Human sat III and Drosophila hsr omega transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells.

Caroline Jolly, Subhash C. Lakhotia

To cite this version:

HAL Id: inserm-00334259
https://inserm.hal.science/inserm-00334259
Submitted on 24 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
ABSTRACT
Exposure of cells to stressful conditions elicits a highly conserved defense mechanism termed the heat shock response, resulting in the production of specialized proteins which protect the cells against the deleterious effects of stress. The heat shock response involves not only a widespread inhibition of the ongoing transcription and activation of heat shock genes, but also important changes in post-transcriptional processing. In particular, a blockade in splicing and other post-transcriptional processing has been described following stress in different organisms, together with an altered spatial distribution of the proteins involved in these activities. However, the specific mechanisms that regulate these activities under conditions of stress are little understood. Non-coding RNA molecules are increasingly known to be involved in the regulation of various activities in the cell, ranging from chromatin structure to splicing and RNA degradation. In this review, we consider two non-coding RNAs, the hsr transcripts in Drosophila and the sat III transcripts in human cells, that seem to be involved in the dynamics of RNA-processing factors in normal and/or stressed cells, and thus provide new paradigms for understanding transcriptional and post-transcriptional regulations in normal and stressed cells.

INTRODUCTION
The heat shock response is a highly conserved cellular response to a variety of stresses and principally involves a transient reprogramming of transcriptional and translational activities (1) besides other physiological changes in cellular organization. The transcriptional reprogramming includes not only a widespread inhibition of transcription of most genes and activation of the heat shock or stress genes, but also significant changes in post-transcriptional processing. Several of the pioneering studies on heat shock response in eukaryotes established that heat and other cellular stresses blocked splicing, other post-transcriptional processing and transport of products of most of the nuclear genes that were active prior to the stress (2–4).

Post-transcriptional processing of most gene products is elaborate and requires multitudes of proteins and other factors (5), which are organized into a variety of distinct nuclear and cytoplasmic compartments (6–8). The blockage of RNA-processing events following stress affects the organization and composition of these sub-structures and functional compartments. The significance of the blockage of post-transcriptional processing following heat shock lies in the fact that many of the RNA-processing proteins may themselves be denatured or mis-folded or otherwise affected by the elevated temperature (and other cellular stresses) and thus may not function optimally and precisely. An error in RNA processing would generate a cascade of deleterious consequences for the cell and, therefore, it is a good survival strategy to block most of the routine RNA processing and translational activities during stress.

Molecular and cell biological studies, typically involving immunofluorescent staining, have documented a variety of changes in cellular sub-structures in diverse eukaryotic cells following heat shock or other stresses. Notable among these stress-induced changes are (i) aggregation and round-up of the various speckled domains like interchromatin granules or splicing speckles (9–11), Lamin A/C speckles (10,14,15), coiled or Cajal bodies (16,17), paraspeckles (18) or other hnRNP speckles (19) etc., (ii) release of snRNPs from the splicing speckles (9,20), (iii) formation of novel nuclear stress bodies in human cells (21–24), (iv) formation of cytoplasmic stress granules in...
plant and several animal cell types (25–28) and (v) changes in translocation of acetylcholinesterase splice variants in neuritic cells (29,30).

Although the actual movement of several RNA-processing proteins from one compartment to another under conditions of stress (or conditions where transcription is inhibited) has been documented in many cases, the specific mechanisms that regulate and actually bring about these changes under conditions of stress are little understood. In this review, we consider two non-coding RNAs, the hsrω transcripts in Drosophila and the sat III transcripts in human cells that appear to be involved in the dynamics of some of the RNA processing factors in normal and/or stressed cells. These non-coding transcripts seem to provide new paradigms for understanding dynamics of transcriptional and post-transcriptional regulations in normal and stressed cells.

hsrω transcripts and omega speckles in Drosophila

Among the stress-induced loci in Drosophila melanogaster, the 93D or hsrω (heat shock RNA omega) gene is unique (31). Although it produces several transcripts, it does not code for any protein. The 93D locus is conserved among Drosophila species. In D. melanogaster, it is one of the most active genes after heat exposure. This locus, however, is not sensu stricto a heat shock gene, since it is constitutively expressed at relatively high levels in different cell types and it is also uniquely responsive to amides like benzamide, colchicine etc. Genomic organization of the gene, rather than its base sequence, is conserved among species. The locus, which spans 10–20 kb, is composed of two small exons (∼475 and 700 bp in D. melanogaster) separated by a 700 bp intron, and followed by a long stretch (5–15 kb) of short (280 bp in D. melanogaster) tandem repeat units. The hsrω loci in different species of Drosophila share a common organization with two exons and one intron and a long stretch of tandem repeats at the 3′ end of the gene. However, it is interesting that in spite of the comparable organization in different species, the base sequence of all the regions in general shows high divergence (32). In all the species of Drosophila examined so far, this gene produces three transcripts using alternative polyadenylation sites and splicing. A first polyadenylation site located at the 3′ end of exon 2 generates a precursor nuclear RNA of ∼1.9 kb termed hsrω-pre-c, which by splicing of the intron gives rise to the mature 1.2 kb cytoplasmic hsrω-c transcript. The third nucleus-limited transcript, the hsrω-n, covers the entire length of the gene including the tandem repeats, and is thus typically larger than 10 kb. It is also polyadenylated, but curiously the intron is not spliced out. The three transcripts display distinct cellular localization, with only the short hsrω-c being cytoplasmic.

A short translatable open reading frame, coding for 23–27 amino acids in different Drosophila species, is present in the sequence of the hsrω-c transcript; however no corresponding protein has been detected (33), confirming the earlier suggestions that these transcripts are indeed non-coding (32,34,35).

Although not coding for any protein, this gene is essential not only under conditions of stress but for normal development as well. The multiple hsrω transcripts are expressed in nearly all cells from embryo to adult stages in a developmentally regulated pattern (36–38). Hsrω-nullosomics mostly die at embryonic stage; the few surviving flies are weak, sterile and die within a few days of emergence. In addition, cells of hsrω nullosomics fail to acquire thermo-tolerance and the organisms cannot survive at 31°C. Recent studies (M. Mallik and S. C. Lakhotia, unpublished data) using conditional and targeted over- or under-expression of the hsrω transcripts further underscore the vital importance of the hsrω non-coding transcripts in normal development as well as following heat shock. In agreement with this gene’s highly regulated and dynamic expression in different cell types, the promoter region of hsrω is rather long and includes multiple regulatory elements (37,39,40).

An interesting feature of the hsrω-n transcripts is that these are present at the gene locus and within specific nuclear structures termed omega speckles, which are seen in varying numbers in different cell types (12,13). Analyses of the proteins associated with the omega speckles have suggested possible function of the hsrω-n transcripts. The following proteins are known to be associated with the nucleus-limited hsrω-n transcripts: (i) various hnRNPs (12,13); Hrp40 (hnRNP A) (41), Hrb87F (hnRNP A1/A2) (42), Hrb57A (hnRNP K) (43), S5 (hnRNP M, H. Saumweber, personal communication), Hrb98DE (44), etc., (ii) other nuclear RNA-binding proteins like NonA, PEP (45), Sxl (46), nuclear non-histone proteins recognized by Q14, Q16, T29, P75 antibodies (12,47), (iii) Sfn (45), (iv) Hsp83 (48) and (v) Tpr (45), etc. Most of these proteins (in groups i–iii) are normally associated, in addition to their presence in the nucleoplasmic omega speckles, with many sites on chromosomes, specially those that are transcriptionally active. However, following heat shock, most of them disappear from almost all the chromosome sites, except the hsrω gene site (at 93D4 band in polytene chromosomes). Interestingly, the omega speckles also disappear following heat shock and after ~30–60 min of exposure of cells to 37°C, the only site where the hsrω-n and most of the above proteins are seen is the hsrω gene locus (12,13).

It is believed that the omega speckles are dynamic storage sites for the various RNA-processing and related proteins from which the different proteins are released as required by the state of nuclear activities at any given moment. It seems that the very high turnover of the hsrω-n transcripts (49) is related to the release of the various proteins from these storage sites. Under conditions of heat shock and other cellular stresses, which affect new transcription and processing of the transcripts, the hsrω-n RNA-binding proteins are released from their chromosomal locations and are quickly sequestered by the concomitantly elevated levels of hsrω-n transcripts. With increasing levels of sequestration, the omega speckles themselves coalesce, initially forming larger nucleoplasmic clusters and finally all the nuclear hsrω-n RNA and the associated proteins get restricted to the hsrω gene locus itself. As the cells recover from stress, the hnRNPs and other RNA processing proteins resume their chromosomal locations and at the same time the fine omega speckles also appear in the nucleoplasm (13,31). An altered organization of the omega speckles has been shown to be associated with specific phenotypes. For example, clustering of omega speckles in cyst cells in testis is associated with formation of unindividualized sperm bundles (30). Likewise,
altered organization of omega speckles and conditional over-expression of hsrω also dominantly enhance the neurodegeneration caused by expression of proteins with expanded polyglutamine in developing eye imaginal disks (51).

Sat III transcripts and nSBs in human cells

No homologue of the hsrω gene has been described so far in mammalian cells. However, new heat-induced non-coding transcripts, which display intriguing similarities with the hsrω transcripts, have recently been described in human cells. These are transcribed from satellite III repeated sequences, present mainly in the pericentromeric region of human chromosome 9 (9q12 region), described as one of the largest heterochromatic blocks of the genome (23,52). The sat III transcripts are not expressed constitutively, at least in different cell types examined so far. However, it has been shown that a conditional loss-of-function mutation of Dicer in chicken DT40-derived hybrid cells containing human chromosome 21, leads to an accumulation of transcripts from both centromeric α-satellite and pericentromeric satellite III human sequences, suggesting that these sequences may be normally expressed but barely detectable due to their very rapid degradation by Dicer (53). The expression of sat III sequences is strongly induced upon heat shock (23). They are transcribed by the RNA polymerase II, and this transcription is strictly dependent on the heat shock transcription factor HSF1 (23). The sat III transcripts are polyadenylated (52), but their size is not definitively known. Indeed, while the group of Claire Vourc’h found very large sat III transcripts (23), Giuseppe Siamonti’s group reported smaller transcripts of 2–5 kb size (52). Because of its repeated nature, the genomic organization of the 9q12 locus remains unknown. In a recent work, some of the sat III transcripts have been cloned, and they appear to contain numerous repetitions of the GGAAT motif typical of the satellite III repeats (54).

The sat III transcripts remain associated with the 9q12 locus after synthesis, even after transcription is arrested (23). These sites of accumulation of sat III transcripts are visible at the light microscopic level, and appear in heat-shocked cells as large nuclear foci, each one being associated with a 9q12 locus. These structures were first described as massive accumulation sites of HSF1, appearing transiently during heat exposure (21), and were recently renamed as nuclear stress bodies or nSBs (55). Like the hsrω transcripts, several, but not all, splicing factors and hnRNPs associate with sat III transcripts in the nSBs (Table 1). Heat shock induces the accumulation of these proteins within nSBs and this relocalization is strictly dependent on the presence of sat III transcripts (56,57) similar to the restriction of hnRNPs to the hsrω locus in heat shocked cells being dependent upon the presence of the hsrω transcripts (13).

One suggested possible function of sat III transcripts is a role in chromatin structure (23,55). This hypothesis is mainly based on the fact that several non-coding RNAs are now recognized as major players in establishing specific chromatin structure. For instance, small RNAs are involved in the initiation of heterochromatin formation at centromeric regions, while the nucleus-restricted, non-coding Xist transcripts appear essential for the establishment of the heterochromatic state of the inactive X chromosome (58). Interestingly, the 9q12 locus appears to have a particular chromatin structure, displaying features of both heterochromatin and euchromatin.

Indeed, despite the highly repeated nature of this chromosomal region, epigenetic marks typical of heterochromatin are not detected at this locus (23,52). In addition, separation of nuclease digested chromatin fragments on sucrose gradient has shown that the 9q12 contains both compact and open chromatin fibers, representative of heterochromatin and euchromatin, respectively (59). It may also be noted in this context that the 9q12 is a region of the genome sensitive to DNA breakage and chromosome rearrangements in tumor cells, for example. One can thus imagine a role for these transcripts in the maintenance of the particular chromatin structure of the 9q12 locus during stress, as has been shown for other tandem repeats (60), or in protecting the region against stress-induced damage. In this hypothesis, the various splicing factors and hnRNPs associated with the nSBs would be involved in the processing of sat III transcripts.

Sat III transcripts as functional analogues of hsrω transcripts

The sat III and hsrω-n transcripts display many interesting similarities (Table 1). Besides their repeated and non-coding nature, both transcripts massively accumulate at the sites of synthesis during stress and associate with a variety of RNA-processing factors, several of which are common. This suggests that these transcripts may perform comparable functions. A tempting hypothesis is that formation of the nSBs and the clustering of omega speckles at the hsrω gene locus following stress reflect similar roles in sequestering key regulatory factors involved in transcriptional and splicing events. Stress exposure, particularly heat shock, is well known to inhibit transcription, transcript maturation and transport (2,3). Association of several splicing factors and other RNA-processing proteins with nSBs and the hsrω locus (see Table 1) is significant in the context of this inhibition. The inhibition of synthesis, processing and transport of the various non-heat shock nuclear transcripts during stress may be related, at least in part, to the sequestration of the various RNA-binding proteins and/or transcription factors within nSBs or at the hsrω locus.

This interpretation finds support in the observed alternative splicing of some viral transcripts (52) or of some endogenous transcripts in heat shocked HeLa cells (A. Metz, C. Vourc’h and C. Jolly, unpublished data). It is known that alternative splicing is regulated, at least in part, by two classes of factors with opposite functions, namely the SR proteins and the hnRNPs (61–63). As noted above, during heat stress, only certain, but not all, splicing factors are retained within nSBs or at the hsrω locus (see Table 1). For example, members of the SR family such as SF2/ASF or SRp30c are retained in nSBs while hnRNPs such as the hnRNP A1 are not (24), although HAP1, an hnRNP A1 associated protein, accumulates in the nSBs (64). Such a differential sequestering of the processing factors is likely to not only largely inhibit RNA processing in general in stressed cells but may also be responsible for a modification in the alternative splicing pattern of some transcripts as noted above. It is likely that
the retention of some of the RNA processing factors within specific nuclear structures during heat shock is part of the cellular strategy to modulate splicing patterns, to restrain splicing factors from action under non-physiological conditions and to safeguard the splicing machinery components themselves from heat/stress-induced damage. In addition to a role in regulating the availability of transcription and splicing factors, these sites of accumulation could also serve as storage sites for active factors and/or as places where the different proteins are post-translationally modified and recycled. For example, it is well known that the activity of at least certain splicing factors is regulated by phosphorylation (65,66), and one can thus imagine that certain steps in the regulation of the phosphorylation status occurs within the nSBs or the omega speckles.

It is interesting that the various nuclear speckled domains, which are distinguished by their structure, their proteome and their distinct roles in RNA processing, show comparable aggregation following stress (10,12,14,16–18). Under normal cellular conditions, the speckled domains are believed to represent storage sites for the different components of RNA synthesis and processing. It is logical to presume that the conditions, which block these activities, would necessitate greater and more stringent conditions for storage of these components. Apparently, aggregation of the individual speckles provides a physical basis for this.

An open question for the moment is how hsr0 and sat III transcripts are retained at the site of transcription in stressed cells? This may occur via interaction with protein complexes, in particular those involving RNA-binding proteins such as splicing factors. However, one could also imagine that a physical interaction exists between the sat III transcripts and DNA in the corresponding gene loci, perhaps in the form of triplex structures. Indeed, sat III transcripts are retained at the site of transcription in stressed cells, exclusively at the site of transcription (12,13)

Table 1. Comparison of hsr0 and human sat III transcripts and their associated proteins

<table>
<thead>
<tr>
<th>Feature</th>
<th>hsr0 locus/hsr0-n transcripts</th>
<th>9q12 locus/sat III transcripts/nSBs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of base sequence</td>
<td>5' unique sequence followed by long stretch (&gt;5 kb) of short tandem repeats at the 3' end (31)</td>
<td>Size and specific sequence not known but transcripts contain repetitions of the 158 bp sat III consensus sequence (82)</td>
</tr>
<tr>
<td>Intranuclear distribution of transcripts</td>
<td>At the site of transcription and in varying numbers of nucleoplasmic omega speckles in unstressed cells but in stressed cells, mostly at the site of transcription (12,13)</td>
<td>Transcripts apparently absent in unstressed cells but in stressed cells, exclusively at the site of transcription in form of nuclear stress bodies (23,52)</td>
</tr>
<tr>
<td>Acetylated core histones</td>
<td>Not known</td>
<td>Yes (23)</td>
</tr>
<tr>
<td>Methylated H3 K9</td>
<td>Not known</td>
<td>Not present (52)</td>
</tr>
<tr>
<td>HP1</td>
<td>Not known</td>
<td>Not present (23)</td>
</tr>
<tr>
<td>Transcription by RNA pol II</td>
<td>Present at the site of transcription; presence in omega speckles not known</td>
<td>RNA pol II (23)</td>
</tr>
</tbody>
</table>

Proteins associated with the transcripts/nuclear structures

- hnRNPs: Yes (12,13)
- HSF: Binds the hsr0 locus transiently after heat shock (84)
- RNA pol II: Present at the site of transcription; presence in omega speckles not known
- Hsp90: Yes in heat shocked cells at the site of transcription (48) and to some extent also in the omega speckles (R. Fatima and S. C. Lakhotia, unpublished data)
- SWI/Snf: Yes (45)
- CBP: Present in stressed cells at the site of transcription (87); binding to hsr0-n transcripts not known
- Sam68: Not known
- Srp30c: Not known
- SF2/ASF: Not known
- C2PA: Not known
- Simplekin: Not known
- SC35: Not present
- Hsp70: Reported to be present at the hsr0 locus after heat shock in D.hydei but not in D.melanogaster (48)
- snRNPs: Not present in omega speckles or at the hsr0 locus (13)
- Sxl: Present in omega speckles and exclusively at the hsr0 locus after heat shock (46)
- Bx34C (Tpr): Present after heat shock at the site (45)
- Nona, PEP: Present after heat shock at the site (45)

The hnRNP M and the hnRNP A1-associated protein HAP are present in the nSBs (57), but not the hnRNPA1, A2, C1/C2, I and U (64). Association with the sat III transcripts is not known.

Yes for HSF1 (21), HSF2 (85) and HSF4b (86).

No evidence that any of these factors is associated with the sat III transcripts.

Present in the nSBs in addition to thousands remaining nuclear foci (23)

Not known

Not known

Present in nSBs in addition to a diffuse population in the nucleoplasm (23)

Present in the nSBs (24)

Present in the nSBs (24)

Present in the nSBs (24) and associated to sat III transcripts (56,57)

Yes (88)

Yes (89)

Not present (24)

Not present (23,85)

Not detected in the nSBs but associated with the sat III transcripts (56)

Not applicable
and retain proteins in a particular place, rather than to degrade them when transiently not in use. Indeed, the later solution may actually lead to a saturation of the degradation machinery in stressed cells, which are already loaded with stress-induced mis-folded proteins awaiting degradation. In addition, once the environmental conditions become normal, the different functions need to be rapidly recovered. A rapid release of the sequestered proteins would facilitate this recovery without any requirement of de novo transcription and protein synthesis. Since transcripts like hsrw are extremely rapidly produced in response to changes in the environment in the cell, they can effectively provide a dynamic sink for proteins that need to be transiently withdrawn from active compartment. This rapid response is another attractive feature of the RNAs that function as regulatory molecules.

Several examples are known where specific RNA molecules bind with specific proteins and thereby regulate their activities through sequestration. The transcripts of the mutated DMPK gene, responsible for myotonic dystrophy of types 1 and 2 (67,68) contain numerous extra-CUG repeats; these transcripts remain in the nucleus and accumulate as foci. Interestingly, specific splicing factors of the muscleblind (MBNL) family, which display a high affinity for CUG repeats, are sequestered within these transcript foci, thereby altering the splicing pattern of several other transcripts and cumulatively these lead to the disease phenotype (69).

This hypothesis can be extended further to a model of regulation of transcription factor activity via association with particular chromatin loci. Indeed, similar to the stress-induced accumulation of HSF1 on the 9q12 locus, several transcription factors have been shown to accumulate under certain circumstances on specific heterochromatic regions, through a direct DNA–protein interaction with repeated sequences. This is the case for example for the C/EBP proteins during adipocyte differentiation in mouse (70,71) or for the GAGA factor in Drosophila (72,73). Thus the accumulation of certain key regulatory factors within nSBs or at the hsr site in stressed cells also appear to represent a way to regulate their availability and their function at particular time points of cell life.

Functional analogues of sat III and hsrw transcripts in other species?

What about other species? All species of the genus Drosophila that have been examined have a functional homolog of the hsrw locus (32,34,74). A bioinformatics analysis of the hsrw homologues in recently sequences genomes of several species of Drosophila reveals remarkable similarities in the locations of the ORF-omega, the 5' and 3' splice sites, although the base sequence itself shows high variability (E. Mutt and S. C. Lakhota, unpublished data). In view of the high sequence divergence at this locus in different species, it is suggested that the structure of the transcripts rather than their sequence is conserved. One of the major heat shock puffs in Chironomus genus also produces non-coding transcripts reminiscent of Drosophila hsrw transcripts (48). In mammalian cells, a functional analog of sat III and/or hsrw transcripts may be expected to exist, even though nSBs have only been observed in human cells so far. In monkey cells, focal concentration sites of HSF1 are also observed in the nucleus of heat-shocked cells, strongly suggesting the existence of similar structures in these species (C. Jolly and R. I. Morimoto, unpublished data). It is interesting to note that the expression of several non-coding transcripts is induced by stress. In mouse for example, heat shock induces the transcription of SINE B2 repetitive elements (75). Stress also induces an increase in Alu and SINE transcription in human, mouse, rabbit and silkworm (75–79). However, unlike hsrw and sat III transcripts, these stress-induced repeated transcripts do not display a particular nuclear focal localization. Nevertheless, they can have comparable functions in titrating key regulatory factors to control their activity. This is the case for example for the stress-induced B2 transcripts in mouse, which interact directly with RNA polymerase II to inhibit transcription during heat exposure (80,81).

CONCLUDING REMARKS

The acts of transcription and post-transcriptional processing of nascent RNA involve myriad of proteins in specific combinations to produce the highly regulated transcriptome unique to a given cell at a given time. These highly dynamic processes obviously necessitate the existence of specific and dynamic systems that can precisely withdraw or release the given sub-set of the protein factors for a particular act. It appears that the sat III and the hsrw transcripts are but two examples of a large variety of non-coding RNA species that actually exist in different cell types to provide dynamic systems to sequester and release specific protein factors in normal and/or stressed cells. It is likely that specific non-coding RNA species act as platforms for the association of specific proteomes with each of the various nuclear sub-structures. A proactive search for the existence of such RNA species and detailed characterization of their nature and regulation will help us understand how cells can rapidly adapt to the ever-changing internal and external environments.

ACKNOWLEDGEMENTS

We thank Dr Claire Vourc’h for critical reading of the manuscript. Research on the hsrw locus in S.C.L.’s laboratory has been supported by research grants from the Department of Science & Technology, Government of India, New Delhi. C.J. is a member of the laboratory of Dr Vourc’h, who is granted by the French Ministère de la Recherche, the Région Rhône-Alpes, and the Association pour la Recherche sur le Cancer (grant 3449). Funding to pay the Open Access publication charges for this article was provided by the Association pour la Recherche sur le Cancer. Conflict of interest statement. None declared.

REFERENCES


