JOURNAL OF BACTERIOLOGY, Jan. 1992, p. 30–34
0021-9193/92/010030-05$02.00/0
Copyright © 1992, American Society for Microbiology
Vol. 174, No. 1

Transcriptionally Active Regions in the Genome of the Archaebacterium *Haloferax volcanii*

BRUCE A. TRIESELMANN AND ROBERT L. CHARLEBOIS*

Department of Biology, University of Ottawa, 30 Marie Curie/priv., Ottawa, Ontario K1N 6N5, Canada

Received 28 August 1991/Accepted 24 October 1991

Transcriptionally active regions of the *Haloferax volcanii* genome were mapped by hybridization of radiolabeled cDNA to Southern blots of our minimal set of overlapping cosmid clones covering 96% of the 4.1-Mbp genome. Transcription during exponential growth occurred in nearly every region of the 2,920-kbp chromosome. Large parts of the 690- and the 86-kbp plasmids were transcribed, but the 440-kbp plasmid showed little expression. Transcription after a 40-min heat shock at 65°C was generally reduced, apart from a small set of strongly expressed loci all situated on the chromosome.

A high-resolution physical map of the 4.1-Mbp genome of the archaebacterium *Haloferax volcanii* DS2 has been recently completed (3). The genome is partitioned into five circular replicons: a chromosome of 2,920 kbp and four plasmids of 690, 440, 86, and 6.4 kbp. The detailed study of halobacterial genome structure is useful in that it provides an archaebacterial data base for comparative genome analysis, and it addresses the issue of the dynamics and evolution of a genome rich in insertion elements (4, 8, 15, 16) and plasmids.

We and our collaborators are interested in mapping genes to the physical framework of overlapping cosmid clones which currently cover 96% of the *H. volcanii* genome. The following three approaches have been used: (i) hybridization of Southern blots of digested cosmid clones by using bulk probes (tRNA or insertion elements) has allowed the mapping of families of genes (4), (ii) individual cloned halobacterial genes have been mapped by homologous or heterologous hybridization (3), and (iii) complementation of auxotrophic mutants with cosmid clone DNA has mapped 140 mutations to 35 loci (4, 5, 10).

With the goal of further understanding the genome of *H. volcanii*, we designed an experiment to find those loci most transcriptionally active under different physiological conditions by synthesizing radiolabeled cDNA and hybridizing it to Southern blots of digests of our minimal set of overlapping cosmid clones. This report describes the transcriptional organization of the genome of a culture in exponential growth (in chemically defined medium [CDM] and in complex medium) and compares it with that of a culture experiencing heat shock.

MATERIALS AND METHODS

Growth media and conditions. CDM was a modification of that described by Kauri et al. (9) in which glycerol and succinate were replaced with 30 mM trisodium citrate. Rich medium was as described in reference 6. Cultures (50 ml) were grown at 45°C with shaking in 250-ml Erlenmeyer flasks. Cells were harvested during exponential growth (A_{540}, 1.6) in rich medium, during exponential growth (A_{540}, 1.0) in CDM, or after exponential growth (A_{540}, 1.0) in CDM followed by a 40-min heat shock at 65°C with occasional agitation.

RNA extraction. Cells were quick-chilled to −15°C by immersing and swirling the flask for 2 min in a salted ice-water bath before the culture was centrifuged at −10 to −15°C for 10 min at 6,000 × *g*. The cell pellet was suspended in 0.25 ml of cold CDM. Cells were lysed with 10 ml of cold lysis buffer (20 mM Tris [pH 7.5], 50 mM EDTA, 100 mM NaCl, 0.5% [wt/vol] sodium dodecyl sulfate [SDS]) with vortexing and then extracted twice with 5 ml of phenol equilibrated with 50 mM Tris (pH 7.5) and 1 mM EDTA. Most of the DNA was removed by spooling during ethanol precipitation. RNA was stored in 95% ethanol at −20°C until needed.

High-molecular-weight RNA was purified by LiCl-ethanol precipitation (1). The nucleic acid pellet was suspended in 0.9 ml of 0.2% SDS–50 mM EDTA (pH 8)–110 U of RNase inhibitor (Boehringer Mannheim) and then salt precipitated with 3.8 μl of 2 M acetic acid plus 0.5 ml of 3:2 5 M LiCl–ethanol on ice for 2 h. Salt precipitation was repeated once or twice until contaminating DNA had disappeared, as judged by gel electrophoresis. High-molecular-weight RNA was retrieved by centrifugation, redissolved in autoclaved distilled water, and kept on ice or at −20°C. Yield was determined by measuring the A_{260} of a sample. The RNA was precipitated with ethanol, diluted to 1 mg/ml in water, and immediately frozen to −20°C for future use.

Southern hybridizations. Gels containing *MluI* digests of the minimal set of overlapping clones covering the *H. volcanii* genome (3) were depurinated briefly in 0.25 M HCl, transferred to GeneScreen membrane (DuPont) in 0.4 M NaOH, rinsed in 2× SSC (0.3 M NaCl, 35 mM trisodium citrate), and subjected to UV binding. A probe was prepared by synthesizing a first-strand cDNA (14) from 10 μg of salt-precipitated RNA by using avian myeloblastosis virus reverse transcriptase (Boehringer Mannheim), random hexamer primers, and [α-^{32}P]dCTP (6,000 Ci/mmol, 20 mCi/ml; DuPont). Blots were prehybridized for several hours and then hybridized in 1 M NaCl–50 mM Tris (pH 7.5)–5% SDS–50% formamide–5 mg of herring sperm DNA per ml at 37°C overnight. Blots were washed in 2× SSC–1% SDS at 70°C for 3 × 30 min after a rinse in 2× SSC for 60 min at room temperature.

* Corresponding author.

RESULTS AND DISCUSSION

We isolated high-molecular-weight RNA from cells grown in rich medium to the exponential phase, in CDM to the exponential phase, and in CDM to the exponential phase and then subjected to a 40-min heat shock at 65°C ($\Delta T° = +20°C$) prior to RNA extraction. Randomly primed, radiolabeled cDNA was used to probe Southern blots of the collection of overlapping cosmid clones covering 96% of the *H. volcanii* genome.

General issues. Three general considerations of the methods used are important to discuss at the outset. It was necessary that the RNA extracted from the cells be representative of the physiological state in question rather than representative of cells experiencing the rigors of centrifugation and preparation for lysis. We do not know how these manipulations affect gene expression and mRNA stability in halobacteria. We minimized any chance for new transcription or RNA degradation by chilling the cells to −15°C, a drop of 60 to 80°C, in 2 min. The salty medium does not freeze at this temperature, and the cells remain intact. Further manipulations were done in the cold. We assumed for this work that *H. volcanii* cells are transcriptionally inactive below 0°C and that their RNases do not function at significant rates within the 20 min needed to reach the lysis step.

The second consideration was that the RNA be free of DNA and tRNA. DNA would generate a general background signal, and tRNAs have already been mapped (4). Removal of DNA and tRNA was accomplished by repeated LiCl-ethanol precipitation. As judged by gel electrophoresis, the DNA was effectively removed. An RNase-treated RNA preparation was unable to support sufficient radiolabeled cDNA synthesis to result in hybridization signals against the *H. volcanii* cosmid clone collection (data not shown). We also concluded that tRNAs were not important contributors to our cDNA probe, since at least some *Mlu*I fragments that code for abundant tRNAs (16) did not hybridize. We are therefore confident that our probe represents high-molecular-weight RNA.

We made little effort to quantitate signal intensity, beyond such designations as "weak," "strong," and "strongest." Aside from the main factor of transcript abundance, signal strength could vary with transcript length, bias in reverse transcription, blot irregularities, etc. This does not diminish the value of our results. The main purpose of this work was to examine the genomic distribution of transcription in *H. volcanii*. The signals range continuously from nearly invisible to most obvious, and they were ranked roughly by intensity to point out the most likely candidates for highly expressed or differentially expressed genes.

Transcription during exponential growth. The genome of *H. volcanii* is partitioned into a chromosome and four plasmids. The genome is also partitioned into FI and FII DNAs (3). Halobacterial FII DNA, the lower-moles-percent G+C type (11), tends to be rich in insertion elements which are believed to be particularly active there (12, 13). One explanation for the survival of such insertion element-rich halobacteria as *Halobacterium halobium* is that essential genes are spared from disruption by the preferred targeting of FII DNA (2). These ideas imply that one should find few genes within FII DNA, apart from functions specific to insertion elements. Fortunately, the locations of members of the major insertion element family of *H. volcanii*, ISH51 (8), are known (4, 16).

Figure 1 shows a map of the *H. volcanii* genome and the distribution of transcriptionally active loci for the three physiological conditions presented in this report. Only the strongest signals are indicated (see Fig. 2 for an explanation and examples). Major transcripts map to nearly every part of the chromosome under exponential growth in both CDM and rich medium, with two notable gaps centered around positions 1,480 and 2,070 kbp. In contrast, transcripts map more sparsely within the 690-kbp plasmid pHV4 and especially poorly to the large, FI-like plasmid pHV3. We observed several hybridization signals mapping to the 86-kbp FII plasmid pHV1 but nothing for pHV2 (6 kbp) above the chosen threshold level of signal intensity.

FII regions within the chromosome and within pHV4 and pHV1 are expressed (Fig. 2 and data not shown). In some cases, but not all, known ISH51-positive bands (16) are seen as strong signals. We do not know whether any signals are

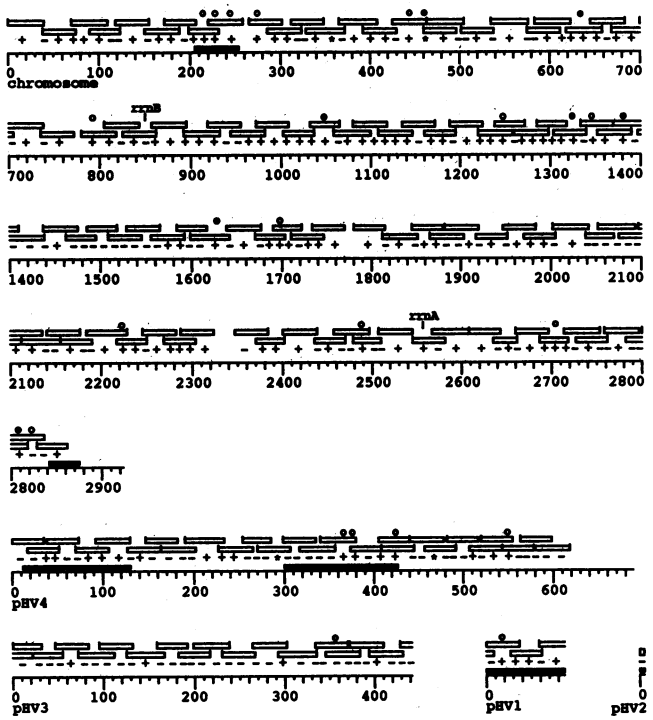

FIG. 1. Transcription map of the cloned regions of the *H. volcanii* genome. Each of the five replicons is a circular molecule. Cosmid clones representing the genome (3) are indicated by open boxes. Regions of FII DNA (lower moles percent G+C) are indicated by solid black bars above the kilobase-pair distance scale. Positions of observed transcription, judged by hybridization of radiolabeled cDNA to Southern blots of cosmid clone DNA, are shown above and below the cosmids (respectively) for heat shock and for exponential growth. Symbols: ●, major heat shock signals; ○, minor heat shock signals; + and *, signals from exponential-growth cDNA; −, no signal above the threshold for exponential-growth cDNA. Asterisks denote loci where the transcription pattern of cultures at the exponential phase in rich medium differs from that of cultures grown in CDM (see the text and Table 1). Resolution of this map is to the nearest third of a cosmid clone, i.e., within its left overlap, center portion, or right overlap. Each hybridizing locus within such an interval represents one or more positive *Mlu*I fragments spanning a length from several kilobase pairs to several tens of kilobase pairs. Where minor heat shock signals map to the same locus as signals from exponential growth, the hybridizing *Mlu*I fragments are not necessarily the same. The two rRNA operons are indicated by rrnA and rrnB.

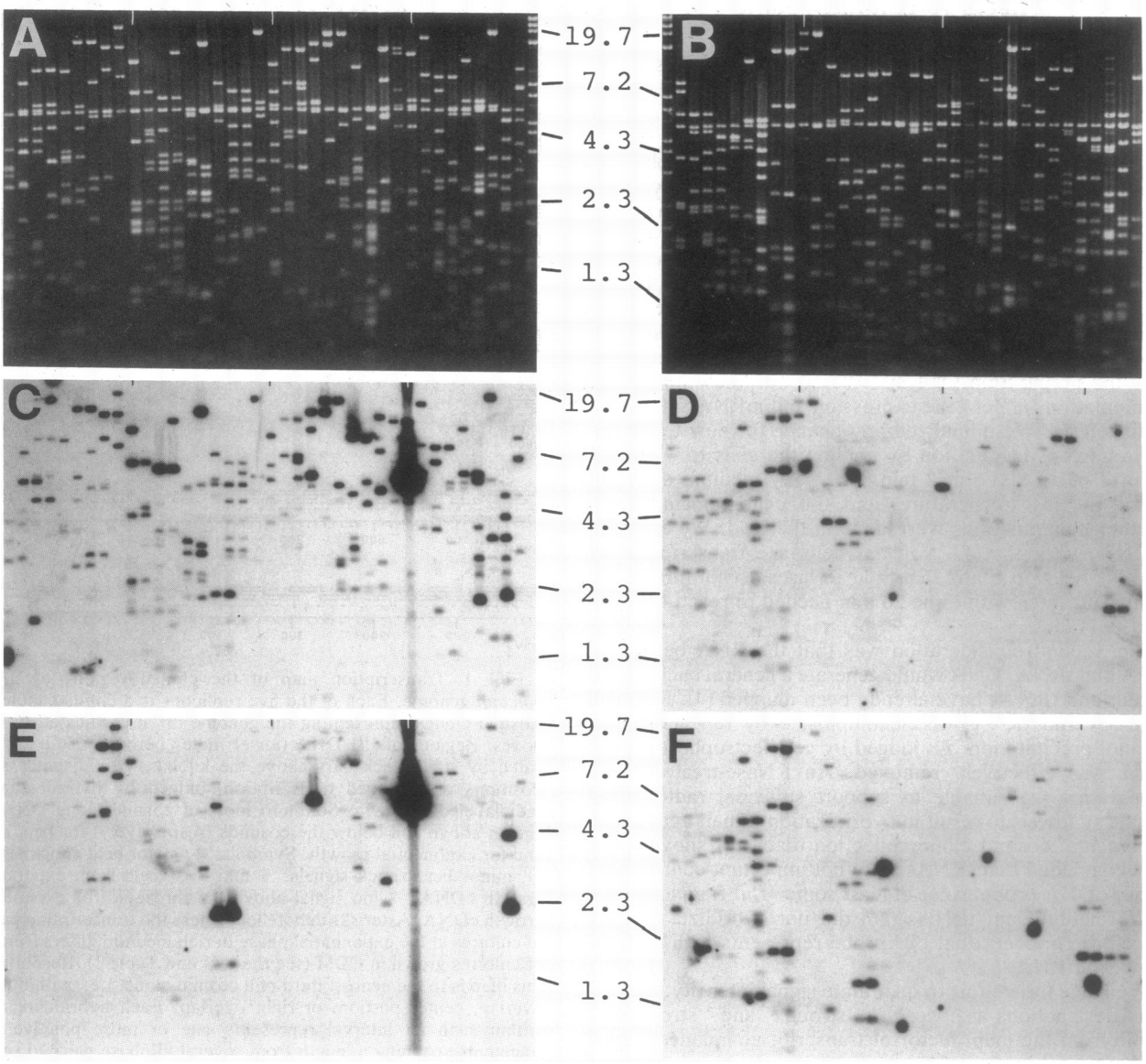

FIG. 2. Examples of mapping of transcripts to the genome. (A) Ethidium bromide-stained gel showing *Mlu*I digests of overlapping cosmids from chromosomal positions 0 to 1,100 kbp. Lanes 10, 20, and 30 are marked in all panels; marker sizes are in kilobase pairs. (B) Same as panel A but showing pHV4 positions 227 to 618 kbp (lanes 1 to 17) and all 442 kbp of pHV3 (lanes 18 to 34). (C and D) Southern hybridization of the gels above with radiolabeled cDNA representing transcription from cells growing exponentially in CDM. We called a hybridization signal positive if it had an intensity greater than or equal to that of the 7.5-kbp band in lane 2 of panel C. (E and F) Southern hybridization of gels similar to those in panels A and B with radiolabeled cDNA made from heat shock RNA. We decided that a major heat shock signal was one whose intensity was at least as strong as the major signal (5.9 kbp) in lane 23 of panel E and that a minor signal was one at least as strong as the 18.2-kbp band in lane 9 of panel E. Regions of FII DNA are represented in lanes 8 and 9 of panels C and E and in lanes 4 to 9 of panels D and F. The intense signal in lane 30 of panels C and E is due to an rRNA operon located there.

caused by hypothetical ISH51 transcription, since the fragments could contain other genes as well. What is important here is that many fragments within FII DNA known not to contain members of the ISH51 family of elements hybridized with our cDNA probe (Fig. 3). The presence of abundant transcripts in the FII regions means either that there are many undiscovered and transcriptionally active insertion elements in *H. volcanii* or that there are numerous genes there. One of the three strongest hybridization signals that we observed maps to the overlap between cosmids A316 and D165 in FII DNA of pHV4.

We do not understand, in contrast, why so little of the 440-kbp plasmid pHV3 is expressed although its DNA resembles chromosomal DNA in physical respects (3). One could speculate that pHV3 has some specialized physiological role and thus is not used in exponential growth. Only two putative tRNA genes and several insertion elements have been mapped to pHV3 (3, 4).

The differences in transcriptional patterns between rich medium and single-carbon-source CDM at the exponential phase are slight. The pictures shown in Fig. 2 for the exponential phase in CDM are virtually identical to those for the exponential phase in rich medium (data not shown because of this). Only four *Mlu*I fragments in the whole

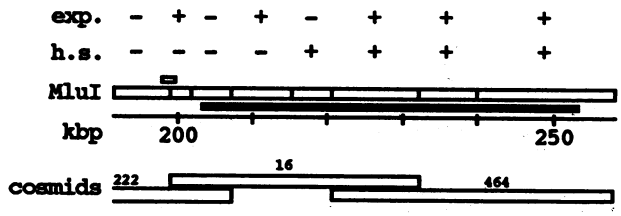

FIG. 3. Higher-resolution transcription map of the 50-kbp chromosomal FII region. The FII DNA is shown as a solid black bar above the distance scale. The position of the lone ISH51 element in this region (near position 199 kbp) is shown as a small box above the local *Mlu*I map. Abbreviations: exp., transcription in the exponential phase; h.s., minor heat shock signals. Compare with Fig. 1 and 2.

genome are differentially expressed in an obvious manner (Table 1). Although these four loci are now interesting, this result suggests that most of the signals observed do not represent activities directly involved in the interaction with medium.

Transcription during heat shock. The heat shock response in *H. volcanii* has been previously described at the protein level (6). Roughly 10 proteins were shown to accumulate under heat shock (a shift from 37 to 60°C), whereas synthesis of the general protein pool was lowered transiently, in agreement with the heat shock response in other systems (7). We saw the opportunity to find putative heat shock genes by mapping the locations of their transcripts. As expected, the transcription observed during heat shock was generally depressed relative to that in the exponential phase, apart from seven loci which remained (or became) particularly strong signals (Fig. 2 and Table 1). They all map to the chromosome (Fig. 1) but are not confined to any particular region of it. All of the heat shock fragments listed in Table 1 also showed strong hybridization with cDNA from exponential-phase cells growing at 45°C in CDM, but Daniels et al. (6) had shown that the major heat shock proteins were expressed at normal (37°C) temperatures. (It is worth noting that one of these heat shock fragments is not active in rich medium, one of the four differences between transcription in rich medium and that in CDM.) The signals that we observed may result from increased expression at a particular locus during heat shock, or they may simply be transcripts which are not degraded during heat shock. Determination of whether or not noninduced but stable transcripts encode heat shock functions will require further characterization of the signals found in this work. The assignment of these loci as heat shock genes is thus tentative.

TABLE 1. The best of the differentially expressed loci in the *H. volcanii* genome

Locus name[a]	Size(s) of *Mlu*I fragment(s) (kbp)	Map position (kbp)[b]
RE-416	1.1	348–365
RE-567	2.1	pHV4 285–299
CE-531/D57	2.2	459–464
CE-G317/483	17.3	pHV4 457–478
HS-531/D57	2.2	459–464
HS-H11	5.9	624–644
HS-A199	4.1, 2.2	1037–1058
HS-268	3.2	1318–1333
HS-456/A210	0.42	1370–1390
HS-452	3.9	1616–1638
HS-10D2/470	4.7	2799–2819

[a] RE, loci expressed more strongly in rich medium than in CDM; CE, loci expressed more strongly in CDM than in rich medium; HS, putative heat shock loci. Only the seven strongest heat shock hybridization signals are listed. Locus names include the name(s) of the cosmid(s) on which the signals reside.

[b] Unless otherwise noted, map positions refer to the chromosome and are consistent with the coordinates in reference 3.

Besides the seven signals discussed above, the next strongest class of heat shock cDNA signals (Fig. 2) shows an interesting genomic distribution (Fig. 1). Within these 19 minor heat shock loci, there are seven *Mlu*I fragments which do not correspond to strongly hybridizing fragments from the CDM exponential-phase hybridization. One of these fragments is located in the cluster of signals in and around the chromosomal FII island located between positions 203 and 253 kbp (Fig. 2C and E, lanes 8), two are in pHV4 (Fig. 2D and F, lanes 8 and 9 and 14 and 15), and one is in pHV3 (Fig. 2D and F, lanes 30 and 31). We also saw obvious differential expression among weaker signals (Fig. 2). The question of the function of these transcripts is open to speculation. Half of the minor heat shock signals originate from plasmid or chromosomal FII island DNA, sequences which may obey different rules of transcriptional control than the bulk of the chromosome. This idea is plausible if it is true that the origin of some or all of the plasmids found in *H. volcanii* was through horizontal transfer from another species, a hypothesis supported by the compositional differences between FI and FII DNAs (3).

Implications and prospects. In this investigation, we studied the genomic distribution of transcription in *H. volcanii*. Transcription in FII DNA, in both exponential growth and during heat shock, suggests a functional role for these regions. The dense and widespread expression of chromosomal FI DNA was no surprise, but we did not expect such transcriptional quiescence in the FI-like plasmid pHV3. FI regions of the other megaplasmid, pHV4, displayed moderate activity. Clearly, we do not understand the nature and purpose of the numerous subgenomic compartments present in halobacteria. Sequencing of some of the signals found here may answer many questions.

The approach that we took, mapping bulk-labeled cDNA to cosmid clones, seems to have sufficient sensitivity to be of use in the study of differential gene expression in this archaebacterium. Although we did not see as many differences as expected between the rich medium and CDM transcription patterns, four interesting loci were found. Heat shock RNA, on the other hand, produced many differences relative to the control. Higher-resolution mapping using RNA from other physiological states should provide a powerful means of finding additional genes and getting a first look at the regulation of their expression.

ACKNOWLEDGMENTS

This work was funded by NSERC.

We thank L. C. Schalkwyk, L. Bonen, and C. J. Daniels for helpful discussions and I. M. Pilote for technical assistance.

REFERENCES

1. **Birnboim, H. C.** 1988. Rapid extraction of high molecular weight RNA from cultured cells and granulocytes for Northern analysis. Nucleic Acids Res. **16**:1487–1497.
2. **Charlebois, R. L., and W. F. Doolittle.** 1989. Transposable elements and genome structure in halobacteria, p. 297–307. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
3. **Charlebois, R. L., L. C. Schalkwyk, J. D. Hofman, and W. F. Doolittle.** A detailed physical map and set of overlapping clones covering the genome of the archaebacterium *Haloferax volca-*

nii. J. Mol. Biol., in press.
4. **Cohen, A., W. L. Lam, R. L. Eharlebois, W. F. Doolittle, and L. C. Schalkwyk.** Submitted for publication.
5. **Conover, R. K., and W. F. Doolittle.** 1990. Characterization of a gene involved in histidine biosynthesis in *Halobacterium* (*Haloferax*) *volcanii*: isolation and rapid mapping by transformation of an auxotroph with cosmid DNA. J. Bacteriol. **172:** 3244–3249.
6. **Daniels, C. J., A. H. Z. McKee, and W. F. Doolittle.** 1984. Archaebacterial heat-shock proteins. EMBO J. **3:**745–749.
7. **Georgopoulos, C., D. Ang, A. Maddock, S. Raina, B. Lipinska, and M. Zylicz.** 1990. Heat shock response of *Escherichia coli*, p. 405–419. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
8. **Hofman, J. D., L. C. Schalkwyk, and W. F. Doolittle.** 1986. ISH51: a large, degenerate family of insertion sequence-like elements in the genome of the archaebacterium, *Halobacterium volcanii*. Nucleic Acids Res. **14:**6983–7000.
9. **Kauri, T., R. Wallace, and D. J. Kushner.** 1990. Nutrition of the halophilic archaebacterium, *Haloferax volcanii*. Syst. Appl. Microbiol. **13:**14–18.
10. **Lam, W. L., A. Cohen, D. Tsouluhas, and W. F. Doolittle.** 1990. Genes for tryptophan biosynthesis in the archaebacterium *Haloferax volcanii*. Proc. Natl. Acad. Sci. USA **87:**6614–6618.
11. **Moore, R. L., and B. J. McCarthy.** 1969. Characterization of the deoxyribonucleic acid of various strains of halophilic bacteria. J. Bacteriol. **99:**248–254.
12. **Pfeifer, F.** 1986. Insertion elements and genome organization of *Halobacterium halobium*. Syst. Appl. Microbiol. **7:**36–40.
13. **Pfeifer, F., K. Ebert, G. Weidinger, and W. Goebel.** 1982. Structure and function of chromosomal and extrachromosomal DNA in halobacteria. Zentralbl. Bakteriol. Hyg. I Abt. Orig. C **3:**110–119.
14. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed., p. 8.3–8.67. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
15. **Sapienza, C., and W. F. Doolittle.** 1982. Unusual physical organization of the *Halobacterium* genome. Nature (London) **295:**384–389.
16. **Schalkwyk, L. C.** 1991. Ph.D. thesis. Dalhousie University, Halifax, Nova Scotia, Canada.