ISSN 0006-2979, Biochemistry (Moscow), 2008, Vol. 73, No. 8, pp. 839-852. © Pleiades Publishing, Ltd., 2008.
Original Russian Text © R. I. Dmitriev, M. I. Shakhparonov, N. B. Pestov, 2008, published in Biokhimiya, 2008, Vol. 73, No. 8, pp. 1045-1060.

REVIEW

Structure and Function of MYST1 Histone Acetyltransferase in the Interactome of Animal Cells

R. I. Dmitriev*, M. I. Shakhparonov, and N. B. Pestov

Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, 117997 Moscow, Russia; E-mail: enisseisk@gmail.com

Received September 18, 2007
Revision received January 22, 2008

Abstract—The major function of protein MYST1 is acetylation of histone H4 at the K16 residue. This modification is essential for chromatin remodeling and is used for regulation of gene expression in eukaryotes. MYST1 is a part of multiprotein complexes that accomplish functions of male X-chromosome activation and thereby functions of dosage compensation in drosophila and, in mammals, global acetylation of histone H4 K16. Recently, novel functional links between MYST1 and proteins ATM and p53 have been observed, and it is recognized that MYST1 plays a role in tumor suppression mechanisms. In the present review, we examine novel data about functional composition and mechanisms of MYST1-containing complexes. Interplay between MYST1 and other components of the animal cell interactome is also discussed.

DOI: 10.1134/S0006297908080014

Key words: hampin, MYST1, MSL, compensasome, chromatin structure rearrangement

Fundamental structural units of eukaryotic chromatin are nucleosomes consisting of DNA and histones. Nucleosomes are integral participants of such nuclear processes as genomic DNA transcription and replication [1]. Changes in nucleosome structure result in alteration of chromosomal DNA accessibility [2]. Changes in chromatin structure may be ATP-dependent due to replacement of histones, incorporated in nucleosomes, or due to numerous covalent modifications of the histone N-terminal sequences. A great number of such modifications are known including acetylation, phosphorylation, methylation, ubiquitinylation, sumoylation, and poly(ADP-ribosyl)ation [3-5]. Acetylation is one of the best studied covalent modifications of N-terminal histone sequences [3]. Enzymes catalyzing this reaction are called histone acetyltransferases. In the course of this reaction the enzyme transfers acetyl group from acetyl-CoA to ε-amino group of Lys residue in the histone N-terminal part.

The nucleosome is a disc-shaped octamer consisting of two heterotetramers formed by histones H3/H4 and histones H2A and H2B. About 147 bp of genomic DNA are wound around this octamer [1]. Each histone contains a globular region that interacts with DNA and other histones within the nucleosome. At the same time N-terminal histone fragments are exposed outside the globular region and probably interact with DNA or mediate interaction between nucleosomes. Histone acetylation as a rule disturbs these interactions and causes nucleosome destabilization [3, 4] which facilitates nucleosome association with such transcription systems as the RNA polymerase II complex. Indeed, in most cases histone acetylation is associated with activation of gene transcription [3]. This covalent modification is reversible due to the existence of enzymes called histone deacetylases [6]. It should be also mentioned that some histone acetyltransferases may acetylate Lys residues in non-histone proteins.

It is assumed to classify numerous known histone acetyltransferases to different types, distinguishing, for example, GNAT, MYST, coactivators of nuclear receptors p300/CBP, $TAF_{II}250$, and TFIIIC [3].

All human autosomes are susceptible to histone H4 acetylation by Lys16 residue (H4 AcK16) and acetyltransferase of MYST family called MYST1 (alternative name is MOF) [7, 8]. It has been recently found that the disturbance of normal acetylation of K16 in histone H4 together with trimethylation of Lys20 in histone H4 is associated with early stages of tumor development [9]. Besides, it was shown that these two modifications are necessary for control of chromatin structure and probably they are most important in regulation of the high order

Abbreviations: ATM) ataxia-telangiectasia-mutated protein.

* To whom correspondence should be addressed.

chromatin structure, namely, in the nucleosome decondensation or condensation [10].

The decrease in the MYST1 protein content using RNA interference results in accumulation of cells in the G2/M stage and the cell cycle arrest [8]. It was shown that MYST1 protein is involved in the DNA repair processes due to interaction with ATM protein (ataxia-telangiectasia-mutated protein), the key link in initiation of the cell signal pathways triggered after effects of ionizing radiation or other agents causing DNA damage. The decrease of intracellular content of functionally active MYST1 correlates with lowering the ATM functional activity and results in lowered cell survival in response to ionizing radiation [11].

However, the MYST1 specificity to Lys16 of histone H4 is not absolute, because in experiments *in vitro* the protein is also able to acetylate histones H3 and H2A, whereas *in vivo* only modification of histone H4 is specific [8]. Apparently, interaction with certain proteins is necessary to observe such specificity.

A probable candidate for the role of such protein is the recently identified protein hampin [12] (alternative name MSL1 from male specific lethal 1) due to at least two reasons. First, it was found that proteins MYST1 and hampin are able to interact with each other in the yeast two-hybrid system [13] and second, this interaction has been recently confirmed by affinity chromatography [14] that showed the existence of a multisubunit complex "human MSL" (see below) exhibiting acetyltransferase activity towards Lys16 of histone H4. It was shown using RNA interference that the decrease in the MYST1 content in human cells results in lowered Lys16 acetylation in histone H4. Besides, the decrease in hampin content caused a similar and even more pronounced effect. These results allow us to suppose that hampin is responsible either for modulation of MYST1 activity or for its correct association with chromosomes. Most interesting is that the hampin–MYST1 interaction is conserved from drosophila, where it takes place within MSL complex, to human being [14, 15]. This review deals with the protein MYST1 interactions with different proteins and discussion of their functional significance for its regulation.

INTERACTIONS BETWEEN MSL1 AND MOF WITHIN MSL COMPLEX IN DROSOPHILA

The MSL1 protein is a component of a ribonucleoprotein complex compensasome whose function is equalizing expression of X chromosome genes between males and females of the fruit fly drosophila [16]. The drosophila male cells have one X chromosome while cells of female fly have two such chromosomes. To obtain equal total expression level of male and female X chromosome genes it is necessary to have double transcription activity of male X chromosome. The compensasome function is twofold increase in transcription level of male X-chromosome genes [16].

Genetic studies of the mechanism of the gene dose compensation in drosophila resulted in identification of a number of mutations leading to nonviable males. The "male-specific lethal" mutations resulted in death of males already at early stages of embryonic development but had no essential effect on the female viability [17]. Already by the beginning of 90s it became clear that all these mutations involve genes, products of which are important for gene dose compensation in males [18]. According to current concepts, the compensasome includes proteins MSL1, MSL2, MSL3, MLE, MOF (histone acetyltransferase homologous to MYST1), JIL-1, and two non-coding RNA: roX1 and roX2 [16].

Gene encoding MSL1 of drosophila was cloned in 1993 [19] and it was shown that its product is a protein consisting of 1039 a.a. and devoid of any significant homology with known proteins. The peculiarity of its primary structure is the presence in N-terminal part of two rather acidic clusters, almost completely consisting of aspartic and glutamic acid residues. Immunofluorescence labeling of drosophila polytene chromosomes has shown that MSL1 is localized mainly on the male X chromosome. At the same time in much lower amounts this protein is also detected in females [20].

Investigations carried out by different research groups during the next decade allowed them to characterize in detail the structure and function of MSL1 protein. MSL1 appeared to be one of main components of compensasome and it interacts with MSL2, MSL3, and MOF proteins which is necessary for assembly of compensasome. In this case "main" regions of MSL1 molecule are its N- and C-terminal parts: N-terminal part is responsible for interaction with MSL2 [21], while C-terminal domain is responsible for interaction with MSL3 and MOF [22].

It was shown that the MSL1 fragment from 766 to 939 a.a. containing the conserved PEHE domain interacts with MOF, whereas the fragment between 973 and 1039 a.a. is responsible for interaction with MSL3 [23].

Li et al. [21] studied in detail the structure of the MSL1 N-terminal fragment. It was found that this fragment contains three arranged in-line regions with different structural motifs: "basic", "Gly-rich", and the leucine-zipper like. The "basic" motif is responsible for MSL1 binding to so-called "high affinity" sites on the male X chromosome, i.e. to such chromatin sites that are targets for binding the MSL complex proper. This region of the sequence is conserved among all members of *Drosophila* genus [21]. The leucine-zipper like motif is also necessary for interaction with male X chromosome, but already via interaction with MSL2. At the same time, the "Gly-rich" motif stimulates the MSL1 molecule dimerization both *in vitro* and *in vivo*.

Protein MSL2 was identified in 1995 [24], its gene is expressed in drosophila of both sexes but only male

mRNA is translated. In the norm translation of mRNA encoding this protein is under control of SXL (sex-lethal) protein present only in female drosophila. The RNA-binding protein SXL interacts with 5′- and 3′-nontranslated regions of the MSL2 mRNA and inhibits translation of this protein [25]. This inhibition was shown to be carried out in two ways: the SXL binding to 3′-nontranslated region disturbs mRNA interaction with 43S ribosomal pre-initiation complexes, while SXL bound to 5′-nontranslated region prevents these complexes from finding the site of the MSL2 mRNA translation initiation [25]. Ectopic expression of MSL2 in drosophila females causes a number of defects in the development and lowers the female viability [26].

MSL2 within compensasome interacts with MSL1 via its own RING-finger domain. It was shown using the yeast two-hybrid system that the first site of Zn^{2+} binding is responsible for this interaction. It is supposed that the second Zn^{2+}-binding site within the MSL2 RING-finger domain is involved in interactions with other proteins, because mutations in this region of protein molecule do not influence its binding to MSL1 [27].

The polytene chromosome labeling by antibodies to MSL1 and MSL2 has shown that both proteins have practically similar localization on the male X chromosome. It is interesting that in the absence of other compensasome components MSL3 and MLE proteins MSL2 and MSL1 retain their localization of X chromosome [28]. However, the MSL2 protein sequence does not contain any known DNA- or chromatin-binding sites and has only the MSL1-binding site within its own RING-finger domain. Thus, it can be supposed that MSL2 is associated with chromatin via interaction with MSL1. At the same time it appeared that the MSL2 protein removal using RNA interference in cell culture disturbs histone H4 acetylation by Lys16 residue and following double increase in the transcription level of X chromosome genes [28]. But this does not cause the disturbance in their viability [29]. Thus, the AcK16 modification of histone H4 is not necessary for cell viability and is important for development of drosophila males.

Like in the case of MSL1, expression of the MSL3 protein gene [20] is detected both in males and females but it is higher in males. The 57 kD protein (512 a.a.) has two domains, N-terminal CRD domain (chromo-related domain) and MRG domain (MORF4-related gene) [30]. It was supposed earlier [20] that just CRD domain, homologous to chromodomain of many DNA-binding proteins, is responsible for MSL3 binding to X chromosome, but it turned out later that MSL3 binds chromatin only via interaction with MSL1 [31]. Mutation in MSL3 with deletion of MSL3 N-terminal fragment including CRD domain does not cause the disturbance of the compensasome functional activity [31]. MSL3 is able to bind single-stranded nucleic acids, preferably RNA, such as roX2. Practically the whole N-terminal half of MSL3, including both CRD and a part of MRG domain, is necessary to bind RNA [31].

The MSL3 association with MSL1 in the complex with MOF results in increase of acetyltransferase activity of MOF protein, although MSL3 and MOF either do not interact with each other or affinity of their interaction is very weak. Moreover, MOF is able to acetylate MSL3 (by the Lys116 residue) and MSL1 to the much lower extent, as shown in [22]. The MSL3 acetylation lowers its ability to bind roX2 [32]. MSL3 also interacts with histone deacetylase RPD-3 able to deacetylate MSL3 [33].

The MRG domain is important for MSL3 binding to X chromosomes. Mutations in CBD domain of MSL3 protein disturbed expression of X chromosome genes [34]. It was shown that CBD domain is also able to interact with chromatin. This interaction may be important for directed acetylation by H4 K16 of specific sites on X chromosome, probably, containing genes whose expression should be doubled.

The MLE (maleless) protein is historically the first identified protein from the compensasome complex (1991) [18]. MLE belongs to the DEAH helicase superfamily. This protein has the RNA/DNA helicase and ATPase activities like human RNA helicase A [34]. Although MLE is relatively weakly associated with the MSL complex, its involvement in the complex interaction with roX2 RNA is supposed. It was shown that MLE is able to activate transcription of roX2 gene by binding to its 5′-region [35].

It was noticed long ago that the compensasome localization on X chromosomes (maybe its functional activity as well) positively correlates with H4 histone acetylation by the Lys16 residue [36]. It was shown in 1997-2000 that the histone acetyltransferase MOF (males absent on the first) is responsible for this acetylation [37-39]. MOF protein of drosophila consists of 827 a.a., two domains can be distinguished in its structure—the "chromodomain"-like (CD) domain and the histone-acetyltransferase one (HAT) that includes C2HC-Zn-finger motif and the acetyl-CoA-binding site. The MOF protein structure and function were characterized in detail [22, 39, 40] and it was shown that CD domain is able to bind roX RNA, but this interaction is not necessary for MOF localization on X chromosome. Evidently MOF localization on X chromosome requires interaction of its C2HC-Zn-finger region with MSL1. Moreover, MOF interaction with the MSL1–MSL3 complex is accompanied by the increase in MOF acetyltransferase activity towards nucleosomes or free histones [39]. MOF is present both in females and males approximately at the same level, which means that the involvement in compensasome activity is not its unique function.

It was supposed in 1999 that the compensasome functioning also requires RNA of roX1 and roX2 (RNA on the X), co-localized on male X chromosomes together with compensasomes: deletion of these RNA genes disturbed the MSL complex binding to X chromosomes in

embryos [41]. Both known roX RNAs are non-coding. They have different length (3.7 and 0.6 kb) and are almost non-homologous by their primary structure. At the same time both types of these RNA are interchangeable, i.e. the presence of at least one of these roX types is necessary for normal development of males. The use of roX deletion mutants has shown that the roX1 fragment approximately of 3.35 kb in length is insignificant for its functioning [42]. Probably, the MSL complex has low affinity to X chromosomes in the absence of roX. It is rather interesting that MSL is also able to bind to roX genes even after their transfer to autosomes [43].

The 220 bp long region within the roX gene sequence, necessary for binding the MSL complex, forms the male chromosome-specific DNase 1 hypersensitive sites (DHS). The MSL complex requires roX RNA for DHS recognition [44]. At the same time it is not known how the MSL complex distinguishes X chromosome from autosomes [45].

The chromosome kinase JIL-1 (able to phosphorylate histone H3) may be also a protein of MSL complex [46]. JIL-1 is known to co-localize on male X chromosome with MSL complex and interact with it. JIL-1 accumulation on X chromosome depends on the presence of MSL. However, JIL-1 is also associated with female X chromosomes and even with male autosomes [46]. Drosophilae with mutant JIL-1 have defects in MSL functioning such as eye pigmentation defects [47].

It is possible to imagine roughly the mechanism of the complex activity. Probably the key link in the complex is MSL1, interacting with MSL2, MSL3, and MOF. The MSL2 binding is the limiting stage for complex assembly. MSL1 has sites for interaction with MSL3 and MOF. Probably reversible acetylation of MSL3 plays a certain role in MOF binding. MLE, roX RNA, JIL-1, and possibly RPD-3 histone deacetylase are weaker bound to the complex. Binding of the complex to DNA/chromatin may occur via N-terminal region of MSL1, but selective anchoring of the complex on X chromosome will happen only due to interaction with the roX transcript/gene.

Searching targets for the MSL complex binding to X chromosome using chromatin immunoprecipitation and following hybridization on microchips [48] has shown that the complex binding is gene-specific and happens mainly within gene exons. Authors also tried to reveal consensus motifs in the gene base sequence responsible for the complex binding: all above-mentioned motifs contain dinucleotides GA and CA. After localization on X chromosome the complex activates transcription of X chromosome genes due to the MOF acetyltransferase activity to Lys16 of histone H4. It was supposed previously that individual regions of X chromosome undergo decompaction and become accessible to association with RNA polymerase II and following transcription. This hypothesis is supported by the fact that the H4 histone acetylation by Lys16 residue exhibits the effect opposite to that of ATPase ISWI that is necessary for creation of the chromatin condensed state [49].

Recently Furuhashi et al. [50] obtained additional data in favor of this hypothesis of compensasome functioning. The so-called SCF protein (DNA supercoiling factor) has been known for rather a long time now. It is necessary for formation together with topoisomerase II of negative supercoiled regions in the genomic DNA molecule. It was shown earlier that protein SCF is localized (in drosophila cells) on polytene chromosomes, i.e. it is probably involved in transcription. Furuhashi et al. studied functional aspects of this protein, as applied to the drosophila male cells, and they found that the knockdown results in the death of males due to lowering transcription levels of X chromosome-specific genes. Besides, it was shown that SCF is co-localized on X chromosome together with compensasome, and the SCF localization depends on the presence of MOF protein and is disturbed in its absence. Apparently, correct localization (or co-localization) on male X chromosome requires the presence of functionally active MOF, namely, the presence of modified K16 residues of histone H4.

Thus, the chain of events, resulting in the gene dose compensation in drosophila males with the help of compensasome, can be described as following:

– the MSL complex assembly, beginning from association of MSL1 and MSL2 proteins, their further association with MSL3, MOF, roX, and (possibly) MLE, and finishing by compensasome localization in certain sites on male X chromosome;

– exhibition of MOF protein functional activity—acetylation of K16 residue of histone H4, and as a result, the SCF protein correct localization next to MSL complex;

– decondensation of certain chromatin regions, containing X-specific genes, resulting in improvement of their accessibility for transcription system of RNA polymerase II.

SIMILARITIES AND DISTINCTIONS OF MSL1 AND ITS HOMOLOG HAMPIN

Analysis of databases of expressed nucleotide sequences shows that proteins related to MSL1 are present both in arthropods (bee *Apis mellifera*, mosquito *Anopheles gambiae*, as well as different species of fruit fly *Drosophila*) and in man [12, 15]. At the same time, no protein homologs of MSL1 are detected in such organisms as roundworms, fungi, plants, and prokaryotes.

Owing to obvious structural differences and supposed protein function the name “hampin” was proposed for MSL1 homologs in other animals [12]. However some research groups still use the name MSL1. In our opinion, this fact seems curious because mammals have no the “male-specific complex” (MSL).

It was proposed that the MSL1 gene originated from the hampin gene and underwent strong divergence during evolution of *Drosophila* genus [12]. The comparison of primary structures shows that hampin and MSL1 have only 26% identity in amino acid sequence. Moreover, MSL1 is poorly conserved even within Diptera family [12, 15]. Multiple alignment of amino acid sequences of hampin from insects and vertebrates made possible to reveal two conserved regions in its molecule which comprise the so-called PEHE domain having in its sequence the highly conserved residues P, E, H, E, and the coiled-coil region [15]. However, a shown for mammalian proteins, some hampin isoforms may have no such regions. An interesting peculiarity is the high variability of N-terminal hampin sequences, the structure of which is known only for mammals.

The main peculiarity of mammalian hampin is alternative splicing of the encoding mRNA resulting in the existence of at least five isoforms in mice [12]. This variability is the result of different exon combinations. Hampin A was first identified during a search for MSL1 homologs in mammals [15]. The protein consisting of 616 a.a. in mice, 615 a.a. in rat, and 614 a.a. in man is highly conserved (about 95% identity of amino acid sequence). Four domains can be distinguished in its sequence: a highly variable and proline-rich domain I, domain II containing a coiled-coil region, a moderately conserved domain III, and C-terminal PEHE domain IV (Fig. 1) [12]. Hampin B is 16 a.a. residues shorter than hampin A and also contains PEHE and the coiled-coil domains, whereas hampin C has no PEHE domain.

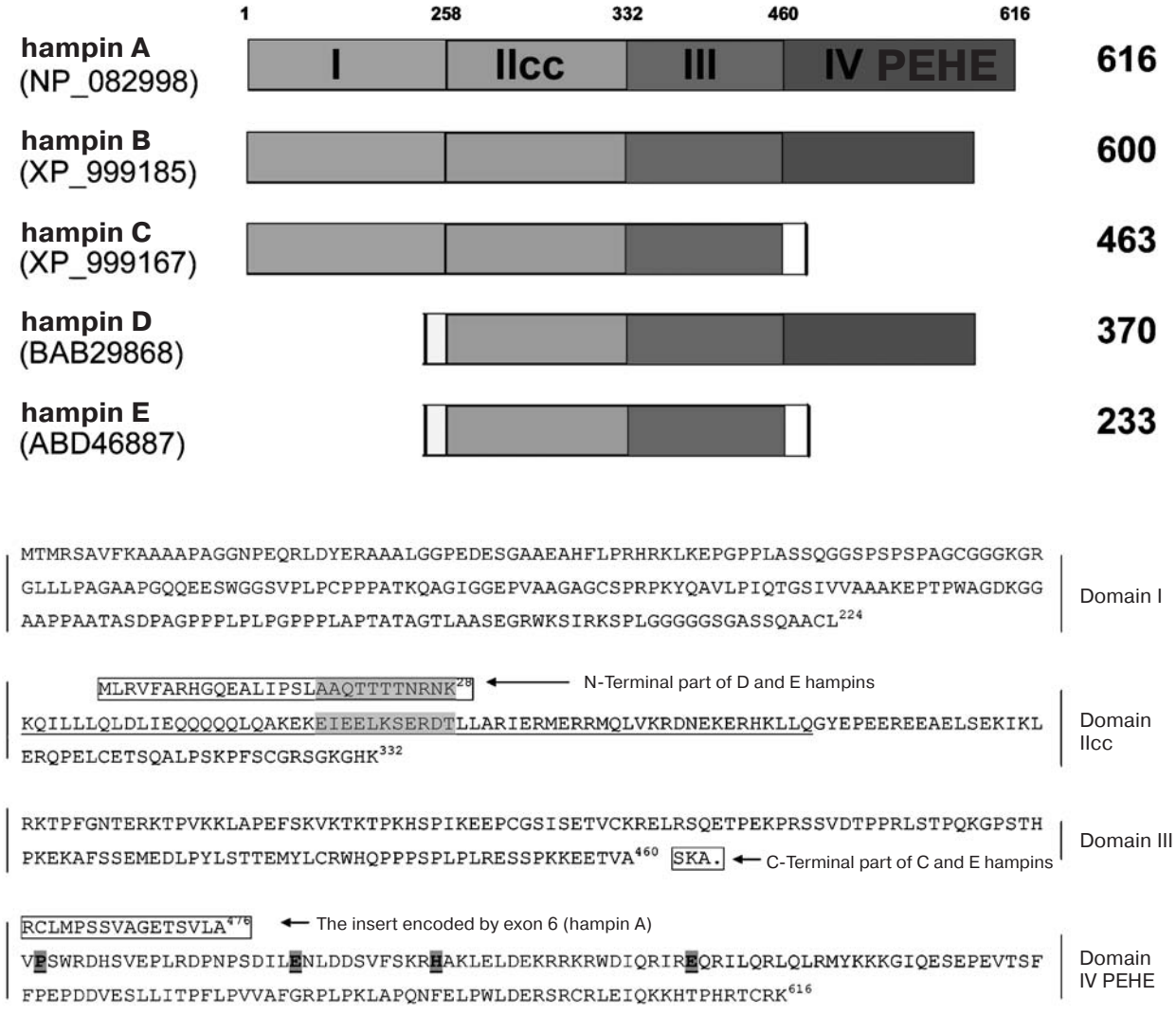


Fig. 1. Primary structure of mouse hampin isoforms. Top: schematic structures of all mouse isoforms (different domains are marked with different shades of gray) with deposit numbers in the GenBank database. The number of amino acid residues making up each protein isoform is shown on the right. Bottom: a more detailed primary structure; arrangement of domains is marked on the right. Polypeptide fragments encoded by exons 2, 5, and 6 are marked separately. The coiled-coil region is underlined. Highly conserved residues P, E, H, and E are shown in gray and in bold underlined type.

Hampin isoforms D and E differ by the presence of a highly shortened amino terminal part (it is encoded in them by exon 2 instead of exon 1 as in hampins A-C), domain 1 is absent from these isoforms, while the coiled-coil region is also shortened. The most interesting is the hampin E structure: besides domain I, it also has no PEHE domain, i.e. it is not a structural homolog of MSL1, although it is encoded by the homologous gene.

The composition of human hampin isoforms is not completely studied, but analysis of expressed genomic sequences shows that human hampin is no less diverse: at the present time the existence of hampin A is known, besides components with C-terminal sequence of mouse hampin C (SKA tripeptide), i.e. having no PEHE domain IV are also detected.

Hampin isoforms have different tissue specificity: hampins A-C are present in all mouse tissues, whereas D and E are testis-specific [12].

It is quite possible that the isoform diversity is necessary for different regulatory mechanisms of protein–protein interactions involving hampin. Identification of proteins interacting with different isoforms is undoubtedly an important step in understanding the physiological function of hampins.

ARE THERE MSL COMPLEXES IN OTHER ORGANISMS?

The compensasome and the gene dose mechanism carried out by it, namely, twofold activation of X chromosome, are unique only for the *Drosophila* genus. Other animals perform these processes in different ways: it is known that nematodes reduce by one half the activity of each of two X chromosomes, whereas one human X chromosome is inactivated [51]. Even in other Diptera, like members of *Nematocera* genus, some proteins of the complex (MLE) are evidently not involved in gene dose compensation [52]. The low level of MSL1, which is lowly conserved in insects, also points to the uniqueness of compensasomes [12]. Nevertheless, DNA sequences related to other compensasome elements (except MSL1) are detected in very different organisms.

It is supposed that the MSL3 protein homologs are present in all eukaryotes from yeasts to mammals [53]. In this case in animal (including drosophila) the MSL3 paralog MRG15 is present [54]. The ortholog of human MSL3 (MSL3L1) has about 30% identity with the drosophila protein [55]. As in the case of drosophila MSL3, RNA of MSL3L1 undergoes alternative splicing resulting in the existence of a protein with shortened N-terminus. In fact, within the human MSL complex two variants of MSL3 protein were detected—long and short forms, and the latter has no chromodomain [14]. The functional role of the shortened protein is unknown.

Mammalian protein homologous to MSL2 has only a similar RING-finger domain and a new CXC domain revealed during comparison of their amino acid sequences [15]. The low percentage of identical amino acid residues in drosophila and mammalian MSL2 (36% identity between human and drosophila proteins) points to the high variability of the *MSL2* gene during evolution and simultaneously to the importance of their conserved regions.

Homologs of *MLE* (RNA/DNA helicases of the DEAH family) and *MOF* genes are also found in all eukaryotes [56]: in mammals it is the above-described gene of RNA helicase A for MLE and MYST1 (MOF).

The presence of *MOF* and *MSL3* genes in all eukaryotic cells indirectly points also to the conservation of their interaction. In fact, the yeast protein complex NuA4 involved in rearrangement of chromatin structure contains homologs of MOF (Esa1p) and MSL3 (Eaf3p) interacting with each other. It was shown that in yeast cells the interaction with Eaf3p is important for Esa1p acetyltransferase activity [57]. Also, this complex in yeasts has no absolute specificity towards K16 of H4, and in this cell type this modification plays an absolutely different role compared to animals. However, it is known that the yeast complex NuA4 also contains the ATM-related protein Tra1p [57].

Human proteins, homologs of drosophila MOF (MYST1), MSL1 (hampin A), MSL2, and MSL3 are able to form complexes in human erythroleukemia K562 cells [14]. Interactions between the complex components hMOF, hMSL3, hMSL2, and hMSL1 (hampin A) were confirmed by co-immunoprecipitation and binding *in vitro*. In this case mammalian MSL complex has a number of peculiarities: it does not contain the known non-coding RNA and homologs of RNA helicase MLE and JIL-1 chromosome kinase. Different isoforms of MSL3 protein are present within the complex and their role remains unclear. It should be remembered that the hampin gene involves alternative splicing. Keeping in mind that hampin, like MSL1, has to interact with acetyltransferases of the histone family MYST1 via the PEHE domain, one can suppose that at least two kinds of MSL complexes with hampin isoforms A and B should exist in mammalian cells. In addition, the existence of hampin D containing complexes specific for individual tissues such as testes is also possible. Knock-down of either hampin A (hMSL1) or MYST1 (hMOF) gives identical results, namely, the loss of acetylation of Lys16 in histone H4 at the level of all chromosomes of a cell and cell arrest at the G2/M stage as the result of knock-down [14]. MYST1 and hMSL3 knock-down in human cells results in similar consequences: the emergence of morphological deformations in the cell nuclei and defects of their structure. However, it seems that knock-down of hMSL3 alone does not result in lowering K16 acetylation of H4. An additional confirmation of the involvement of MYST1 in signal cascades, using ATM, is the observation that cells with knocked-down MYST1

have an increased content of phospho-ATM [8]. However, it is not clear what the role of hampin in signal cascades launched in response to DNA damage is.

Thus, it can be concluded that the quaternary structure of individual components of MSL complex is highly conserved, at least for hampin/MSL1, MSL3, and MYST1/MOF proteins. This is even more surprising because there is a difference in primary structures between the drosophila compensasome elements and their homologs in mammals.

OTHER COMPLEXES INVOLVING HAMPIN, MSL1, AND MYST1 ACETYLTRANSFERASE

MYST1 is incorporated into the so-called MLL1–WDR5 complex existing in human cell nuclei [58]. MLL1 protein is a methyltransferase of histone H3 K4 [58]. Protein WDR5, having a WD40 domain, is the conserved subunit of methyltransferase complexes [59]. Altogether 29 proteins have been identified within the complex (Table 1) purified on Flag-labeled WDR5. It is noteworthy that hampin was not detected among them. This complex exhibits two types of activities—acetyltransferase towards H4 K16 and methyltransferase towards H3 K4.

In a recent work on identification of additional proteins interacting with the MSL complex components in flies and humans, it has been shown that MSL is associated with components of nuclear pore complex [60].

The authors of that report used affinity chromatography on drosophila recombinant MOF and showed that the MSL complex contains additional protein factors (Table 2): Z4, Chriz/Chromator, MBD-R2, Nup153,

Table 1. Proteins interacting with WDR5 [56]

Identified partner	Accession No. in the GenBank database	Function (known or hypothetical)
Max dimerization protein 5	NP_001074010	transcription regulation
CHD8	NP_065971	transcription regulation, rearrangement of chromatin structure
TAF1	NP_004597	transcription initiation by RNA polymerase II
PELP1	NP_055204	co-activator of estrogen receptors
TAF4	NP_003176	transcription initiation by RNA polymerase II
PHF20	NP_057520	transcription regulation
LOC284058	NP_056258	no
HCF1/HCFC1	NP_005325	cell cycle control, transcription regulation
FLJ12525/LAS1L	NP_112483	no
TAF6	EAW76592	transcription initiation by RNA polymerase II
TEX10	NP_060216	no
ASH2L	NP_004665	structural component of methyltransferase complexes
HSP70	NP_005337	heat-shock protein
RbBP5	NP_005048	structural component of methyltransferase complexes
SENP3	NP_056485	SUMO1/sentrin/SMT3-specific protease 3
CAB43677	CAB43677	no
MOF/MYST1	NP_115564	histone (H4 K16) acetyltransferase
MLL1	NP_005924	histone (H3 K4) methyltransferase
MCRS2	AAQ84517	shortening telomere length
Tip49a/b/RUVBL1	NP_003698	DNA helicase
TAF7	AAH32737	transcription initiation by RNA polymerase II
Ring2	NP_005735	E3 ubiquitin ligase
E2F6 isoform 1	AAT02641	transcription regulation
TAF9	NP_001015892	transcription initiation by RNA polymerase II
C18ORF37	NP_919257	no
Max protein isoform b	NP_660087	transcription regulation
MGC49942	NP_777553	no

Table 2. Proteins identified by affinity chromatography on immobilized drosophila MOF [58]

Identified partner	Accession No. in the GenBank database	Function
MSL1	NP_476896	compensasome component
MSL2	NP_523467	compensasome component
MSL3	NP_523951	compensasome component
Mtor	NP_477067	intracellular transport of proteins
Nup153	NP_573136	intracellular transport of proteins
Z4	NP_649297	rearrangement of chromatin structure
Chromator/ Chriz	NP_730762	rearrangement of chromatin structure
MBD-R2	NP_731688	transcription regulation
NSL1	CG4699	no
NSL2	CG18041	no
NSL3	CG8233	no
WDS/ WDR5	NP_438172	MLL1 binding
dMCRS2	CG1135	no
Dis3	CG6413	ribonuclease, the exosome component

Table 3. Proteins identified by affinity chromatography on immobilized human MYST1 [58]

Identified partner	Accession No. in the GenBank database	Function
Hampin A	XP_937190	no
hMSL2	NP_060603	no
hMSL3	NP_523353	no
PHF20L1	NP_057102	transcription regulation
TPR	NP_003283	protein import into the cell nucleus
PHF20	NP_057520	transcription regulation
hNSL1	KIAA1267	no
hNSL2	FLJ20436	no
hNSL3	FLJ10081	no
WDR5	NP_438172	MLL1 binding
MCRS2	NP_001012300	no
Dis3	NP_055768	ribonuclease, the exosome component
HCF1	NP_005325	cell cycle control, transcription regulation
OGT	NP_858058	O-glycosyl transferase
ILF-1	NP_004505	transcription regulation

wds, α-tubulin, MCRS2, as well as novel proteins NSL1-3 ("nonspecific lethal"). At the same time, chromatography on MSL3 protein revealed the following set of proteins: Mtor, MBD-R2, Nup160, Nup154, Dis3, Rrp6, α-tubulin, and EIF-4B. A number of proteins specific for individual cell types were also found. It is seen that proteins MSL3 and MOF are associated with the partially overlapped set of proteins. However, it is not clear whether MSL3 and MOF interact directly or indirectly with these proteins. Additional affinity chromatography experiments on human HeLa cell extract on immobilized MYST1 were carried out (Table 3). It appeared that in addition to the previously known human homologs of MSL1-MSL3 proteins, the eluate contained proteins homologous to those obtained in similar experiments on drosophila: protein PHF20 (homolog of MBD-R2), TPR

(ortholog of Mtor), human NSL1-3, WDR5 (homolog of wds), MCRS2, and Dis3. A number of proteins absent from drosophila were also identified, among them being co-activator of transcription HCF1, glycosyl transferase OGT, and transcription factor ILF-1/FOXK2. It is noteworthy that some of these proteins were identified earlier within the complex based on MLL1 and WDR5 proteins. It was shown using co-immunoprecipitation that practically the whole set of "novel" proteins is present in the complex with MYST1 but not with hampin A (MSL1). It is supposed that its function in this complex (also exhibiting acetyltransferase activity towards H4 K16) is carried out by NSL1 protein also incorporating a PEHE domain.

Protein NSL1 was first identified by Marin [15] in work on MSL1 and MSL2 homologs in mammals, and it was shown that the drosophila protein NSL1 is related to MSL1 and also contains a modified PEHE domain (Fig. 2). Marin supposed that NSL1 is a paralog of MSL1, and both genes encoding these proteins were generated by duplication of a certain ancestor gene that emerged together with coelomate animals. It is interesting that the MSL1/NSL1 homologs were found in some nematodes, whereas they are absent from *C. elegans*. Unlike hampin, mammalian NSL1 consists of about 1000 a.a. but also has no presently known domains except the C-terminal PEHE and coiled-coil region. Its primary structure identity with that in mammalian hampin A is about 27%. In addition to interaction with MYST1, in the context of the human interactome investigation by screening in the yeast two-hybrid system, a number of additional interactions were identified for this protein such as that with transcription factor SMAD3 and with nuclear protein DIPA (delta-interacting protein A), which was also confirmed by *in vitro* binding [61]. At the present time the functional significance of these interactions is not known.

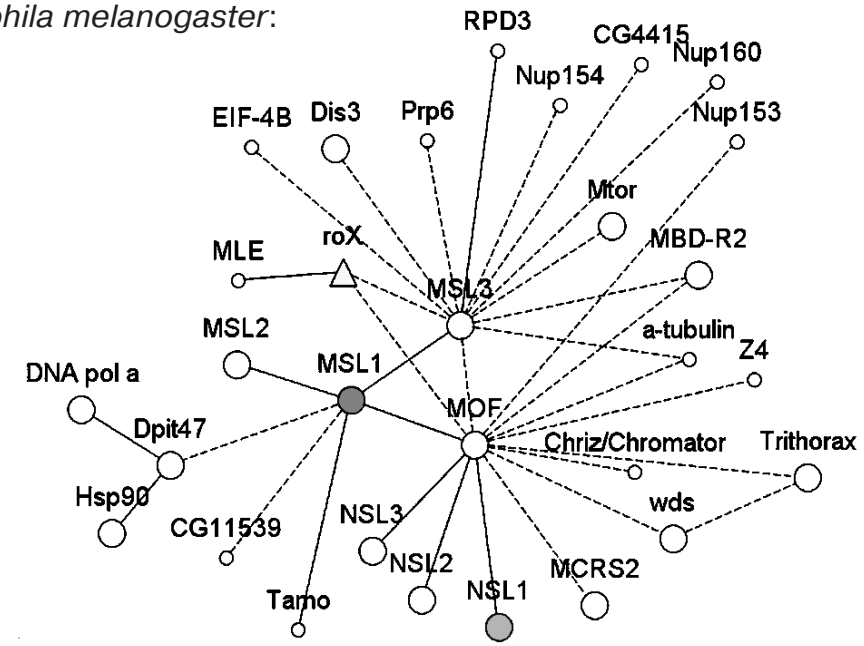
Investigation of the effect of Mtor and Nup153 protein knock-down in drosophila cells has shown that these components of nuclear pore complex are important for X-chromosomal localization of MSL complex and its activity towards gene dose regulation [60]. However, the following question is still open: if there are two types of MOF-incorporating complexes, which of them is associated with Mtor and Nup153 proteins? The authors give no adequate answer to this question. It is supposed that just MSL complex is associated with nuclear pore complex, because knock-down of Mtor and Nup153 proteins is accompanied by disappearance of X chromosome labeling revealed by staining with antibodies to MSL1, MSL3, and MOF proteins. At the same time, the role of so-called NSL complex (also partly identified [60]) is completely unclear.

```
Mus_hampin          --------SSVAGETSVLAVPSWRDHSVEPLR--DPNPSDLLENLDDSVFSKRHAKLELD 513
Homo_hampin         --------SSVAGETSVLAVPSWRDHSVEPLR--DPNPSDLLENLDDSVFSKRHAKLELD 513
Gallus_hampin       --------------------IPSWRDHVVEPLR--DPNPSDILENLDDSVFSKRHAKLELD 295
Xenopus_hampin      --------------------IPSWREHHIEPLQ--LAEPCDIPENLDDAVFAKRHAKLELD 394
Fugu_hampin         --------THMSKNSSSLLVPSWRENHIEPLD--EEPSFVPPELLEDNVYLKRHMKLELD 284
Danio_hampin        --------------------IPSWRENIMEPLK--EEDAADLPESLDDSVFLKRHAKLELD 400
Homo_NSL1           PMSVAATTRVEKLQYKEILTPSWREVDLQSLKGSPDEENEEIEDLSDAAFAALHAKCEEM 927
Mus_NSL1            PMSVAATTRVEKLQYKEILTPSWREVDVQSLKGSPDEENEEIEDLSDAAFAALHAKCEEM 858
Tetraodon_NSL1      PMSVAATTRVEKLQYKEILTPSWRAVDVFSQPITEEEDEREVEDLSDAAFTQLHQPYEDQ 986
Xenopus_NSL1        PMSLIAPTKLEKLKYKEIITPSWKEVVLEPLESPAHDMP---EDLSDEAYISRHEKYELK 788
Drosophila_NSL1     PMSLIAPTKLEKLKYKEIITPSWKEVVLEPLESPAHDMP---EDLSDEAYISRHEKYELK 788
Drosophila_MSL1     NEFFLEEADELLADNPSLEIPKWRDVPVPPSS-----DKIDTELLSDATFERRHQKYVKD 903
                       .             :  *.*:   : .           * *.* .:   *

Mus_hampin          EKRRK-RWDIQRIREQRILQRLQLRMYKKK--GIQESEPEVTSFFPEPDDVESLMITPFL 570
Homo_hampin         EKRRK-RWDIQRIREQRILQRLQLRMYKKK--GIQESEPEVTSFFPEPDDVESLMITPFL 570
Gallus_hampin       EKRRK-RWDIQRIREQRILQRLQLRMYKRK--GIQESEPEVTSFFPEPDDVESLLITPYL 352
Xenopus_hampin      EKRRK-RWDIQRIREQRLLQRLQLRMCKKK--GNQESEPEVTSFFPDVDDVESVMITPYL 451
Fugu_hampin         EKRRK-RWDIQRIREQRMFQRLQQRMNRKK--VITEAEPELSSFYPDTEDVETIVITPFL 341
Danio_hampin        EKRRK-RWDIQRIREQRMLQRLQQRMEKKKPNVVQESEPEVSSFHPDLENVEAIMVTPFL 459
Homo_NSL1           ERARW-LWTTSVPPQRRGSRSYRSSDGRTTPQLGSANPSTPQPASPDVSSSHSLSEY-SH 985
Mus_NSL1            ERARW-LWTTSVPPQRRGSRSYRSSDGRTTPQLGSANPSTPQPASPDVSSSHSLSEF-SH 916
Tetraodon_NSL1      ERSRW-TWMALAPAKRRG----RINPWTAGPPLCCAGPTLPSPP-ARLSRPGSL------ 1034
Xenopus_NSL1        EKARWSLWDHSKRPKRNRSSSYSFGTSPRTVLLSCECSCSPNSQAPSEALPSDTGGYRTL 848
Drosophila_NSL1     EKARWSLWDHSKRPKRNRSSSYSFGTSPRTVLLSCECSCSPNSQAPSEALPSDTGGYRTL 848
Drosophila_MSL1     EVDRK-CRDARYMKEQIRLEQLRMRRNQDE-VLVALDPLRASTFYPLPEDIEAIQFVNEV 961
                    *  *          ::                              .  .
```

Fig. 2. Multiple alignment of PEHE domains in primary structures of hampin and its paralog NSL1 from different organisms. The animal species are shown on the left. Coordinates of amino acid residues in the primary structure are shown on the right and identical residues in all species are marked by gray color and designated by asterisks. Residues identical in primary structures of only separate species are marked by dots. The alignment was carried out using the ClustalW (http://www.ebi.ac.uk/clustalw/) program.

Drosophila melanogaster:



Mammals:

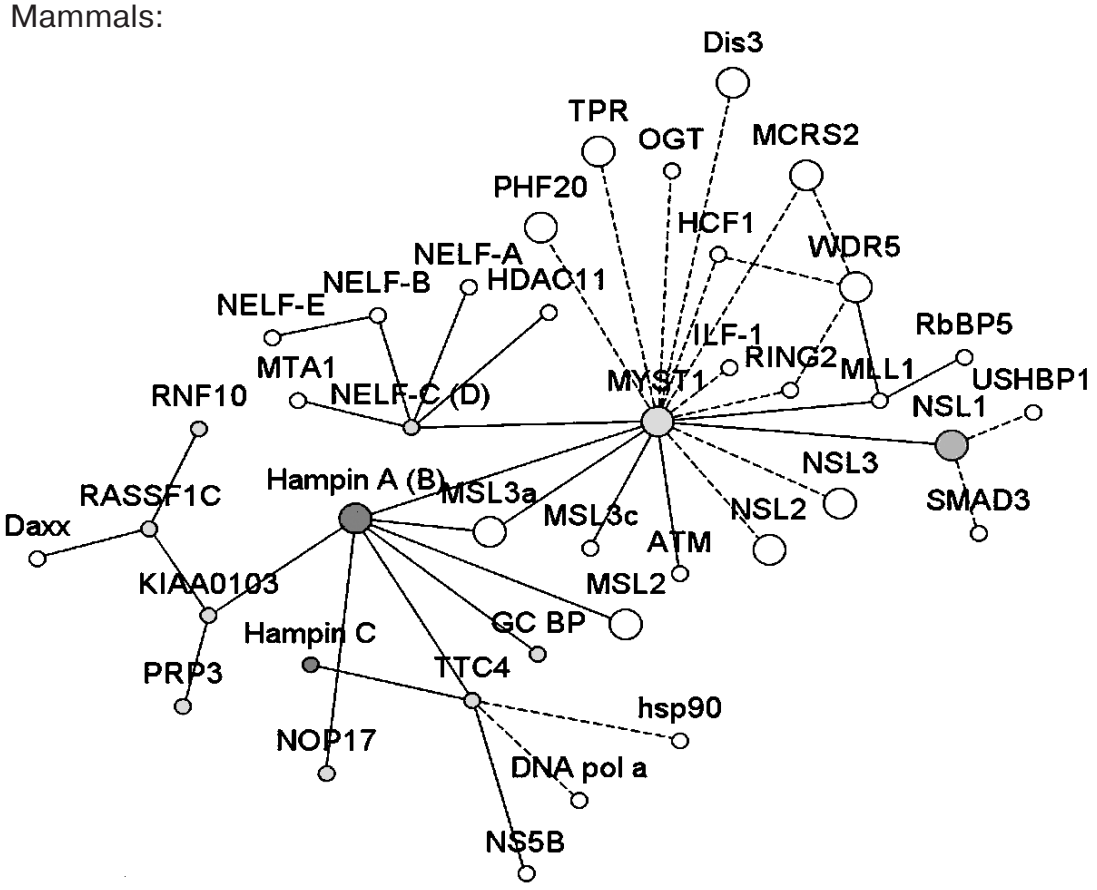


Fig. 3. Scheme of the drosophila and mammalian cell interactome fragments including protein homologs of hampin and MYST1. The scheme was plotted using the results obtained in [13, 14, 58, 60] and on the basis of information taken from databases of DIP, BioGrid, etc. using the yEd 2.4.1 program (www.yWorks.com). Solid lines show interactions confirmed by several (two and more) independent approaches, while dashed lines show interactions obtained in "high-throughput" experiments or those considered as hypothetical.

Taking into account data concerning the existence of a cellular homolog of MSL1 and hampin in the form of NSL1, the presence of additional hampin-associated protein complexes in mammals seems quite reasonable. In fact, it was shown [13] that in the yeast two-hybrid system mouse hampin is able to interact with five proteins—MYST1, TPR-motif-containing proteins TTC4 and mouse homolog of TTC35 protein, transcription factor GC BP, and a homolog of the yeast nucleolar protein NOP17 (Fig. 3). The identified interactions are confirmed using alternative approaches such as *in vitro* binding. Comparative analysis of hampin isoforms has shown that interaction with TTC4 is evidently possible for all hampin isoforms, whereas only hampins with C-terminal IV PEHE domain are able to interact with NOP17 and TTC35 proteins. An extensive search for protein interactions between hampin partners was carried out in the yeast two-hybrid system, which resulted in the discovery of new interactions for proteins MYST1 (with a component of the negative transcription elongation factor NELF-C), TTC35 (with tumor suppressor RASSF1C and splicing factor PRP3), but at the same time the authors failed in identification of new interactions for proteins TTC4 and NOP17.

It should be noted that the MYST1 partner protein NELF-C is involved in regulation of transcription elongation by RNA polymerase II, because it is a component of NELF complex participating in the "promoter-proximal pausing" phenomenon [62] and in regulation of histone mRNA polyadenylation [63]. So, this interaction directly connects the histone H4 K16 acetylation (and chromatin decompaction) with transcription regulation. Although nothing is yet known about the functional significance of this interaction, it can be suggested that MYST1 influences the function of the whole NELF complex via activation/repression of gene transcription. It is interesting that protein NELF-C has the NELF-D isoform with a slightly shorter N-terminal region that is formed via an alternative site of translation initiation in the NELF-C/D mRNA [62]. Both isoforms are interchangeable within the NELF complex, and the necessity of existence of two such protein variants is unknown. At the same time, another research group has shown that the NELF-D protein (called by this group TH1) is able to interact with A-Raf kinase, is present in testes, and is localized exclusively in the cytoplasm [64]. This contradicts data showing that protein NELF-C/D is able to participate in regulation of transcription elongation, because in this case it should be localized in the nucleus. Isoforms C and D may have different intracellular localization, or they change it depending on the cell cycle stages.

The possible acetylation of non-histone proteins in drosophila using MYST1 protein has been already mentioned in this review. Recently it has also been shown that human MYST1 is able to acylate non-histone proteins, in particular p53, known by its important role in regulation of apoptosis launching in response to damage to cellular DNA. Sykes et al. [65] demonstrated the existence of a new posttranslational modification in p53 protein—acetylation of the K120 residue which is carried out by acetyltransferases MYST1 and TIP60 to approximately equal extent. This covalent modification of p53 takes place immediately after damage of genomic DNA and evidently enhances the capability of p53 protein for association near promoter regions of proapoptotic genes *BAX*, *PUMA*, and (possibly) others. However, in human tumors containing mutant p53 with R120 instead of K120, no p53-dependent apoptosis is observed. It should be noted that this modification of p53 does not concern its functions in cell cycle arrest. The authors suppose that the ability to carry out this modification by two enzymes is necessary for different types of genotoxic stress or different types of tissues in which it takes place. However, it is not clear how proteins MYST1 and TIP60 "learn" about damage to genomic DNA. It seems that the detailed study of this mechanism of activation of apoptosis (from DNA damage to the launching of cell death) may have perspectives in the development of new anticancer drugs.

Recent investigations of expression levels of the MYST1 and TIP60 acetyltransferase genes reveal the problem from a new point of view [66]. Analysis of expression levels of these acetyltransferases in mice has shown that mRNA of TIP60 has extremely low expression level restricted mainly to testes. However, the level of *MYST1* gene expression is much higher and is detected in all tissues, also with predominance in testes. Based on these results, one can assume that just MYST1 is the "main" acetyltransferase that acetylates p53 *in vivo*. Moreover, one should expect the prevalence of histone acetyltransferase complexes on the basis of this protein over those on the basis of TIP60.

Thus, analysis of complexes incorporating homologs of MYST1, hampin, and other proteins allows one to imagine the general evolution of complexes acetylating histone H4. Possibly, the yeast NuA4 complex can be considered as the most primitive example of such complexes, but in the course of evolution multicellular organisms needed additional components of this complex such as hampin and NSL1 (Fig. 4). As a result, in some species like *Drosophila* one of complexes acquired specific functions and began to take part in dosage compensation. Thus, at least two complexes were formed that are involved in acetylation of K16 in histone H4, and possibly in other chromatin modifications, in particular in histone methylation. In highly organized animals further divergence of the components of the complex occurred, which was revealed in alternative splicing of individual components and existence of different isoforms (hampin, MSL3, etc.). During evolution complexes changed both in functional (because the significance of histone H4 K16 changed) and structural aspects. New data on the interaction of

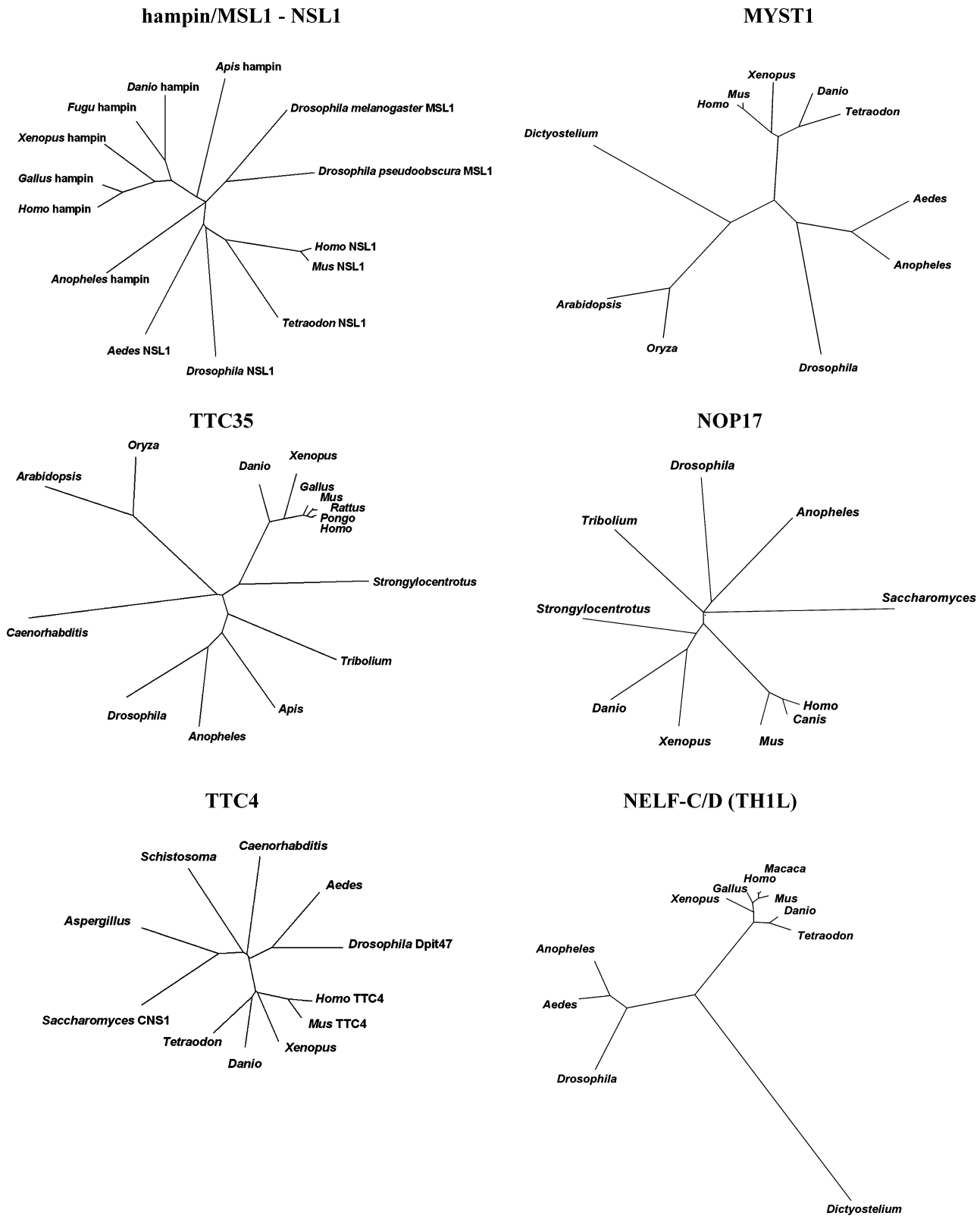


Fig. 4. Family trees of hampin/NSL1, MYST1 and their partner proteins plotted on the basis of primary structures of proteins from different organisms. The ClustalW program (http://www.ebi.ac.uk/clustalw/) was used for alignment of amino acid sequences, and the tree was built using the Treeview 1.6.6. program (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html). Designations: *Arabidopsis*, garden cress; *Oryza*, rice; *Apis*, bee; *Aedes* and *Anopheles*, mosquito; *Drosophila*, fruit fly; *Dictyostelium*, parasitic amoeba; *Tetraodon* and *Fugu*, puffer fish; *Danio*, zebrafish; *Gallus*, hen; *Xenopus*, frog; *Mus*, mouse; *Homo*, man.

MYST1 with non-histone proteins (NELF-C, ATM, p53) point to a more extended function of this protein in the interactome; evidently it plays not the last role in activation of programmed cell death. Undoubtedly, many aspects of MYST1 functioning are poorly known. Thus the results obtained to date are incentives for further investigation of MYST1 protein and its cofactors.

This work was supported by the Russian Foundation for Basic Research (project Nos. 04-04-49413 and 07-04-01497a).

REFERENCES

1. Cosgrove, M. S., and Wolberger, C. (2005) *Biochem. Cell Biol.*, **83**, 468-476.
2. Turner, B. M. (2002) *Cell*, **111**, 285-291.
3. Sterner, D. E., and Berger, S. L. (2000) *Microbiol. Mol. Biol. Rev.*, **64**, 435-459.
4. Kouzarides, T. (2007) *Cell*, **128**, 693-705.
5. Berger, S. L. (2007) *Nature*, **447**, 407-412.
6. Clayton, A. L., Hazzalin, C. A., and Mahadevan, L. C. (2006) *Mol. Cell*, **23**, 289-296.
7. Neal, K. C., Pannuti, A., Smith, E. R., and Lucchesi, J. C. (2000) *Biochim. Biophys. Acta*, **1490**, 170-174.
8. Taipale, M., Rea, S., Richter, K., Vilar, A., Lichter, P., Imhof, A., and Akhtar, A. (2005) *Mol Cell Biol.*, **25**, 6798-6810.
9. Fraga, M. F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., et al. (2005) *Nat. Genet.*, **37**, 391-400.
10. Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) *Science*, **311**, 844-847.
11. Gupta, A., Sharma, G. G., Young, C. S., Agarwal, M., Smith, E. R., Paull, T. T., Lucchesi, J. C., Khanna, K. K., Ludwig, T., and Pandita, T. K. (2005) *Mol. Cell. Biol.*, **25**, 5292-5305.
12. Dmitriev, R. I., Pestov, N. B., Korneenko, T. V., Gerasimova, A. V., Zhao, Kh., Modyanov, N. N., Kostina, M. B., and Shakhparonov, M. I. (2005) *Bioorg. Khim.*, **31**, 363-371.
13. Dmitriev, R. I., Korneenko, T. V., Bessonov, A. A., Shakhparonov, M. I., Modyanov, N. N., and Pestov, N. B. (2007) *Biochem. Biophys. Res. Commun.*, **355**, 1051-1057.
14. Smith, E. R., Cayrou, C., Huang, R., Lane, W. S., Cote, J., and Lucchesi, J. C. (2005) *Mol. Cell. Biol.*, **25**, 9175-9188.
15. Marin, I. (2003) *J. Mol. Evol.*, **56**, 527-539.
16. Taipale, M., and Akhtar, A. (2005) *Progr. Mol. Subcell. Biol.*, **38**, 123-149.
17. Belote, J., and Lucchesi, J. (1980) *Genetics*, **96**, 165-186.
18. Kuroda, M. I., Kernan, M. J., Kreber, R., Ganetzky, B., and Baker, B. S. (1991) *Cell*, **66**, 935-947.
19. Palmer, M. J., Mergner, V. A., Richman, R., Manning, J. E., Kuroda, M. I., and Lucchesi, J. C. (1993) *Genetics*, **134**, 545-557.
20. Gorman, M., Franke, A., and Baker, B. S. (1995) *Development*, **121**, 463-475.
21. Li, F., Parry, D. A., and Scott, M. J. (2005) *Mol. Cell. Biol.*, **25**, 8913-8924.
22. Scott, M. J., Pan, L. L., Cleland, S. B., Knox, A. L., and Heinrich, J. (2000) *EMBO J.*, **19**, 144-155.
23. Morales, V., Straub, T., Neumann, M. F., Mengus, G., Akhtar, A., and Becker, P. B. (2004) *EMBO J.*, **23**, 2258-2268.
24. Zhou, S., Yang, Y., Scott, M. J., Pannuti, A., Fehr, K. C., Eisen, A., Koonin, E. V., et al. (1995) *EMBO J.*, **14**, 2884-2295.
25. Beckmann, K., Grskovic, M., Gebauer, F., and Hentze, M. W. (2005) *Cell*, **122**, 529-540.
26. Kelley, R. L., Solovyeva, I., Lyman, L. M., Richman, R., Solovyev, V., and Kuroda, M. I. (1995) *Cell*, **81**, 867-877.
27. Copps, K., Richman, R., Lyman, L. M., Chang, K. A., Rampersad-Ammons, J., and Kuroda, M. I. (1998) *EMBO J.*, **17**, 5409-5417.
28. Lyman, L. M., Copps, K., Rastelli, L., Kelley, R. L., and Kuroda, M. I. (1997) *Genetics*, **147**, 1743-1753.
29. Straub, T., Neumann, M. F., Prestel, M., Kremmer, E., Kaether, C., Haass, C., and Becker, P. B. (2005) *Chromosoma*, **114**, 352-364.
30. Gorman, M., Franke, A., and Baker, B. S. (1995) *Development*, **121**, 463-475.
31. Koonin, E. V., Zhou, S., and Lucchesi, J. C. (1995) *Nucleic Acids Res.*, **23**, 4229-4233.
32. Morales, V., Regnard, C., Izzo, A., Vetter, I., and Becker, P. B. (2005) *Mol. Cell. Biol.*, **25**, 5947-5954.
33. Buscaino, A., Kocher, T., Kind, J. H., Holz, H., Taipale, M., Wagner, K., Wilm, M., and Akhtar, A. (2003) *Mol. Cell.*, **11**, 1265-1277.
34. Buscaino, A., Legube, G., and Akhtar, A. (2006) *EMBO Rep.*, **7**, 531-538.
35. Lee, C. G., Chang, K. A., Kuroda, M. I., and Hurwitz, J. (1997) *EMBO J.*, **16**, 2671-2681.
36. Lee, C. G., Reichman, T. W., Baik, T., and Mathews, M. B. (2004) *J. Biol. Chem.*, **279**, 47740-47745.
37. Bone, J. R., Lavender, J., Richman, R., Palmer, M. J., Turner, B. M., and Kuroda, M. I. (1994) *Genes Dev.*, **8**, 96-104.
38. Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A., and Lucchesi, J. C. (1997) *EMBO J.*, **16**, 2054-2060.
39. Gu, W., Szauter, P., and Lucchesi, J. C. (1998) *Dev. Genet.*, **22**, 56-64.
40. Smith, E. R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R. G., Allis, C. D., and Lucchesi, J. C. (2000) *Mol. Cell. Biol.*, **20**, 312-318.
41. Franke, A., and Baker, B. S. (1999) *Mol. Cell*, **4**, 117-122.
42. Kelley, R. L. (2004) *Dev. Biol.*, **269**, 18-25.
43. Oh, H., Park, Y., and Kuroda, M. I. (2003) *Genes Dev.*, **17**, 1334-1339.
44. Fagegaltier, D., and Baker, B. S. (2004) *PLoS Biol.*, **2**, e341.
45. Deng, X., Rattner, B. P., Souter, S., and Meller, V. H. (2005) *Mech. Dev.*, **122**, 1094-1105.
46. Jin, Y., Wang, Y., Johansen, J., and Johansen, K. M. (2000) *J. Cell Biol.*, **149**, 1005-1010.
47. Lerach, S., Zhang, W., Deng, H., Bao, X., Girton, J., Johansen, J., and Johansen, K. M. (2005) *Genesis*, **43**, 213-215.
48. Gilfillan, G. D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P. B. (2006) *Genes Dev.*, **20**, 858-870.
49. Corona, D. F., Clapier, C. R., Becker, P. B., and Tamkun, J. W. (2002) *EMBO Rep.*, **3**, 242-247.
50. Furuhashi, H., Nakajima, M., and Hirose, S. (2006) *Development*, **133**, 4475-4483.

51. Lucchesi, J. C., Kelly, W. G., and Panning, B. (2005) *Annu. Rev. Gen.*, **39**, 615-651.
52. Ruiz, M. F., Esteban, M. R., Donoro, C., Goday, C., and Sanchez, L. (2000) *Genetics*, **156**, 1853-1865.
53. Marin, I., and Baker, B. S. (2000) *Mol. Biol. Evol.*, **17**, 1240-1250.
54. Tominaga, K., Kirtane, B., Jackson, J. G., Ikeno, Y., Ikeda, T., Hawks, C., Smith, J. R., Matzuk, M. M., and Pereira-Smith, O. M. (2005) *Mol. Cell. Biol.*, **25**, 2924-2937.
55. Prakash, S. K., van den Veyver, I. B., Franco, B., Volta, M., Ballabio, A., and Zoghbi, H. Y. (1999) *Genomics*, **59**, 77-84.
56. Sanjuan, R., and Marin, I. (2001) *Mol. Biol. Evol.*, **18**, 330-343.
57. Eisen, A., Utley, R. T., Nourani, A., Allard, S., Schmidt, P., Lane, W. S., Lucchesi, J. C., and Cote, J. (2001) *J. Biol. Chem.*, **276**, 3484-3491.
58. Dou, Y., Milne, T. A., Tackett, A. J., Smith, E. R., Fukuda, A., Wysocka, J., Allis, C. D., Chait, B. T., Hess, J. L., and Roeder, R. G. (2005) *Cell*, **121**, 873-885.
59. Li, D., Burch, P., Gonzalez, O., Kashork, C. D., Shaffer, L. G., Bachinski, L. L., and Roberts, R. (2000) *Biochem. Biophys. Res. Commun.*, **274**, 117-123.
60. Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J., Wilm, M., Stunnenberg, H. G., Saumweber, H., and Akhtar, A. (2006) *Mol. Cell*, **21**, 811-823.
61. Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., et al. (2005) *Nature*, **437**, 1173-1178.
62. Narita, T., Yamaguchi, Y., Yano, K., Sugimoto, S., Chanarat, S., Wada, T., Kim, D. K., Hasegawa, J., Omori, M., Inukai, N., Endoh, M., Yamada, T., and Handa, H. (2003) *Mol. Cell. Biol.*, **23**, 1863-1873.
63. Narita, T., Yung, T. M., Yamamoto, J., Tsuboi, Y., Tanabe, H., Tanaka, K., Yamaguchi, Y., and Handa, H. (2007) *Mol. Cell.*, **26**, 349-365.
64. Liu, W., Shen, X., Yang, Y., Yin, X., Xie, J., Yan, J., Jiang, J., Liu, W., Wang, H., Sun, M., Zheng, Y., and Gu, J. (2004) *J. Biol. Chem.*, **279**, 10167-10175.
65. Sykes, S. M., Mellert, H. S., Holbert, M. A., Li, K., Marmorstein, R., Lane, W. S., and McMahon, S. B. (2006) *Mol. Cell.*, **24**, 841-851.
66. Thomas, T., Loveland, K. L., and Voss, A. K. (2007) *Gene Expression Patterns*, **7**, 657-665.