree of sequence similarity to AraC protein from *Escherichia coli*, have been classified as members of the AraC family of transcription factors (Tobin and Schleif, 1987; Webster *et al.*, 1987; Mermod *et al.*, 1987; Henikoff *et al.*, 1990; Ramos *et al.*, 1990a; Gallegos *et al.*, 1993). Most members of this protein family are transcriptional activators and contain the highly conserved motif I-DIA- - -GF-S- -YF- - -F- - - -G-TPS- -R, identified as the consensus sequence for the family, within their C-terminal part (Gallegos *et al.*, 1993). A common structure has been proposed for dimeric proteins of the family (Ramos *et al.*, 1990a; Gallegos *et al.*, 1993). The unconserved N-terminal part carries protein dimerization and effector binding capabilities while the conserved C-terminal part carries a site-specific DNA-binding capability (Brunelle and Schleif, 1989; Lauble *et al.*, 1989; Ramos *et al.*, 1990b; Bustos and Schleif, 1993; Eustance *et al.*, 1994; Michán *et al.*, 1995). Furthermore, these functional parts have been shown to form separable domains, connected by a flexible hinge region in the AraC protein (Bustos and Schleif, 1993; Eustance *et al.*, 1994). The C-terminal domain contains two regions which have been predicted to form the helix-turn-helix DNA-binding units (Pabo and Sauer, 1984; Francklyn and Lee, 1988; Brunelle and Schleif, 1989). The binding sites for AraC family proteins, characterized thus far (reviewed in Egan and Schleif, 1994) show some similarities: the half-sites are approximately 17 bp in length, they include two major-groove regions and an intervening minor-groove region on the protein-contacting surface, and they can be arranged as either direct or inverted repeats, placed on the same face of DNA (Tobin and Schleif, 1990; Lu *et al.*, 1992; Caswell *et al.*, 1992; Carra and Schleif, 1993; Egan and Schleif, 1994).

XylS belongs to the AraC family of bacterial transcriptional activators. It positively regulates transcription of the *Pseudomonas putida* TOL plasmid *meta*-cleavage operon, which specifies enzymes for the metabolism of alkylbenzoates. XylS is activated by a variety of *meta*-pathway substrates but overproduction of the protein can compensate for the lack of an effector (for a review see Ramos *et al.*, 1987). The *P. putida* TOL plasmid encodes enzymes required for the catabolism of toluene, xylenes and related aromatic hydrocarbons (Worsey and Williams, 1975). The genes coding for these enzymes are grouped into two

Received 9 November, 1995; revised 4 January, 1996; accepted 25 January, 1996. *For correspondence. E-mail ustav@ebc.ee; Tel. (7) 420194; Fax (7) 420286/420194.

operons which are positively regulated by two regulatory genes, *xylR* and *xylS*. The upper pathway operon specifies oxidation of toluene to benzoate and xylenes to toluates. The *meta*-cleavage pathway operon specifies further oxidation of these carboxylic acids to Krebs-cycle precursors, whereas the aromatic ring in catechols (the pathway intermediates) is cleaved in *meta*-fission. In the presence of upper pathway substrates, *Pu* and *Ps*, the σ^{54}-dependent promoters of the upper operon and *xylS* gene, are activated by XylR. Subsequently, an excess of XylS or XylS in combination with effectors, i.e. the degradation products of the upper pathway and substrates for the *meta*-pathway, activates *Pm*, the *meta*-pathway promoter (Harayama and Timmis, 1989; Nakazawa *et al.*, 1990).

The DNA sequences which confer XylS responsiveness have been mapped within the *Pm* promoter/operator region. The structure of the *Pm* operator *Om* resembles *ara1*, the region required for activation of transcription by AraC at the *araBAD* operon promoter in *E. coli*. Two direct repeats 5′-TGCAAPuAAPuPyGGNTA-3′, partially overlapping with the −35 hexamer of *Pm*, are essential for activation of transcription by XylS. Proper phasing and integrity of the repeats have proved to be necessary for the functioning of XylS. Point mutations of nucleotides which are identical between the half-sites resulted in a dramatic decrease in *Pm* induction, whereas substitutions in the variable bases of the repeats or in the inter-repeat region had no effect. The exact 6 bp spacing between the repeats, which enables the spacing of 21 bp, i.e. two helical turns, between the equivalent bases of the half-sites, turns out to be important, since changes of the spacing result in the loss of promoter activity (Kessler *et al.*, 1993). Although this tandem of repeats could behave as a putative binding site for an XylS dimer (Kessler *et al.*, 1994), similarly to the *ara1* site which is occupied by an AraC dimer in the presence of arabinose (Ogden *et al.*, 1980), studies of DNA-binding properties of purified XylS *in vitro* have been impeded because it has been difficult to obtain sufficient quantities of the soluble active XylS protein (de Lorenzo *et al.*, 1993).

In this report we describe expression and purification of the functionally active XylS protein and provide detailed characterization of interactions of XylS with DNA in binding to *Pm* operator *Om*.

Results

Expression and purification of XylS

We fused the influenza virus haemagglutinin (HA) epitope (Field *et al.*, 1988) to the N-terminus of XylS for monitoring and purification of the expressed protein. Mutational studies have indicated that XylS tolerates modifications in the very N- and C-termini of the protein without apparent changes in activity *in vivo*, whereas point mutations or deletions in other parts of the molecule are deleterious to the functioning of the protein (Kessler *et al.*, 1994).The resulting tagged N-XylS protein was identified and immunopurified using the monoclonal antibody 12CA5 (Field *et al.*, 1988; Zhou *et al.*, 1992).

The *xylS* sequence, amplified by the polymerase chain reaction (PCR) and verified by sequencing, was inserted into several expression plasmids. Both wild-type XylS and epitope-tagged N-XylS were expressed from the phage T7 promoter of the pET11 vector in *E. coli* strain BL21(DE3) (Studier *et al.*, 1990) under variable conditions of cultivation and induction. The inducible expression of polypeptides of the appropriate molecular weight for XylS and N-XylS was verified while the tagged protein was recognized by 12CA5 antibodies. The large amounts of overexpressed protein were found exclusively in inclusion bodies (Fig. 1, a and b). We tried several methods for solubilization and renaturation of the aggregated protein, but all of the procedures were unsuccessful.

It has been reported that the expression of XylS from the *tet*-promoter of pBR322 facilitated activation of the responsive promoter *in vivo* (Mermod *et al.*, 1987), indicating that a substantial amount of correctly folded and functionally active XylS is produced. Therefore, N-*xylS* was cloned into pBR322 under the *tet*-promoter and transformed into *E. coli* strain DH5α. Monitoring of the tagged protein showed that a considerable fraction of the expressed N-XylS protein remained soluble in bacteria grown at 37°C or 20°C. The level of expression was not sufficient to determine the appearance of a novel band during SDS–PAGE, but the tagged protein was readily detectable as a single band on immunoblots (Fig. 1, c and d).

The epitope-tagged N-XylS was purified by single-step, batchwise immunoaffinity binding to the 12CA5 monoclonal antibody, coupled to Protein A–Sepharose (Pharmacia) or Protein A–Fractogel (Merck) beads (Harlow and Lane, 1988). Purity and intactness of the N-XylS preparations were estimated by silver staining of SDS–PAGE and immunoblotting (Fig. 1, c and d). Two contaminating bands could be seen in SDS–PAGE: one corresponded to the light chain of the antibody and another corresponded to a co-purified protein (whose molecular mass was approximately 65 kDa) which could be detected on the Western blots of the bacterial lysates. No products of degradation of N-XylS could be detected on immunoblots.

N-XylS specifically binds to Om

The ability of N-XylS to bind specifically to the putative recognition sequence within the *Pm* region was tested using a DNA immunoprecipitation assay, both in the presence and absence of *meta*-toluate as an effector. The plasmid pUPM190, containing the cloned *Pm* region, was

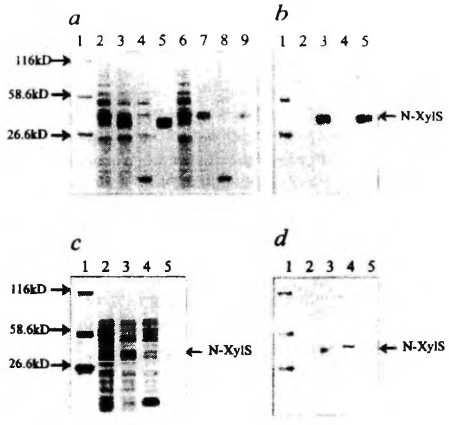

Fig. 1. Expression and purification of XylS and N-XylS proteins.
a. Coomassie-stained SDS–PAGE analysis of expression of XylS (lanes 2–5) and N-XylS (lanes 6–9) from the T7 promoter-driven expression plasmids pETS117 and pETSN117, respectively, in *E. coli* strain BL21(DE3). Lane 1, molecular weight markers; lane 2, total lysate of uninduced bacteria, containing plasmid for expression of XylS; lane 3, total lysate of induced bacteria; lane 4, crude lysate of the induced bacteria, soluble fraction; lane 5, insoluble fraction of the induced bacteria (inclusion bodies); lane 6, total lysate of the uninduced bacteria, containing plasmid for expression of N-XylS; lane 7, total lysate of the induced bacteria; lane 8, crude lysate of the induced bacteria, soluble fraction; lane 9, insoluble fraction of the induced bacteria (inclusion bodies).
b. Immunoblot analysis of expression of N-XylS from T7 promoter of plasmid pETNS117 in *E. coli* strain BL21(DE3). Lane 1, molecular weight markers; lane 2, total lysate of the uninduced bacteria; lane 3, total lysate of the induced bacteria; lane 4, crude lysate of the induced bacteria, soluble fraction; lane 5, insoluble fraction of the induced bacteria (inclusion bodies).
c and d. Silver-stained SDS–PAGE analysis and immunoblot analysis of expression of N-XylS from the *tet* promoter of plasmid pBRSN117, respectively, in *E. coli* strain DH5α. Lane 1, molecular weight marker; lane 2, total protein of pBR322 containing *E. coli* DH5α; lane 3, total protein of pBRSN117 containing bacteria; lane 4, crude lysate, soluble fraction; lane 5, immunopurified N-XylS, retained on the 12CA5 antibody beads.

digested with *Hpa*II and *Hin*fI restriction endonucleases and generated fragments were end-labelled using the Klenow fragment of *E. coli* DNA Polymerase I. Plasmid pUC19, treated in the same way, served as a control for the specificity of DNA immunoprecipitation. We demonstrated that N-XylS-coupled beads bind a single 115 bp *Hpa*II (−97) to *Hin*fI (+19) *Om*-containing fragment of the pUPM190 digest, whereas binding of any other fragment of this plasmid or any fragment of the pUC19 digest could not be observed (Fig. 2, lanes 1 and 2). Beads which were incubated with the lysate of the pBR322-containing bacteria, expressing no N-XylS, were unable to bind any DNA, including the *Om* fragment (data not shown). The data provided indicate that N-XylS binds to a specific site within the *Pm* region. Furthermore, we studied the influence of *meta*-toluate, an XylS-activating effector, on the specificity and affinity of the interaction. According to our results, *meta*-toluate clearly facilitated the specific DNA binding by N-XylS (Fig. 2, lanes 1 and 3), but had no effect on the non-specific DNA binding.

*DNase I footprinting of the N-XylS-*Om *complexes*

In order to map the binding site of XylS in more detail, we analysed the interaction of N-XylS with its putative binding site *Om* by modified protocol of DNase I footprinting (Ustav *et al.*, 1991). The *Om*-containing DNA fragment was end-labelled with the Klenow fragment of *E. coli* DNA polymerase I and the N-XylS-*Om* complex was formed on the beads. Non-bound DNA was removed by washing of the beads and retained fragment was treated with DNase I. After cleavage, the bound DNA was extracted from the beads, purified, and analysed on sequencing gels. The DNase I footprinting data are presented in Fig. 3 and summarized later in Fig. 5. Both strands of the *Om* show a 44 bp area of protection, centred at the two direct repeats. The footprint extends from position −72 to −28 on the

Fig. 2. Immunoprecipitation assay showing that N-XylS binds specifically to *Om*. Radiolabelled *Hpa*II/*Hin*fI digests of pUPM190 (lane 5) and pUC19 (lane 6) were incubated with N-XylS-immunoaffinity beads in the absence (lanes 1, 2) or in the presence (lanes 3, 4) of *m*-toluate as an effector in the binding buffer. Unbound DNA was removed by washing the beads with the buffer. Bound DNA was released and analysed on non-denaturing TBE–PAGE. Lanes 1 and 2, DNA retained on the N-XylS beads from the pUMP190 and pUC19 digests, respectively, in the absence of *m*-toluate; lanes 3 and 4, DNA retained on the N-XylS beads from the pUMP190 and pUC19 digests, respectively, in the presence of *m*-toluate.

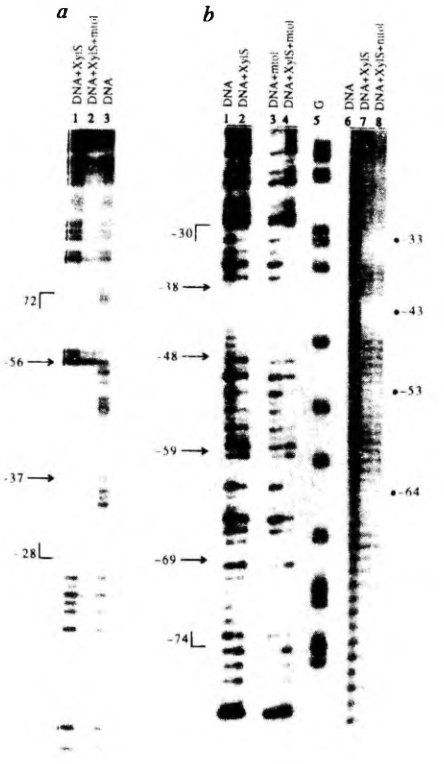

Fig. 3. DNase I and hydroxyl radical footprints of N-XylS at *Om*. DNA templates containing the *Om* region were end-labelled either in the upper strand (a) or the lower strand (b), and incubated with N-XylS-beads for 1 h in binding buffer. Unbound DNA was removed by washing. Both unbound and retained templates were subjected to DNase I or ·OH cleavage in the presence or absence of *m*-toluate as described in the *Experimental procedures*. Brackets indicate regions protected by N-XylS from DNase I cleavage. Arrows mark the hypersensitive sites within the protected region. Dots indicate central positions of regions protected from hydroxyl radicals.

a. Lanes 1 and 2, DNase I digest of the N-XylS-bound (*Xho*I (−117) to *Eco*RI (+68)) fragment from pUPM191, ^{32}P-labelled in the upper strand at +68, in the absence and presence of *m*-toluate, respectively; lane 3, DNase I digest of the same fragment when unbound.

b. DNase I (lanes 1–4) and hydroxyl radical (lanes 6–8) cleavage of the (*Xho*I (−117) to *Eco*RI (+242)) fragment from pUPM190, ^{32}P-labelled in the bottom strand at −117. Lanes 1 and 3, DNase I digest of the unbound template in the absence and presence of the *m*-toluate, respectively; lanes 2 and 4, DNase I digest of the N-XylS-bound fragment in the absence and presence of *m*-toluate, respectively; lane 5, G-specific DNA sequence marker; lane 6, hydroxyl radical cleavage of the unbound fragment; lanes 7 and 8, hydroxyl radical cleavage of the N-XylS-bound fragment in the absence and presence of *m*-toluate, respectively.

upper strand (Fig. 3a) and from −74 to −30 on the lower strand (Fig. 3b). Two DNase I hypersensitive sites appear within the footprint at nucleotides −58/−56 and −37 on the upper strand while four hypersensitive sites, separated by one helical turn, occur at positions −69, −58/−59, −48 and −38 on the lower strand. The DNase I hypersensitive sites appear at equivalent positions within both direct repeats, but the cleavage in these sites is far more effective in the left repeat than in the promoter-proximal repeat. The presence of *meta*-toluate in the binding buffer did not change the protection pattern to any detectable extent, indicating that the effector does not change the mode of interaction between N-XylS and its target DNA.

Hydroxyl radical footprinting of the N-XylS-Om complexes

Hydroxyl radicals are useful probes for studying phosphoribose-backbone contacts with proteins in DNA protein complexes because of their small size and sequence-independent cleavage reactions (Tullius and Dombroski, 1986). In order to prevent hydroxyl radical scavenging by the polysaccharide matrix of Sepharose we used the N-XylS-Fractogel affinity beads in these experiments. Results of the hydroxyl radical footprinting of the non-coding strand are presented in Fig. 3b and summarized later in Fig. 5. Four protected 5 bp regions with equal 5 bp spacing, centred at positions −64, −53, −43 and −33 appear. Less apparent protection emerges around position −72. The protection pattern of *Om* looks almost symmetrical although the recognition sequence has the unsymmetrical direct repeat arrangement. Both half-sites are protected equally, but protections in the promoter-distal ends of the half-sites, near −64 and −43, are less obvious. A three-dimensional representation of the hydroxyl radical footprinting data is presented later (Fig. 6). Four protected regions of a half-helical turn extent appear, separated by a half turn. Consequently, N-XylS binds to *Om* probably as a dimer and contacts the sugar-phosphate backbone on one face of the helix in four adjacent helical turns. As in the case of DNase I footprinting, we did not see any influence of *meta*-toluate on the hydroxyl radical footprints.

Base-specific interactions in the N-XylS-Om complexes

Dimethyl sulphate (DMS) methylates guanine residues at the N^7 position in the major groove and adenine residues at the N^3 position in the minor groove. This observation has been used for studying sequence-specific interactions or close contacts of the proteins with DNA (Johnsrud, 1978; Siebenlist and Gilbert, 1980). We used matrix-bound N-XylS to identify the critical residues for the DNA-binding of XylS in methylation interference and

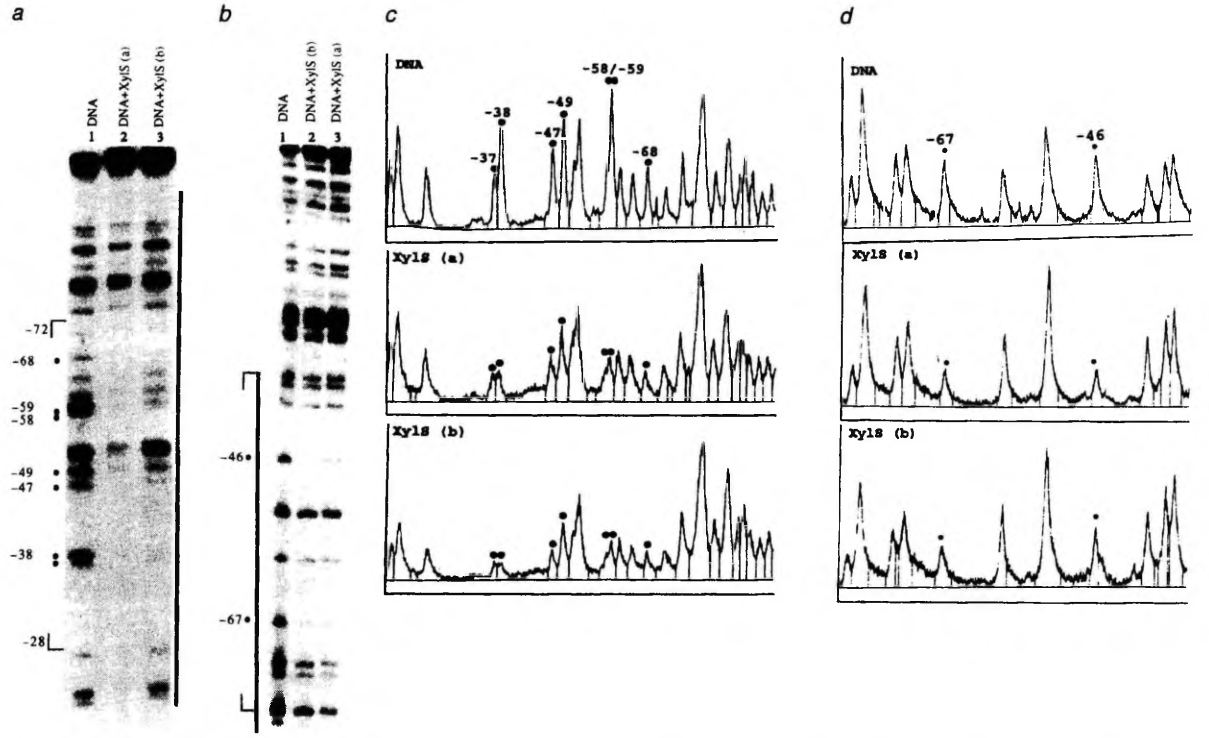

Fig. 4. Methylation interference with N-XylS binding at *Om*. End-labelled DNA templates (Fig. 3) containing *Om* were premethylated and incubated with N-XylS-beads. Bound DNA was separated from unbound template. Both unbound and retained DNA were cleaved at the methylated guanines and analysed on sequencing gels. The relative intensities of the bands were measured using PhosphoImager. Dots indicate guanines which interfere with N-XylS binding when methylated. Brackets indicate regions, protected by N-XylS from the DNase I cleavage. Vertical lines indicate the extent of PhosphoImager scans.

a. The (*Xho*I (−117) to *Eco*RI (+68)) fragment from pUPM191, ^{32}P-labelled in the upper strand at +68 , methylated and treated for G-specific cleavage. Lane 1, the input fragment; lanes 2 and 3, the N-XylS-bound fragment (two independent methylation and binding experiments).

b. The (*Xho*I (−117) to *Eco*RI (+242)) fragment from pUPM190, ^{32}P-labelled in the bottom strand at −117, methylated and cleaved. Lane 1, the input fragment; lanes 2 and 3, the N-XylS-bound fragment (two independent methylation and binding experiments).

c and d. The PhosphoImager analysis of the methylation interference on the upper strand and the bottom strand, respectively.

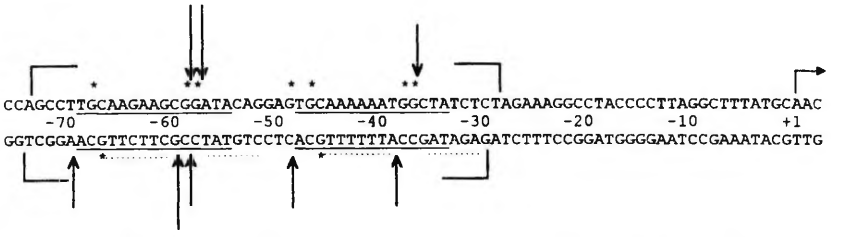

Fig. 5. Summary of footprinting and methylation interference data. The sequences of the two direct repeats are underlined with solid lines. The numbered bases correspond to their positions from the transcription the start site, indicated by a filled arrow. Brackets indicate the outer limits of DNase I protection. Open arrows indicate the DNase I hypersensitive sites. The height of the arrows represents the relative intensity of cleavage. Asterisks indicate guanines which interfere with N-XylS binding when methylated. The regions, protected from hydroxyl radicals are underlined with dotted lines.

methylation protection experiments. This approach enabled us to omit the gel-retardation step for separation of the protein–DNA complex from the unbound DNA and replace it by means of a simple procedure involving washing of the beads with the buffer. The retained DNA was extracted from the beads, cleaved with piperidine and analysed on sequencing gels (Fig. 4, a and b). The relative intensity of the bands was quantified on PhosphorImager (Fig. 4, c and d). Methylation of guanines at positions −68, −59, −58, −49, −47, −38 and −37 on the top strand and at −67 and −46 on the bottom strand interferes with the binding of N-XylS. All these guanines except G −49 are identical bases between the *Om* half-sites. At the same time methylation of the three guanines in variable positions of the left half-site (Om_L) and those in the inter-repeat region does not influence DNA binding by N-XylS. The spacing between the interfering guanines is approximately 10 bp on the upper strand and 20 bp on the lower strand, i.e. one or two helical turns, respectively (Figs 5 and 6). From these results we conclude that XylS binds to its target site probably as a dimer, covering four helical turns, which is the length of *Om*. The N-XylS protein makes base-specific contacts in four adjacent major-groove regions on the same face of DNA, i.e. in two neighbouring major-groove regions per monomer. The pattern of methylation interference is repeated over one major-groove region, reflecting the direct repeat structure of *Om*.

To confirm that indicated bases are engaged in the interaction with XylS, we performed methylation protection experiments of the N-XylS-*Om* complexes. All guanines in the recognition sequence, methylation of which lead to reduction of the specific affinity to N-XylS, were readily protected by the bound N-XylS (data not shown).

Discussion

A number of DNA-binding proteins aggregate during overexpression and form insoluble inclusion bodies. The phenomenon has been interpreted to reflect the coupling of local protein folding to DNA binding by the 'induced fit' model (Spolar and Record, 1994). Local unfolding of DNA-binding regions may result in increased solvent exposure of hydrophobic amino acids and consequent aggregation of the protein. Several proteins of the AraC family turned out to be completely insoluble when overexpressed at high levels (de Lorenzo *et al.*, 1993; Egan and Schleif, 1994), while some pose solubility problems but retain a soluble fraction sufficient for biochemical analysis (Schleif

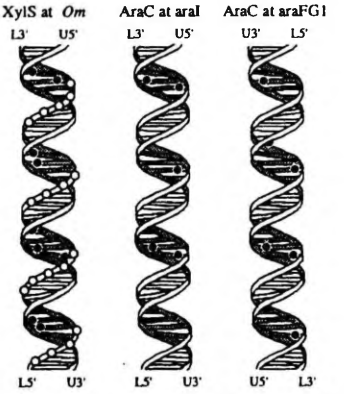

Fig. 6. Comparison of N-XylS contacts at *Om* site with AraC contacts at *araI* and *araFG1* sites. Dots on the three-dimensional representations of the DNA helices indicate the guanines which interfere with protein binding when methylated. Open circles indicate phosphodiesters protected from hydroxyl radicals by N-XylS. U, upper strand; L, lower strand. The data for the *araI* site are from Hendrickson and Schleif (1985) and the data for *araFG1* are from Lu *et al.* (1992).

and Favreau, 1982; Tobin and Schleif, 1990; Caswell *et al.*, 1992). The conventional procedures of renaturation failed to restore the functional activity of these proteins but the problem has been solved for RhaS as well as for AraC by renaturation in the presence of DNA (Egan and Schleif, 1994). In this work, we show that reduction of the level of expression along with affinity purification of the recombinant protein could constitute, in principle, an alternative approach. The processes of protein folding and assembly are controlled by molecular mechanisms involving chaperones (Hartl *et al.*, 1992; Craig *et al.*, 1993). Deficiency of chaperones has been suggested as a possible cause of aggregation of overexpressed proteins and could be overcome by a lower level of expression as well as the simultaneous overexpression of chaperones, which has been used successfully to increase the yield of soluble protein in some cases (reviewed by Wall and Plückthun, 1995). We describe the procedure for expression and purification of XylS, using a moderate level of overexpression of recombinant N-XylS from the *tet* promoter of pBR322. Combination of the moderate level of expression with the use of a monoclonal antibody for single-step purification of the tagged N-XylS protein enabled us to obtain sufficient quantities of recombinant protein for biochemical studies. The use of matrix-bound protein (Ustav *et al.*, 1991) excluded aggregation of N-XylS but also permitted quick and efficient separation of the protein–DNA complex from unbound template. In our assays the cleavage reactions were initiated when the great majority of the DNA template was bound to the protein, without titration of the protein of interest or a risk of non-specific DNA binding.

Our experiments convincingly demonstrated that XylS binds directly to *Om*. The N-XylS-coupled beads bind only one restriction fragment from a total of 23 from the plasmid pUPM190 digest: this fragment was the one that carries the *Om* sequence. This result completely conforms with the data from *in vivo* experiments (Kessler *et al.*, 1993; 1994). N-XylS protects 44 bp, i.e. four helical turns of DNA, in the *Om* region on both strands from DNase I cleavage. Hydroxyl radical footprinting clearly shows that XylS binds to one face of the DNA helix in the *Om* site and the identical bases between the repeats are positioned in the major-groove regions on the protected face of helix. In contrast, the variable bases of the repeats and inter-repeat area lie in the minor-groove regions on the same face of the helix (Fig. 6). Methylation of guanines identical at the two *Om* half-sites at the major-groove-exposed N^7 positions interferes with the N-XylS-DNA interaction. The interfering guanines occur at regular intervals of one helical turn, and are protected from methylation by N-XylS. Therefore, these bases are likely candidates for the specific interactions with XylS, which take place in four adjacent major-groove regions (Fig. 6). The moderate interference by methylated G −49 in the inter-repeat region is an exception and could be caused by the close proximity to the contact site at −47/−46.

The results of methylation interference are in accordance with the prediction that XylS binds DNA by two putative helix-turn-helix (HTH) motifs at positions 230–249 and 282–301 (Mermod *et al.*, 1987; Ramos *et al.*, 1990a). If XylS binds to its target site as a dimer, a XylS monomer should contain two separate DNA-binding units for the base-specific interactions in two adjacent major-groove regions of a half-site. Genetic data are consistent with that prediction, as mutations in both putative HTH forming regions alter XylS activities *in vivo* (Ramos *et al.*, 1986; 1990b; Michán *et al.*, 1992; Gallegos *et al.*, 1993). However, the DNA-binding properties of these mutant proteins have not been studied *in vitro*. The HTH structure of the binding domains has been supported by mutagenesis and subsequent missing-contact probing for one of the two putative sites in the AraC protein (Brunelle and Schleif, 1989). It has been shown that residues 208 and 212 in a proposed recognition helix of AraC, which correspond to amino acids 242 and 246 in XylS, apparently contact DNA. Therefore the stretch of amino acids between 196 and 215 in AraC probably forms an HTH structure. Mutations in the second putative HTH region at positions 245–264 of AraC, equivalent to the 282–301 region in XylS, were inconclusive.

A comparison of the mode of N-XylS binding to the recognition sequence *Om* with that of the binding of AraC from *E. coli* to the *araI* and *araFG1* sequences is presented in Fig. 6. In spite of similar structure and location, the footprint patterns at *Om* and *araI* are different (Fig. 6). In the activating mode, i.e. in the presence of arabinose (Lobell and Schleif, 1990), an AraC dimer contacts three promoter-distal helical turns of *araI*. The polymerase-distal *araI1* half-site is fully occupied, whereas only a half of the polymerase-proximal *araI2* contacts AraC protein, as shown by footprinting, interference assays and missing-contact probing (Hendrickson and Schleif, 1985; Brunelle and Schleif, 1989). In binding to *Om* N-XylS equally protects both half-sites from the cleaving agents and methylation of the critical bases of both half-sites equally interferes with the binding of N-XylS. Therefore, the mode of XylS binding to *Om* resembles that of AraC binding to *araFG1* and *araFG2* binding sites, where both half-sites are evenly occupied by the bound protein (Lu *et al.*, 1992). Similar results have been obtained for another AraC family protein, RhaR, in binding to its recognition site at the p_{sr} promoter (Tobin and Schleif, 1990). Both *araFG* and p_{sr} binding sites are required for transcriptional activation of respective promoters, but the arrangement of the half-sites is completely different from that of *Om* or *araI* (Hendrickson *et al.*, 1990; Tobin and Schleif, 1990). The extension of the DNase I footprint by N-XylS to −28 and methylation

interference of the guanines at −38 and −37 with the N-XylS binding suggest that XylS makes base-specific contacts close to the −35 hexamer of Pm, required for the base-specific interactions by RNA polymerase. As the simultaneous base-specific interactions of both proteins within the same region could not take place, consecutive contacts or binding of the RNA polymerase to the promoter without interaction with the −35 hexamer could be proposed. Base-specific contacts in the region, overlapping with promoter, are not unique for XylS but have also been reported for RhaR by means of missing-contact probing (Tobin and Schleif, 1990). However, these aspects of XylS–DNA interaction remain to be demonstrated.

Intrinsic and protein-induced DNA bending both play important roles in the initiation of transcription (Harrington, 1992; Pérez-Martín *et al.*, 1994) and probably also occur at the *Om* site, as deduced from the sequence of *Om* and changes in the DNase I cleavage pattern of bound N-XylS. Two DNase I-hypersensitive sites on the upper strand and four on the lower strand appear in positions at which N-XylS base-specifically interacts with DNA. Comparison of the footprinting data of different crystallographically characterized protein–DNA complexes justifies a link between the appearance of DNase I-hypersensitive sites and protein-induced bending of DNA (Ansari *et al.*, 1995). Enhanced cleavage is observed at the phosphodiester bond immediately 3′ to the protein-induced kinks in which DNA is bent towards the major groove (Gaston *et al.*, 1990; Schultz *et al.*, 1991). The promoter-proximal Om_R half-site contains an A / T tract 6 bp long which has been shown to bend the DNA helix 17–21 degrees at the junctions between the A / T tract and adjacent DNA (Koo *et al.*, 1986; 1990). The bending is directed towards the minor groove, viewed from the centre of the A6 tract (Koo and Crothers, 1988). Therefore, XylS and the A6 tract could bend DNA co-operatively and into the same direction, mutually strengthening the bending effect. It is notable that DNase I-hypersensitive sites occur at equal intervals of approximately 10 bp. Hence, the XylS-induced kinks should occur on the same helical face and produce a planar curve.

It has been shown that the presence of effectors plays a significant role in regulation of transcription by both AraC and XylS. Arabinose switches between the repressing and the activating mode of AraC binding at the *araBAD* promoter (Lobell and Schleif, 1990). The effector discriminates in this system between the different possibilities for binding-site occupation and causes substantial changes in the conformation of the protein–DNA complex (Carra and Schleif, 1993). For XylS, however, it has been shown that a variety of aromatic compounds can induce the activation of transcription as well as overexpression of XylS in bacterial cells which can substitute the effectors (Mermod *et al.*, 1987). Study of the N-XylS-Om complexes, developed both in the presence of *meta*-toluate and without effector, did not reveal any differences in the patterns of DNA cleavage or protection. DNA-immunoprecipitation experiments, however, indicated enhancement of the binding of the *Pm* fragment by N-XylS in the presence of *m*-toluate. These data support the hypothesis that effectors increase the cellular concentration of XylS in the DNA-binding active conformation which is in equilibrium with an inactive form (Mermod *et al.*, 1987). Once the complex of XylS with *Om* has formed, the effector has no further effect on the interaction, or at least it does not influence the DNA-contact pattern in the complex. An analogous effect has been described for RhaR. The DNA-binding activity of that protein is higher in the presence of L-rhamnose, but the pattern of DNA contacts made both in the presence and absence of effector appears to be identical (Tobin and Schleif, 1990).

Experimental procedures

Plasmids

Vector plasmids pET11c (Studier *et al.*, 1990), pUC19, and pBR322 were used in cloning experiments. pUSR112 is pUC19 containing the TOL plasmid *Eco*RI E fragment bearing *xylS* and *xylR*. pUPM190 is pUC19 containing a 338 bp *Pst*I–*Nae*I fragment of P_L between polylinker *Pst*I and *Ecl*136II binding sites with a 8-meric *Xho*I linker inserted into the *Hin*dIII site. pUPM191 is pUC19 with a *Xho*I site instead of an *Hin*dIII site containing a 158 bp fragment of P_L, generated by spontaneous deletion downstream of +56. pETS117 is pET11c containing a 998 bp PCR-amplified fragment of *xylS*. pETSN117 is pET11c containing the coding sequence for the 12CA5 epitope upstream of *xylS*. pBRSN117 is pBR322 containing the coding sequence for N-XylS inserted between the *Nhe*I and *Bam*HI sites.

Standard recombinant-DNA procedures and PCR

DNA cloning and other common DNA manipulations were performed according to standard protocols (Sambrook *et al.*, 1989). The *xylS*-containing fragment was amplified from pUSR112 DNA by standard PCR methods (Sambrook *et al.*, 1989). We used the following primers: 5′-CCTCTA-GAATGGATTTTTGCTTATTG-3′ (for *xylS*), 5′-CCCCTCT-AGAGATTTTTGCTTATTGAACGAG-3′ (for N-*xylS*), and 5′-GGGGATCCCTTCTTCGGCTACG-3′ for PCR reactions. Target plasmid was linearized by restriction and denatured by alkali. DNA was amplified in 21 cycles of 96°C for 1.3 min, 37°C for 1.6 min and 71°C for 2.5 min.

Growth of bacteria and overproduction of the N-XylS protein

For cloning and plasmid propagation, *E. coli* strain DH5α was grown in Luria–Bertani (LB) medium or on L agar at 37°C (Sambrook *et al.*, 1989). For overproduction of N-XylS and XylS proteins *E. coli* was grown in LB medium. For

expression from T7 promoter pETSN117 or pETS117-containing BL2(DE3) strain (Studier *et al.*, 1990) cells were grown at 20°C or 37°C to an OD_{600} of 0.7, 0.5 mM IPTG was added to induce T7 RNA polymerase, and cells were incubated for an additional 2 h. For constitutive expression, pBRSN117 containing DH5α strain cells were grown at 37°C and harvested at an OD_{600} value of ~1.0. Cells were washed with TBS and stored in lysis buffer (50 mM Tris-HCl, pH 7.8, 100 mM KCl, 2 mM EDTA, 10% (w/v) sucrose) at −70°C.

Cell-free extracts and purification of N-XylS

Cells were thawed, and DTT (to 10 mM), PMSF (to 50 μg ml^{-1}), leupeptin (to 2 μg ml^{-1}) and aprotinin (to 1 μg ml^{-1}) were added to lysis buffer. Cells were incubated with lysozyme at a concentration of 0.5 mg ml^{-1} on ice for 20 min. CHAPS (3-((3-cholamidopropyl) dimethylammonio) 1-propanesulfonate) was added up to 0.2% and KCl to 220 mM final concentrations and incubation was continued for an additional 10 min. Cells were disrupted by sonication and clarified lysate was applied for batchwise immunopurification of N-XylS with continuous end-over-end agitation at 4°C overnight. Affinity beads were prepared by coupling 12CA5 monoclonal antibodies to Protein A–Sepharose (Pharmacia) or Protein A–Fractogel (Merck) according to the standard direct coupling protocol (Harlow and Lane, 1988). Immunopurified N-XylS was stored as bound on beads at −20°C in storage buffer containing 10 mM Tris-HCl, pH 7.8, 100 mM KCl, 10 mM $MgCl_2$, 80 μM EDTA, 80 μM DTT, 50% (v/v) glycerol, and protease inhibitors as in lysis buffer without detectable loss of activity. Protein preparations were analysed by Tricine–SDS–PAGE (Schägger and von Jagow, 1987) and immunoblotting as described earlier (Harlow and Lane, 1988).

DNA immunoprecipitation

pUPM190 and pUC19 were digested with *Hpa*II and *Hin*fI and 3' end-labelled with the Klenow fragment of *E. coli* DNA polymerase I and α-^{32}P. N-XylS beads were suspended in 10 volumes of storage buffer and 50 μl was mixed with radiolabelled DNA and 0.5 ml of binding buffer containing 50 mM Tris-HCl, pH 7.8, 100 mM KCl, 2 mM EDTA, 10% (w/v) sucrose, 0.1% CHAPS, 700 μg ml^{-1} BSA, 2 μg ml^{-1} leupeptin, and 1 μg ml^{-1} aprotinin; 5 mM *meta*-toluate was added in some experiments. Binding mixture was incubated for 1 h at 22°C with end-over-end agitation. Beads were washed with three changes of binding buffer and bound DNA was released by incubation with 50 μg ml^{-1} Proteinase K at 37°C for 30 min in stop solution containing 200 mM NaCl, 5 mM EDTA and 1% SDS. DNA was phenol-extracted, precipitated, and analysed by 5% non-denaturing PAGE.

Footprinting

DNA templates were the *Xho*I (−117) to *Eco*RI (+68) fragment from pUPM191, ^{32}P-labelled in the upper strand at +68, and *Xho*I (−117) to *Eco*RI (+242) fragment from pUPM190, ^{32}P-labelled in the bottom strand at −117 with the Klenow fragment of *E. coli* DNA polymerase I. Complexes were prepared as described for DNA immunoprecipitation but PEG 6000 at a final concentration of 4% was added into binding buffer. After washing, the beads were suspended in 100 μl of DNase I buffer containing 10 mM Tris-HCl, pH 7.8, 100 mM KCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 100 μM DTT, 100 μM EDTA, 100 μg ml^{-1} BSA and 2 μg ml^{-1} sonicated salmon-sperm DNA. DNase I was added to the 0.6–3 μg ml^{-1} concentrations and the reaction mixture was incubated at 30°C for 1 min. DNase I cleavage of unbound template was carried out in the same conditions but DNase I concentrations of 30–300 ng ml^{-1} were used. The reaction was stopped by 200 μl of stop solution containing 600 mM NaCl, 120 mM EDTA, 3% SDS and 150 μg ml^{-1} dextran, followed by incubation with 50 μg ml^{-1} Proteinase K at 37°C for 30 min, phenol-extraction and precipitation. Complexes for hydroxyl radical footprinting were prepared as for DNase I footprinting but the last washing was done with buffer containing 50 mM Tris-HCl, 100 mM KCl and 10 μM EDTA instead of the lysis buffer. DNA retained on the beads was subjected to •OH cleavage as described earlier (Tullius *et al.*, 1987) in a 90 μl volume of reaction mixture, and the reaction was stopped with 100 μl of 87% glycerol. DNA was released and purified as in DNase I footprinting. Complexes for DMS footprinting were prepared as for DNase I footprinting. The methylation reaction was performed as described by Shaw and Stewart (1994) and DNA was released and purified as described above. Cleavage of methylated DNA probe with piperidine was performed according to a standard protocol (Shaw and Stewart, 1994). Cleavage products of footprinting assays were analysed on sequencing gels.

Methylation interference assay

DNA templates were as for footprinting. Methylation by DMS was performed as described by Shaw and Stewart (1994). Complexes were prepared and N-XylS-bound DNA was released as in footprinting. DNA was cleaved as in DMS footprinting and was analysed on sequencing gels. The relative intensity of bands was measured using PhosphorImager (Molecular Dynamics). Part of the pre-methylated DNA probe was not subjected to the binding reaction but was cleaved and used as a negative control for methylation interference and in DMS footprinting assays or as a DNA sequence marker.

Acknowledgements

We are gratefui to Jüri Parik and Tago Sarapuu for technical advice, and to Ain Heinaru for his continuous support. We wish to thank Richard Villems, Maia Kivisaar, Juhan Sedman, Jaanus Remme, Maido Remm, and Marko Piirsoo for helpful comments and suggestions on the manuscript. We also thank Andres Tennus and Anne Märdimäe for their help with artwork. This work was supported, in part, by grants 1134, 287 and 295 from the Estonian Science Foundation, grants LDL 6000 and LKL 100 from the International Science Foundation, and a Tumour Biology grant from the Independent Estonian Cancer Research Foundation (IECRF) to M.U.

References

Ansari, A.Z., Bradner, J.E., and O'Halloran, T.V. (1995) DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature* **374:** 371–375.

Brunelle, A., and Schleif, R.F. (1989) Determining residue-base interactions between AraC protein and *araI* DNA. *J Mol Biol* **209:** 607–622.

Bustos, S.A., and Schleif, R.F. (1993) Functional domains of the AraC protein. *Proc Natl Acad Sci USA* **90:** 5638–5642.

Carra, J.H., and Schleif, R.F. (1993) Variation of half-site organization and DNA looping by AraC protein. *EMBO J* **12:** 35–44.

Caswell, R., Webster, C., and Busby, S. (1992) Studies on the binding of the binding of the *Escherichia coli* MelR transcription activator protein to operator sequences at the *melAB* promoter. *Biochem J* **287:** 501–508.

Craig, E.A., Gambill, B.D., and Nelson, R.J. (1993) Heat shock proteins: molecular chaperones of protein biogenesis. *Microbiol Rev* **57:** 402–414.

Egan, S.M., and Schleif, R.F. (1994) DNA-dependent renaturation of an insoluble DNA binding protein. Identification of the RhaS binding site at *rhaBAD*. *J Mol Biol* **243:** 821–829.

Eustance, R.J., Bustos, S.A., and Schleif, R.F. (1994) Reaching out. Locating and lengthening the interdomain linker in AraC protein. *J Mol Biol* **242:** 330–338.

Field, J., Nikawa, J., Broek, D., MacDonald, B., Rodgers, L., Wilson, I.A., Lerner, R.A., and Wigler, M. (1988) Purification of a RAS-responsive adenylyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. *Mol Cell Biol* **8:** 2159–2165.

Francklyn, C.S., and Lee, N. (1988) AraC proteins with altered DNA sequence specificity which activate a mutant promoter in *Escherichia coli*. *J Biol Chem* **263:** 4400–4407.

Gallegos, M.-T., Michán, C., and Ramos, J.L. (1993) The XylS/AraC family of regulators. *Nucl Acids Res* **21:** 807–810.

Gaston, K., Bell, A., Kolb, A., Buc, H., and Busby, S. (1990) Stringent spacing requirements for transcription activation by CRP. *Cell* **62:** 733–743.

Harayama, S., and Timmis, K.N. (1989) Catabolism of aromatic hydrocarbons by *Pseudomonas*. In *Genetics of Bacterial Diversity*. Hopwood, D., and Chater, K. (eds) London: Academic Press, pp. 151–174.

Harlow, E., and Lane, D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Harrington, R.E. (1992) DNA curving and bending in protein–DNA recognition. *Mol Microbiol* **6:** 2549–2555.

Hartl, F.U., Martin, J., and Neupert, W. (1992) Protein folding in the cell: the role of molecular chaperones Hsp70 and Hsp60. *Annu Rev Biophys Biomol Struct* **21:** 293–322.

Hendrickson, W., and Schleif, R.F. (1985) A dimer of AraC protein contacts three adjacent major groove regions of the *araI* DNA site. *Proc Natl Acad Sci USA* **82:** 3129–3133.

Hendrickson, W., Stoner, C., and Schleif, R.F. (1990) Characterization of the *Escherichia coli araFGH* and *araJ* promoters. *J Mol Biol* **215:** 487–510.

Henikoff, S., Wallace, J.C., and Brown, J.P. (1990) Finding protein similarities with nucleotide sequence databases. *Meth Enzymol* **183:** 111–132.

Johnsrud, L. (1978) Contacts between *Escherichia coli* RNA polymerase and a *lac* operon promoter. *Proc Natl Acad Sci USA* **75:** 5314–5318.

Kessler, B., de Lorenzo, V., and Timmis, K.N. (1993) Identification of a *cis*-acting sequence within the *Pm* promoter of the TOL plasmid which confers XylS-mediated responsiveness to substituted benzoates. *J Mol Biol* **230:** 699–703.

Kessler, B., Herrero, M., Timmis, K.N., and de Lorenzo, V. (1994) Genetic evidence that the XylS regulator of the *Pseudomonas* TOL *meta* operon controls the *Pm* promoter through weak DNA–protein interactions. *J Bacteriol* **176:** 3171–3176.

Koo, H.S., and Crothers, D.M. (1988) Calibration of DNA curvature and a unified description of sequence-directed bending. *Proc Natl Acad Sci USA* **85:** 1763–1767.

Koo, H.S., Wu, H.M., and Crothers, D.M. (1986) DNA bending at adenine-thymine tracts. *Nature* **320:** 501–506.

Koo, H.S., Drak, J., Rice, J.A., and Crothers, D.M. (1990) Determination of the extent of DNA bending by an adenine–thymine tract. *Biochemistry* **29:** 4227–4234.

Lauble, H., Georgalis, Y., and Heinemann, V. (1989) Studies on the domain structure of the *Salmonella typhimurium* AraC protein. *Eur J Biochem* **185:** 319–325.

Lobell, R.B., and Schleif, R.F. (1990) DNA looping and unlooping by AraC protein. *Science* **250:** 528–532.

de Lorenzo, V., Eltis, L., Kessler, B., and Timmis, K.N. (1993) Analysis of *Pseudomonas* gene products using *lacI*^{Q}/*Ptrp–lac* plasmids and transposons that confer conditional phenotypes. *Gene* **123:** 17–24.

Lu, Y., Flaherty, C., and Hendrickson, W. (1992) AraC protein contacts asymmetric sites in the *Escherichia coli araFGH* promoter. *J Biol Chem* **267:** 24848–24857.

Mermod, N., Ramos, J.L., Bairoch, A., and Timmis, K.N. (1987) The *xylS* gene positive regulator of TOL plasmid pWWO: identification, sequence analysis and overproduction leading to constitutive expression of *meta* cleavage operon. *Mol Gen Genet* **207:** 349–354.

Michán, C., Kessler, B., de Lorenzo, V., Timmis, K.N., and Ramos, J.L. (1992) XylS domain interactions can be deduced from intraallelic dominance in double mutants of *Pseudomonas putida*. *Mol Gen Genet* **235:** 406–412.

Michán, C.M., Busby, S.J.W., and Hyde, E.I. (1995) The *Escherichia coli* MelR transcription activator: production of a stable fragment containing the DNA-binding domain. *Nucl Acids Res* **23:** 1518–1523.

Nakazawa, T., Inouye, S., and Nakazawa, A. (1990) Regulatory systems for expression of *xyl* genes on the TOL plasmid. In Pseudomonas *Biotransformations, Pathogenesis and Evolving Biotechnology*. Silver, S., Chakrabarty, A.M., Iglewski, B., and Kaplan, S. (eds). Washington, DC: American Society for Microbiology, pp. 133–140.

Ogden, S., Haggerty, D., Stoner, C., Kolodrubetz, D., and Schleif, R. (1980) The *Escherichia coli* L-arabinose operon: binding sites of the regulatory proteins and a mechanism of positive and negative regulation. *Proc Natl Acad Sci USA* **77:** 3346–3350.

Pabo, C.A., and Sauer, R.T. (1984) Protein–DNA recognition. *Annu Rev Biochem* **53:** 293–321.

Pérez-Martín, J., Rojo, F., and de Lorenzo, V. (1994) Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. *Microbiol Rev* **58:** 268–290.

Ramos, J.L., Stolz, A., Reineke, W., and Timmis, K.N. (1986)

Altered effector specificities in regulators of gene expression: TOL plasmid *xylS* mutants and their use to engineer expansion of the range of aromatics degraded by bacteria. *Proc Natl Acad Sci USA* **83:** 8467–8471.

Ramos, J.L., Mermod, N., and Timmis, K. (1987) Regulatory circuits controlling transcription of TOL plasmid operon encoding *meta*- cleavage pathway for degradation of alkylbenzoates by *Pseudomonas*. *Mol Microbiol* **1:** 293–300.

Ramos, J.L., Rojo, F., Zhou, L., and Timmis, K.N. (1990a) A family of positive regulators related to the *Pseudomonas putida* TOL plasmid XylS and *Escherichia coli* AraC activators. *Nucl Acids Res* **18:** 2149–2152.

Ramos, J.L., Michán, C., Rojo, F., Dwyer, D., and Timmis, K.N. (1990b) Signal-regulator interactions: genetic analysis of the effector binding site of XylS, the benzoate-activated positive regulator of *Pseudomonas* TOL plasmid *meta*-cleavage pathway operon. *J Mol Biol* **211:** 373–382.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Schägger, H., and von Jagow, G. (1987) Tricine-sodium dodecyl sulphate-polyacrylamide gel electrophoresis for proteins in the range from 1 to 100 kDa. *Anal Biochem* **166:** 368–379.

Schleif, R.F., and Favreau, M.A. (1982) Hyperproduction of araC protein from *Escherichia coli*. *Biochemistry* **21:** 778–782.

Schultz, S.C., Shields, G.C., and Steitz, T.A. (1991) Crystal structure of a CAP–DNA complex: the DNA is bent by 90°. *Science* **253:** 1001–1007.

Shaw, E.P., and Stewart, F.A. (1994) Identification of protein–DNA contacts with dimethyl sulphate: methylation protection and methylation interference. In *DNA–Protein Interactions: Principles and Protocols*. Kneale, G.G. (ed.). Totawa, New Jersey: Humana Press, pp. 79–87.

Siebenlist, U., and Gilbert, W. (1980) Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. *Proc Natl Acad Sci USA* **77:** 122–126.

Spolar, R.S., and Record, Jr, M.T. (1994) Coupling of local folding to site-specific binding of proteins to DNA. *Science* **263:** 777–784.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Meth Enzymol* **185:** 60–89.

Tobin, J.F., and Schleif, R.F. (1990) Purification and properties of RhaR, the positive regulator of the L-rhamnose operons of *Escherichia coli*. *J Mol Biol* **211:** 75–89.

Tullius, T.D., and Dombroski, B.A. (1986) Hydroxyl radical 'footprinting': high-resolution information about DNA–protein contacts and application to lambda repressor and Cro protein. *Proc Natl Acad Sci USA* **83:** 5469–5473.

Tullius, T.D., Dombroski, B.A., Churchill, M.E.A., and Kam, L. (1987) Hydroxyl radical footprinting: a high-resolution method for mapping protein–DNA contacts. *Meth Enzymol* **155:** 537–558.

Ustav, M., Ustav, E., Szymanski, P., and Stenlund, A. (1991) Identification of the origin of replication of bovin papillomavirus and characterization of the viral origin recognition factor E1. *EMBO J* **10:** 4321–4329.

Wall, J.G., and Plückthun, A. (1995) Effects of overexpressing folding modulators on the *in vivo* folding of heterologous proteins in *Escherichia coli*. *Curr Op Biotechnol* **6:** 507–516.

Webster, C., Kempseil, K., Booth, I., and Busby, S. (1987) Organisation of the regulatory region of the *Escherichia coli* melibiose operon. *Gene* **59:** 253–263.

Worsey, M.J., and Williams, P.A. (1975) Metabolism of toluene and xylenes by *Pseudomonas putida (arvilla)* mt-2: evidence for a new function of the TOL plasmid. *J Bacteriol* **134:** 757–764.

Zhou, Q., Lieberman, P.M., Boyer, T.G., and Berk, A.J. (1992) Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes Dev* **6:** 1964–1974.

II

Kaldalu, N., Lepik, D., Kristjuhan, A. and Ustav, M. (2000) Monitoring and purification of proteins using bovine papillomavirus E2 epitope tags. Biotechniques 28(3), 456–462.

Short Technical Reports

Bovine Papillomavirus Type 1 E2 Epitope Tagging System for Monitoring and Purification of Proteins

BioTechniques 28:_-_(March 2000)

ABSTRACT

We describe here the use of two newly mapped bovine papillomavirus type 1 (BPV-1) E2 proteins as epitope tags. We constructed several vector plasmids for overexpression as well as for moderate expression of single- or double-tagged proteins in either Escherichia coli *or eukaryotic cells. The new tags were fused to several proteins, and the activity of the tagged proteins was tested in different assays. The tags were shown not to interfere with the function of these proteins in vivo and in vitro. Interaction of the monoclonal antibodies 3F12 and 1E2 with their respective epitopes was specific and had high affinity in a variety of conditions. We have demonstrated that the 3F12 antibody-epitope interaction tolerates high salt concentrations up to 2 M. This permits immunoprecipitation and immunopurification of the tagged proteins in high-salt buffers and reduction of the nonspecific binding of the contaminating proteins. We also provide a protocol for DNA binding and DNase I footprinting assays using the tagged, resin-bound DNA-binding proteins. The BPV-1 E2-derived tags can be recommended as useful tools for detection and purification of proteins.*

INTRODUCTION

Epitope tagging is a recombinant DNA technique by which a protein is made immunoreactive to a preexisting antibody. This technique simplifies detection, characterization, purification and in vivo localization of proteins and has become a standard method of molecular genetics (3). However, some tags are not useful in certain applications due to high background binding. In some cases, affinity purification and immunoprecipitation of a tagged protein is problematic due to co-precipitation of contaminating proteins.

Here, we describe the use of two recently mapped bovine papillomavirus type 1 (BPV-1) E2 protein as epitope tags. We constructed several vector plasmids for over-expression as well as for moderate-level expression of either single- or double-tagged proteins in *Escherichia coli* and eukaryotic cells. The new tags were fused to functionally different proteins: a bacterial transcriptional activator, XylS, that agregates and becomes nonfunctional at high levels of expression, several mutants of the tumor suppressor protein p53 for overexpression in *E. coli* and rat glyceraldehyde-3-phosphate dehydrogenase (GAPDH) for expression in eukaryotic cells. Hereby, we report the use of BPV-1 E2-derived tags and respective monoclonal antibodies for the identification of proteins on Western blots, in the immunofluorescence staining of cells and DNA band-shift assay. In addition, we describe an application of the resin-bound tagged protein in DNA binding and DNase I footprinting assays.

MATERIALS AND METHODS

Plasmid Constructions

For the construction of pBR-3F12 and pBR-1E2, we inserted the coding sequence of XylS with the N-terminally fused influenza virus hemagglutinin (HA) epitope (2) from the pET11c-based parent plasmid pETSN117 (4) between the *Nhe*I and *Bam*HI sites of pBR322. The subcloned fragment contained in addition to a ribosome binding site, an extra *Nde*I site and a start codon preceding the tag sequence. The resultant plasmid pBRSN 117 contains *Xba*I and *Bam*HI sites for cloning of a recombinant coding sequence (4). Then, the coding sequence of the HA epitope between *Nde*I and *Xba*I sites was replaced with double-stranded synthetic oligonucleotides encoding peptides GVSSTSSDFRDR and TTGHYSVRD, recognized by anti-BPV-1 E2 monoclonal antibodies 3F12 and 1E2, respectively (7).

For the construction of pBR-NC, we amplified *xylS* sequence by PCR using the 3′-end primer containing a coding sequence for the peptide TSSDFRDR. The peptide recognized by 3F12 Mab was fused in frame to the C-terminus of XylS and flanked by *Kpn*I and *Bam*HI sites. The resultant PCR fragment was cloned into pBR-1E2 generating the expression plasmid with cloning sites *Xba*I and *Kpn*I.

pET-3F12 was generated by cloning the *Bam*HI/*Nde*I fragment from pBR-3F12-*xylS* into the corresponding sites in pET-11c. Then, the *XylS* gene was removed by cleavage with *Xba*I and *Bam*HI and replaced with coding sequences for different mutant p53 proteins bearing *Xba*I and *Bgl*II sites at the ends.

For the construction of pCG-3F12, double-stranded synthetic oligonucleotide encoding 3F12 epitope and incorporating multicloning sites was inserted between the *Xba*I and *Bgl*II sites of pCG (9). Plasmid pCG-3F12-GAPDH contains rat glyceraldehyde-3-phosphate dehydrogenase (GAPDH) inserted between *Bam*HI and *Bgl*II sites of pCG-3F12. Plasmid pCG-3F12-GAPDH-NLS contains nuclear localization signal of human p53 (amino acids 305–327) fused to the C-terminus of the GAPDH protein.

Immunoaffinity Binding of 3F12-XylS

pBR-3F12-*xylS* bearing *E. coli* DH5α cells were grown in LB medium supplemented with 100 μg/mL ampicillin at 20°C to an A_{600} of approximately 1.0. Cells were harvested, washed with TBS and resuspended in 1/10 volume of high-salt lysis buffer containing 100 mM Tris-HCl, pH7.5, 1.5 M NaCl and 5 mM EDTA, 20% (wt/vol) glycerol. Cell suspension was frozen in liquid nitrogen and stored at -70°C.

For the preparation of crude extracts, cells were thawed and dithiothreitol (DTT; 10 mM), PMSF (100 μg/mL), aprotinin (1 μg/mL), and CHAPS (0.2%) were added. Cells were incubated with lysozyme (0.5 mg/mL) on ice for 20 min and disrupted by sonication. The lysate was clarified by centrifugation at 40000× *g* at 4°C for 30 min.

Affinity beads were prepared by coupling 3F12 anti-BPV E2 monoclonal antibody to divinylsulfon-activated

Table 1. Main Characteristics of Expression Vectors

Plasmid	Promoter	Tag location	Tag sequence
pBR-1E2	*Ptet*	N-terminus	M* T T G H Y S V R D S R... 5'-CAT ATG ACA ACA GGG CAT TAC TCT GTA AGA GAT TCT AGA GGA TCC -3' *NdeI* **XbaI** **BamHI**
pBR-3F12	*Ptet*	N-terminus	M* G V S S T S S D F R D R S R... 5'-CAT ATG GGT GTC TCA TCC ACC TCT TCT GAT TTT AGA GAT CGC TCT ACA GGA TCC -3' *NdeI* **XbaI** **BamHI**
pBR-NC	*Ptet*	N-terminus 1E2 C-terminus 3F12	M* T T G H Y S V R D S R... 5'-CAT ATG ACA ACA GGG CAT TAC TCT GTA AGA GAT TCT AGA-3'... *NdeI* **XbaI** ...G T S S D F R D R STOP ...5'-GGT ACC TCT TCT GAT TTT AGA GAT CGC TGA GGA TCC-3' **KpnI** **BamHI**
pET-3F12	T7	N-terminus	M* G V S S T S S D F R D R S R... 5'-CAT ATG GGT GTC TCA TCC ACC TCT TCT GAT TTT AGA GAT CGC TCT ACA-3' **NdeI** **XbaI**
pCG-3F12	CMV IE	N-terminus	M* G V S S T S S D F R D R... 5'-TCT AGA ATG GGT GTC TCA AGT ACT TCT TCT GAT TTT AGA GAT CGC... **XbaI** *ScaI* ...G A K L P G L Q G T R S ...GGA TCC AAG CTT CCC GGG CTG CAG GGT ACC AGA TCT **BamHI** **HindIII** **SmaI** **PstI** **KpnI** **BglII**

In amino acid sequences, the epitopes are underlined and the initiator methionine codons are marked with asterisks
In nucleotide sequences, cleaving sites for restriction endonucleases are underlined.
Restriction enzymes which cut once per plasmid are marked in bold. pET-3F12 contains an additional *Xba*I site which is sensitive to *dam* methylation and is not cleaved when the plasmid is isolated from *dam*$^+$ strains

Toyopearl HW65 TSK-gel. The 3F12-TSK affinity beads were incubated with crude lysate at 4°C for 1 h with end-over-end agitation. Beads were washed extensively with high-salt lysis buffer on glass filter. These 3F12-XylS beads were stored at -20°C in approximately 10 gel volumes of storage buffer containing 10 mM Tris-HCl, pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 80 µM EDTA, 80 µM DTT, 50% (vol/vol) glycerol, 100 µg/mL PMSF and 1 µg/mL aprotinin. These beads were used for DNA binding and "footprinting" assays.

DNA Precipitation Assay

End-labeled restriction fragments of pUPM190 *Hpa*II/*Hin*fI digest and 50 µL of the suspension of 3F12-XylS beads were mixed with 0.3 mL of DNA binding buffer containing 50 mM Tris-HCl, pH 7.5, 100 mM KCl, 2 mM EDTA, 10% (wt/vol) sucrose, 0.1% CHAPS, 700 µg/mL BSA, 1 µg/mL aprotinin and 1 mM *meta*-toluate. The mixture was incubated for 30 min at room temperature with end-over-end agitation. Beads were washed 3× with 1 mL of DNA binding buffer and incubated with Proteinase K (50 µg/mL) at 37°C for 30 min in stop solution containing 200 mM NaCl, 5 mM EDTA and 1% SDS. The retained DNA was released by phenol-extraction, ethanol-precipitated, and identified in 5% PAGE under non-denaturing conditions.

DNase I Footprinting

The *Om*-containing *Eco*RI/*Xho*I fragment of pUPM190 was end-labeled at *Xho*I site in the lower strand with Klenow fragment. The labeled probe

was incubated with 50 µL of the suspension of 3F12-XylS beads in DNA binding buffer as for the DNA precipitation. The beads were washed 3× and suspended in 100 µL of DNase I buffer containing 10 mM Tris-HCl, pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 100 µM DTT, 100 µM EDTA, 100 µg/mL BSA and 2 µg/mL sonicated salmon-sperm DNA. DNase I was added to concentrations of 0.6–3.0 µg/mL, and the reaction mixture was incubated for 1 min at 30°C. The reaction was stopped by 200 µL of stop solution containing 600 mM NaCl, 120 mM EDTA, 3% SDS and 150 µg/mL dextran, and the beads were treated with Proteinase K (50 µg/mL) at 37°C for 30 min. DNA was extracted from the beads with phenol/chloroform, precipitated with ethanol and analyzed on sequencing gel. DNase I cleavage of the unbound template was carried out in the same buffer at DNase I concentrations of 30–300 ng/mL.

Gel-Shift Assay

Gel-shift assay was performed as previously described (1). We used the crude lysate preparations of p53 mutants. The double-stranded p53 con oligo (5'-GAT CCG GAC ATG CCC GGG CAT GTC CGG ATC-3') was used as a probe. Protein-DNA complexes were separated from the unbound DNA on 5% PAGE (55:1).

Immunodetection of Proteins

For Western blot analysis, cells expressing the tagged proteins were lyzed in SDS sample buffer, total protein extracts were separated on 12% SDS-PAGE and electroblotted onto nitrocellulose membrane filters. Blots were blocked with 1% non-fat dry milk for 1 h, incubated with anti-BPV E2 1E2, 3F12 or anti-p53 pAb 240 monoclonal antibodies, followed by incubation with anti-mouse IgG alkaline phosphatase-conjugated secondary antibody.

For immunofluorescence staining, Saos-2 cells were transfected with plasmids expressing 3F12-GAPDH-NLS or 3F12-GAPDH proteins. Cells were grown on microscopy cover glasses and fixed 24 h after transfection with methanol at -20°C. Proteins were detected with 3F12 and anti-mouse Ig antibody conjugated with FITC as a primary and secondary antibody, respectively.

RESULTS AND DISCUSSION

Moderate Level Bacterial Expression of Epitope-Tagged Proteins

The XylS protein, a transcriptional activator from the TOL plasmid of the soil bacterium *Pseudomonas putida*, like some other AraC/XylS family transcription factors, does not tolerate high level of over-expression and is prone to aggregation both inside the cell as well as in the solution, in the course of purification (4,5). We expressed XylS and several truncated variants of the protein at a near to native level in *E. coli* and tested their physiological activities in vivo. For that, we constructed vectors for the moderate level of expression of epitope-tagged fusion proteins. The vectors were based on pBR322, and the *tet* promoter of this plasmid was used for expression of recombinant proteins. The vectors pBR-3F12 and pBR-1E2 were constructed for the expression of proteins with N-terminally fused peptides GVSSTSSDFRDR and TTGHYSVRD, recognized by anti-BPV E2 monoclonal antibodies 3F12 and 1E2, respectively (7). Both vectors contain *Xba*I and *Bam*HI sites for cloning of the recombinant sequence. We also constructed the vector pBR-NC for the

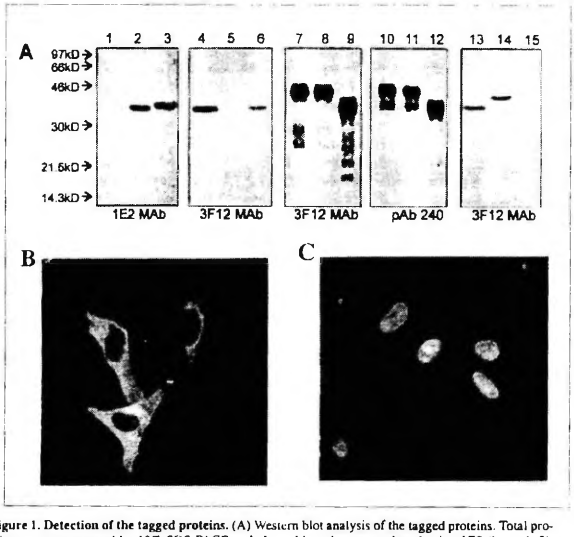

Figure 1. Detection of the tagged proteins. (A) Western blot analysis of the tagged proteins. Total protein extracts, separated by 12% SDS-PAGE and electroblotted, were analyzed using 1E2 (lanes 1–3), 3F12 (lanes 4–9 and 13–15) or anti-p53 pAb 240 (lanes 10–12) primary antibodies and alkaline phosphatase-conjugated secondary antibody. Lanes 1–6: *E. coli* DH5α cells producing tagged XylS protein, bearing pBR-3F12 (lanes 1 and 4), pBR-1E2 (lanes 2 and 5) and pBR-NC (lanes 3 and 6) derived expression plasmids. Lanes 7–12: *E. coli* BL21 (DE3) cells bearing pET-3F12 derived expression plasmids, producing 3F12-tagged p53 variants ΔN39ΔC362 (lanes 7 and 10), ΔN39ΔC362trp248 (lanes 8 and 11) and ΔN61ΔC362 (lanes 9 and 12). Lanes 13 and 14: Saos-2 cells transfected with pCG-3F12 derived expression plasmids producing 3F12-tagged GAPDH (lane 13) and GAPDH-NLS (lane 14). Saos-2 cells expressing untagged p53 were used for a negative control (lane 15). (B and C) Subcellular localization of 3F12-GAPDH and 3F12-GAPDH-NLS. Saos-2 cells were transfected with pCG-3F12-GAPDH (B) or pCG-3F12-GAPDH-NLS (C) and tagged proteins were detected with immmunofluorescence analysis.

expression of proteins with different epitope tags in both N- and C-termini. The N-terminal tag was TTGHYSVRD as in pBR-1E2 and the C-terminal tag was TSSDFRDR, a shorter version of the epitope recognized by 3F12 Mab. The cloning sites for a coding sequence in pBR-NC are *Xba*I and *Kpn*I (Table 1).

We transformed *E. coli* DH5α with plasmids expressing the tagged versions of XylS and analyzed the expression of the tagged XylS proteins by Western blotting. The 3F12 antibody recognized both 3F12-XylS and NC-XylS proteins, while the 1E2 antibody recognized 1E2-XylS and NC-XylS proteins as single bands on the Western blot (Figure 1A, lanes 1–6). No cross-reaction with cellular proteins was observed. However, when 3F12 Mab was used for the detection of the double-tagged NC-XylS protein, a much weaker signal was detected when compared with the 1E2 Mab signal. (Figure 1A, lanes 4 and 6). This could be explained by the use of the shorter version of the 3F12-specific epitope in the double-tagged protein. We studied the effect of these tags on the activity of XylS protein in *E. coli* strain CC118*Pm-lacZ*, which carries a chromosomal copy of the XylS responsive *Pm* promoter fused to the *lacZ* gene (6). The tags had no effect on the transcriptional activation by XylS (data not shown).

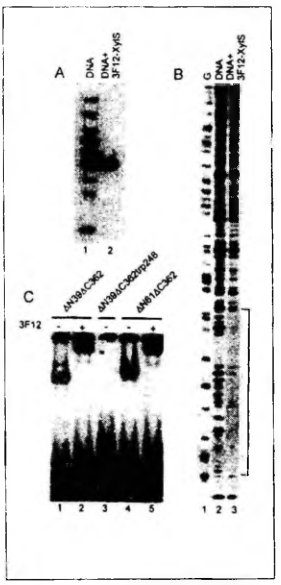

Figure 2. Use of the tagged proteins in DNA binding assays. (A) DNA precipitation by 3F12-tagged XylS. Radiolabeled *HpaII*/*Hin*fI digest of the *Om*-containing plasmid pUPM190 was incubated with 3F12-beads containing 3F12-XylS. Unbound DNA was removed by washing. Bound DNA was extracted from the beads and analysed on non-denaturing TBE/PAGE (5%), using the fragment mixture as a marker (lane 1). Only a single fragment, which contained the XylS binding site *Om*, was retained on the beads (lane 2). (B) DNase I footprinting by resin-bound 3F12-tagged XylS. *Xho*I/*Eco*RI *Om* containing fragment from pUPM190 was end-labeled in the lower strand at 3' terminus and incubated with 3F12-beads containing 3F12-XylS. Unbound DNA was removed by washing. Both free and protein-bound templates were subjected to DNase I cleavage. Lane 1, G-specific DNA sequence marker; lane 2, DNase I digest of the unbound DNA fragment; lane 3, DNase I digest of the fragment bound to 3F12-XylS. Brackets indicate the region protected from the DNase I cleavage by the DNA-bound XylS. (C) Band-shift assay of the 3F12-tagged p53 proteins. The DNA-binding activity of the p53 proteins was studied by separating the protein-DNA complexes from the unbound DNA on 5% PAGE (55:1). Crude *E. coli* lysates containing p53 mutants ΔNΔC362 and ΔNΔC362 produce a shifted band (lanes 1 and 4). Supershift with tag-specific monoclonal antibody 3F12 was used to show that the protein-DNA complexes contain p53 (lanes 2 and 5). The ΔN39ΔC362trp248 p53 protein carrying a point mutation in its DNA-binding domain was used as a negative control (lane 3).

Study of Specific DNA Binding by the Matrix-Attached Epitope-Tagged Protein

XylS is a DNA-binding protein, which specifically binds to the *Om* operator sequence and activates the responsive *Pm* promoter. However, suitable conditions to study the soluble XylS in vitro have not been found, as the protein tends to aggregate. To demonstrate the site-specific DNA binding of epitope-tagged XylS in vitro, we used immunobound 3F12-XylS protein that was attached to the TSK beads through the N-terminal tag and 3F12 antibody. Western blot analysis showed that high salt concentrations, up to 2 M NaCl, do not hinder the interaction of 3F12 MAb with the specific epitope (data not shown). Therefore, to avoid aggregation of XylS and coimmunoprecipitation of contaminating proteins, the crude lysate was prepared in a high-salt lysis buffer, containing 1.5 M NaCl, and 3F12-XylS was bound to the affinity resin carrying 3F12 Mab by a single-step, batchwise procedure.

The DNA binding assay was carried out by mixing the 3F12-XylS beads with a mixture of end-labeled restriction fragments of the *Om*-containing plasmid pUPM190 (4). After several washes, the bound DNA was released and identified by gel electrophoresis, using the input mixture of fragments as a marker. Figure 2A shows that only a single, the *Om*-containing fragment of the pUPM190 digest was retained on the beads, and binding of any other fragment could not be detected.

Further, we analyzed the interaction of the immobilized 3F12-XylS protein with *Om* by DNase I footprinting. The specific complexes of the *Om*-containing DNA fragment and TSK-bound 3F12-XylS were prepared identically as for DNA precipitation and were treated with DNaseI. After the cleavage, DNA was extracted from the beads and analyzed on a sequencing gel. As a control, the unbound template was cleaved at lower concentrations of DNaseI to obtain the equal rate of cleavage. Figure 2B shows that 3F12-XylS protects a 44 bp area on the lower strand and four DNase I hypersensitive sites occur within the protected region. We obtained an identical DNaseI footprint earlier with HA-epitope tagged XylS, using similar technical approach (4). DNA precipitation and DNase I footpriting with a resin-bound epitope-tagged protein (10) are the methods of choice for proteins that are prone to aggregation and are difficult to purify.

High Level Bacterial Expression of Epitope-Tagged Proteins

For high-level bacterial expression of N-terminally tagged proteins, we constructed the vector pET-3F12, a derivative of pET-11c that contains the tag-encoding sequence and the cloning sites identical to pBR-3F12 (Table 1). The coding sequences for mutant p53 proteins were cloned into the vector. Mutant proteins ΔN39ΔC362 and

ΔN61ΔC362 had their transactivation and regulatory parts deleted, but maintained the ability to bind DNA, while the mutant protein ΔN39ΔC362trp248 had lost its DNA-binding ability due to the point mutation in its DNA-binding domain (8). Plasmids, generated for the T7 promoter-directed expression of 3F12-tagged p53 fusion proteins were transformed into *E. coli* strain BL21 (DE3).

Expression of the tagged p53 proteins was monitored by Western blot analysis. The tag-specific 3F12 antibodies and p53-specific pAb240 antibodies were used for the detection of the proteins. Using this vector system resulted in enormous overexpression of the protein detected with 3F12 Mab as well as with pAb240 antibodies (Figure 1A, lanes 7–12).

The DNA-binding activity of the tagged p53 proteins was studied in a band-shift assay (Figure 2C). The p53 mutant ΔN39ΔC362trp248, carrying a point mutation in its DNA-binding domain, was used as a negative control. The mutants ΔN39ΔC362 and ΔN61ΔC362 are functional in DNA binding and produced a shifted band. To show that the produced complex really contains p53, the protein-DNA complexes were supershifted with the tag-specific monoclonal antibody 3F12. This way, epitope tags can be used to verify specificity of the shifted complex in a DNA band-shift assay without purification of the protein.

Expression of Epitope-Tagged Proteins in Eukaryotic Cells

For eukaryotic expression, two coding sequences of rat glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene were cloned into pCG-3F12 plasmid (Table 1). The first had 3F12 epitope fused in frame to GAPDH amino acids 2–333 (pCG-3F12-GAPDH) and the second also contained a nuclear localization signal of p53 (amino acids 305–327) fused to the C-terminus of the GAPDH protein (pCG-3F12-GAPDH-NLS).

Saos-2 cells were transfected with 1 μg of expression plasmids and analysed 24 h after transfection. Expression and localization of the proteins were determined by Western blotting and immunofluorescence analysis, respectively. The 3F12 antibody recognized both proteins as single bands on the Western blot (Figure 1A, lanes 13 and 14) and no cross-reaction with cellular proteins was observed (Figure 1A, lane 15). Immunofluorescence staining of transfected cells with 3F12 antibody indicated that both proteins were localized in the appropriate compartment of the cell: The 3F12-GAPDH in the cytoplasm and 3F12-GAPDH-NLS in the nucleus (Figure 1, B and C). These results indicate that 3F12 epitope-tag can also be used for the detection and determination of the localization of proteins expressed in eukaryotic cells.

Advantages of the BPV E2-Derived Epitope Tags

We analyzed the expression of proteins tagged with the BPV E2-derived epitopes in *E. coli* and eukaryotic cells. Detection of the tagged proteins both on immunoblots and by immunofluorescence staining of cells indicates low background activity, sensitivity and good signal-to noise ratio of the used epitope-antibody combinations. We did not observe any cross-reaction with cellular proteins. Because of the high specificity of the epitope-antibody interaction, our tagging system is especially useful for the studies of protein localization in the cells. In addition, we have shown that interaction of 3F12 MAb with the specific epitope is not hindered by high salt concentrations. That allows to immunoprecipitate and immunopurify the tagged proteins in high-salt conditions and to avoid coimmunoprecipitation of contaminating proteins as well as to avoid aggregation of the protein of interest in the course of purification.

REFERENCES

1. **Abroi, A., R. Kurg and M. Ustav.** 1996. Transcriptional and replicational activation functions in the bovine papillomavirus type 1 E2 protein are encoded by different structural determinants. J. Virol. *70*:6169-6179.
2. **Field, J., J. Nikawa, D. Broek, B. MacDonald, L. Rodgers, I.A. Wilson, R.A. Lerner and M. Wigler.** 1988. Purification of a RAS-responsive adenylyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. Mol. Cell. Biol. *8*:2159-2165.
3. **Jarvik, J. W. and C. A. Telmer.** 1998. Epitope tagging. Annu. Rev. Genet. *32*:601-618.
4. **Kaldalu, N., T. Mandel and M. Ustav.** 1996. TOL plasmid transcription factor XylS binds specifically to the *Pm* operator sequence. Mol. Microbiol. *20*:569-579.
5. **Kessler, B., M. Herrero, K.N. Timmis and V. de Lorenzo.** 1994a. Genetic evidence that the XylS regulator of the Pseudomonas TOL meta operon controls the Pm promoter through weak DNA-protein interactions. J. Bacteriol. *176*:3171-3176.
6. **Kessler, B., K.N. Timmis and V. de Lorenzo.** 1994b. The organization of the *Pm* promoter of the TOL plasmid reflects the structure of its cognate activator protein XylS. Mol. Gen. Genet. *244*:596-605.
7. **Kurg, R., J. Parik, E. Juronen, T. Sedman, A. Abroi, I. Lilv, U. Langel and M. Ustav.** 1999. Effect of bovine papillomavirus E2 protein-specific monoclonal antibodies on papillomavirus DNA replication. J. Virol. *73*:4670-4677.
8. **Lepik, D., I. Ilves, A. Kristjuhan, T. Maimets and M. Ustav.** 1998. p53 protein is a suppressor of papillomavirus DNA amplificational replication. J. Virol. *72*:6822-6831.
9. **Tanaka, M. and W. Herr.** 1990. Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. Cell *60*:375-386.
10. **Ustav M., E. Ustav, P. Szymanski and A. Stenlund.** 1991. Identification of the origin of replication of bovine papillomavirus and characterization of the viral origin recognition factor E1. EMBO J. *10*:4321-4329.

The study was supported by the Estonian Science Foundation Grant Nos. 2496, 2497, 2315 and 2316. *Address correspondence to Dr. Mart Ustav, Department of Microbiology and Virology, Institute of Molecular and Cell Biology, Tartu University, 23 Riia Street, Tartu, 51010, Estonia. Internet: ustav@ebc.ee*

Received 14 July 1999; accepted 11 November 1999.

Nillo Kaldalu, Dina Lepik, Arnold Kristjuhan and Mart Ustav
Tartu University
Tartu, Estonia

III

Kaldalu, N., Toots, U., de Lorenzo, V. and Ustav, M. (2000)
Functional domains of the TOL plasmid transcription factor XylS.
J Bacteriol 182(4), 1118–1126.

JOURNAL OF BACTERIOLOGY, Feb. 2000, p. 1118–1126 Vol. 182, No. 4
0021-9193/00/$04.00+0

Functional Domains of the TOL Plasmid Transcription Factor XylS

NIILO KALDALU,[1] URVE TOOTS,[1] VICTOR DE LORENZO,[2] AND MART USTAV[1]*

Department of Microbiology and Virology, Institute of Molecular and Cell Biology, Tartu University, Estonian Biocentre, 51010 Tartu, Estonia,[1] and Centro Nacional de Biotecnología–CSIC, Campus de Cantoblanco, 28049 Madrid, Spain[2]

Received 16 September 1999/Accepted 17 November 1999

The alkylbenzoate degradation genes of *Pseudomonas putida* TOL plasmid are positively regulated by XylS, an AraC family protein, in a benzoate-dependent manner. In this study, we used deletion mutants and hybrid proteins to identify which parts of XylS are responsible for the DNA binding, transcriptional activation, and benzoate inducibility. We found that a 112-residue C-terminal fragment of XylS binds specifically to the *Pm* operator in vitro, protects this sequence from DNase I digestion identically to the wild-type (wt) protein, and activates the *Pm* promoter in vivo. When overexpressed, that C-terminal fragment could activate transcription as efficiently as wt XylS. All the truncations, which incorporated these 112 C-terminal residues, were able to activate transcription at least to some extent when overproduced. Intactness of the 210-residue N-terminal portion was found to be necessary for benzoate responsiveness of XylS. Deletions in the N-terminal and central regions seriously reduced the activity of XylS and caused the loss of effector control, whereas insertions into the putative interdomain region did not change the basic features of the XylS protein. Our results confirm that XylS consists of two parts which probably interact with each other. The C-terminal domain carries DNA-binding and transcriptional activation abilities, while the N-terminal region carries effector-binding and regulatory functions.

The TOL plasmid of *Pseudomonas putida* encodes a pathway for the catabolism of toluene and xylenes (35, 36). The genes, which encode enzymes for catabolism of these hydrocarbons, are grouped into two operons. The upper pathway operon specifies oxidation of toluene to benzoate and xylenes to alkylbenzoates. The *meta*-pathway operon specifies further oxidation of these compounds, whereas the aromatic ring in catechols, the pathway intermediates, is cleaved in *meta* fission. Two regulatory proteins, XylR and XylS, positively regulate the catabolic operons (28). In the presence of upper pathway substrates, XylR activates the *Pu* promoter of the upper pathway operon and the *Ps1* promoter of the *xylS* gene. Subsequently, overproduced XylS protein activates the *Pm* promoter of the *meta*-cleavage operon (10, 26) via binding to the operator-sequence *Om* (9, 12, 13). Furthermore, XylS protein is constitutively expressed at a low level from the weak *Ps2* promoter (6) and in the presence of benzoates, i.e., the degradation products of the upper pathway and substrates for the *meta* pathway, it activates the *Pm* promoter at low protein concentrations. Therefore, transcriptional activation by XylS is stimulated by alkylbenzoates and modulated by the intracellular level of the protein. Expression of XylS from strong promoters has shown that overproduction of the protein, naturally mediated by XylR, is sufficient for activation of *Pm* in the absence of benzoate effector (11, 19, 34). On that basis, Mermod et al. (19) have suggested a hypothesis about a dynamic equilibrium between inactive and active, DNA-binding conformations of the protein in the cell. The putative role of effector would be to shift the equilibrium toward the active conformation. In vitro studies from our lab support the idea, as we have shown that alkylbenzoate slightly facilitates site-specific DNA binding but does not change the pattern of the DNA contacts made by XylS (12).

To identify which parts of the protein mediate the effect of benzoates, extensive mutagenesis of the *xylS* gene has been carried out. Several amino acid substitutions have been found that altered effector specificity of XylS (20, 25, 27) or produced a semiconstitutive phenotype. The latter was characterized by an increased basal level of transcriptional activity which was still inducible by benzoate (20, 37). Mutations of both types were scattered all over the *xylS* gene; however, many of them were clustered in a small glycine-rich N-terminal region P37-R45 (Fig. 1). It has been shown that some mutations in the C terminus are intra-allelically dominant over substitutions in the N terminus and, in contrast, a mutation in the N terminus (R45T) can restore the effector control that has been lost due to these C-terminal mutations. All these data suggest that the N and C termini of XylS may interact, and benzoate effectors may regulate the activity of the protein by modulation of that interaction (21).

XylS belongs to the AraC-XylS family of bacterial transcriptional activators (7). Proteins of this family are characterized by significant homology over a 100-residue stretch, a region that is proposed to be necessary for DNA binding and stimulation of transcription. Several small monomeric activators in the AraC-XylS family (e.g., MarA and SoxS from *Escherichia coli*) match the conserved sequence and do not contain additional domains (1, 3). A crystal structure for one of these proteins, MarA, in complex with its cognate binding site has been determined recently (29). AraC, the model protein of the family consists of two functional domains. The conserved C-terminal domain carries sequence-specific DNA-binding capability while the nonconserved N-terminal domain mediates effector responsiveness and carries dimerization capability (2, 16). The DNA-binding domain most probably contains all of the determinants necessary for transcriptional activation, since the separately expressed C-terminal domain of AraC has re-

* Corresponding author. Mailing address: Department of Microbiology and Virology, Institute of Molecular and Cell Biology, Tartu University, Estonian Biocentre, 23 Riia St., 51010 Tartu, Estonia. Phone: 372-7-375047. Fax: 372-7-420286. E-mail: ustav@ebc.ee.

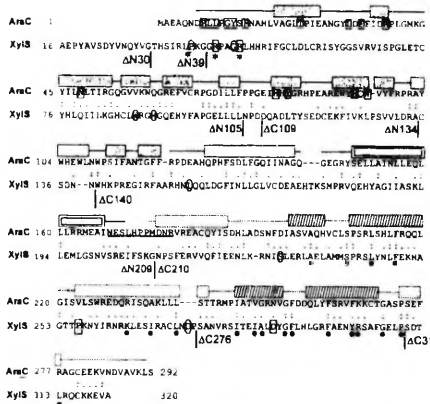

FIG. 1. Sequence alignment of *E. coli* AraC (accession no. P03021) and *P putida* XylS (P07859) proteins. Identical residues are marked by colons and similar ones by periods. An 18.5% identity and a 51.2% similarity were found in a 248-residue overlap (amino acids 45 to 284 in AraC and 79 to 319 in XylS). A 21.2% identity and a 58.6% similarity were found in a 104-residue C-terminal region (amino acids 183 to 284 in AraC and 218 to 319 in XylS). Sequences were aligned by using the LALIGN program. Vertical lines indicate the endpoints of XylS truncations generated in this study. Mutations of boxed XylS residues affect effector specificity. Mutations of rounded XylS residues produce semiconstitutive transcriptional activation, and mutations of residues marked by an asterisk increase XylS stability. Mutations of R41 and R45 in XylS cause both semiconstitutive and altered effector specificity phenotypes (20, 27, 37). Dots under the sequence of XylS indicate conserved residues in the AraC-XylS family (7). The AraC residues, which interact directly with arabinose, are printed on a gray background, and those interacting indirectly are boxed. The elements of the AraC NTR secondary structure are shown above the sequence. Gray boxes indicate beta-sheets, and transparent boxes indicate helices. The dimerization helix, participating in a coiled coil, is shown by a double box (32). The linker region of AraC is underlined (4, 5). The secondary structure elements of homologous MarA protein are shown above the AraC CTD sequence with a dotted outline; the helix-turn-helix motifs are shaded (29).

sidual ability to activate transcription without the arabinose effector (18). In AraC, the N-terminal arm of the regulatory domain binds to the C-terminal domain and acts as an intramolecular modulator of transcriptional activation (31). Structural data show that binding of arabinose causes rearrangement of the arm, thereby releasing the protein to bind the target sites which are necessary for activation of transcription (32, 33).

Like the other regulators of carbon metabolism in the family, XylS contains the conserved region in its C terminus and could have the AraC-like modular organization. However, the modular structure of the XylS protein had not been demonstrated up to now, and the truncated variants of the protein have been reported to be completely inactive (14). XylS and AraC show some sequence homology not only in the C-terminal portion but also in the N-terminal region (Fig. 1). Therefore, we set a goal to specify which parts of XylS are responsible for the DNA-binding, transcriptional activation, and effector responsiveness of the protein and to clarify whether XylS and AraC proteins use similar mechanisms under which the effector controls the activity of the transcription factor. The results of our work show that XylS is a modular protein with a C-terminal DNA-binding–transcription activation domain and an N-terminal effector-binding–regulatory region.

MATERIALS AND METHODS

Construction of plasmids and strains. For cloning and plasmid propagation, *E. coli* DH5α was grown in Luria-Bertani (LB) medium or on L agar at 37°C (30). Media were supplemented with 100 μg of ampicillin ml^{-1} and 7 μg tetracycline (Tc) ml^{-1} when required. DNA cloning and other common DNA manipulations were performed according to standard protocols (30).

Plasmid pBRSN217 for expression of XylS variants was constructed as follows. The *lac* operator sequence was generated by annealing of oligonucleotides 5′-AGCTTTAATGCGGTAATTGTGAGCGGATAACAATT-3′ and 5′-AGCTAATTGTTATCCGCTCACAATTACCGCATTAA-3′ (the underlined sequences denote the binding site of the *lac* repressor) and inserted into the single *Hin*dIII site in pBRSN117 (12). For expression of the N-terminal tag recognized by anti-BPV E2 monoclonal antibody 3F12 (15), oligonucleotides 5′-TATGGGTG TCTCATCCACCTCTTCTGATTTTAGAGATCGCT-3′ (coding strand) and 5′-CTAGAGCGATCTCTAAAATCAGAAGAGGTGATGAGACACCCA-3′ were annealed and inserted between the *Nde*I and *Xba*I sites to replace the sequence encoding the hemagglutinin epitope. The coding sequences of *xylS* variants were amplified by PCR with the primers listed in Table 1 and inserted into pBRSN217 between the *Xba*I and *Bam*HI sites.

The *xylS* variants were amplified from pUSR112 (12) DNA by standard PCR methods (30). For amplification of the *xylS* variants with N-terminal deletions and XylS(L5R,L6K), we used the *xylS* reverse primer and a suitable forward primer. For amplification of the *xylS* variants with C-terminal deletions, the *xylS* forward primer and a suitable reverse primer were used. Termination of polypeptides ΔC276 and ΔC310 appeared to be ineffective. Therefore, a *Bcl*I linker containing an additional stop codon was inserted into the *Bam*HI site.

For construction of the *xylS* variants Δ140–209 and Δ174–209, we amplified fragments of *xylS* containing nucleotides (nt) 3 to 416 and 3 to 521, with the *xylS* forward primer and reverse primers ΔC140-*Xba*I and ΔC174-*Xba*I, respectively. The amplified fragments were cloned into the *Xba*I site of pBRSN217 with the coding sequence for ΔN209 between the *Xba*I and *Bam*HI sites. For construction of Δ39–47, we amplified a *xylS* fragment containing nt 141 to 962 with ΔN47 forward primer and *xylS* reverse primer. pBRSN217 with the coding sequence for *xylS* was cut with *Sty*I and *Bam*HI, and the amplified fragment was inserted between these two sites.

xylS variants with insertions were constructed as following. The coding sequence for the λCI dimerization domain was amplified from λ phage DNA. The coding sequence for the bovine papillomavirus type 2 (BPV1) E2 protein hinge region was amplified from pET11 bearing the E2 gene. For construction of the GA/XylS expression plasmid, we amplified a sequence encoding a peptide, $(GS-(GAGGGAGGAGAGARS)_4$, the tetrameric Gly-Ala (4GA) repeat, which was inserted into the BPV-1 E2 gene to replace the region encoding residues 192 to 311 (D. Örd, R. Kurg, and M. Ustav, unpublished data). Next, we amplified fragments of *xylS* containing nt 3 to 416 and 3 to 626, with forward primer *xylS-Spe*I and reverse primers ΔC140-*Xba*I and ΔC209-*Xba*I, respectively. The amplified *xylS* fragments were cloned into the *Xba*I site of pBRSN217 with the coding sequence for ΔN209 between the *Xba*I and *Bam*HI sites. Thereafter, the heterologous sequences were inserted into the *Xba*I site of the resultant plasmids.

pBRSN317 is pBRSN217 with the *Ps2* promoter inserted instead of modified *Ptet*. For construction of pBRSN317, first, a 568-bp *Ps2*-containing *Bgl*II fragment was inserted into the pUC18 polylinker, and the *Eco*RI-*Hin*dIII fragment of the resultant plasmid was cloned into pBRSN217 between the *Eco*RI and *Hin*dIII sites.

Characterization of XylS variants in vivo. *E. coli* CC118*Pm-lacZ* (14) was transformed with plasmids producing XylS variants. The plasmids bearing bacteria were grown in LB medium overnight with appropriate antibiotics. Then, bacteria were diluted 1:100 in the same medium, in the presence or absence of 1 mM *meta*-toluate, and β-galactosidase levels were determined after 4 h according to the standard protocol (23). Cells were permeabilized with toluene.

DNA immunoprecipitation and DNase I footprinting. Plasmid-bearing *E. coli* DH5α cells were grown at 20°C to an optical density at 600 nm (OD_{600}) of ~0.6, 1 mM IPTG was added, and the cells were incubated for an additional 2 h. Cells were washed with TBS and stored in high-salt lysis buffer (100 mM Tris·HCl [pH 7.5], 1.5 M NaCl, 5 mM EDTA, 20% [wt/vol] glycerol) at −70°C. For preparation of lysates, cells were thawed, and dithiothreitol (to 10 mM), phenylmethylsulfonyl fluoride (to 100 μg ml^{-1}), aprotinin (to 1 μg ml^{-1}), and CHAPS (3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate) (up to 0.2%) were added. Cells were incubated with lysozyme (0.5 mg ml^{-1}) on ice for 20 min and then disrupted by sonication. Clarified lysate was applied for batchwise affinity enrichment with end-over-end agitation at 4°C for 1 h. Affinity beads were prepared by coupling 3F12 anti-BPV E2 monoclonal antibody to divinylsulfon-activated Toyopearl HW65 TSK-gel. The purified proteins were stored, the pUPM190 probes for DNA immunoprecipitation and footprinting were made, and the experiments were carried out as described earlier (12), except that polyethylene glycol 6000 was omitted from the binding buffer when complexes were formed for DNase I footprinting. For footprinting, the wild-type (wt) N-XylS-*Om* complexes were formed in the presence of 1 mM *meta*-toluate, and the ΔN209-*Om* complexes were formed without *meta*-toluate.

Immunoblots. *E. coli* DH5α cells bearing the pBRSN217 series expression plasmids were grown in LB medium with ampicillin at 37°C to an OD_{600} of 0.6,

TABLE 1. Nucleotide sequences and locations of PCR primers

Primer	Sequence (5'→3')[a]	Position in gene[b]	Added restriction site
Forward			
xylS	CCCCTCTAGAGATTTTTGCTTATTGAACGAG	4–24 (*xylS*)	*Xba*I
*xylS-Spe*I	GGACTAGTGATTTTTGCTTATTGAACGAG	4–24 (*xylS*)	*Spe*I
xylS(L5R,L6K)	GCTCTAGAGATTTTTGCCGAAAGAACGAGAAAA	4–28 (*xylS*)	*Xba*I
ΔN8	GCTCTAGAAAAAGTCAGATCTTCGTCC	25–43 (*xylS*)	*Xba*I
ΔN30	CCCCTCTAGAACGCACTCTATTCGCCTGC	90–118 (*xylS*)	*Xba*I
ΔN39	CCCCTCTAGAGGGCGCCCGGCAGG	117–130 (*xylS*)	*Xba*I
ΔN47	*GCC*CACAGAATCTTCGGATGCC[c]	141–159 (*xylS*)	
ΔN105	CCCCTCTAGAAATCCGGATGACCAAGCC	315–332 (*xylS*)	*Xba*I
ΔN134	CCCCTCTAGATGCAGTGACAACAATTGG	402–419 (*xylS*)	*Xba*I
ΔN209	CCCCTCTAGAAACCCGTCTTTCGAGCGAG	627–645 (*xylS*)	*Xba*I
λ *c*I CTD	GCTCTAGAACCTTTACCAAAGGTGATGCG	360–369 (λ *c*I)	*Xba*I
4GA	GCTAGCGTAAGAGATCAGGACCGATCC	564–572 (BPV E2)	*Nhe*I
E2 hinge	GCTCTAGAGATCGCCCAGACGGAG	618–636 (BPV E2)	*Xba*I
Reverse			
xylS	GGGGATCCCTTCTTCGGCTACG	988–973 (*xylS*)	*Bam*HI
ΔC109	GGGGGGATCCTCAATCCGGATTGAGCAG	323–309 (*xylS*)	*Bam*HI
ΔC140	GGGGGGATCCTCAATTGTTGTCACTGCATGCCC	416–397 (*xylS*)	*Bam*HI
ΔC140-*Xba*I	GCTCTAGATCAATTGTTGTCACTGCATGC	416–399 (*xylS*)	*Xba*I
ΔC174-*Xba*I	GCTCTAGACATACAAAGTCGATGCCT	521–504 (*xylS*)	*Xba*I
ΔC210	GGGGGATCCTCAACCTTTGCTGAAAATTTCACGG	626–605 (*xylS*)	*Bam*HI
ΔC209-*Xba*I	GCTCTAGAACCTTTGCTGAAAATTTCACGG-3'	623–605 (*xylS*)	*Xba*I
ΔC276	GGGGGGATCCTCAGGGATCGTTCAAGCAGGC	824–807 (*xylS*)	*Bam*HI
ΔC310	GGGGGGATCCTCAAGGCAACTCGCCGAACGC	926–909 (*xylS*)	*Bam*HI
λ *c*I CTD	GCTCTAGAGCCAAACGTCTCTTCAGGCCAC	711–690 (λ *c*I)	*Xba*I
4GA	ACTAGTATCATTGGTGGTGCGCCTG	900–883 (BPV E2)	*Spe*I
E2 hinge	GCTCTAGACGGTACCGTGCCCTGCACG	858–840 (BPV E2)	*Xba*I

[a] Restriction sites in noncomplementary overhangs are underlined.
[b] Locations refer to positions within indicated open reading frames based on the first nucleotide of the initiation codon.
[c] Bases of an added Ala codon are italicized.

1 mM IPTG was added, and cells were incubated for additional 2 h. Equal amounts of cells, judged by the OD of the bacterial culture, were suspended in sodium dodecyl sulfate (SDS) sample buffer, supplemented with β-mercaptoethanol (30). Samples were boiled for 5 min, and proteins were separated by SDS–12% polyacrylamide gel electrophoresis (PAGE). The proteins were electroblotted onto nitrocellulose membrane filters. The filters were blocked for 1 h with 1% nonfat dry milk, probed with 3F12 anti-BPV E2 monoclonal antibody, and treated with goat anti-mouse alkaline phosphatase-conjugated secondary antibody.

RESULTS

Stimulation of *Pm* by truncated XylS proteins. To examine whether the N and C termini of XylS constitute separable functional domains, we constructed two sets of progressing terminal deletions, from both ends of the coding sequence (Fig. 1). We truncated the protein in putative loop regions, indicated by a prediction of the secondary structure (not shown). The deletion mutants of *xylS* were generated by PCR and verified by sequencing. A peptide, GVSSTSSDFRDR, from BPV E2 protein was fused to the N terminus of wt XylS and the deletion mutants for monitoring and purification of the proteins. The tagged proteins were identified (Fig. 2A) by using the anti-BPV E2 monoclonal antibody 3F12 (15). The tag did not affect the *meta*-toluate responsiveness of wt XylS nor did it affect its ability to stimulate transcription from *Pm*. Therefore, we applied the tagged full-size XylS as a wt control (wt N-XylS). We have used previously the moderate *tet* promoter of pBR322 for production of the N-terminally tagged XylS protein (12). Since expression of several truncated XylS proteins (ΔN209 and ΔC310 in particular) from the *tet* promoter was apparently toxic to *E. coli* and caused plasmid instability, we inserted the *lac* operator sequence downstream of the promoter to reduce expression.

For assay of *Pm* activation, plasmids containing these *xylS* variants were transformed into *E. coli* strain CC118*Pm-lacZ* containing a chromosomal copy of the *Pm* promoter fused to the *lacZ* gene (14). Figure 3 shows β-galactosidase (β-Gal) levels, mediated by wt N-XylS and various deletion mutants, in the presence or absence of *meta*-toluate. Note that the level of expression of wt N-XylS from the modified *tet* promoter (labeled as *Ptet** below) mimics the XylR mediated overexpression of XylS in *P. putida*—it produces full activation of *Pm* without the effector, and addition of the ligand has no further effect on the promoter activity (Fig. 3, line 2). We found that deletion mutant ΔN209, which corresponds to the putative DNA-binding domain, stimulates transcription from *Pm* as efficiently as wt N-XylS. When expressed from *Ptet**, both mediated β-Gal levels close to 10^4 Miller units (Fig. 3, lines 2 and 15).

To probe whether ΔN209 is inducible by benzoates, ΔN209 and wt N-XylS were expressed from the weak *Ps2* promoter (Fig. 4). The level of expression from *Ps2* was so low that both proteins remained undetectable by the Western blotting of crude lysates even by enhanced chemiluminescence (ECL) detection. However, in the presence of *meta*-toluate, such a small amount of wt N-XylS was enough to produce the same β-Gal level as that produced by wt N-XylS overexpressed from *Ptet**. We found that ΔN209 is not inducible by the effector and provides the phenotype of constitutive activator. When activator proteins were expressed from *Ps2*, the stimulation of *Pm* caused by ΔN209 was almost fivefold higher than that pro-

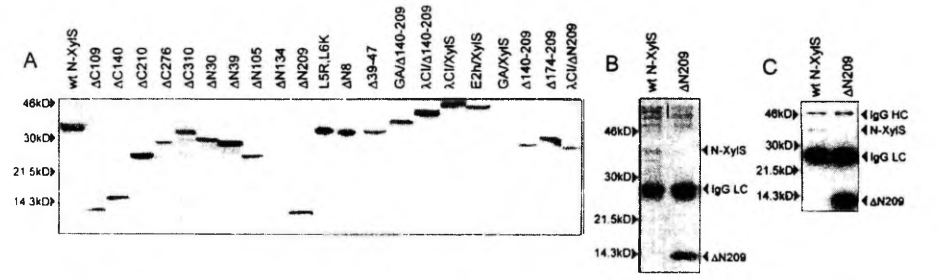

FIG. 2. Expression and partial purification of tagged XylS variants. (A) Immunoblot analysis of expression of XylS proteins in *E. coli* DH5α. Equal amounts of bacteria bearing the pBRSN217 series expression plasmids were used for preparation of the total lysates in Laemmli sample buffer after 2 h of induction with IPTG (isopropyl-β-D-thiogalactopyranoside). Total protein extracts were separated by SDS-PAGE (12%) and analyzed by immunoblotting with the 3F12 monoclonal antibody. The proteins are schematically depicted on Fig. 3. (B) Silver-stained SDS-PAGE (12%) analysis of wt N-XylS and ΔN209 proteins retained on the 3F12 beads. (C) Immunoblot analysis of wt N-XylS and ΔN209 proteins retained on the 3F12 beads. Proteins were separated by SDS-PAGE (12%) and analyzed by immunoblotting by using the 3F12 monoclonal antibody. wt N-XylS and ΔN209 proteins, as well as the light chain (IgG LC) and heavy chain (IgG HC) of the 3F12 monoclonal antibody eluted from the TSK beads are indicated with arrows.

duced by wt N-XylS without effector. However, it reached only 4% of that produced by wt activator in the presence of *meta*-toluate (Fig. 4).

Furthermore, we found that the other deletion mutants which retained the putative DNA-binding domain (ΔN8, ΔN30, ΔN39, ΔN105, and ΔN134) were able to stimulate transcription from *Pm* to some extent when overproduced from *Ptet** (Fig. 3, lines 10 to 14). These N-terminally truncated proteins mediated 4- to 16-fold-higher β-Gal levels than the uninduced basal level of the strain. However, it makes only 0.6 to 2.6% of the β-Gal level produced by wt N-XylS. The XylS variants with longer N-terminal deletions than that of ΔN209 were unable to stimulate *Pm* and could not be detected by immunoblotting, obviously due to instability (data not shown). All deletions from the C terminus produced proteins which could not activate the *Pm* promoter (Fig. 3, lines 3 to 7). In conclusion, these results suggest that the C terminus of XylS indeed forms a DNA-binding domain and contains all the elements necessary for activation of transcription.

DNA binding by XylS CTD. All deletion mutants which activate transcription from *Pm* in vivo should specifically bind to *Om*. In order to confirm the site-specific DNA binding of the C-terminal domain (CTD) of XylS in vitro, we purified both epitope-tagged ΔN209 and wt N-XylS proteins by single-step, batchwise immunoaffinity binding to the 3F12 monoclonal antibody, coupled to TSK beads. The DNA-binding activity was studied with the immobilized protein because purified XylS tends to aggregate in solution, making it impossible to use the regular gel-shift assays for this purpose. The cell extracts were prepared in a high-salt lysis buffer containing 1.5 M NaCl. High concentrations of sodium chloride do not hinder the specific interaction of 3F12 antibody with the epitope, while they reduce coimmunoprecipitation of the contaminating proteins. The protein preparations were examined by silver staining of SDS-PAGE and immunoblotting (Fig. 2B and C). We succeeded in isolating both wt N-XylS and ΔN209 that were functionally active in DNA binding; however, the yield of wt N-XylS was lower, and preparations contained some degradation products or contaminating proteins.

We used the matrix-bound N-XylS and ΔN209 in the specific DNA-binding and DNase I-footprinting assays. The mixture of end-labeled restriction fragments of the *Om*-containing plasmid pUPM190 (12) was incubated with the protein-loaded TSK beads in the presence or absence of *meta*-toluate. After the removal of free probe, only DNA that was bound to the immunopurified protein was retained on the beads. Since estimation of the amounts of proteins used in the assay was complicated, the experiment was carried out at oversaturating levels of the DNA probe so that less than 2% of the input labeled probe were retained on the beads. Figure 5A shows that both ΔN209 and wt N-XylS bind specifically the *Om*-containing 115-bp fragment of the pUPM190 *Hpa*II-*Hin*fI digest, whereas binding of any other fragment could not be observed. DNA binding by wt N-XylS was strongly (up to 100 times) induced by *meta*-toluate. Such a strong effect of *meta*-toluate could be observed only when wt N-XylS was purified from the high-salt lysate (1.5 M NaCl) and was not observed earlier when we used low-salt buffer conditions (220 mM KCl) (12). We did not detect any effect of *meta*-toluate on DNA binding by ΔN209. The 3F12 beads lacking XylS did not bind any DNA (data not shown).

Further, we analyzed the interaction of both protein variants with *Om* by DNase I footprinting. The *Om*-containing DNA fragment was end labeled, and the DNA-protein complex was formed on the beads. With wt N-XylS the binding was done in the presence of *meta*-toluate. Again, the experiments were carried out at oversaturating levels of the DNA probe. After the removal of unbound DNA, the complex was subjected to DNase I cleavage. The treated DNA was extracted from the beads and analyzed on the sequencing gels. We found that both ΔN209 and wt N-XylS protect a 44-bp area, extending from positions −74 to −30 on the lower strand (Fig. 5B). The DNase I footprints of these two protein variants were almost identical, indicating that both have a similar mode of interaction with the binding site. The only difference is hypersensitive to DNase I at nucleotide −75, observed in the complex with wt N-XylS. Enhanced cleavage at that position was detected when the N-XylS-*Om* complex was formed in the presence of *meta*-toluate. That slight effect has been previously overlooked by us (12).

The results of DNA immunoprecipitation and DNase I footprinting confirm that the DNA-binding domain of XylS is located within the C-terminal 112 residues of the protein.

The complete N-terminal domain is required for effector responsiveness of XylS. The N-terminal portion of XylS is believed to be necessary for effector binding and ligand-depen-

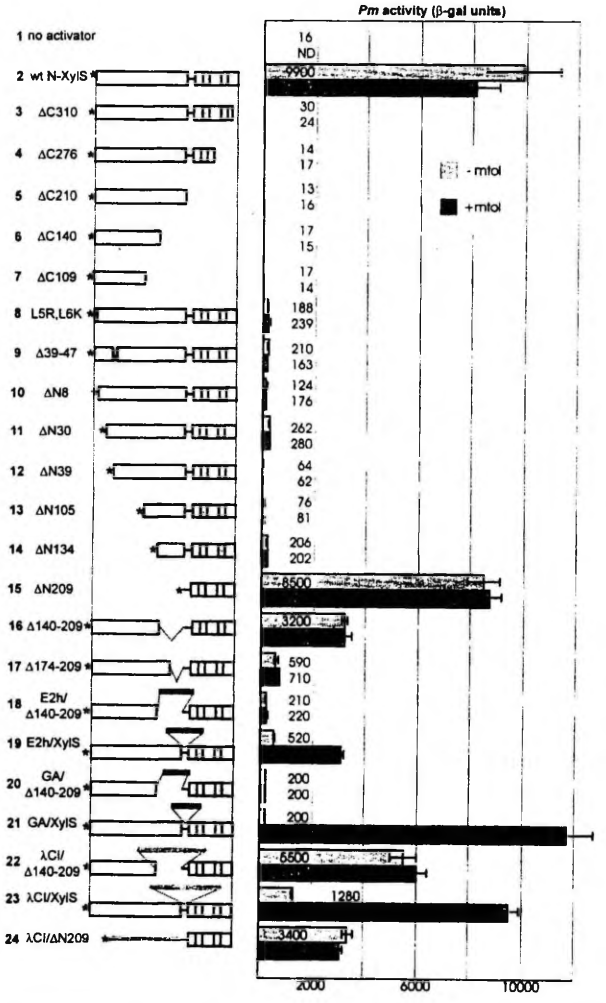

FIG. 3. Activity of *Pm* when mediated by different XylS variants, overexpressed from the *Ptet** promoter, in the presence or absence of *meta*-toluate. Transparent boxes on the diagram indicate the putative domains of XylS. Asterisks indicate the N-terminal tags, and gray boxes within the C-terminal domain indicate the putative helix-turn-helix regions. The hybrid XylS variants contain insertions of the hinge region of BPV E2 protein (lines 18 and 19), a synthetic Gly-Ala-rich region (lines 20 and 21), and the dimerization domain of the λCI protein (lines 22, 23, and 24) of *E. coli* CC118*Pm-lacZ* was transformed with plasmids for expression of XylS variants from the *Ptet** promoter. Bacteria were grown in LB medium overnight, diluted 1:100 in the same medium in the absence (−mtol) or presence (+mtol) of 1 mM *meta*-toluate, and β-Gal levels were determined after 4 h. The values in Miller units are the averages of results from three to six assays. Error bars indicate the standard deviations.

dent regulation of the activity of the protein. We demonstrated that the truncated proteins ΔN8, ΔN30, ΔN39, ΔN105, and ΔN134, which contain a part of the putative regulatory domain in addition to the complete DNA-binding domain, are noninducible by effector and, when expressed from *Ptet**, produce a much lower β-Gal activity than wt N-XylS or ΔN209 (Fig. 3, lanes 2 and 10 to 15). When placed under the control of *Ps2* promoter, these deletion mutants were unable to stimulate the

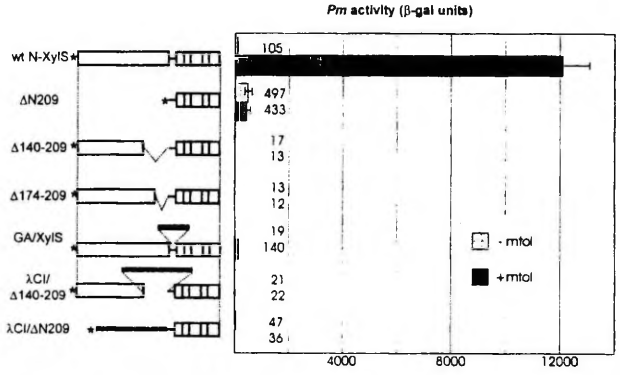

FIG. 4. Activity of *Pm* when mediated by different XylS variants, expressed from the *Ps2* promoter in the presence or absence of *meta*-toluate. *E. coli* CC118*Pm-lacZ* was transformed with plasmids for expression of XylS variants from the *Ps2* promoter. Bacteria were grown in LB medium overnight, diluted 1:100 in the same medium in the absence (−mtol) or presence (+mtol) of 1 mM *meta*-toluate, and β-Gal levels were determined after 4 h. The values in Miller units are the averages of results from three to six assays. Error bars indicate the standard deviations.

Pm promoter (data not shown). The deletion mutants ΔN8, ΔN30, ΔN39, ΔN105, and ΔN134 are expressed at the different intracellular concentrations, probably due to different stabilities of the truncated proteins (Fig. 2A). However, the activity of the N-terminal deletion mutants does not merely correlate with the protein levels seen in Fig. 2A. As the levels of these proteins are readily detectable, they must be far more abundant than wt N-XylS and ΔN209, expressed from the *Ps2* promoter. As we have mentioned above, wt N-XylS and ΔN209 produce substantial *Pm* activation even at the levels of expression which are undetectable by Western blot analysis (Fig. 4). Thus, in vivo transcriptional activation data reflect, at least partially, the intrinsic properties of the truncated proteins and not only different levels of expression. Consequently, N-terminal deletions in the putative regulatory domain cause the loss of the regulatory function and reduce the activity of the C-terminal domain of XylS.

For further characterization of the regulatory portion of XylS and the requirements for its proper functioning, we constructed additional deletion and insertion mutants and tested their ability to stimulate Pm (Fig. 3, lines 16 to 23). Most of the mutants described below were expressed from *Ptet** at a level comparable with that of wt N-XylS (Fig. 2A). Only GA/XylS and E2h/Δ140–209 were apparently very unstable. The latter could be detected on the Western blot only by ECL and is not shown on Fig. 2A.

We deleted two different portions of the central part of XylS from the other end of the putative regulatory domain. The deletion mutants Δ140–209 and Δ174–209 were more active than N-terminal truncations but also had lost their effector responsiveness and showed a reduced ability to stimulate *Pm* (Fig. 3, lines 16 and 17; Fig. 4). To ascertain whether that effect could be caused by the loss of hinge flexibility, which does not allow the domains to interact properly, we replaced the deleted residues 140 to 209 with long unstructured regions: the 80-residue hinge region of BPV1 E2 protein (8) and a 69-residue region, consisting mainly of glycine and alanine, which has been shown to replace effectively the BPV1 E2 hinge (Örd et al., unpublished). As we found that the region containing residues 140 to 209 is responsible for dimerization of XylS (N. Kaldalu and M. Ustav, unpublished data), it was also replaced with the dimerization domain of λCI repressor. The resultant proteins E2h/Δ140–209, GA/Δ140–209, and λCI/Δ140–209 remained unresponsive to *meta*-toluate. E2h/Δ140–209 and GA/Δ140–209 were much less active than wt N-XylS, whereas λCI/Δ140–209, when overexpressed, mediated about one-half of the β-Gal level produced by wt N-XylS (Fig. 3, lines 18, 20, and 22). Expression from the *Ps2* promoter shows, however, that λCI/Δ140–209 is a much weaker transcription activator than wt N-XylS and ΔN209 (Fig. 4). When the same heterologous protein portions were inserted into the putative hinge region of XylS, retaining the N-terminal region intact, we found that these hybrid XylS activators were responsive to the effector. λCI/XylS, E2h/XylS, and GA/XylS, which carry the interdomain insertions, were weaker transcription activators than wt N-XylS but were clearly inducible by *meta*-toluate (Fig. 3, lines 19, 21, and 23; Fig. 4). Since XylS variants with the inserted λCI dimerization domain were relatively more active than the others, we substituted the entire N-terminal region with the λCI domain and expressed the resultant fusion protein λCI/ΔN209 from both *Ptet** and *Ps2* promoters. β-Gal levels mediated by λCI/ΔN209 remained several times lower than those produced by the C-terminal domain ΔN209 itself (Fig. 3, line 24; Fig. 4). These data suggest that the heterologous dimerization domain, as well as the incomplete N-terminal regions of XylS, interfered with transcriptional activation by the XylS CTD. Therefore, we do not have evidence of whether the λCI domain-containing proteins were more active due to their dimeric state or due to the lower level of interdomain interference.

In AraC protein, the N-terminal arm of the regulatory domain has been shown to interact with the C-terminal domain as an intramolecular repressor of binding to the adjacent binding sites, which is necessary for transcription activation. Leucines in the N-terminal arm were crucial for that activity (31). Using mutational analysis, we tried to specify a subdomain of a similar function within the N-terminal domain of XylS. Several mutations that caused the semiconstitutive behavior of XylS are clustered in a small glycine-rich N-terminal region, P37 to R45 (20, 37) (Fig. 1), which was expected to form a loop from

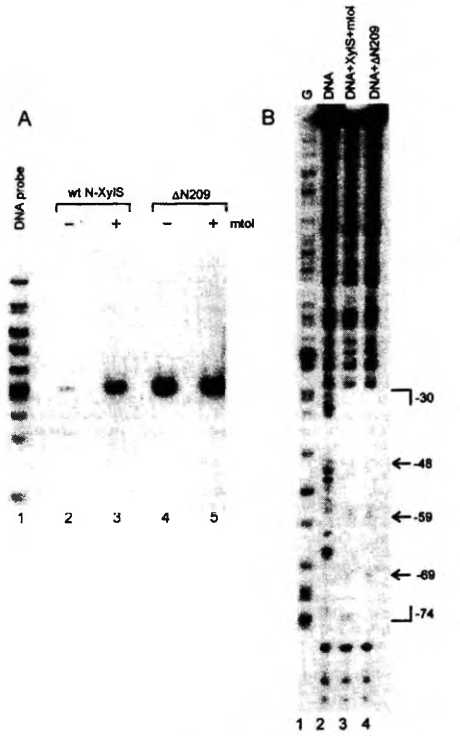

FIG. 5. DNA immunoprecipitation (A) and DNase I footprinting (B) showing that wt N-XylS and XylS CTD bind specifically to *Om*. (A) A radiolabeled *Hpa*II-*Hin*fI digest of pUPM190 (lane 1) was incubated with 3F12 beads containing wt N-XylS (lanes 2 and 3) and ΔN209 (lanes 4 and 5) either in the absence (lanes 2 and 4) or in the presence (lanes 3 and 5) of 1 mM *meta*-toluate. Unbound DNA was removed by washing. Bound DNA was analyzed on nondenaturating TBE-PAGE (5%). (B) The *Xho*I (−117) to *Eco*RI (+242) *Om*-containing fragment from pUPM190 was end labeled in the lower strand at −117 and incubated with 3F12 beads containing wt N-XylS or ΔN209. Unbound DNA was removed by washing. Both free and protein-bound templates were subjected to DNase I cleavage. Lane 1, G-specific DNA sequence marker; lane 2, DNase I digest of the unbound DNA fragment; lane 3, DNase I digest of the fragment bound to wt N-XylS in the presence of *meta*-toluate; lane 4, DNase I digest of the ΔN209-bound DNA fragment. Brackets indicate a region protected by DNA-bound protein from DNase I cleavage. Arrows mark the hypersensitive sites within the protected region.

a prediction of the secondary structure. To remove the putative loop, we produced a deletion mutant, Δ39–47. We also constructed a double mutant, XylS(L5R,L6K), to test whether the leucines in the extreme N terminus have a role in regulation of the XylS activity. Both Δ39–47 and XylS(L5R,L6K), as well as the N-terminal deletion mutant ΔN8, were unresponsive to the effector. However, these mutants were very weak and not constitutive activators (Fig. 3, lines 8 to 10). Therefore, we did not find any region in the XylS N-terminal region (NTR) which could be deleted to produce a constitutive phenotype, as has been shown in the case of AraC.

Thus, the data of the mutational analysis indicate that 210 N-terminal residues of XylS provide the ligand responsiveness to the protein, and the entire region is necessary for that activity. Since it has not been demonstrated that this part of XylS folds independently or functions (binds benzoate) independently of the CTD, it would be correct to use the term the NTR of XylS. All the examined deletions in the XylS NTR reduced the activity of the XylS CTD. Since the level of the *Pm* activation by the mutant proteins positively correlates with the strength of the promoter used for the expression of these proteins, the reduced activation must be an intrinsic characteristic of the mutants and not caused by aggregation due to overexpression.

DISCUSSION

We have shown here that XylS is comprised of a C-terminal DNA-binding–transcriptional-activation domain and an N-terminal regulatory region. Existence of separable domains has been confirmed earlier for AraC and MelR from the same protein family (2, 16, 18, 22) but not for XylS. Kessler et al. (14) have characterized several deletion mutants of XylS in vivo and found that all of them were unable to activate *Pm* or modulate the activity of wt XylS. These mutants were expressed as a result of the readthrough transcription at a very low level, one much lower than those produced from the *Ptet** or *Ps2* promoter, as deduced from the β-Gal activities mediated by wt XylS. Therefore, stimulation of *Pm* by the deletion mutants presumably remained undetectable by Kessler et al. (14) due to the low levels of expression, and their results principally match those obtained by us in the case of the XylS variants with deletions in the NTR.

In the present work, we demonstrated that the C-terminal DNA-binding domain lacks effector responsiveness and that to modulate the activity of the protein in ligand-dependent manner, the complete N-terminal region is necessary. It is possible that mutations in the N terminus which make XylS noninducible by effector and strongly reduce its ability to activate *Pm* disrupt the native structure of the N terminus of the protein. We suggest that the DNA-binding domain may contain a surface or surfaces that interact with the regulatory part of XylS. That idea is in agreement with the knowledge that mutations in both the N and C termini of XylS can yield the semiconstitutive phenotype (27) and that mutations in one domain can be suppressed by mutations in the other (21). It is possible that the misfolded N terminus provides a new site or sites for irreversible interaction with the CTD and, in that way, inhibits DNA binding and/or transcriptional activation.

Currently, we cannot make final conclusions about the mechanism by which ligand regulates the XylS activity. We have shown that *meta*-toluate strongly facilitates DNA binding by matrix-attached N-XylS but does not affect DNA binding by the XylS CTD in vitro (Fig. 4A). Therefore, stimulation of DNA binding must be at least one, if not the single, major effect of ligand. However, additional, cooperative effects should not be excluded. Stimulation of dimer formation by effector is possible, but this should not enhance DNA binding in our assay since we used immobilized wt N-XylS which was attached to the beads through the N terminus and presumably had not enough freedom to change its multimeric state. To account for our data, we propose that the regulatory region reversibly interacts with the DNA-binding–activator domain. The fact that ΔN209, presumably with a monomeric XylS CTD, mediated almost fivefold-higher *Pm* activation than wt N-XylS without effector, when these proteins were expressed from *Ps2*, suggests that the NTR works as an intramolecular

repressor. Binding of the effector to the NTR causes a conformational change in it and releases the intramolecular repression. A similar mechanism has been validated for AraC (31). In the presence of effector, the XylS NTR may either passively release the inhibition or actively facilitate the function of CTD, e.g., by assistance in local folding of the DNA-binding regions.

In spite of the similarity of footprints produced by XylS in the presence or absence of ligand, we cannot conclude that liganded and unliganded XylS form identical complexes with *Om* or that these complexes behave identically in transcriptional activation. The data of in vivo DMS footprinting by Miura et al. (24) suggest that the complexes are not transcriptionally identical and that the benzoate effector modifies the interaction of XylS with RNA polymerase at the *Pm* promoter. RNA polymerase is retained on the *Pm* promoter by XylS in the absence of benzoate inducer and released by effector binding to XylS, with concomitant initiation of transcription (24). Recently, Marqués et al. (17) suggested that alkylbenzoates may have also an indirect role in *Pm* activation. They showed that transcription from the *Pm* promoter, which does not show similarity to the −10/−35 consensus sequence for binding of σ^{70} RNA polymerase, but matches in the −10 region with the consensus sequence of σ^{32} RNA polymerase, is dependent on σ^{32} in exponential-growth-phase *E. coli*. They also demonstrated activation of a σ^{32}-dependent heat shock promoter as an indication of induction of heat shock response by *meta*-toluate. However, since *meta*-toluate did not affect the activation of *Pm* by the XylS CTD and other N-terminal deletion mutants of XylS, we did not observe any indirect effect of *meta*-toluate beyond its well-known role as an allosteric effector of XylS.

As we found, the XylS CTD, which is proposed to be a monomer, can recognize the *Om* binding site and activate transcription. Therefore, occupation of both half-sites, not necessarily the dimeric structure of the activator, appears to be essential for transcriptional activation. The DNase I footprinting results of the XylS CTD-*Om* complexes, as well as the earlier footprinting and methylation interference experiments with XylS (12), demonstrated equal occupation of both *Om* half-sites and suggest that the half-sites have similar affinities for XylS. The proposed dimerization of XylS certainly facilitates recognition of the operator by producing the dimeric protein with increased affinity. That may serve for an explanation of why wt N-XylS, in the presence of *meta*-toluate, stimulates the *Pm* promoter more effectively than ΔN209 when these proteins are expressed at low concentrations (Fig. 4).

ACKNOWLEDGMENTS

We are grateful to Jüri Parik for technical advice and assistance, to Daima Örd for the BPV E2 gene containing Gly-Ala repeats, and to Tanel Tenson, Reet Kurg, Maia Kivisaar, and Aare Abroi for critical reading of the manuscript.

This work was supported by grants 2496 and 2497 from the Estonian Science Foundation, grant HHMI 75195-541301 from the Howard Hughes Medical Institute, and EU grant CIPA-CT94-0154. N.K. was the recipient of an FEMS Young Scientist Award.

REFERENCES

1. **Amábile-Cuevas, C. F., and B. Demple.** 1991. Molecular characterization of the *soxRS* genes of *Escherichia coli*: two genes control a superoxide stress regulon. Nucleic Acids Res. **19**:4479–4484.
2. **Bustos, S. A., and R. F. Schleif.** 1993. Functional domains of the AraC protein. Proc. Natl. Acad. Sci. USA **90**:5638–5642.
3. **Cohen, S. P., H. Hachler, and S. B. Levy.** 1993. Genetic and functional analysis of the multiple antibiotic resistance (*mar*) locus in *Escherichia coli*. J. Bacteriol. **175**:1484–1492.
4. **Eustance, R. J., S. A. Bustos, and R. F. Schleif.** 1994. Reaching out: locating and lengthening the interdomain linker in AraC protein. J. Mol. Biol. **242**: 330–338.
5. **Eustance, R. J., and R. F. Schleif.** 1996. The linker region of AraC protein. J. Bacteriol. **178**:7025–7030.
6. **Gallegos, M. T., S. Marques, and J. L. Ramos.** 1996. Expression of the TOL plasmid *xylS* gene in *Pseudomonas putida* occurs from a σ^{70}-dependent promoter or from σ^{70}- and σ^{54}-dependent tandem promoters according to the compound used for growth. J. Bacteriol. **178**:2356–2361.
7. **Gallegos, M. T., R. Schleif, A. Bairoch, K. Hofmann, and J. L. Ramos.** 1997. AraC/XylS family of transcriptional regulators. Microbiol. Mol. Biol. Rev. **61**:393–410.
8. **Giri, I., and M. Yaniv.** 1988. Structural and mutational analysis of E2 *trans*-activating proteins of papillomaviruses reveals three distinct functional domains. EMBO J. **7**:2823–2829.
9. **González-Pérez, M. M., J. L. Ramos, M. T. Gallegos, and S. Marqués.** 1999. Critical nucleotides in the upstream region of the XylS-dependent TOL *meta*-cleavage pathway operon promoter as deduced from analysis of mutants. J. Biol. Chem. **274**:2286–2290.
10. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1987. Expression of the regulatory gene *xylS* on the TOL plasmid is positively controlled by the *xylR* gene product. Proc. Natl. Acad. Sci. USA **84**:5182–5186.
11. **Inouye, S., A. Nakazawa, and T. Nakazawa.** 1987. Overproduction of the *xylS* gene product and activation of the *xylDLEGF* operon on the TOL plasmid. J. Bacteriol. **169**:3587–3592.
12. **Kaidalu, N., T. Mandel, and M. Ustav.** 1996. TOL plasmid transcription factor XylS binds specifically to the *Pm* operator sequence. Mol. Microbiol. **20**:569–579.
13. **Kessler, B., V. de Lorenzo, and K. N. Timmis.** 1993. Identification of a *cis*-acting sequence within the *Pm* promoter of the TOL plasmid which confers XylS-mediated responsiveness to substituted benzoates. J. Mol. Biol. **230**:699–703.
14. **Kessler, B., K. N. Timmis, and V. de Lorenzo.** 1994. The organization of the *Pm* promoter of the TOL plasmid reflects the structure of its cognate activator protein XylS. Mol. Gen. Genet. **244**:596–605.
15. **Kurg, R., J. Parik, E. Juronen, T. Sedman, A. Abroi, I. Liiv, U. Langel, and M. Ustav.** 1999. Effect of bovine papillomavirus E2 protein-specific monoclonal antibodies on papillomavirus DNA replication. J. Virol. **73**:4670–4677.
16. **Lauble, H., Y. Georgalis, and U. Heinemann.** 1989. Studies on the domain structure of the *Salmonella typhimurium* AraC protein. Eur. J. Biochem. **185**:319–325.
17. **Marqués, S., M. Manzanera, M. M. González-Pérez, M. T. Gallegos, and J. L. Ramos.** 1999. The XylS-dependent Pm promoter is transcribed in vivo by RNA polymerase with σ^{32} or σ^{38} depending on the growth phase. Mol. Microbiol. **31**:1105–1113.
18. **Menon, K. P., and N. L. Lee.** 1990. Activation of *ara* operons by a truncated AraC protein does not require inducer. Proc. Natl. Acad. Sci. USA **87**:3708–3712.
19. **Mermod, N., J. L. Ramos, A. Bairoch, and K. N. Timmis.** 1987. The *xylS* gene positive regulator of TOL plasmid pWWO: identification, sequence analysis and overproduction leading to constitutive expression of *meta* cleavage operon. Mol. Gen. Genet. **207**:349–354.
20. **Michán, C., L. Zhou, M. T. Gallegos, K. N. Timmis, and J. L. Ramos.** 1992. Identification of critical amino-terminal regions of XylS. The positive regulator encoded by the TOL plasmid. J. Biol. Chem. **267**:22897–22901.
21. **Michán, C., B. Kessler, V. de Lorenzo, K. N. Timmis, and J. L. Ramos.** 1992. XylS domain interactions can be deduced from intraallelic dominance in double mutants of *Pseudomonas putida*. Mol. Gen. Genet. **235**:406–412.
22. **Michán, C. M., S. J. Busby, and E. I. Hyde.** 1995. The *Escherichia coli* MelR transcription activator: production of a stable fragment containing the DNA-binding domain. Nucleic Acids Res. **23**:1518–1523.
23. **Miller, J. H.** 1992. A short course in bacterial genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. **Miura, K., S. Inouye, and A. Nakazawa.** 1998. Protein binding in vivo to OP2 promoter of the *Pseudomonas putida* TOL plasmid. Biochem. Mol. Biol. Int. **46**:933–941.
25. **Ramos, J. L., A. Stolz, W. Reineke, and K. N. Timmis.** 1986. Altered effector specificities in regulators of gene expression: TOL plasmid *xylS* mutants and their use to engineer expansion of the range of aromatics degraded by bacteria. Proc. Natl. Acad. Sci. USA **83**:8467–8471.
26. **Ramos, J. L., N. Mermod, and K. N. Timmis.** 1987. Regulatory circuits controlling transcription of TOL plasmid operon encoding *meta*-cleavage pathway for degradation of alkylbenzoates by *Pseudomonas*. Mol. Microbiol. **1**:293–300.
27. **Ramos, J. L., C. Michán, F. Rojo, D. Dwyer, and K. Timmis.** 1990. Signal-regulator interactions. Genetic analysis of the effector binding site of *xylS*, the benzoate-activated positive regulator of *Pseudomonas* TOL plasmid *meta*-cleavage pathway operon. J. Mol. Biol. **211**:373–382.
28. **Ramos, J. L., S. Marqués, and K. N. Timmis.** 1997. Transcriptional control of the *Pseudomonas* TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. Annu. Rev. Microbiol. **51**:341–373.
29. **Rhee, S., R. G. Martin, J. L. Rosner, and D. R. Davies.** 1998. A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. Proc. Natl. Acad. Sci. USA **95**:10413–10418.
30. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a

laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
31. **Saviola, B., R. Seabold, and R. F. Schleif.** 1998. Arm-domain interactions in AraC. J. Mol. Biol. **278**:539–548.
32. **Soisson, S. M., B. MacDougall-Shackleton, R. Schleif, and C. Wolberger.** 1997. Structural basis for ligand-regulated oligomerization of AraC. Science **276**:421–425.
33. **Soisson, S. M., B. MacDougall-Shackleton, R. Schleif, and C. Wolberger.** 1997. The 1.6 Å crystal structure of the AraC sugar-binding and dimerization domain complexed with D-fucose. J. Mol. Biol. **273**:226–237.
34. **Spooner, R. A., M. Bagdasarian, and F. C. Franklin.** 1987. Activation of the *xylDLEGF* promoter of the TOL toluene-xylene degradation pathway by overproduction of the *xylS* regulatory gene product. J. Bacteriol. **169**:3581–3586.
35. **Williams, P. A., and K. Murray.** 1974. Metabolism of benzoate and the methylbenzoates by *Pseudomonas putida* (*arvilla*) mt-2: evidence for the existence of a TOL plasmid. J. Bacteriol. **120**:416–423.
36. **Worsey, M. J., and P. A. Williams.** 1975. Metabolism of toluene and xylenes by *Pseudomonas putida* (*arvilla*) mt-2: evidence for a new function of the TOL plasmid. J. Bacteriol. **124**:7–13.
37. **Zhou, L. M., K. N. Timmis, and J. L. Ramos.** 1990. Mutations leading to constitutive expression from the TOL plasmid meta-cleavage pathway operon are located at the C-terminal end of the positive regulator protein XylS. J. Bacteriol. **172**:3707–3710.

CURRICULUM VITAE

NIILO KALDALU

Date and place of birth:	November 19, 1963 in Valga, Estonia
Citizenship:	Estonian
Family status:	married, two children
Work address:	Technology Center of Tartu University Riia Str. 23 51010 Tartu, Estonia
Phone:	+372 7 375 040
Fax:	+372 7 420 286
E-mail	*nkaldalu@ebc.ee*
Home address:	Kalda tee 42–46, 50707 Tartu

Education and professional employment

1971–1982	First Secondary School of Valga
1982–1987	University of Tartu, Faculty of Biology and Geography Graduated as biologist-geneticist
1987–1994	University of Tartu, Department of Genetics, junior research associate
1994–1999	University of Tartu, Department of Virology and Microbiology, Ph.D. student
1996	receiver of FEMS Young Scientist Award, working 3 months at Spanish National Center for Biotechnology, CSIC, Madrid
1999–present	Technology Center of Tartu University, senior laboratory assistant

Scientific work

I have been involved in studies of bacterial biodegradation since my work at Department of Genetics. During my work at Department of Virology and Technology Center I have been involved in studies of bacterial transcriptional regulation and development of expression vectors. The published results of these studies have concomitantly developed into the present thesis.

CURRICULUM VITAE

NIILO KALDALU

Sünniaeg ja -koht:	19. november 1963, Valgas
Kodakondsus:	Eesti
Perekonnaseis:	abielus, kaks last
Aadress tööl:	Tartu Ülikooli Tehnoloogiakeskus Riia 23, 51010 Tartu
Tel.:	(27) 375 040
Faks:	(27) 420 286
E-post:	*nkaldalu@ebc.ee*
Kodune aadress:	Kalda tee 42–46, 50707 Tartu

Haridus ja erialane teenistuskäik

1971–1982	Valga 1. Keskkool
1982–1987	Tartu Ülikool, bioloogia-geograafiateaduskond, geneetika eriala
1987–1994	Tartu Ülikooli geneetika kateedri nooremteadur
1994–1999	Tartu Ülikooli viroloogia ja mikrobioloogia õppetooli doktorant
1996	FEMS-i noore teadlase stipendium, kolm kuud Hispaania Rahvuslikus Biotehnoloogia Keskuses (CSIC) Madridis
Alates 1999	Tartu Ülikooli Tehnoloogiakeskuse vanemlaborant

Teadustegevus

Toksiliste ainete lagundamist mikroobide poolt hakkasin uurima, töötades geneetika õppetoolis. TÜ viroloogia ja mikrobioloogia õppetoolis ja Tehnoloogiakeskuses osalesin bakterite transkriptsiooni regulatsiooni uurimises ja uute ekspressioonivektorite väljatöötamises. Sama tööd jätkan praegugi ja käesolev väitekiri on kokku pandud selle tulemuste põhjal.

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand**. The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.

19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.
20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous crassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.

40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.
41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.
42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.

ISSN 1024–6479
ISBN 9985–56–482–0