JOURNAL OF BACTERIOLOGY, July 1996, p. 3967–3970 Vol. 178, No. 13
0021-9193/96/$04.00+0

Use of PCR To Isolate Genes Encoding σ^{54}-Dependent Activators from Diverse Bacteria

R. ILENE KAUFMAN AND B. TRACY NIXON*

Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania 16802

Received 16 February 1996/Accepted 29 April 1996

Degenerate PCR probes were used to amplify gene fragments encoding the catalytic domain of σ^{54}-dependent transcription activators. The procedure should be widely applicable, as it recovered both known and novel gene fragments: 5 from *Rhizobium meliloti*, 13 from *Myxococcus xanthus*, and 3 from *Bacillus subtilis*. No fragments were obtained from *Synechococcus* sp. strain PCC 7002 or *Saccharomyces cerevisiae*.

The prokaryotic RNA polymerase derives its promoter specificity from its sigma factor. One group of sigma factors, known as the σ^N or σ^{54} family, differs from all other sigma factors. First, members of this family do not share amino acid sequence similarity with other sigma proteins. Second, RNA polymerase σ^{54} ($E\sigma^{54}$) forms stable closed complexes that do not initiate transcription unless assisted by a transcriptional activator. These activators bind to upstream activation sequences and interact with $E\sigma^{54}$-promoter closed complexes via a DNA loop (for reviews see references 8 to 10; also reference 14 and references therein). This ability to act at a distance has earned them the name "enhancer-binding proteins," drawing attention to similarities between $E\sigma^{54}$ and eukaryotic RNA polymerases. Most of the known σ^{54}-dependent activators appear to be nonessential for growth in rich media, being involved instead in various accessory metabolic functions. Some limited examples include NtrC/NR_1, NifA, DctD, and XylR, which are involved in nitrogen assimilation, nitrogen fixation, dicarboxylate transport, and degradation of aromatic hydrocarbons, respectively.

In 1993, Morett and Segovia reviewed the mechanism of action and phylogenetic relationships of the functional domains in σ^{54}-dependent transcriptional activators (10). At that time, there were 30 sequences of 13 different σ^{54}-dependent activators from 20 bacterial species. These genes had been identified through traditional genetic screens and by hybridization experiments. A current search of the databases shows that 26 additional σ^{54}-dependent activators have since been identified in several laboratories by similar means. Here we report using degenerate oligonucleotide primers and PCR to identify gene fragments from 17 novel σ^{54}-dependent activators from only three eubacterial genomes, those of *Rhizobium meliloti*, *Myxococcus xanthus*, and *Bacillus subtilis*. These gene fragments contain ~400 bp that are predicted to encode a region essential for σ^{54}-dependent activator function. In organisms for which 400 bp is sufficient to mediate homologous recombination, these fragments might be used to construct site-directed insertion mutations. This strategy, together with techniques for identifying the genes that are being regulated, can be used to rapidly investigate the role of σ^{54}-dependent activators in the regulation of diverse bacterial activities. While this approach is diametrically opposed to the conventional wisdom of requiring a well-defined phenomenon before embarking on genetic and molecular biological investigations, its broad applicability and rapid success justify its use. Indeed, this type of cloning strategy has identified many genes, including novel helicases (1), serine proteases (12), and tyrosine kinases (16).

The family of σ^{54}-dependent activators is defined by a highly conserved domain (2, 5). It has been referred to as the catalytic domain, as it must hydrolyze ATP or GTP to stimulate transcription by $E\sigma^{54}$ (11, 15). When DNA encoding this σ^{54}-dependent catalytic domain is used to probe genomic DNA, several gene fragments are recognized by the probe; however, many of these fragments simply encode nucleotide binding sites, not the desired catalytic domain. In addition to this common nucleotide binding site, the catalytic domains of σ^{54}-dependent activators share two highly conserved motifs spanned by ~150 residues. These latter motifs are unique to σ^{54}-dependent activators. We reasoned that it might be possible to use highly degenerate oligonucleotide primers to amplify ~450-bp DNA fragments encoding this catalytic domain, which could then be used to generate null mutants by homologous recombination. By using *Taq* DNA polymerase, PCR was carried out on chromosomal DNA from *R. meliloti*, *M. xanthus*, *B. subtilis*, *Synechococcus* sp. strain PCC 7002, and *Saccharomyces cerevisiae* (in 100 μl of 10 mM Tris-HCl [pH 8.3]–1.5 mM $MgCl_2$–50 mM KCl–15 mM $(NH_4)_2SO_4$–0.2 mM deoxynucleoside triphosphates–1 μM primers–0.8 to 1.0 μg of DNA–2.5 U of *Taq* DNA polymerase). One primer was 5′-TCAAGCTTACRTTNCCNGGVMA-3′, and the second was 5′-GAGAATTCGARYTNTTYGGNCAYG-3′ (M = A or C; N = G, A, T, or C; R = A or G; V = C, A, or G; Y = C or T). Each primer has two additional nucleotides and either a *Hin*dIII or an *Eco*RI site to facilitate cloning, plus a sequence that encodes one of the highly conserved amino acid sequences (W/F)PGNV or ELFGH(V/A/D/E/G). Before the addition of *Taq* DNA polymerase, the reaction mixtures were overlaid with 100 μl of Nujol mineral oil (Plough, Inc.), and incubated at 94°C for 10 min to denature the DNA. Table 1 shows the typical thermal cycling profile.

The ~450-bp PCR products were isolated from 1.6% agarose gels. Initially, these fragments were cloned into M13mp19 after digestion with *Hin*dIII and *Eco*RI. However, neither of these enzymes was able to digest the fragments well enough to allow for efficient cloning. Therefore, in subsequent experiments, the PCR products were first ligated into a T-tailed vector (pCRII, Invitrogen). Subsequent transformants were pooled and used to prepare plasmid DNA that was digested

* Corresponding author. Mailing address: Department of Biochemistry and Molecular Biology, Pennsylvania State University, 327 S. Frear Lab, University Park, PA 16802. Phone: (814) 865-3679. Fax: (814) 863-7024. Electronic mail address: btn1@psu.edu.

TABLE 1. PCR conditions

Cycle	Temp and/or time			
	Denaturation	Annealing	Ramp time	Extension
1	94°C for 1 min	37°C for 1 min	1 min	72°C for 2 min
2	94°C for 1 min	41°C for 1 min	1 min	72°C for 2 min
3	94°C for 1 min	45°C for 1 min	1 min	72°C for 2 min
4–34	94°C for 1 min	50°C for 1 min	0 min	72°C for 2 min
35	94°C for 1 min	50°C for 1 min	0 min	72°C for 10 min

with *Eco*RI, and the resulting fragments were isolated and cloned into M13mp19 for further analysis. Phage from individual plaques were subjected to complementation assays (C tests) to sort them into groups containing unique inserts. (Ten-microliter portions from each of two phage supernatants were combined along with 3 μl of 10× sample buffer [20% Ficoll 400; 0.1 M Na_2 EDTA, pH 8; 2% sodium dodecyl sulfate; and 0.25% bromphenol blue], incubated at 65°C for 30 min, and then electrophoresed through an agarose gel (0.7%) containing ethidium bromide until the bromphenol blue migrated 5 cm). This step was rapid and greatly reduced the amount of sequencing required to identify novel gene fragments.

Experiments using DNA from the gram-negative bacterium *R. meliloti* yielded five different C-test groups of ~450-bp PCR fragments. Of the 300 M13 clones obtained by this method, 50% were fragments of *ntrC*, 30% were fragments of *nifA*, 6% were fragments of *dctD*, and 14% were fragments of two novel genes (19_Rme and 4_Rme [Fig. 1]). Of 60 clones selected from experiments with DNA from the gram-negative bacterium *M. xanthus*, we obtained 13 different groups of ~450-bp PCR fragments (Fig. 1). Only one of these, 15_Mxa, was concurrently found by other means (17). DNA from the gram-positive bacterium *B. subtilis* yielded 44 M13 clones, one of which contained a fragment of *levR*. The remaining cloned fragments were from two novel genes (Fig. 1); one (70_Bsu) occurred 40 times, and one (81_Bsu) occurred 3 times. The frequency with which a given gene was amplified was clearly biased in these PCR products. Such bias might result from the 3′-most G base in primer 2 being inappropriate for annealing, as it is now clear that 11 of the 75 known genes (including *levR*) do not encode a V, E, A, D, or G in that position. Other reaction conditions relevant to successful annealing of a given oligonucleotide are also likely to have contributed to this bias. No gene fragments having the appropriate size or containing the appropriate conserved sequences were amplified from genomic DNA of either the cyanobacterium *Synechococcus* sp. strain PCC 7002 or the yeast *Saccharomyces cerevisiae*. These two failures might simply reflect a limitation of the approach, as the homologous regions used to design the primers might

FIG. 1. Deduced amino acids for catalytic domains of new σ^{54}-dependent activators. Sequences for 73 catalytic domains were aligned by using ClustalW with the full scoring method. Shown are sequences from the genomes of *R. meliloti* (Rme), *M. xanthus* (Mxa), and *B. subtilis* (Bsu), together with PilR and potential AlgB homologs from *Pseudomonas aeruginosa* (Pae).

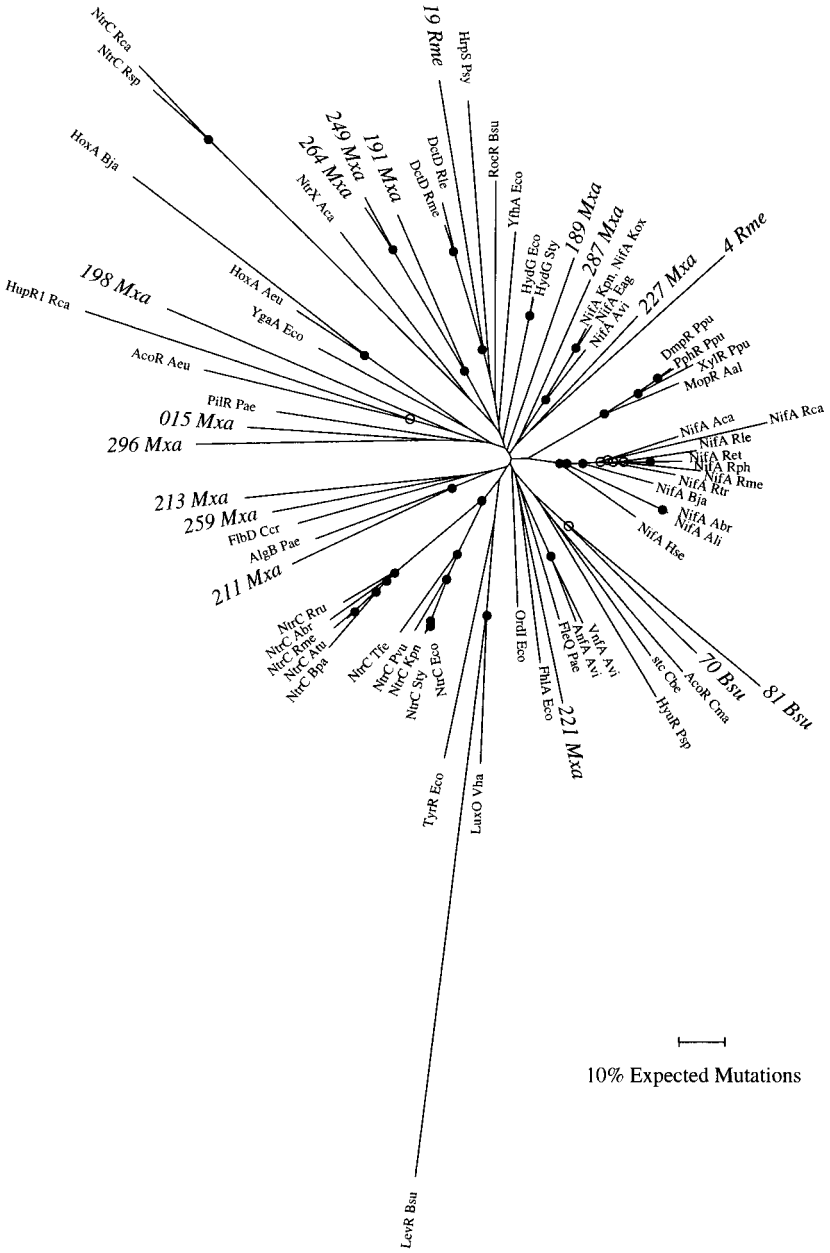

FIG. 2. Phylogenetic map of σ^{54}-dependent catalytic domains. The complete alignment, partially shown in Fig. 1, was used with the PHYLIP programs Protdist, Neighbor, and Drawtree (3) to make a distance-based, neighbor-joined tree. Distances were determined by using the Dayhoff PAM 001 matrix and ranged from 0.0065 to 2.8282, where 0.1000 denotes an expected 10% change (see scale bar). Groups were further analyzed with the PHYLIP program Seqboot to generate 500 resampled data sets that were each analyzed by Protdist and Neighbor. The program Consense was then used to obtain the frequencies with which all groups occur in the 500 bootstrap trees; some of this information was included in the tree (solid dot, >50% frequency; open dot, 40 to 50% frequency; no dot, <40% frequency). The newly discovered genes are in a larger, italic font. Abbreviations: Abr, *Azospirillum brasilense*; Aal, *Acinetobacter calcoaceticus*; Aca, *Azorhizobium caulinodans*; Aeu, *Alcaligenes eutrophus*; Ali, *Azospirillum lipoferum*; Atu, *Agrobacterium tumefaciens*; Avi, *Azotobacter vinelandii*; Bja, *Bradyrhizobium japonicum*; Bpa, *Bradyrhizobium parasponia*; Bsu, *Bacillus subtilis*; Cbe, *Clostridium beijerinckii*; Ccr, *Caulobacter crescentus*; Cma, *Clostridium magnum*; Eag, *Enterobacter agglomerans*; Eco, *Escherichia coli*; Hse, *Herbaspirillum seropedicae*; Kox, *Klebsiella oxytoca*; Kpn, *Klebsiella pneumoniae*; Pae, *Pseudomonas aeruginosa*; Ppu, *Pseudomonas putida*; Psp, *Pseudomonas* sp. strain NS671; Psy, *Pseudomonas syringae* pv. *phaseolicola*; Pvu, *Proteus vulgaris*; Rca, *Rhodobacter capsulatus*; Ret, *Rhizobium etli*; Rle, *Rhizobium leguminosarum*; Rme, *Rhizobium meliloti*; Rph, *Rhizobium phaseoli*; Rru, *Rhodospirillum rubrum*; Rsp, *Rhodobacter sphaeroides*; Rtr, *Rhizobium trifolii*; Sty, *Salmonella typhimurium*; Tfe, *Thiobacillus ferrooxidans*; Vha, *Vibrio harveyi.*

not be adequately conserved. However, when this cyanobacterial genome was probed at low stringency for hybridization to DNA encoding the catalytic domain of NtrC from *Bradyrhizobium parasponia*, the strongest signal observed was subsequently shown to correspond only to Walker ATP binding motifs present in both the probe and a cyanobacterial homolog of the *B. subtilis oppD* gene (13).

A distance-based phylogenetic tree was created to see if any functions could be predicted for the new, putative σ^{54}-dependent activators (Fig. 2). The significance of groupings within the tree topology was further analyzed by bootstrapping 500 resampled data sets. The tree topology and the bootstrap analysis suggest that gene 19 of *R. meliloti* might be a homolog of *dctD*, but no predictions about potential functions for gene 4

are supported. Homologous recombination has been used to make null mutants of both of the newly discovered *R. meliloti* genes, demonstrating that together they complement *dctD* mutants to sustain nitrogen fixation but not free-living utilization of 4-carbon dicarboxylates (7). Gene 15 of *M. xanthus* has been found to be a homolog of the *Pseudomonas* gene *pilR*, controlling pilin synthesis (17). While this might be suggested from the tree topology, the bootstrap test does not provide support for such a hypothesis. The tree and bootstrap frequencies do suggest that gene 211 of *M. xanthus* is a homolog of the *Pseudomonas* gene *algB*, which participates in controlling alginate biosynthesis. Otherwise, the *M. xanthus* genes encode several new groups of putative σ^{54}-dependent activators whose functions remain to be characterized, though early studies have shown genes 191, 211, 213, 259, and 287 to be important in fruiting-body development (4). The two new genes from *B. subtilis* do not group with the characterized gene *levR* or *rocR*; however, they show some tendency to group with each other. In summary, these early results clearly demonstrate that this cloning approach can be widely applied to identify potential σ^{54}-dependent regulator genes. The predictive power of the phylogenetic analysis is likely to improve as the target genes that are controlled by these and other σ^{54}-dependent activators become known.

This work was supported by U.S. Department of Agriculture grant 91-37305-6804 and award GM40404 from the Public Health Service to B.T.N.

ADDENDUM IN PROOF

Gene 70_BSU is required for isoleucine utilization (R. Gardan and M. Debarbouille, personal communication).

REFERENCES

1. **Chang, T.-H., J. Arenas, and J. Abelson.** 1990. Identification of five putative yeast RNA helicase genes. Proc. Natl. Acad. Sci. USA **87:**1571–1575.
2. **Drummond, M., P. Whitty, and J. Wootton.** 1986. Sequence and domain relationships of *ntrC* and *nifA* from *Klebsiella pneumoniae*: homologies to other regulatory proteins. EMBO J. **5:**441–447.
3. **Felsenstein, J. (http://evolution.genetics.washington.edu/phylip.html [Internet]).** 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genetics, University of Washington, Seattle.
4. **Gorski, L., and D. Kaiser.** 1995. Personal communication.
5. **Huala, E., and F. M. Ausubel.** 1989. The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from the *R. meliloti nifH* promoter. J. Bacteriol. **171:**3354–3365.
6. **Jiang, J., B. H. Gu, L. M. Albright, and B. T. Nixon.** 1989. Conservation between coding and regulatory elements of *Rhizobium meliloti* and *Rhizobium leguminosarum dct* genes. J. Bacteriol. **171:**5244–5253.
7. **Kaufman, R. I., and B. T. Nixon.** Unpublished data.
8. **Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss.** 1989. Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism. Microbiol. Rev. **53:**367–376.
9. **Magasanik, B.** 1993. The regulation of nitrogen utilization in enteric bacteria. J. Cell. Biochem. **51:**34–40.
10. **Morett, E., and L. Segovia.** 1993. The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. J. Bacteriol. **175:**6067–6074.
11. **Popham, D. L., D. Szeto, J. Keener, and S. Kustu.** 1989. Function of a bacterial activator protein that binds to transcriptional enhancers. Science **243:**629–635.
12. **Sakanari, J. A., C. E. Staunton, A. E. Eakin, C. S. Craik, and J. H. McKerrow.** 1989. Serine proteases from nematode and protozoan parasites: isolation of sequence homologues using generic molecular probes. Proc. Natl. Acad. Sci. USA **86:**4863–4867.
13. **Wagner, S. J., M. McKinstry, and B. T. Nixon.** Unpublished observations.
14. **Wang, J. T., A. Syed, M. Hsieh, and J. D. Gralla.** 1995. Converting *Escherichia coli* RNA polymerase into an enhancer-responsive enzyme: role of an NH_2-terminal leucine patch in s54. Science **270:**992–994.
15. **Weiss, D. S., J. Batut, K. E. Klose, J. Keener, and S. Kustu.** 1991. The phosphorylated form of the enhancer-binding protein NtrC has an ATPase activity that is essential for activation of transcription. Cell **67:**155–168.
16. **Wilks, A. F.** 1989. Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction. Proc. Natl. Acad. Sci. USA **86:**1603–1607.
17. **Wu, S. S., and D. Kaiser.** 1995. Personal communication.