Critical Nucleotides in the Upstream Region of the XylS-dependent TOL meta-Cleavage Pathway Operon Promoter as Deduced from Analysis of Mutants*

M. Mar González-Pérez, Juan L. Ramos‡, María-Trinidad Gallegos, and Silvia Marqués

From the Department of Biochemistry and Molecular Cellular Biology of Plants, Estación Experimental del Zaidín, Consejo Superior de Investigaciones Científicas, Apartado de Correos 419, E-18008 Granada, Spain

The Pm promoter, dependent on TOL plasmid XylS regulator, which is activated by benzoate effectors, drives transcription of the meta-cleavage pathway for the metabolism of alkylbenzoates. This promoter is unique in that in vivo transcription is mediated by RNA-polymerase with different sigma factors. In vivo footprinting analysis shows that XylS interacts with nucleotides in the −40 to −70 region. In vivo and in vitro methylation of Pm shows extensive methylation of T at position −42 in the bottom strand, suggesting that it represents a key distortion point that may favor XylS/RNA polymerase interactions. Methylation of T−42 was highest in cells bearing XylS and in the presence of an effector. Gs in the −47 to −61 region appeared to be more protected in cells harboring XylS in the presence than in the absence of the effector. Almost 100 mutants in the Pm region between −41 and −78 were generated; transcriptional analysis of these mutants defined the XylS target as two direct repeats with the sequence TGCAN₆GGNCA. These motifs cover the −70 to −56 and the −49 to −35 regions. Single point mutations revealed that nucleotides located at −49 to −46 and at −59, −60, −62, and −70 are the most critical for appropriate XylS-Pm interactions.

The TOL plasmid pWW0 of Pseudomonas putida specifies a meta-cleavage pathway for the oxidative catabolism of benzoate and toluates. Genes encoding the TOL meta-cleavage pathway are grouped in a single operon, the expression of which is positively regulated at the level of transcription by the xylS gene product, which is activated by benzoate effectors (1–4). Stimulation of transcription from the Pm promoter requires a DNA sequence extending to about 80 bp1 upstream of the transcription initiation point (5–7). On the basis of genetic data, two regions can be distinguished in the architecture of the transcription initiation point. The oligonucleotide 5′-GGTCTAAGAAACCATTATTATCAG-3′ was complementary to the noncoding strand upstream from the Pm promoter region. The first C at the 5′-end was located 218 bp from the +1 of the transcription initiation point. The oligonucleotide 5′-GGGTCGGTTGAAACATCTGCCGCGTCC-3′ was complementary to the coding strand downstream from the Pm promoter. The first G at the 5′-end was located 124 bp from the +1 of the start of the transcript. Extension products were separated by electrophoresis on urea-polyacrylamide sequencing gels.

MATERIALS AND METHODS

Bacterial Strains, Culture Medium, and Plasmids—Escherichia coli MC4100 was grown at 30 °C in Luria-Bertani medium supplemented, when required, with 100 μg/ml ampicillin, 25 μg/ml kanamycin, or 50 μg/ml streptomycin.

The plasmids used in this study, and previously constructed were: pERD103, which is an IncQ plasmid encoding kanamycin resistance (7); pJLR100, which is a pEMBL9 derivative bearing the Pm promoter cloned between the EcoRI and HindIII sites (3); pMD1405, which carries a promoterless lacZ gene and encodes resistance to ampicillin; and pJLR107, which is a pMD1405 derivative bearing the Pm promoter in front of lacZ (3).

DNA Techniques—DNA preparation, digestion with restriction enyzmes, analysis by agarose gel electrophoresis, isolation of DNA fragments, ligations, transformations, and sequencing reactions were done according to standard procedures (10).

The oligonucleotide 5′-CGTCTAAGAAACCATTATTATCAG-3′ was complementary to the noncoding strand upstream from the Pm promoter region. The first C at the 5′-end was located 218 bp from the +1 of the transcription initiation point. The oligonucleotide 5′-GGGTCGGTTGAAACATCTGCCGCGTCC-3′ was complementary to the coding strand downstream from the Pm promoter. The first G at the 5′-end was located 124 bp from the +1 of the start of the transcript. For primer extension with Taq-DNA-polymerase, about 2 × 10⁶ cpm of the corresponding oligonucleotide end-labeled with ³²P was used. The extension products were separated by electrophoresis on urea-polyacrylamide sequencing gels.

Construction of Pm Mutant Promoters by Polymerase Chain Reac-

*This study was supported by European Commission Grant BIO-CT97-2313 and Comisión Interministerial de Ciencia y Tecnología of Spain Grant BIO97-0641. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡To whom correspondence should be addressed. Tel.: 34-955-121011; Fax: 34-955-129600; E-mail: jlramos@eez.csic.es.

1The abbreviations used are: bp, base pair(s); HTH, helix-turn-helix; 3MB, 3-methylbenzoate; Pm, promoter for the TOL meta-pathway.
The proposed +1 nucleotide is based on reverse transcriptase extension of primers that overlap the mRNA synthesized from this promoter (2, 31). The two arrows over the −70 to −35 region are nucleotides important for XylS recognition in Pm, as proposed by Kessler et al. (6); the two dashed arrows at the bottom represent the motif proposed by Gallegos et al. (5).

\[
AGTCAGCCTGTCAAGAAGCGGATACAGGATCAGAAAATGGCTATCTGAGACAGGCCTATAGCCTATAGC
\]

**Fig. 1. Nucleotide sequence of the Pm promoter.** The proposed +1 nucleotide is based on reverse transcriptase extension of primers that overlap the mRNA synthesized from this promoter (2, 31). The two arrows over the −70 to −35 region are nucleotides important for XylS recognition in Pm, as proposed by Kessler et al. (6); the two dashed arrows at the bottom represent the motif proposed by Gallegos et al. (5).

**Fig. 2. In vivo footprinting of XylS.** E. coli MC4100 cells bearing only plasmid pJLR107 (Pm::lacZ) or this plasmid plus pERD103 (xylS) were grown in the presence and in the absence of 3MB, and then exposed to dimethyl sulfate. The panel on the left shows the region of the top strand between −77 and −35, and the panel on the right shows the region of the top strand between −40 and +3 with respect to the main transcription initiation point. Lane V is the same DNA methylated in vitro. The G bases at −77, −68, −61, −56, −47, −39, −37, −24 and −3 are indicated, and the position corresponding to T at −42 is also shown.
Tandem Repeats as XylS Targets

Table I
Block scanning mutagenesis of the upstream region of the Pm promoter and its effect on its transcriptional activity

E. coli MC14100 bearing the indicated Pm*: lacZ fusion and xylS in p的利益103 were grown on LB medium with 3MB as described under “Materials and Methods.” In the absence of 3MB, basal activity was 50–150 Miller units. The data correspond to induced levels and are

| Promoter | β-Galactosidase activity | Activity of wild-type |
|----------|-------------------------|-----------------------|
| Pm wild-type | 8900 | 100 |
| Pm 250 (44–41 AAAA → CCCT) | 3850 | 43 |
| Pm 251 (45–42 AAAA → CCAC) | 5600 | 62 |
| Pm 252 (45–42 ACAA → CCAC) | 5550 | 62 |
| Pm 206 (49–46 TGCA → GTCA) | 1700 | 19 |
| Pm 207 (49–46 TGCA → GAGG) | 550 | 6 |
| Pm 208 (49–46 TGCA → GTGT) | 350 | 4 |
| Pm 209 (49–46 TGCA → ACAG) | 500 | 6 |
| Pm 253 (49–46 TGCA → GGCC) | 1200 | 13 |
| Pm 254 (49–46 TGCA → GGCC) | 600 | 7 |
| Pm 255 (49–46 TGCA → TCGA) | 500 | 5 |
| Pm 257 (53–50 GAGG → CACC) | 8130 | 91 |
| Pm 212 (57–54 GACCA → TACG) | 2100 | 24 |
| Pm 213 (57–54 TACG → GTTT) | 800 | 89 |
| Pm 259 (61–58 CGGA → CCAG) | 3500 | 39 |
| Pm 260 (61–58 CGGA → TCTC) | 2700 | 30 |
| Pm 261 (61–58 CGGA → CTCA) | 4500 | 51 |
| Pm 262 (66–63 GAAA → GAGT) | 4400 | 50 |
| Pm 267 (66–63 GAAA → CTAA) | 6500 | 73 |
| Pm 268 (66–63 GAAA → GGAA) | 9900 | 111 |
| Pm 222 (70–67 TGCA → AGCT) | 5600 | 63 |
| Pm 219 (70–67 TGCA → AACAG) | 2200 | 25 |
| Pm 220 (70–67 TGCA → AGCA) | 2900 | 33 |
| Pm 233 (74–71 GCCT → TCCG) | 8900 | 100 |
| Pm 226 (78–75 TCGA → GATT) | 7400 | 83 |
| Pm 245 (49–46, TGCA → AGGA; 57–54 TACA → TTGG) | 60 | 1 |
| Pm 244 (57–54 TACA → GTTT; −67−70, TGCA → ACCT) | 1500 | 16 |

The mutations that resulted in the largest reduction in transcription, i.e. a decrease equal to or greater than 80% of the wild-type activity, were any random substitution of the TGCA sequence between −46/−49 (Table I). This suggests that these nucleotides are critical for XylS-dependent transcription activation of Pm. The substitution by random sequences of the AAAA sequence located at −41/−45, the TACA sequence between −54/−57, the CGGA sequence between −58/−61, and the TGCA sequence between −67/−70 resulted in a significant decrease in XylS-dependent transcription activation of Pm. The activity of most of the mutant Pm promoters at these locations ranged from approximately 25 to about 65% of the activity of the wild-type promoter (Table I), although certain substitutions had little effect. These results suggest that these sets of bases are less critical than those at the −46/−49 region; however, they may play a direct role in the recognition of the Pm DNA sequences by XylS, or they may contribute to the overall affinity for Pm. We cannot rule out other effects.

The third group of mutations, i.e. those that had no effect (or little effect) on transcription from Pm, were found to correspond to the locations of −50/−53, −62/−65, −63/−66, −71/−74, and −75/−78.

We also investigated whether the combination of different blocks of mutations had a synergistic effect on XylS-dependent transcription activation from the mutant promoters. The combination of a block of mutations in −46/−49 (TGCA → AGGA) with a block of mutations at −54/−57 (TACA → TTGG) resulted in a mutant Pm promoter (Pm 245) that had no activity at all (Table I).

At the −54/−57 block some substitutions had little effect on XylS-dependent transcription from Pm (i.e. Pm 213, 11% reduction), whereas other substitutions at the same block had a clear effect (i.e. Pm 212, 78% reduction). In Pm 244, we combined the block of mutations in Pm 213 with a mutant block that had a moderate effect on the level of expression from Pm (i.e. −67/−70 (TGCA → AGCT)). This combination had cumulative effects (Table I). These results confirmed the essential role of the nucleotides at these positions in XylS-dependent transcription activation from Pm.

The transcription initiation point of the mRNA generated from a number of the above Pm mutant promoters was the same that that determined for the wild-type promoter (not shown). This suggests that the mutations analyzed affected the strength of transcription from the mutant promoters.

Single Point Mutations within the Set of Blocks That Showed or Did Not Show Any Effect on Transcription Activation from Pm—The above series of assays suggested critical, less critical, and irrelevant blocks for the XylS-dependent transcription activation from Pm. We expected that single mutations in irrelevant sets of sequences would have no effect at all on transcriptional activity from Pm, whereas single mutations in the critical and important blocks of sequences would have an effect. A number of bases at the noncritical region were selected to introduce single point mutations: A → G, A → C, C → G, G → T, A → C, C → G, and C → T. These mutations, as expected, had little or no effect (<20%) on XylS-dependent transcription activation from Pm (not shown).

Of the less critical boxes (−41/−44, −54/−57, −58/−61, and −67/−70), some of the substitutions (for example, A → T, A → G, A → G, A → G, A → T, G → C, G → T) did not significantly affect activity from Pm (Table II); others had an intermediate effect, reducing the induced XylS-dependent activation of Pm by 20–50%. This effect was observed for the changes G → A, C → G, A → C, A → T, and C → T (Table II). Within these sequences, the change T → G resulted in loss of almost 90% of the activity (Table II).

In the −58/−61 box, the G → C, G → T, and G → A
changes had a significant effect on transcription, as shown by the finding that β-galactosidase activity was less than 20% of that seen with the wild-type Pm promoter. A surprising finding was that the G→C change resulted in a mutant Pm promoter that lacked activity.

In the critical –46/–49 set of bases, single bp substitutions had a significant effect on activity. The changes A→T, C→G→T, G→T, G→G→C, and T→G resulted in a 60–85% decrease in activity. However, the change G→A had little or no effect on transcriptional activity (Table II).

Role of the Bases at –54/–57—When we compared the critical sequences proposed by Kessler et al. (6) and Gallegos et al. (5), we found that they had the two TGCA submotifs at –46/–49 and –67/–70 in common (Fig. 1). The hypothesis of Gallegos et al. (5) suggested that the TACA submotif between –54/–57 was critical for transcription activation from Pm, whereas Kessler et al. (6) suggested that only the –57/–56 nucleotides were of importance (Fig. 1). Our results with single point mutations in the –54 to –57 region showed some effect in some positions, but in no case was the effect large enough to fully impede transcription (Table II). To elucidate the possible role of these four nucleotides, we generated the set of changes (30) and galactosidase activity was less than 20% of that seen with the wild-type Pm promoter, between –70 and –56 and between –49 and –35 (Fig. 1). In favor of this proposal is the observation that a tagged XylS-protein immunoabsorbed onto glass beads produced footprints in vitro, which showed protection of the Gs within the above direct repeat at –48, –59, –60, and –69 (22). This is in agreement with our in vivo results.

The AraC protein—the best characterized regulator of the family—stimulates transcription as a dimer (23–25). The consensus sequence for AraC-activable promoters is a direct repetition of the TAGCN7TCCATA motif; each AraC monomer recognizes one of the direct repeats (26–29).

At the C-terminal end, the regulators of the AraC/XylS family show a highly conserved stretch of about 100 amino acids that seems to be involved in DNA binding and probably in interactions with RNA polymerase (11). One characteristic of members of this family is that they exhibit two possible HTH DNA binding motifs (located at 228–251 and 281–305 in XylS and at 198–217 and 246–264 in AraC). Brunelle and Schleif (15) analyzed these possible HTH motifs with substitutions of several amino acids that should contact DNA in an HTH structure and found evidence for the first one. Niland et al. (30), using synthetic oligonucleotides, systematically substituted the bases in the AraC recognition sequence and did gel retardation assays with mutant AraC in each of the possible HTH elements. They showed that the mutant AraC in each of the HTH motifs exhibited altered DNA binding properties. On the basis of their results, these authors proposed that each AraC monomer binds the 5’-TAGC submotif with one of the HTH motifs, and the 3’-TCCATA submotif with the second HTH.

If XylS contacts DNA via the two possible HTH elements, all mutant Pm promoters generated in different laboratories can be explained by the following model (Fig. 3): the XylS protein recognizes two submotifs in Pm, TGCA and GGNTA, which are separated from each other by six nucleotides. Each submotif is recognized by the recognition helix of one of the HTH elements of XylS. For the wild-type protein, recognition of direct sequences leads to the formation of a dimer. One monomer recognizes the upstream motif (from –70 to –56) with the two HTH DNA binding elements; the second monomer recognizes mainly the TGCA submotif and interacts with the downstream sequence, where it may compete for binding with the RNA-polymerase (Fig. 3). Mutations at the –46/–49 TGCA submotif result in mutants in which the capacity to activate transcrip-
tion is impaired because they cannot be contacted properly by XylS. Failure of one of the XylS monomers to interact with this submotif prevents dimer formation and leads to a nonactivable mutant Pm promoter. Pm mutants at the distal TGCA (−67/−70) submotif can be activated weakly by the wild-type XylS protein as a result of the formation of an unstable dimer; however, they can still be induced to a high level of activation by mutant XylS proteins with higher affinity for target sequences than the wild-type regulator (5). This is because one of the XylS mutant monomers binds to the downstream motif (−49/−35); and because of the mutation in the XylS, the second monomer is still able to interact well with the GGNTA submotif, and this suffices for dimer formation. The transcriptional activity of mutant Pm promoters with altered GGNTA at −60/−56 sequences is seriously impaired, because failure of one of the monomers to bind correctly prevents dimer stabilization at Pm.

REFERENCES

1. Franklin, F. C. H., Bagdasarian, M., Bagdasarian, M. M., and Timmis, K. N. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 7458–7462
2. Inouye, S., Nakazawa, A., and Nakazawa, T. (1981) J. Bacteriol. 148, 413–418
3. Ramos, J. L., Stolz, A., Reineke, W., and Timmis, K. N. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 8467–8471
4. Ramos, J. L., Marqués, S., and Timmis, K. N. (1997) Annu. Rev. Microbiol. 51, 341–374
5. Gallegos, M. T., Marqués, S., and Ramos, J. L. (1996) J. Bacteriol. 178, 6427–6434
6. Kessler, B., de Lorenzo, V., and Timmis, K. N. (1993) J. Mol. Biol. 230, 699–703
7. Ramos, J. L., Mermod, N., and Timmis, K. N. (1987) Mol. Microbiol. 1, 297–300
8. Marqués, S., Gallegos, M. T., and Ramos, J. L. (1995) Mol. Microbiol. 18, 851–857
9. Marqués, S., Manzanera, M., González-Pérez, M. M., Gallegos, M. T., and Ramos, J. L. (1999) Mol. Microbiol., in press
10. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (ed.) (1991) Current Protocols in Molecular Biology, John Wiley & Sons, New York
11. Gallegos, M. T., Schleif, R., Bairach, A., Hofmann, K., and Ramos, J. L. (1997) Microbiol. Mol. Biol. Rev. 61, 393–410
12. Inouye, S., Nakazawa, A., and Nakazawa, T. (1986) Gene 44, 235–242
13. Mermod, N., Ramos, J. L., Bairach, A., and Timmis, K. N. (1987) Mol. Gen. Genet. 207, 349–354
14. Spooner, R. A., Bagdasarian, M., and Franklin, F. C. H. (1987) J. Bacteriol. 169, 3581–3586
15. Brunelle, A., and Schleif, R. (1989) J. Mol. Biol. 209, 607–622
16. Caswell, R., Williams, J., Lyddiatt, A., and Busby, S. (1992) Biochem. J. 287, 493–499
17. Egan, S. M., and Schleif, R. F. (1993) J. Mol. Biol. 234, 87–98
18. Eustance, R. J., Bustos, S. A., and Schleif, R. (1994) J. Mol. Biol. 242, 330–338
19. Fawcett, W. P., and Wolf, R. E., Jr. (1995) J. Bacteriol. 177, 1742–1750
20. Li, Z., and Dempo, B. (1996) Mol. Microbiol. 20, 937–945
21. Lobell, B. R., and Schleif, R. F. (1996) Science 270, 528–532
22. Kaldalu, N., Mandel, T., and Ustav, M. (1996) Mol. Microbiol. 20, 569–579
23. Bustos, S. A., and Schleif, R. F. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 5638–5642
24. Menon, K. P., and Lee, N. L. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 3708–3712
25. Schleif, R. (1992) in Transcriptional Regulation (McKnight, S. L., and Yamamoto, K. R., eds) Vol. 2, pp. 643–665, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
26. Brunelle, A., and Schleif, R. (1999) J. Mol. Biol. 209, 607–622
27. Carra, J. H., and Schleif, R. F. (1993) EMBO J. 12, 35–44
28. Lobell, B. R., and Schleif, R. F. (1991) J. Mol. Biol. 218, 45–54
29. Reeder, T., and Schleif, R. (1993) J. Mol. Biol. 231, 205–218
30. Nishida, P., Hube, R., and Muller-Hill, B. (1996) J. Mol. Biol. 264, 667–674
31. Marqués, S., Ramos, J. L., and Timmis, K. N. (1993) Biochim. Biophys. Acta 1216, 227–236
