

Linking Global Histone Acetylation to the Transcription Enhancement of X-chromosomal Genes in *Drosophila* Males*

Received for publication, June 25, 2001,
and in revised form, July 2, 2001
Published, JBC Papers in Press, July 9, 2001, DOI
10.1074/jbc.C100351200

Edwin R. Smith^{‡§}, C. David Allis[§],
and John C. Lucchesi^{‡¶}

From the [‡]Department of Biology, Emory University,
Atlanta, Georgia 30322 and the [§]Department of
Biochemistry and Molecular Genetics, University of
Virginia Health System, Charlottesville, Virginia 22908

It has become well established for several genes that targeting of histone acetylation to promoters is required for the activation of transcription. In contrast, global patterns of acetylation have not been ascribed to any particular regulatory function. In *Drosophila*, a specific modification of H4, acetylation at lysine 16, is enriched at hundreds of sites on the male X chromosome due to the activity of the male-specific lethal (MSL) dosage compensation complex. Utilizing chromatin immunoprecipitation, we have determined that H4Ac16 is present along the entire length of X-linked genes targeted by the MSL complex with relatively modest levels of acetylation at the promoter regions and high levels in the middle and/or 3' end of the transcription units. We propose that global acetylation by the MSL complex increases the expression of X-linked genes by facilitating transcription elongation rather than by enhancing promoter accessibility. We have also determined that H4Ac16 is absent from a region of the X chromosome that includes a gene known to be dosage-compensated by a MSL-independent mechanism. This study represents the first biochemical interpretation of the very large body of cytological observations on the chromosomal distribution of the MSL complex.

Many post-translational modifications of the highly conserved histone N-terminal tails are linked to transcriptional regulation (1), with the most widely studied modification being acetylation of histones H3 and H4. Genetic and biochemical data have demonstrated that histone acetyltransferases such as yGcn5p and yEsa1p are recruited to the promoters of specific genes for the activation of transcription (2–6). Loss of function mutations of the genes encoding these enzymes, however, reveal that they are also responsible for a broad pattern of acetylation with relatively modest effects on the level of transcription (2, 4, 6).

* This work was supported by National Institutes of Health Grants GM15961 (to J. C. L.) and GM53512 (to C. D. A.) and a grant from the Human Frontiers of Science Program (to C. D. A.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: Dept. of Biology, Emory University, 1510 Clifton Rd., Atlanta, GA 30322. Tel.: 404-727-4234; Fax: 404-727-2880; E-mail: Lucchesi@biology.emory.edu.

In *Drosophila*, evidence for special roles for particular histone modifications was revealed with antisera to specific isoforms of H4. While H4 isoforms acetylated at lysine 5 or 8 are associated with numerous sites throughout the genome, H4 acetylated at lysine 12 is enriched in chromocentric heterochromatin, and H4 acetylated at lysine 16 (H4Ac16) is exclusively associated with the male X chromosome (7, 8). Male-specific acetylation of H4 at lysine 16 is mediated by males absent on the first (MOF),¹ a MYST-family histone acetyltransferase present in the male-specific lethal (MSL) complex (9–11). This histone acetyltransferase-containing complex is responsible for the dosage compensation of many X-linked genes by increasing their transcription in males to achieve a level of gene product equivalent to that generated by the two X chromosomes in females. In addition to its protein subunits, the MSL complex contains one of two nontranslated RNAs, roX1 and roX2 (11–15). Current models suggest that functional complexes form at the sites of transcription of these RNAs, then access the X chromosome at a small number of additional entry sites and subsequently spread to the hundreds of other locations along the X chromosome where they are normally found (16). This spreading step requires both the histone acetyltransferase activity of MOF and the ATP-dependent RNA/DNA helicase activity of another subunit, MLE (17, 18).

In this paper, we use chromatin immunoprecipitation to map the distribution of H4Ac16 acetylation on known dosage-compensated genes in *Drosophila* embryos and find that this isoform is not restricted to the promoter region but is distributed throughout the length of these genes. Importantly, we show that *runt*, an X-linked gene known to be dosage-compensated independently of the MSL complex, escapes H4Ac16 acetylation and resides in a large chromosomal region that lacks this histone isoform, suggesting that broad or "global" patterns of acetylation are targeted to specific chromosomal or nuclear domains. Finally, we demonstrate an enhanced binding of the MSLs at the roX2 locus; this observation supports the hypothesis that the initial X chromosome targets of the MSL complex are likely to be entry sites, including the roX loci where the complex is expected to form.

EXPERIMENTAL PROCEDURES

Embryo Chromatin IP—Mixed-stage embryos (0–12 h old) were cross-linked and sonicated according to Orlando *et al.* (19). The average size of DNA fragments was ~500 base pairs. After sonication, debris was removed by centrifugation, and the supernatant was adjusted to 0.5% Sarkosyl and mixed for 10 min at 22 °C and respun. The supernatant was adjusted to RIPA (0.14 M NaCl, 1% Triton X-X100, 0.1% SDS, 0.1% sodium deoxycholate), aliquoted, and frozen.

Antiserum that is specific for histone H4 acetylated at lysine 16 was obtained from Serotec. In immunofluorescence experiments, this antiserum paints the *Drosophila* male X chromosome as described previously (7). Core histone monoclonal MAB052 (Chemicon) was chosen to control for nucleosome concentration differences based on its high efficiency in ChIP; other general H4 antibodies tested showed little enrichment above the background obtained using naive rabbit serum.

3 μ l of antisera was bound to 12.5 μ l of protein A-agarose and incubated with 500 μ l of extract (representing roughly 0.2-g embryos). Beads were washed several times in RIPA buffer, then washed in 10 mM Tris, 1 mM EDTA, pH 8.0. RNase was added at a concentration of 50 μ g/ml for 10 min at 22 °C, adjusted to 0.5% SDS and 0.5 mg/ml pro-

¹ The abbreviations used are: MOF, males absent on the first; MSL, male-specific lethal; IP, immunoprecipitation; ChIP, chromatin immunoprecipitation; PCR, polymerase chain reaction; kb, kilobase pair(s).

teinase-K, and incubated 42 °C for 5 h. Formaldehyde cross-links were reversed by a further incubation at 65 °C overnight. DNA was purified by phenol-chloroform extraction and ethanol precipitation.

Real-time PCR—Primer pairs were designed to amplify ~300-base pair fragments (278–318) at regular intervals along genes of interest. Each PCR reaction contained 4 pmol of each primer, 3 mM MgCl₂, 0.2 mM concentration of each dNTP, 1:10,000 dilution of Sybr Green1 (BioWhittaker), 5% Me₂SO, 0.1% bovine serum albumin, 0.1% Tween 20, 50 mM Tris, pH 8, and 1.6 units of HotStar Taq polymerase (Qiagen). PCR was performed and monitored in a Bio-Rad iCycler: 20 min activation of Taq at 94 °C, followed by 40 cycles of 94 °C 30 s, 55 °C 30 s, 72 °C 45 s. Product formation was detected at 72 °C in the fluorescein isothiocyanate channel. Because Sybr Green1 binds to any double-stranded DNA, we checked for primer-dimer formation by performing control reactions without substrate, as well as using agarose gel electrophoresis to check for the desired product.

Quantitation of Real-time PCR—Calculations of fold enrichment of an X sequence relative to an autosomal sequence utilize relative differences in the threshold cycle, the cycle of PCR at which the fluorescence reaches a given value or threshold that is in the log-linear range of amplification.

The fold enrichment is calculated as: fold enrichment = $(2^{\Delta(\text{core IP}_X - \text{Ac16 IP}_X)}) / (2^{\Delta(\text{core IP}_A - \text{Ac16 IP}_A)})$, where core IP_X, core IP_A, Ac16 IP_X, and Ac16 IP_A are the observed threshold cycles for the X or autosomal (A) sequences in the appropriate immunoprecipitation reaction.

RESULTS

Lysine 16 Acetylation Is Enriched along the Length of Dosage-compensated Genes—We performed ChIP experiments on *Pgd* and *Zw*. These two X-linked genes, which encode the pentose-phosphate-shunt enzymes 6-phosphogluconate dehydrogenase and glucose-6-phosphate dehydrogenase, respectively, had been shown previously to be dosage-compensated by the MSL complex (20). An autosomal control for these studies was the *Gpdh* gene that encodes the glycerophosphate shuttle enzyme glycerol-3-phosphate dehydrogenase. Since these three genes are expressed in all cell types, they are particularly good candidates for ChIP studies using mixed-stage embryos as starting material.

Embryos were treated with formaldehyde, the cross-linked chromatin was solubilized by sonication, and the resulting extracts were used in immunoprecipitation reactions with H4Ac16 antiserum (Serotec) or with a monoclonal antiserum recognizing all histones (MAB052, Chemicon). DNA recovered from the precipitates was used in quantitative, real-time PCR with primer sets that amplify 300-base pair fragments at regular intervals across the genes of interest. The relative enrichment of H4Ac16 to the autosomal control was determined for each PCR-amplified fragment (see “Experimental Procedures”).

The *Pgd* gene encodes a 3.2-kb primary transcript (21) and is flanked on its 5' end by a predicted gene, *EG87B1*, and on its 3' end by the *bcn92* gene (Fig. 1). Using the autosomal gene as reference, H4Ac16 is enriched along the entire length of *Pgd*, with greater amounts of this isoform present toward the 3' end. Similar distributions were observed whether input DNA or DNA recovered from immunoprecipitations using core histone antiserum were used for normalization (data not shown).

Because of the larger size of the *Zw* gene, we spaced the 300-base pair PCR-amplified fragments farther apart from one another. As was observed with *Pgd*, H4Ac16 is enriched along the length of *Zw*, with particularly high levels in the middle and 3' end of the transcription unit (Fig. 2).

H4Ac16 Is Not Found in a Chromosomal Region That Includes the *runt* Gene—*runt* is an X-linked gene that is dosage-compensated by a process that is independent of MSL function (22). Using four sets of primers to amplify separate regions of *runt*, we established the absence of any significant enrichment of H4Ac16 at this locus (Fig. 3). The locus of *runt* is in a relatively large gene-poor region of the X chromosome flanked

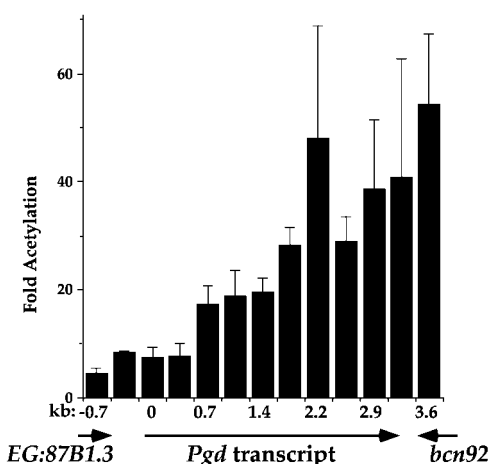


FIG. 1. **Distribution of H4Ac16 at the *Pgd* gene.** An antibody against histone H4AcK16 was used in a ChIP assay from *Drosophila* embryos. PCR amplification of various regions along the known dosage-compensated gene *Pgd* was monitored by Sybr Green real-time PCR. The start site of transcription for *Pgd* is numbered zero, and other regions analyzed are labeled in kilobase pair distances from this point. *EG:87B1.3* and *bcn92* are two genes closely flanking *Pgd*. Enrichment of a sequence in H4Ac16 ChIPs is relative to a segment of autosomal *Gpdh*.

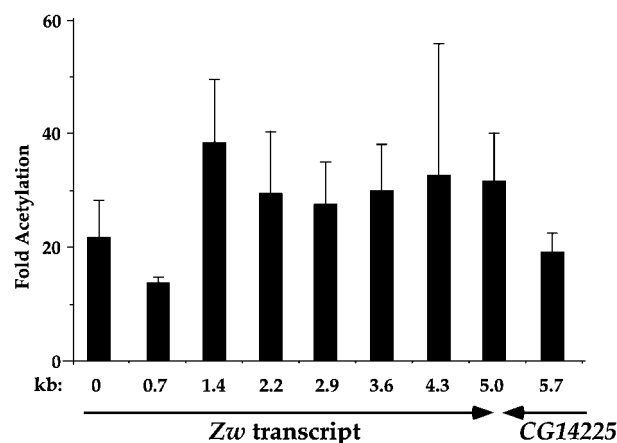


FIG. 2. **Distribution of H4Ac16 along the *Zw* gene that encodes glucose-6-phosphate dehydrogenase.** ChIP-PCR was performed as in Fig. 1, except the regions of *Zw* analyzed by PCR are spaced farther apart. High levels of acetylation are seen along the *Zw* gene and into the adjacent *CG14225* gene.

on either side by heavy clusters of genes. We designed primer sets for several of these flanking genes and failed to find any enrichment for H4Ac16 (data not shown), suggesting that this may be a chromosomal domain to which the MSLs have not spread. The nearest site of enrichment for H4Ac16 that we detected is at the *Tak1/Cg1812* region that is ~100 kb distant from *runt* (Fig. 3).

MSL1 Cross-linking to a Site of Complex Assembly—Our attempts to map the MSL complex along the X-linked genes described above yielded variable results that can be best ascribed to cross-linking efficiency or to antisera affinity problems. To increase the sensitivity of our assay, we attempted to map the complex to a region of the X chromosome where it may be enriched, *i.e.* to a chromosomal entry site. Among the 30–40 entry sites that have been identified by cytoimmunofluorescence, two are known to contain the *roX* genes (14, 16). The MSL complex is found at either of the two *roX* genes when these are moved to ectopic autosomal sites and spreads *in cis* from the transgenes to new sequences. The association of the complex with the *roX* transgenes occurs even when the latter

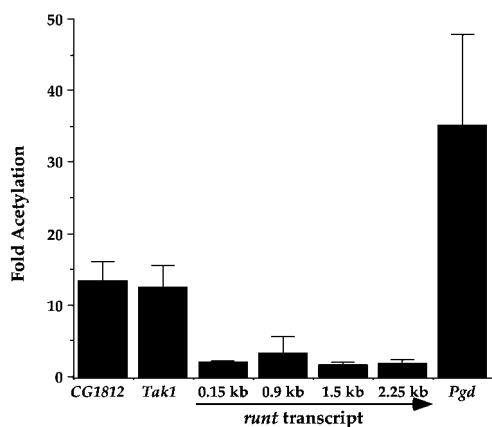


FIG. 3. **H4Ac16 is not enriched at the X-linked *runt* gene.** Four segments of the *runt* gene were tested for enrichment of H4Ac16 relative to autosomal *Gpdh*. *CG1812* and *Tak1* are the two closest genes to *runt* at which we found enrichment for H4Ac16 (~100 kb from *runt*).

are not transcribing, suggesting that the *roX* genes are not only sites of assembly for the complex but are also entry sites (14, 16, 23).

Using an MSL1 antiserum to precipitate the MSL complex (11) and primers to *roX2*, we determined a high enrichment of the complex at this locus (Fig. 4). Recently, a similar enrichment of MSLs was observed at *rox1* (23). A segment of the 3' end of the *Pgd* gene also shows a reproducible but modest enrichment for the MSL complex. Both *Pgd* sequences and *roX2* sequences show similar levels of enrichment for H4Ac16.

DISCUSSION

Global Histone Acetylation by the MSL Complex—Several examples of the *in vivo* recruiting of histone acetyltransferase-containing complexes to specific genes by DNA-binding activator proteins have been reported in yeast (2–5). The resulting localized hyperacetylation of nucleosomes in the promoter domain of these genes can occur in the absence of transcription but is necessary for its inception (2). In addition to this targeted effect, a background level of global histone acetylation is present throughout the yeast genome, leading to the hypothesis that it may function to enhance the competency of genes for the activation of transcription (2, 4, 6). Histone acetyltransferases responsible for global acetylation may perform their function by docking on nucleosomes in a random manner (24) or by associating with an elongating polymerase complex (25).

In contrast to the genome-wide acetylation found in yeast, the global acetylation of X-chromosome domains by the MSL complex in *Drosophila* has a clear role in transcription: a 2-fold enhancement in males. Although MOF can be targeted to a promoter in yeast resulting in a large induction of transcription (9), in *Drosophila*, the MSL complex is unlikely to play a role in gene activation and, therefore, is probably not targeted to promoters. This contention is supported by the fact that the complex is absent in females and by the observation that male embryos lacking the complex will develop up to the third larval instar without exhibiting any specific developmental defects; lethality in these individuals is the result of an imbalance in the relative level of X-linked and autosomal gene products. Concordant with these considerations is our finding of the relatively low level of MSL-dependent acetylation around the promoters of *Pgd*. In addition to the absence of targeting, the complex may be impaired from acetylating these nucleosomes by the presence of other chromatin remodeling modifications to the H4 tail or by the binding of other proteins that would prevent MOF from recognizing it as a substrate. A potentially related phenomenon has been described recently in yeast (26).

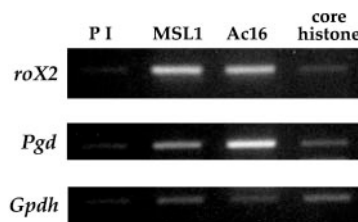


FIG. 4. **MSL complex and H4Ac16 are enriched at an X-chromosomal entry site.** ChIP with an antibody against MSL1 and the corresponding preimmune serum revealed particularly high levels of enrichment at *roX2* and modest enrichment at *Pgd*. H4Ac16 is enriched at similar levels on both genes relative to *Gpdh*, demonstrating the qualitative difference between an entry site and *Pgd*.

Global Histone Acetylation Is Required for the Spreading of the MSL Complex—The histone acetyltransferase activity of MOF is not only directly correlated to the 2-fold enhancement of the rate of transcription of X-linked genes in males, but it is also required for the spreading of the MSL complex. This can be concluded from the observation that an intact complex with an inactive MOF subunit fails to spread beyond the entry sites (17, 18). Furthermore, MOF overexpression results in the presence of H4Ac16 throughout the genome, and the complex is no longer limited to the X chromosome, but is found along all of the autosomes (18). This last observation suggests that broad patterns of acetylation could result if nucleosomes modified by MOF become new sequence-independent binding sites for the complex. A similar property is exhibited by the NuA4 complex of yeast. In an *in vitro* system that is not transcribing, the NuA4 complex can acetylate large domains of chromatin when targeted to a specific site, suggesting that the complex has the intrinsic ability to spread (27). This multiprotein complex includes the only yeast homologs of known MSL subunits: Esa1p (MOF) and Eaf3p (MSL3) (28) and shares with the MSL complex the ability to acetylate histone H4 at lysine 16 (29).

Global Histone Acetylation and Transcription Elongation—A role for global acetylation on transcription was suggested by Tse *et al.* (30) who showed that processivity of RNA polymerase III through an array of nucleosomes was greatly increased by histone acetylation. This effect on transcription elongation correlated with an unfolded state of the highly acetylated nucleosomal arrays. In addition to the effect of histone acetylation on chromatin structure, transcription by T7 RNA polymerase through a single nucleosome was equally efficient whether utilizing tailless or hyperacetylated nucleosomal templates (31). In this context, we suggest that the distribution of H4Ac16-containing nucleosomes along the length of a dosage-compensated gene results in a reduction in the time required for completing a transcript. If reinitiation is the rate-limiting step in production of transcripts, enhancement of elongation could increase gene expression if termination were coupled to reinitiation. Evidence that recycling of terminating polymerases enhances reinitiation rates has been found for polymerase I and polymerase III genes (32, 33).

Targeting Global Histone Acetylation to Chromosomal Domains—Our data provide a molecular confirmation in diploid cells of the discontinuous distribution of the MSL complex and of H4Ac16, as determined in larval salivary gland polytene chromosomes by cytoimmunofluorescence. Antisera to MSL1 (and MSL2, data not shown) are able to efficiently immunoprecipitate *roX2* sequences (Fig. 4) in contrast to the other gene loci tested. This can be explained by the special role of *roX* genes as entry sites: continuous recruitment for assembly and/or the presence of high affinity binding sequences at these loci would favor more efficient cross-linking. The physical proximity of a gene to an entry site may determine whether it

attracts the MSL complex and whether histone H4 of its nucleosomes becomes acetylated at lysine 16.

The nonrandom pattern of global acetylation along the X chromosome reflects specific regulatory differences in X-linked gene transcription. The compensation of the X-linked gene *runx* appears to be under the direct control of the *Sxl* gene and does not involve the MSL complex (22). As expected, since *Sxl* makes a functional product only in females, equal levels of *runx* expression in the two sexes may be achieved by reducing (halving) the expression of the two doses of the gene in females (34). These considerations are fully concordant with the absence of H4Ac16 on *runx* nucleosomes. By analyzing the regions adjacent to the locus of *runx*, we discovered that this gene resides in an extensive H4Ac16-free region.

There is a growing list of examples of global domains of acetylation associated with specific regulatory sequences. An extensive region of acetylation at the human β -globin locus is linked to a sequence 5' of the locus control region and may require relocalization to a particular nuclear domain (35). The entire length of the human growth hormone gene cluster is characterized by an enrichment of H3 and H4 acetylation that appears to emanate from a DNase 1-hypersensitive site in the locus control region (36). In *Drosophila* dosage compensation, a model is emerging where the MSL complex is targeted to specific entry sites within a chromosomal domain and subsequently spreads via histone acetylation throughout this domain to enhance levels of transcription. The mode of targeting and spreading of the MSL complex within nuclear domains may be a paradigm of targeted global remodeling of chromatin responsible for the regulation of large groups of genes.

Acknowledgments—We are grateful to Paul Fisher for his generous gift of embryos and to Krista Fehr for technical assistance. We thank Weigang Gu, Antonio Pannuti, Georgette Sass, Guy Beresford, and Jerry Boss for helpful discussions.

REFERENCES

1. Strahl, B. D., and Allis, C. D. (2000) *Nature* **403**, 41–45
2. Kuo, M. H., vom Baur, E., Struhl, K., and Allis, C. D. (2000) *Mol. Cell* **6**, 1309–1320
3. Kuo, M. H., Zhou, J., Jambeck, P., Churchill, M. E., and Allis, C. D. (1998) *Genes Dev.* **12**, 627–639
4. Reid, J. L., Iyer, V. R., Brown, P. O., and Struhl, K. (2000) *Mol. Cell* **6**, 1297–1307
5. Utley, R. T., Ikeda, K., Grant, P. A., Cote, J., Steger, D. J., Eberharter, A., John, S., and Workman, J. L. (1998) *Nature* **394**, 498–502
6. Vogelauer, M., Wu, J., Suka, N., and Grunstein, M. (2000) *Nature* **408**, 495–498
7. Turner, B. M., Birley, A. J., and Lavender, J. (1992) *Cell* **69**, 375–384
8. Bone, J. R., Lavender, J., Richman, R., Palmer, M. J., Turner, B. M., and Kuroda, M. I. (1994) *Genes Dev.* **8**, 96–104
9. Akhtar, A., and Becker, P. B. (2000) *Mol. Cell* **5**, 367–375
10. Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A., and Lucchesi, J. C. (1997) *EMBO J.* **16**, 2054–2060
11. 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
12. Akhtar, A., Zink, D., and Becker, P. B. (2000) *Nature* **407**, 405–409
13. Amrein, H., and Axel, R. (1997) *Cell* **88**, 459–469
14. Meller, V. H., Gordadze, P. R., Park, Y., Chu, X., Stuckenholtz, C., Kelley, R. L., and Kuroda, M. I. (2000) *Curr. Biol.* **10**, 136–143
15. Meller, V. H., Wu, K. H., Roman, G., Kuroda, M. I., and Davis, R. L. (1997) *Cell* **88**, 445–457
16. Kelley, R. L., Meller, V. H., Gordadze, P. R., Roman, G., Davis, R. L., and Kuroda, M. I. (1999) *Cell* **98**, 513–522
17. Gu, W., Szauter, P., and Lucchesi, J. C. (1998) *Dev Genet* **22**, 56–64
18. Gu, W., Wei, X., Pannuti, A., and Lucchesi, J. C. (2000) *EMBO J.* **19**, 5202–5211
19. Orlando, V., Jane, E. P., Chinwalla, V., Harte, P. J., and Paro, R. (1998) *EMBO J.* **17**, 5141–5150
20. Belote, J. M., and Lucchesi, J. C. (1980) *Nature* **285**, 573–575
21. Scott, M. J., and Lucchesi, J. C. (1991) *Gene (Amst.)* **109**, 177–183
22. Gergen, J. P. (1987) *Genetics* **117**, 477–485
23. Kageyama, Y., Mengus, G., Gilfillan, G., Kennedy, H. G., Stuckenholtz, C., Kelley, R. L., Becker, P. B., and Kuroda, M. I. (2001) *EMBO J.* **20**, 2236–2245
24. Sendra, R., Tse, C., and Hansen, J. C. (2000) *J. Biol. Chem.* **275**, 24928–24934
25. Wittschieben, B. O., Otero, G., de Bizemont, T., Fellows, J., Erdjument-Bromage, H., Ohba, R., Li, Y., Allis, C. D., Tempst, P., and Svejstrup, J. Q. (1999) *Mol. Cell* **4**, 123–128
26. Deckert, J., and Struhl, K. (2001) *Mol. Cell Biol.* **21**, 2726–2735
27. Vignali, M., Steger, D. J., Neely, K. E., and Workman, J. L. (2000) *EMBO J.* **19**, 2629–2640
28. Allard, S., Utley, R. T., Savard, J., Clarke, A., Grant, P., Brandl, C. J., Pillus, L., Workman, J. L., and Cote, J. (1999) *EMBO J.* **18**, 5108–5119
29. Ohba, R., Steger, D. J., Brownell, J. E., Mizzen, C. A., Cook, R. G., Cote, J., Workman, J. L., and Allis, C. D. (1999) *Mol. Cell Biol.* **19**, 2061–2068
30. Tse, C., Wolffe, A. P., and Hansen, J. C. (1998) *Mol. Cell Biol.* **18**, 4629–4638
31. Protacio, R. U., Li, G., Lowary, P. T., and Widom, J. (2000) *Mol. Cell Biol.* **20**, 8866–8878
32. Dieci, G., and Sentenac, A. (1996) *Cell* **84**, 245–252
33. Jansa, P., Burek, C., Sander, E. E., and Grummt, I. (2001) *Nucleic Acids Res.* **29**, 423–429
34. Kelley, R. L., Solovyeva, I., Lyman, L. M., Richman, R., Solovyev, V., and Kuroda, M. I. (1995) *Cell* **81**, 867–877
35. Schubeler, D., Francastel, C., Cimborra, D. M., Reik, A., Martin, D. I., and Groudine, M. (2000) *Genes Dev.* **14**, 940–950
36. Elefant, F., Cooke, N. E., and Liebhaber, S. A. (2000) *J. Biol. Chem.* **275**, 13827–13834