Transcription initiation of stress (heat-shock) genes in archaea

A.J.L. Macario, E. Conway de Macario

Wadsworth Center, New York State Department of Health; and Department of Biomedical Sciences, The University at Albany (SUNY); Empire State Plaza, P.O. Box 509, Albany, New York 12201-0509, USA

ABSTRACT

The purpose of this article is to provide an introductory overview of stress-gene transcription in archaea for those unfamiliar with the topic. Archaea are diverse and inhabit disparate ecosystems (very cold or hot, temperate, etc.). The study of their stress genes should reveal a variety of strategies for cell survival and suggest manifold applications. Survival relies on the mechanisms of transcription initiation and regulation, which are also likely to be diverse and differ in different phylogenetic branches. The basic transcription apparatus of the few archaea examined resembles that of eukaryotes. There is a TATA-binding protein (TBP) instead of the sigma factor characteristic of the bacterial RNA polymerase, and a transcription factor similar to TFIIB. A transcription pre-initiation complex is formed by TBP and TFIIB. Analysis of the stress (heat-shock) genes *hsp23(grpE), hsp70(dnaK),* and *hsp40(dnaJ)* in a methanogen has revealed an eukaryotic type of expression.

Retrospective

Molecular phylogeny

A revolution in biology was set in motion in the late 1970's when it was proposed that prokaryotes may be subdivided into two distinct evolutionary lines: bacteria and archaebacteria (nowadays named archaea) (1,17). Eukaryotes are the third line, today called eucarya.

This classification of all living cells into three phylogenetic domains is based on comparative analyses of 16-18S rRNA sequences, and constitutes a major contribution of molecular phylogeny to modern biology. It provided a new outlook on life on Earth, and on evolution. Old and new data could be re-interpreted, or interpreted in accordance with the novel classification scheme. Microbiologic, immunologic, molecular biologic and genetic, biochemical, and ecologic data were examined under a new light. It thus became evident that bacteria and archaea are indeed different from one another, and from eukaryotes, in several aspects that had been overlooked or misinterpreted in the past. It also became obvious that organisms of each domain share molecular and genetic characteristics with organisms of the other domains to an extent unsuspected before.

Genomes

A second revolution overlapped the last period of the previous one mentioned above, when the first complete genome sequence, that of the bacterium Haemophilus influenzae, was made public in 1995. Others followed soon thereafter, and today there are at least 16

Microbial Biosystems: New Frontiers
Proceedings of the 8th International Symposium on Microbial Ecology
Bell CR, Brylinsky M, Johnson-Green P (eds)
Atlantic Canada Society for Microbial Ecology, Halifax, Canada, 1999.

published (4 from archaea, 11 from bacteria, and 1 from eucarya). Perhaps more than 50 will be completed in the next few years.

Comparative analyses of genomes *in toto,* or partially focussing on limited groups of genes, also revealed differences and similarities between bacteria, archaea, and eucarya, confirming and extending previous results obtained by comparative analyses of individual genes or small groups of genes.

Transcription factors

One of the most remarkable findings produced by these comparative analyses of sequences was that the transcription, translation, and DNA replication and repair machineries in archaea resemble those of eukaryotes, not those of the other prokaryotes, the bacteria.

Bacteria

Bacteria have only one type of DNA-dependent RNA polymerase (RNAP) to transcribe their genes. It is formed by three subunits in the stoichiometry $\beta' \beta\alpha_{2\sigma}$, with an additional subunit (ω) much less well characterized. The σ subunit confers on the enzyme its ability to recognize the promoter of the gene to be transcribed. Different σ subunits in a cell recognize genes important for different functions; for example σ^{32} in Escherichia coli recognizes the promoters for heat-shock genes, hsp70(dnaK) among them, and directs RNAP to initiate their transcription in response to heat shock and other stresses (14).

Eucarya

The eukaryotic transcription mechanism makes use of three different RNAPs, I, II, and III. RNAP II, the one of interest for the purposes of this overview, is composed of 9 subunits but is unable to recognize the promoter, in contrast to the bacterial RNAP which can do so via its σ subunit. RNAP II transcribes genes into mRNA, for ultimate translation into protein. Transcription by RNAP II is initiated when the TATA-binding protein (TBP) binds to the promoter (13). When this happens, other transcription factors gather around TBP and the promoter, and bring RNAP II into position to start transcription. The transcription factors involved are several. TBP is part of TFIID, which has at least 9 components. The other factors are: TFIIA, TFIIB, TFIIF, TFIIE, TFIIH, and TFIIJ. These form the pre-initiation complex (PIC) that starts and pushes transcription along the first steps.

To initiate transcription of the *hsp70(dnaK)* gene, and other stress genes, the eukaryotic cell resorts to specific heat-shock factors (HSF) (15). Upon heat shock, HSF recognize and bind a *cis*-acting signal upstream of the gene, named heat-shock element (HSE), and thus start PIC assembly and transcription.

Archaea

Amazingly, archaea, which superficially look so much like bacteria, have an RNAP similar to the eukaroytic RNAP II, and transcription factors that are homologs of TBP and TFIIB (16). Other factors may also exist but they have not yet been described in detail.

The hsp70*(*dnaK*) Gene in Archaea*

The story of the archaeal *hsp70(dnaK)* gene is characterized by unexpected twists. This gene is widely distributed in nature, and has been found in all organisms except those discussed below. Hsp70(Dnak), the protein encoded in the gene, is considered to be one of

the most conserved in evolution; its amino-acid sequence is at least 40% identical in organisms that otherwise are very far apart. It may be over 90% identical in organisms closely related phylogenetically (6).

The first *hsp70(dnaK)* gene within the domain archaea was cloned and sequenced in 1991 (10). At the time, this finding confirmed the already widely held view that the gene is highly conserved. It was also thought that the gene originated before the bacterial branch separated from the eucaryal-archaeal branch, in accordance with the evolutionary tree of the day (17). However, subsequent findings demonstrated that the situation is not as clear as all that. Many archaea do not have the gene, and those that have it may have received it via lateral transfer (5,9).

The first archaeal *hsp70(dnaK)* gene was cloned from the chromosome of the methanogen *Methanosarcina mazei* S-6. Subsequently, sequencing of the regions adjacent to *hsp70(dnaK)*, and experimental analysis of all the genes in the locus (5'-*orf16-hsp23(grpE)-hsp70(dnaK)-hsp40(dnaJ)-orf11-trkA*-3'), demonstrated a mixture of bacterial and archaeal (eucaryal-type) features (2,4).

Transcription initiation and regulation of the *hsp70*(*dnaK*) gene

Bacteria

In *E. coli*, σ^{32} is the factor that initiates transcription of genes that have a heat-shock promoter for this RNAP subunit. Other bacteria have other ways to regulate transcription initiation of their heat-shock genes. For example, in *Bacillus subtilis* and other bacteria (mostly Gram positives), the *hsp70(dnaK)* gene is negatively regulated via a repressor, HrcA, which binds to a regulatory signal named CIRCE (controlling inverted repeat chaperone expression), usually located close downstream of the transcription-initiation site (14). This *cis*-acting signal is formed by inverted repeats, a palindrome, with 9 base pairs in each repeat and 9 base pairs in between the repeats. Each repeat seems to be the binding site for the repressor protein which has two symmetrical DNA-binding regions. Upon heat shock, the repressor would lose affinity for DNA, release its grip on CIRCE, and become free, thereby allowing transcription initiation. Similar mechanisms, also involving palindromes, operate in other bacteria.

Eucarya

The eukaryotic *hsp70(dnaK)* gene is regulated via an HSF which binds to the HSE, as explained above. Contrary to most prokaryotes (bacteria and archaea), eukaryotes have more than one *hsp70(dnaK)* gene per genome. One or more are active under physiologic conditions, and they are called heat-shock cognates 70 *(hsc70)* genes. In addition, these genes are responsive to physiologic signals such as those involved with, or related to, cell differentiation and development. In contrast, other *hsp70* genes respond to stressors and are active during stress and shortly thereafter, during the post-stress recovery. These genes are regulated by a family of HSF, each being induced by a different stimulus or family of related stimuli, which will result in the specific induction of some but not all the *hsp70* genes in a cell (15).

Archaea

The mechanism responsible for activation of the hsp70(dnaK) gene in archaea has not yet been eluciated. It is likely that it will differ between the archaeal kingdoms, and even within a kingdom, between the major subgroups, perhaps in relation to their preferred ecosystem (e.g., very cold or very hot; very alkaline or very acid). Indeed, archaea constitute a heterogeneous domain with a large variety of species differing in metabolic pathways, living conditions, and other characteristics (1,6,17). For example, extreme halophiles and methanogens (both in the kingdom Euryarchaeota), differ in that the former do not produce methane and live optimally in high-salt environments, namely just the opposite of the methanogens. Extreme halophiles contain extra chromosomal elements, while these elements are rare in methanogens. A common feature though is that both groups, i.e., all extreme halophiles and some of the methanogens, possess the hsp70(dnaK) gene (5,9). Other methanogens, and all the archaeal hyperthermophiles tested, do not have the gene.

There are no data on transcription of the *hsp70(dnaK)* gene in the species in which this gene has been found, except for *M. mazei* S-6 and other *Methanosarcinae* (2,4,9). Transcription initiation and regulation are in fact being investigated at the present time. The data so far obtained show that there are no discernable *cis*-acting signals or a promoter of bacterial type upstream of the *hsp70(dnaK)* gene in *M. mazei* S-6, nor are there these types of signals and promoters upstream of the other genes in the *hsp70(dnaK)* locus. Interestingly, HSEs are not found, either. Thus, one cannot predict from this information the type of regulatory signals and factors that are involved in *M. mazei* S-6 to control expression of *hsp70(dnaK)*. However, since there are inverted repeats that form palindromes in the 5'-flanking region of the gene, one may hypothesize that at least a pair of repeats is a binding site for a regulator with two symmetrical DNA-binding regions, either on a single polypeptide (monomer) or on two (dimer).

Whether regulation will be positive or negative is another question still unanswered. If heat induces binding of the regulatory factor to DNA and thereby induces transcription, it will mean that the *M. mazei* S-6 *hsp70(dnaK)* is positively regulated via an activator, perhaps with functions similar to the eukaryotic HSF, or to some archaeal activators (see below). If on the contrary, heat induces the factor to detach from its site on DNA (similar to a bacterial repressor site or operator), and this derepresses the gene, it will indicate that the gene is negatively regulated in a manner similar to many bacterial genes and operons, including *hsp70(dnaK)* in those species that have CIRCE and use the HrcA repressor.

Archaeal-gene regulatory signals and factors

Although little is known yet about regulatory factors for stress genes in archaea, there is some information pertaining to other genes that can be used to design experiments for elucidating regulation of *hsp70(dnaK)* in *M. mazei* S-6.

Activators

The homolog of the bacterial activator 'leucine-responsive regulatory protein' (Lrp), has been dentified by sequence searches of data bases in the chromosome of the hyperthermophile Pyrococcus furiosus (8). Lrp is a major transcriptional activator in E. coli.

A transcriptional activator named GvpE (for gas vesicle protein E) of eucaryal type has been described in the extreme halophile *Haloferax volcanii* (7). This activator regulates expression of the gas vesicle protein (*gvp*) gene clusters in *H. volcanii*; it resembles the basic leucine zipper proteins that are involved in gene regulation in eukaryotes.

Repressors

The presence of a repressor in the mesophilic archaeon Methanococcus maripaludis has been inferred from mutational analyses (3). A palindromic cis-acting site was found to be essential for repression of the nif (nitrogen fixation) gene by ammonia. The structure of the site and its location with respect to the transcription initiation site suggest that it is an operator, and the regulator is a repressor that binds to the palindrome, which results in nif repression, as repressors for other genes do in bacteria.

More recently, repressor-like proteins, also of bacterial type, were isolated from cell lysates of *Sulfolobus solfataricus* (11). The function(s) of these proteins (7c3 and 7c4) is unknown, but because of their similarity with bacterial repressors (*e.g*., Ecrep6.8 and Ecrep7.3 of *E. coli*; and ACCR of *Agrobactericum tumefaciens*), they are thought to be archaeal repressors of bacterial type.

Histones

Another important point that must be considered when trying to understand transcription initiation in archaea, is that some of these organisms possess histones, and their DNA is organized into nucleosomes (12). Interestingly, true histones and nucleosomes have been observed only in Euryarchaeota, but not in Crenarchaeota, such as Sulfolobus solfataricus (11). The archaeal histones, found in Euryarchaeota, are homologs of the eukaryotic histones H2A, H2B, H3 and H4. H1 has not been found in archaea. The nucleosome would then be formed by four histones in what amounts to a structure similar to that formed by the core histones in eurkaryotes, but smaller. Only 60 base pairs are protected from nuclease digestion in archaeal nucleosomes, in contrast to the 140 base pairs protected in eukaryotes.

Archaea without histones (Crenarchaeota), do possess another type of basic, small, DNA-binding proteins, which do not have the histone fold (11,12). These proteins are 7, 8, or 10kDa in mass. In *S. solfataricus*, a prominent protein of this group is Sso7d, which is 63 amino acids long and very basic. It binds, unwinds, and bends DNA, but its function(s) has not yet been elucidated.

Likewise, the role that these proteins of Crenarchaeota and the histones of Euryarchaeota play in regulation of transcription in archaea has not been determined.

Acknowledgement

Work done in the authors' laboratories was supported by a grant from the Department of Energy, USA.

References

Note: Because of space limitations, only essential references are listed. Others are available upon request.

1. Balch WE, Fox GE, Magrum LJ, Woese CR, Wolfe RS (1979) Methanogens: reevaluation of a unique biological group. Microbiol Rev 43:260-296

2. Clarens M, Macario AJL, Conway de Macario E (1995) The archaeal *dnaK-dnaJ* gene cluster: organization and expression in the methanogen *Methanosarcina mazei*. J Mol Biol 250:191-201
3. Cohen-Kupiec R, Blank C, Leigh JA (1997) Transcriptional regulation in Archaea: *in vivo* demonstration of a repressor binding site in a methanogen. Proc Natl Acad Sci USA 94:1316-1320
4. Conway de Macario E, Clarens M, Macario AJL (1995) Archaeal *grpE*: transcription in two different morphologic stages of *Methanosarcina mazei* and comparison with *dnaK* and *dnaJ*. J Bacteriol 177:544-550
5. Gribaldo S, Lumia V, Creti R, Conway de Macario E, Sanangelantoni A, Cammarano P (1999) Discontinuous occurrence of the *hsp70(dnaK)* gene among archaea and sequence features of HSP70 suggest a novel outlook on phylogenies inferred from this protein. J Bacteriol (in press)
6. Gupta RS, Singh B (1994) Phylogenetic analysis of 70 kD heat shock protein sequences suggests a chimeric origin for the eukaryotic cell nucleus. Curr Biol 4:1104-1114
7. Krüger K, Hermann T, Armbruster V, Pfeifer F (1998) The transcriptional activator GvpE for the halobacterial gas vesicle genes resembles a basic region leucine-zipper regulatory protein. J Mol Biol 279:761-771
8. Kyrpides NC, Ouzounis CA (1995) The eubacterial transcriptional activator Lrp is present in the Archaeon *Pyrococcus furiosus*. Trends Biochem Sci 20:140-141
9. Lange M, Macario AJL, Ahring BK, Conway de Macario E (1997) Heat-shock response in *Methanosarcina mazei* S-6. Curr Microbiol 35:116-121
10. Macario AJL, Dugan CB, Conway de Macario E (1991) A *dnaK* homolog in the archaebacterium *Methanosarcina mazei* S6. Gene 108:133-137
11. Oppermann UCT, Knapp S, Bonetto V, Ladenstein R, Jornvall H (1998) Isolation and structure of repressor-like proteins from the archaeon *Sulfolobus sulfataricus*. FEBS Lett 432:141-144
12. Pereira SL, Reeve JN (1998) Histones and nucleosomes in archaea and eukarya: a comparative analysis. Extremophiles 2:141-148
13. Pugh BF (1996) Mechanisms of transcription complex assembly. Curr Op Cell Biol 8:303-311
14. Ron EZ, Segal G, Sirkis R, Robinson M, Graur D (1999) Regulation of heat-shock response in bacteria (this Book)
15. Satyal SH, Chen D, Fox SG, Kramer JM, Morimoto RI (1998) Negative regulation of the heat shock transcriptional response by HSBP1. Genes & Dev 12:1962-1974
16. Thomm M (1996) Archaeal transcription factors and their role in transcription initiation. FEMS Microbiol Rev 18:159-171
17. Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87:4576-4579