MAR/SAR ELEMENTS FLANK THE RAT HST70 GENE TRANScription UNIT

WIESŁAWA WIDŁAK1* and PIOTR WIDŁAK2
1Department of Tumor Biology and 2Department of Experimental and Clinical Radiobiology, Center of Oncology, Maria Skłodowska-Curie Memorial Institute, Wybrzeże AK 15, 44-100 Gliwice, Poland

Abstract: The rat hst70 gene is specifically expressed in spermatocytes and spermatids. This tissue-specific expression of the gene is primarily mediated through cis-acting elements located within the 0.4 kb segment upstream of the coding region, including two transcription initiation sites. Here, we study the 5' and 3' distal elements flanking the hst70 gene and find that they possess structural motifs characteristic of MAR/SAR elements, and exhibit enhanced affinities for nuclear matrix binding in vitro. Such elements bind efficiently to matrices from either the testis or the liver, i.e. tissues where the gene is either fully active or repressed, although one subfragment in the 5' region was identified as exhibiting testis-specific interactions. Surprisingly, the activity of the CAT reporter gene was repressed in testis-transient transfection assays when the hst70 promoter sequences were extended into the 5’ MAR/SAR.

Key Words: Heat Shock Protein, Testis-specific, Chromatin, MAR/SAR

INTRODUCTION

The rat hst70 gene belongs to the hsp70 multigene family of heat shock or ‘stress’ genes, and is specifically expressed in pachytene spermatocytes [1-5]. The gene encodes a molecular chaperone protein crucial for the development of male germ cells. In some somatic tissues (e.g. the liver) hst70 is fully repressed, while in some other tissues (e.g. the brain), the hst70 promoter is active and low but detectable amounts of the hst70 transcripts are present [6]. The molecular mechanisms responsible for the regulation of hst70 expression are only partially known. In our previous investigations, we determined that the transcription of

*Corresponding author: wwidlak@io.gliwice.pl
Abbreviations used: CAT - chloramphenicol acetyltransferase; hsp - heat shock protein; hst - hsp-related testis-specific; MAR/SAR - matrix/scaffold attached region; UTR - untranslated region.
the hst70 gene is initiated at two main start sites: T1 and T2, located 351 and 116 bp upstream of the ATG codon, respectively (only the latter site is preceded by a canonical TATA box) [7].

Chromatin structure is an important factor involved in the regulation of gene expression. The packaging of DNA into chromatin starts with the creation of the polynucleosomal chain, which is further folded and looped into various higher order structures that form structural and functional domains within interphase chromosomes. The organization of chromatin domains seems to be maintained by the anchorage of specific DNA sequences (MAR/SAR) to a protein network of the nucleoskeleton. Such skeletal structures can be purified after the removal of the majority of the DNA and chromatin proteins (e.g. after the treatment of the nuclei with nucleases and high salt buffers); they are termed the “nuclear matrix”. The nucleoskeleton/nuclear matrix and its interacting DNA elements are postulated to be involved in nuclear organization, the regulation of DNA replication and gene expression. It has been hypothesized that some MAR/SAR elements might be permanently attached, while others may be regulated and dynamic with respect to attachment. The former might act as topological barriers between chromatin domains harboring one or a group of related genes. The dynamic MAR/SAR elements might collaborate with other regulatory elements of replication/transcription/recombination (reviewed in: [8-11]). In fact, MAR/SAR elements frequently co-localize with transcriptional regulatory sequences, such as enhancers and insulators. MAR/SAR elements have been found within either introns or proximal and distal sites flanking the 5’ and 3’ ends of genes (reviewed in: [11, 12]). Interestingly, we earlier provided evidence that MAR/SAR elements encompassed the transcription unit of the rat heat-inducible hsp70.1 gene, another member of the HSP70 family [13]. Here, we describe MAR/SAR elements encompassing the rat hst70 gene transcription unit and delineate their physical map and tissue-specificity for matrix interactions.

MATERIALS AND METHODS

DNA plasmids
A 7.5 kb-long genomic region containing the transcription unit with the 5’ and 3’ end flanking sequences of the rat hst70 gene was cloned as a set of three p121 plasmids [3]. The 3.3 kb-long fragment lying upstream the hst70 transcription unit [HindIII(-3418)-DraII(-62)] was cloned into BlueScript™ as the pHST(3418/62)-BS plasmid. The pHST(3418/62)-CAT, pHST(897/62)-CAT, pHST(368/62)-CAT and pHST(279/62)-CAT plasmids were constructed by inserting defined fragments (HindIII(-3418)-DraII(-62), HindIII(-897)-DraII(-62), BstXI(-368)-DraII(-62) and EcoRV(-279)-DraII(-62), respectively) in front of the promoterless CAT gene cloned in the plasmid pBLCAT6 (GenBank accession no. M80484). The structures of the DNA fragments used for analyses are shown in appropriate figures. The coordinates of the restriction sites refer to the A(+1) in the ATG codon of the hst70 gene.
Electroporation of the testis

Electroporation was done as described in detail elsewhere [14]. Male FVB/N mice (7-9 weeks old) were anesthetized and the testes were exposed under a dissecting stereomicroscope. Supercoiled plasmid DNA (about 150 µg) was injected into the seminiferous tubules (at 3-5 sites in each testis) using an injection glass pipette connected to Eppendorf Microinjector 5242. Immediately after injection, the testis was held between a tweezers-type electrode, and square electric pulses were applied 8 times with an ElectroSquare Porator ECM830 at 40V for 50 msec. After the procedure, the skin was stitched and the mice were maintained for the following 24 h. To verify the efficiency of electroporation, the pGL3 vector (containing the luciferase gene under the control of an SV40 promotor) was co-tranfected with the pHST-CAT vectors.

Analysis of the CAT activity

Tissue lysates were prepared as described earlier [14]. Aliquots of 100 µl of the extracts (10-100 µg of proteins) were added to 100 µl of the reaction mixtures containing 0.25 M Tris/HCl (pH 7.8) with 1 mM EDTA, 4 mM acetyl-CoA (Sigma) and 6.25 µCi of \[^{14}C\] chloramphenicol (2.5 mCi/mmol, ICN), and the samples were incubated for 2 h at 37ºC. The acetylated forms of chloramphenicol were separated via thin layer chromatography. The CAT activity was expressed as a percentage of acetylated products formed per hour per milligram of extract protein, as described previously [15].

Purification of the nuclear matrices

Nuclear matrices from the testes and liver of 2 month-old male Wistar rats were isolated by the DNaseI/"high salt" method, as described in detail elsewhere [13]. Briefly, the tissues were homogenized in a hypotonic buffer, and nuclei were recovered by centrifugation, washed with 1% Triton X-100, stabilized with 5 mM CuCl\(_2\) and washed again with buffer containing 50 mM NaCl and 5 mM MgCl\(_2\). The nuclei were then treated with DNaseI and extracted with 0.5 M NaCl followed by 2 M NaCl to obtain the residual nuclear matrix fraction.

In vitro DNA binding assay

The binding of the DNA to the nuclear matrices was assayed according to the modified procedure of Cockerill & Garrard [16]. The appropriate restriction fragments were \(^{32}P\)-5′-end-labeled using polynucleotide kinase and repurified from polyacrylamide gels. About 50 µg of the matrix proteins was suspended in 100 µl of binding buffer [50 mM NaCl, 2 mM EDTA, 0.25 M sucrose, 10 mM TrisCl (pH 7.4) and 0.25 mg/ml of BSA], then mixed with 20 ng of each of \(^{32}P\)-end-labelled DNA probes and a different amount of non-radioactive competitor DNA (sonicated Escherichia coli DNA, usually 20 µg/ml). After 1 hour’s incubation at 25ºC, matrices were recovered by centrifugation. Both the pellets, containing the matrix-bound fractions, and the supernatants, containing the matrix-unbound fractions, were treated with proteinase K/SDS, extracted with
phenol/chloroform and ethanol precipitated. The purified DNAs were separated electrophoretically. The gels were dried and autoradiographed, and the DNA bands were excised and quantitated in a scintillation counter.

**Electrophoretic mobility-shift assay**

Purified nuclei were incubated for 30 minutes on ice in a buffer consisting of 0.4 M NaCl, 0.2 M Tris/HCl (pH 7.6), 1.5 mM MgCl₂, 0.5 mM EGTA, 1 mM DTT, 5% glycerol and centrifuged for additional 30 minutes at 20,000g. The resulting supernatants were referred to as nuclear extracts. Extracted proteins (5 µg) were incubated for 30 minutes at 4°C with 5’-end-radiolabeled DNA in the presence of 1, 3 and 10 µg of sonicated *E. coli* DNA as a non-specific competitor in a buffer consisting of 20 mM Tris/HCl (pH 7.6), 5 mM MgCl₂, 0.5 mM EDTA, 1 mM DTT, 5% glycerol and 150 mM NaCl (final volume of 20µl). Protein-DNA complexes were separated from free DNA during electrophoresis on a native 4% polyacrylamide gel in a 0.5 x TBE running buffer.

**Computer-based analysis of the DNA sequences**

The nucleotide sequence of 7.5 kb containing the rat *hst70* transcription unit, and its 5’ and 3’ flanking regions is available in the GenBank database (accession no.: X15705). The DNA regions potentially involved in attachment to the nuclear matrix were identified using MAR-VIZ 1.0 software (http://www.futuresoft.org/MAR-Wiz/). Potential transcription factor binding sites were identified using TFSEARCH 1.3 software (http://www.cbrc.jp/research/db/TFSEARCH.html).

**RESULTS AND DISCUSSION**

We first used a computer-based analysis to predict the positions of potential MAR/SAR elements within the 7.5 kb rat *hst70* gene locus. Although there is no rigid MAR/SAR consensus, some specific structural features are common for these elements: AT-richness, topoisomerase II binding/cleavage sites, TG-repeats, polypurine or polypyrimidine stretches and palindromes [12]. These common features have allowed the development of an algorithm for MAR/SAR prediction [17]. As shown in Fig. 1A, we found that potential MAR/SAR elements flank the gene transcription unit (roughly upstream of the -1400 base and downstream of the +2500 base).

To determine if the theoretically predicted MAR/SAR elements indeed possessed their corresponding activities, we next performed *in vitro* nuclear matrix binding assays, which operationally define MAR/SAR elements [16]. We compared the relative efficiencies of matrix binding for the different DNA fragments covering the 7.5 kb locus (Fig. 1B).

Nuclear matrices were purified either from cells that efficiently express the endogenous *hst70* gene (from the testes of adult animals) or from cells where the
Fig. 1. MAR/SAR elements encompass the hst70 gene transcription unit. A – Matrix binding potential in the hst70 gene domain predicted by computer-based analysis. B – The structure of the DNA probes used in the nuclear matrix in vitro binding assay. The positions of particular restriction sites are shown: H, HindIII; B, BamHI; P, PstI; M, MspI; K, KpnI. C – Binding of different DNA probes (corresponding to the hst70 gene transcription unit, 5’ and 3’ end flanking sequences, and pUC19 plasmid DNA) to the nuclear matrix proteins from rat liver and testis. D – Binding of DNA sub-fragments to the nuclear matrix proteins. P, pellet – matrix-bound fractions. SN, supernatant – matrix-unbound fractions. Arrowheads mark the positions of particular probes.
gene is fully repressed (from the liver). Fig. 1C shows the data from an experiment where three DNA fragments encompassing the hst70 gene domain roughly corresponding to the transcription unit, and the 5’ and 3’ flanking sequences (-896 to +2178, -3418 to -897 and +2179 to +4102, respectively) were associated with nuclear matrices. The DNA fragments corresponding to the 5’ and 3’ flanking sequences bound to nuclear matrices from either the testis or liver more efficiently than the hst70 transcription unit (the matrix affinity of the later sequences was similar to that of pUC19 plasmid DNA; bottom part of the panel). This result was further confirmed when the 5’ and 3’ flanking sequences were additionally cleaved into smaller fragments (Fig. 1D). Thus, complete agreement between the theoretical and experimental data was obtained.

To perform a more detailed analysis of the MAR/SAR elements located upstream of the hst70 transcription unit, the 2.5 kb region (-2613 to -120) was cleaved into 4 smaller fragments (Fig. 2A). Fig. 2B shows that matrices from either liver or testis tissues exhibit a low affinity for binding promoter-proximal DNA fragments (fragment δ, -896 to -120) and a higher affinity for binding distal DNA fragments (fragments α and β, -2613 to -1661 and -1660 to -1195, respectively). Interestingly, the matrices from testis tissue bound the short γ fragment (-1194 to -897) more effectively as compared to the matrices from liver tissue.

To quantitate differences in the ability of particular DNA fragments to bind nuclear matrices, an additional assay was performed. Each of four fragments above was separately complexed with matrices in the presence of different amounts of a non-specific competitor (the excess of the competitor ranged from 100-fold to 6000-fold), and then the radioactivity of each of the fragments was directly counted in both the pellet (matrix bound) and supernatant (matrix unbound) fractions. The data presented in Fig. 2C confirms that distal fragments (α and β) bound equally strongly while the proximal fragment (δ) bound equally weakly to the nuclear matrices from either the liver or testis. By contrast, fragment γ (-1194 to -897) bound 2-3 times more efficiently to matrices from the testis of adult animals as compared to matrices from the liver. This indicates the possibility that the DNA sequences located in this fragment, which in fact have a low predicted MAR potency, are involved in tissue-specific interactions with nuclear proteins. Since several regulatory DNA-binding proteins could be found both in a matrix-bound state and in a soluble state [18], we further studied the interactions of this particular DNA with testicular proteins using a gel-retardation experiment. In fact, the protein(s) that bind to the studied fragment γ were clearly more abundant in nuclear extracts obtained from the rat testis than in those from the liver cells. Such testicular factors formed specific complexes with the γ fragment (arrowheads in Fig. 2D).
Fig. 2. The binding of DNA sequences located upstream of the hst70 gene promoter to nuclear proteins from rat liver and testis. A – The structure of the DNA probes used for analyses. B – Binding of the DNA probes to the nuclear matrix proteins in vitro. C – The quantification of interactions between the DNA probes and nuclear matrices. Complexes were formed in the presence of different amounts of a competitor: 1.5 to 60 µg/probe. Fractions bound to nuclear matrices are shown (as a percentage of input DNA). D – The formation of protein-DNA complexes between probe γ and soluble proteins from the nuclear extracts (NE) analyzed via EMSA. Complexes were formed in the presence of different amounts of a competitor.

The tissue-specific and developmentally regulated activity of the hst70 gene promoter has been extensively studied using the CAT and GFP reporter genes in transgenic mice [4-7, 14]. We have shown that the shortest fragment of the hst70 gene promoter conferring high and specific activity was a fragment from -368 to -62 bp upstream of the ATG translation start codon. When this promoter fragment was extended about 0.5 kb 5’ further upstream, similar specific activity was detected [4]. These observations raise the question as to whether even more
activity might be observed if we were to attach sequences extending all the way upstream into the MAR/SAR elements that we identified above. To address this question we used another experimental model to study the effects of 5’ distal sequences on the activity of the hst70 gene promoter. Different fragments of its promoter and 5’ flanking sequences were placed in front of a CAT gene (Fig. 3A); the resulting constructs were electrotransferred into testes in vivo, and then CAT activity was assayed in cellular extracts (Fig. 3B). A strong level of reporter gene activity was detected when the -368 to -62 promoter fragment was analyzed; similar results were obtained with transgenic mice. The extension of the promoter fragment for an additional 0.5 kb (-897 to -62) had little effect on this activity. However, further extension of the analyzed fragment for an additional 2.5 kb (-3418 to -62), to encompass the MAR/SAR elements, drastically reduced the activity of the hst70 promoter. Similar results were obtained when somatic cells (FTO rat hepatoma cells) were electroporated in vitro with the same constructs (data not shown). These results indicate that when the MAR/SAR elements are included in the upstream region, the hst70 gene promoter is unexpectedly significantly downregulated. However, transiently transfected constructs were not integrated into the genome and could be regulated differently to the endogenous hst70 gene.

Fig. 3. The expression of the CAT reporter gene under control of the rat hst70 gene promoter in testes electroporated in vivo. A – The structure of the hst70 gene transcription unit. UTRs (grey boxes), coding sequences (white box), the T1 and T2 transcription initiation sites (arrows) are shown. The bottom part shows the position of particular restriction sites: H, HindIII; E, EcoRV; D, DraI; Bs, BstXI. B – The structure of different constructs and their activity in electroporated testes. The values shown are the means from 4-6 independent experiments.
In an attempt to further understand the effects of various upstream sequences on the hst70 promoter activities, we identified potential transcription factor binding sites within the 5’ end (and the 3’ end) flanking sequences using a computer-based analysis. The most abundant of such potential binding sites were specific for development regulators: homeo-domain factor CdxA, HMG-box containing testis differentiation factor SRY, and myeloid zinc finger protein MZF1. Additional frequent potential TF-binding sites were specific for the GATA-1 and Oct-1 transcription factors. Interestingly, fragment γ (-1194 to -897), which showed enhanced affinity for testicular proteins, in addition to such abundant sites, contained an unique HSE-like element (the 5’ end at -1183). Heat shock elements (HSE) bind heat shock factor (HSF), which activates the transcription of heat shock genes [19]. Among the three members of the HSF family identified in mammals, HSF2 is the most abundantly expressed in mouse testes [20]. HSF2 functions as a regulator of hsp gene expression under non-stress conditions, particularly in cells involved in the differentiation and development process [21], and was reported to interact with the promoter of the hsp70.2 gene, the mouse counterpart of the hst70 [22]. Two HSE-like elements are present upstream of the hst70/hsp70.2 gene: a proximal one between the T1 and T2 sites, and a distal one 0.8 kb upstream of the T1 transcription initiation site. Although in vitro data indicates that the binding of the HSF2 to both HSE-like elements can promote their association and the formation of a DNA loop [23], the functional importance of such potential looping requires further investigation. Several putative regulatory elements have been identified in distal regions flanking the hst70 gene transcription unit. However, their role in the regulation of the hst70 gene expression in vivo remains unclear. There is no evidence for tissue-specific proteins recognizing the MAR/SAR elements present in such regions. In fact, the MAR/SAR elements flanking the hst70 gene transcription unit bind with similar efficiency to the nuclear matrices from the testis and liver, tissues where the gene is either active or repressed, respectively. The regulatory regions of potentially active or active genes are frequently revealed in a form of nuclease hypersensitive sites [24]. We searched for DNaseI hypersensitive sites within the hst70 transcription unit and its 5’ flanking sequences. Hypersensitive sites were detected around the T1 transcription initiation site in cells expressing the gene. In contrast to the promoter region, no such site was detected in the 5’ flanking sequence (up to the -2614 nucleotide), neither in cells with an active or repressed gene (data not shown). This data suggests that possible matrix attachments could rather be involved in the formation of the chromatin domain harboring the hst70 gene than in the tissue-specific regulation of its activity. Interestingly, the analysis of the rat chromosome 6 region (6q24) where the hst70 gene is located (GenBank accession no. NW0477661) revealed two putative transcription units in the immediate surroundings of the hst70 gene. Both ORFs identified using the automated computational analysis GNOMON were placed about 10 kb upstream and downstream of the hst70 transcription unit. In general, the length of the chromatin loops is estimated to range between
The length of the hypothetical domain containing the \textit{hst70} gene is obviously in the lower end of this range. However, the fact that the gene contains only one short intron and the hypothetical necessity of functional separation from neighboring genes might contribute to this phenomenon.

\textbf{Acknowledgements.} The authors wish to thank Prof. W.T. Garrard for his comments and help in the preparation of the manuscript. This study was supported by the Polish Ministry of Scientific Research and Information Technology, Grant No. 3 P04A 033 23.

\textbf{REFERENCES}


