JOURNAL OF BACTERIOLOGY, Nov. 1999, p. 7140–7142 Vol. 181, No. 22
0021-9193/99/$04.00+0

An Improved Transposon for the Halophilic Archaeon *Haloarcula hispanica*

WAYNE G. WOODS, KATRINA NGUI, AND MICHAEL L. DYALL-SMITH*

Department of Microbiology and Immunology, University of Melbourne, Parkville, Victoria 3052, Australia

Received 28 June 1999/Accepted 7 September 1999

An improved transposon (ThD*73*) for *Haloarcula hispanica* is described. Based on the halobacterial insertion sequence ISH*28*, it showed little target sequence specificity but was biased toward a lower G+C content. Twenty randomly selected ThD*73* mutants were analyzed, and the DNA flanking their insertions revealed several recognizable sequences, including two (unrelated) ISH elements.

Prokaryotic genome sequences can now be determined rapidly, but it is commonly found that about 20 to 25% of open reading frames (ORFs) cannot be assigned a likely function by bioinformatics studies alone. Efficient methods for generating and analyzing mutants are needed to determine gene function, and we have been developing transposons (for a review, see references 2–4) for this purpose in extremely halophilic archaea (halobacteria). Our previous haloarchaeal transposon, ThD*28* (6) (Fig. 1), contained the *Haloferax volcanii* mevinolin resistance marker (Mev^r) between two tandem copies of the *Halobacterium salinarum* insertion sequence, ISH*28*. ThD*28* mutants of *Haloarcula hispanica* (9) contained insertions at different positions in the host chromosome, but these were unstable, as the tandem nature of the ISH elements and the presence of active transposase genes allowed further transposition and recombination events (6).

An improved transposon, ThD*73*, is shown in Fig. 1. Construction details can be obtained from the authors. It consisted of a haloarchaeal resistance marker (Mev^r [10]), plasmid pUC19 (19) to allow recovery of the transposon (and flanking DNA) in *Escherichia coli*, and the terminal inverted repeat (TIR) sequences of ISH*28* (18). The ISH*28* transposase gene (without TIRs) was placed outside the transposon. The plasmid was unable to replicate in halobacteria, so simvastatin-resistant transformants should arise only if the transposon integrated into the host genome. Plasmid pMDS73 was introduced into *H. hispanica* cells with a polyethylene glycol-mediated transformation protocol (5) and simvastatin-resistant transformants were selected on solid medium. The transformation frequency was approximately 3×10^3 colonies per μg of plasmid DNA, similar to that of the previous transposon, ThD*28* (6). A replicating plasmid (pWL102 [10]) routinely gave transformation frequencies of 10^4 transformants per μg.

To show that the transformants contained a copy of ThD*73*, 100 simvastatin-resistant colonies were patched onto nylon membranes and grown on selective medium; colony blots were prepared and probed with ^{32}P-labeled pUC19 DNA (a component of ThD*73*). All colonies hybridized strongly, while wild-type *H. hispanica* did not (results not shown). To verify that the transposase gene (which was outside the transposon) had not been inserted, the same colony blot of 100 transformant colonies was probed with an internal *Nae*I-*Dra*III fragment of ISH*28*. This fragment was not present inside the transposon and, as expected, the colonies were all negative (data not shown). A positive control (a *H. hispanica*::ThD*28* mutant, containing a complete ISH28) hybridized strongly.

Twenty randomly selected ThD*73* mutants were grown in liquid culture, chromosomal DNA was extracted, and Southern blots were prepared. Hybridizations of these blots confirmed that the insertions were single copies, and that each one was at a different location in the 20 mutants (data not shown). The same DNA preparations were used to recover the transposons (as plasmids) back into *E. coli*. After digestion with *Mlu*I (which cuts outside the transposon), restriction fragments containing the inserted transposons were circularized (by ligation) and electroporated into *E. coli* DH5α. Ampicillin-resistant transformants were readily obtained for all chromosomal DNA preparations and cells were found to possess plasmids. Sequencing outward from the transposon ends into the flanking DNA gave, on average, about 500 bases of DNA sequence for each plasmid (Table 1). These were compared to the GenBank nucleotide sequence database by the BLAST suite of programs (1), and several flanking sequences showed significant nucleotide or predicted amino acid similarity to genes from halobacteria and other organisms (Table 1).

Target site duplications were usually 8 bp long (15) but two were 9 bp. They varied widely in sequence, although three target sites were similar (e.g., plasmids pWW1, -5, and -16) (Table 1). The percentages of G+C content of the flanking sequences also varied, from 41 to 63% (Table 1). Halobacteria have a high average G+C content in their DNA (about 62 to 68%), but in many cases this can be separated into two fractions termed FI (high G+C) and FII (low G+C) DNA, respectively. *Haloarcula marismortui*, a close relative of *H. hispanica*, has FI and FII DNA fractions of 62 and 55% G+C, respectively. The latter fraction usually represents 10 to 30% of the halobacterial genome and is a mixture of endogenous plasmids and genomic fragments (13, 14). The high G+C fraction is thought to possess most of the important genes (17) and to be relatively stable. The low G+C fraction is quite variable in size and arrangement and seems to contain the majority of insertion sequences (8, 16, 17). The selection pressure maintaining this peculiar distribution is not fully understood (14, 17). The large halobacterial plasmid pNRC100 shows several regions of relatively low percentages of G+C content that contain few genes or long ORFs but have multiple complete or partial ISH elements, suggesting that these regions probably represent favored target areas for ISH element insertion (12).

Two insertions occurred in or near (unrelated) ISH elements, which could reflect a preference either for low G+C DNA or for ISH elements themselves. In IS elements of *Bac-*

* Corresponding author. Mailing address: Department of Microbiology and Immunology, University of Melbourne, Parkville, Victoria 3052, Australia. Phone: 61 3 9344 5693. Fax: 61 3 9347 1540. E-mail: m.dyall-smith@microbiol.unimelb.edu.au.

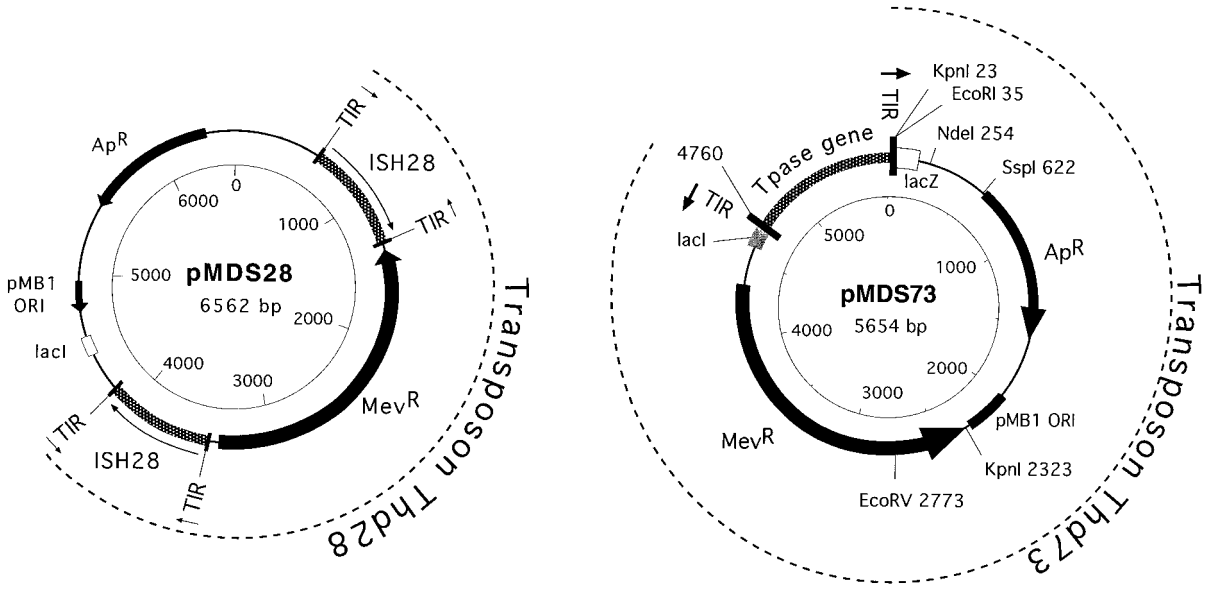

FIG. 1. Maps of plasmids pMDS28, carrying the transposon ThD*28* (previously described in reference 6), and pMDS73, carrying the improved transposon ThD*73* (described in this study). ThD*73* consists of a halobacterial mevinolin-simvastatin resistance gene (Mev^r; accession no. M83531) (10) and the *E. coli* plasmid vector pUC19 (19), which includes parts of the *lacZ* and *lacI* genes, a pMB1 replicon, and an ampicillin resistance gene, Ap^r. These are flanked by ISH*28* TIRs. The ISH*28* transposase (Tpase) gene is located outside the transposon.

teria, target preferences for DNA targets with high or low percentages of G+C content, as well as for other IS elements, have all been observed (7, 11) but the nature of these selection preferences, in terms of DNA structure, is not known. The possible role of histone-like proteins has been discussed recently by Mahillon and Chandler (11).

Previous reports of ISH elements in halobacteria have usually analyzed insertions into specific genes, particularly genes with readily observable phenotypes such as bacteriorhodopsin (*bop*) or gas vesicle production (*gvp*) genes (13, 14). ISH*28* was originally found as an insertion into the *bop* gene of *Halobacterium salinarum* (16). Other more general studies have used

TABLE 1. Properties of the recovered plasmids

Plasmid[a]	Duplicated target sequence	% G+C[b]	Accession no.[c]	Similarity to known sequence[d]
pWW1	GTCTCATC	55	AF150937	*H. salinarum* pHH1 plasmid ORF H0337 (signal-transducing histidine kinase); AAC82814 ($E = 1.2 \times 10^{-6}$)
pWW2	AACGGTTT	54	AF150938	Conserved hypothetical protein in *Pyrococcus horikoshii*; BAA29456 (E = 0.005)
pWW3	ATCGGGAT	41	AF150939	ISH50; X01584, AF016485
pWW4	ATCTGATC	52	AF150940	
pWW5	GTCTCATTC	46	AF150941	
pWW6	AAATAATA	51	AF150942	*H. salinarum* pHH1 plasmid ORF H0761 (putative cell division control-related protein); AAC82854 ($E = 2 \times 10^{-6}$). Conserved hypothetical protein of *Archaeoglobus fulgidus*; AAB90352 ($E = 2 \times 10^{-4}$)
pWW7	AAATGAAG	51	AF150943	
pWW8	CTCGGGAG	63	AF150944	*Candida parapsilosis* (mitochondrion) cytochrome *c* oxidase III; Q34214 ($E = 4 \times 10^{-8}$)
pWW9	GTTGAGTC	60	AF150945	
pWW10	GGGAGAGGG	60	AF150946	
pWW11	GACTTGTT	52	AF150947	ISH27, ISH51; X54432 ($E = 8 \times 10^{-65}$), X54433, X54434, X04389, AF016485
pWW12	GATCATTG	61	AF150948	
pWW13	GACGGTAT	51	AF150949	
pWW14	CTCCTGTT	53	AF150950	*Archaeoglobus fulgidus*, type II secretion system; AAB90245 (E = 0.001)
pWW15	CCTATTTC	54	AF150951	Dehydrogenase-oxidoreductase; S70672, D1031947 ($E = 5 \times 10^{-5}$)
pWW16	GTCTCGTC	46	AF150952	
pWW17	CGTGCGTT	58	AF150953	
pWW18	GTTGAATC	61	AF150954	
pWW19	GTAAATAG	48	AF150955	
pWW20	GGCAAATC	42	AF150956	

[a] Plasmids recovered in *E. coli* from the DNA of *H. hispanica*::ThD*73* transformants.
[b] The percentage of guanine and cytosine content over the region sequenced (approximately 400 to 500 bp surrounding each target site).
[c] The GenBank accession number for each flanking sequence.
[d] Similarity to known sequences was determined by BLAST searches of the GenBank database. The names and accession numbers of significantly similar sequences are given. Expected values (E) of similarity are shown in parentheses.

Southern blot hybridization to establish the distribution of different ISH elements in halobacteria (14). The present study is the first to follow insertions of a haloarchaeal transposon at the molecular level after its introduction into a halobacterial cell. Assuming that the target specificity of ThD*73* is similar to that of ISH*28*, the results indicate that ISH*28* is capable of integrating into a wide variety of target sites.

This work was supported by an Australian Research Council grant to M.D.-S. W.G.W. was supported by a University of Melbourne Postgraduate Scholarship.

We thank W. F. Doolittle and V. Athanasopoulos for reading the manuscript and L. Trantifillou for technical assistance.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Berg, C. M., and D. E. Berg.** 1987. Uses of transposable elements and maps of known insertions, p. 1071–1109. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
3. **Berg, C. M., D. E. Berg, and E. A. Groisman.** 1989. Transposable elements and the genetic engineering of bacteria, p. 879–925. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
4. **Camilli, A., D. A. Portnoy, and P. Youngman.** 1990. Insertional mutagenesis of *Listeria monocytogenes* with a novel Tn*917* derivative that allows direct cloning of DNA flanking transposon insertions. J. Bacteriol. **172:**3738–3744.
5. **Cline, S. W., W. L. Lam, R. L. Charlebois, L. C. Schalkwyk, and W. F. Doolittle.** 1989. Transformation methods for halophilic archaebacteria. Can. J. Microbiol. **35:**148–152.
6. **Dyall-Smith, M. L., and W. F. Doolittle.** 1994. Construction of composite transposons for halophilic Archaea. Can. J. Microbiol. **40:**922–929.
7. **Hallet, B., R. Rezsohazy, J. Mahillon, and J. Delcour.** 1994. IS231A insertion specificity: consensus sequence and DNA bending at the target site. Mol. Microbiol. **14:**131–139.
8. **Hofman, J. D., L. C. Schalkwyk, and W. F. Doolittle.** 1986. ISH*51*: a large, degenerate family of insertion sequence-like elements in the genome of the archaebacterium, *Haloferax volcanii*. Nucleic Acids Res. **14:**6983–7000.
9. **Juez, G., F. Rodriguez-Valera, A. Ventosa, and D. J. Kushner.** 1986. *Haloarcula hispanica* spec. nov. and *Haloferax gibbonsii* spec. nov., two new species of extremely halophilic archaebacteria. Syst. Appl. Microbiol. **8:**75–79.
10. **Lam, W. L., and W. F. Doolittle.** 1989. Shuttle vectors for the archaebacterium *Halobacterium volcanii*. Proc. Natl. Acad. Sci. USA **86:**5478–5482.
11. **Mahillon, J., and M. Chandler.** 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. **62:**725–774.
12. **Ng, W. V., S. A. Ciufo, T. M. Smith, R. E. Bumgarner, D. Baskin, J. Faust, B. Hall, C. Loretz, J. Seto, J. Slagel, L. Hood, and S. DasSarma.** 1998. Snapshot of a large dynamic replicon in a halophilic archaeon: megaplasmid or minichromosome? Genome Res. **8:**1131–1141.
13. **Pfeifer, F.** 1986. Insertion elements and genome organization of *Halobacterium halobium*. Syst. Appl. Microbiol. **7:**36–40.
14. **Pfeifer, F.** 1988. Genetics of halobacteria, p. 105–133. *In* F. Rodriguez-Valera (ed.), Halophilic bacteria. CRC Press, Boca Raton, Fla.
15. **Pfeifer, F., and P. Ghahraman.** 1991. The halobacterial insertion element ISH*28*. Nucleic Acids Res. **19:**5788.
16. **Pfeifer, F., H. Boyer, and M. Betlach.** 1985. Restoration of bacterioopsin gene expression in a revertant of *Halobacterium halobium*. J. Bacteriol. **164:**414–420.
17. **Schalkwyk, L.** 1993. Halobacterial genes and genomes, p. 467–496. *In* M. Kates, D. J. Kushner, and A. T. Matheson (ed.), The biochemistry of the Archaea. Elsevier, Amsterdam, The Netherlands.
18. **Woods, W., and M. Dyall-Smith.** 1996. Revised nucleotide sequence of an archaeal insertion element (ISH*28*) reveals a putative transposase gene. Gene **182:**219–220.
19. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.