FEBS Letters 572 (2004) 154–158 FEBS 28673

A short autonomous repression motif is located within the N-terminal domain of CTCF

Lars Drueppel, Klaus Pfleiderer, Andreas Schmidt, Wolfgang Hillen, Christian Berens*

Lehrstuhl für Mikrobiologie, Institut für Mikrobiologie, Biochemie und Genetik, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstraße 5, D-91058 Erlangen, Germany

Received 10 May 2004; revised 13 July 2004; accepted 13 July 2004

Available online 22 July 2004

Edited by Takashi Gojobori

Abstract The vertebrate transcription factor CTCF is not only involved in transcriptional activation, insulation and genomic imprinting, but also in transcriptional repression. Sequence motifs mediating these activities have not been identified so far. We have mapped a short repression motif to residues 150–170 within the N-terminal domain of CTCF. This motif is active in HeLa, HEK293 and COS-7 cell lines where it is both sufficient and necessary for silencing either an SV40-, or a CMV-enhancer. It also represses the basal activity of an SV40 core promoter. Since this autonomous repression motif displays no sequence similarity to any other regulatory protein, it represents a yet unknown co-repressor recruiting motif.

Keywords: CTCF; Repression; Tet

1. Introduction

The transcription factor CTCF is not only involved in enhancer blocking, but also in transcriptional activation, imprinting of genomic information and in various human malignancies as a putative tumor suppressor [1]. It further participates in transcriptional repression of the chicken *lysozyme* gene locus [2] and the chicken, mouse and human *myc* loci [3]. The chicken CTCF protein consists of 728 aa and has been divided into three domains: an N-terminal domain spanning the first 267 aa, a central domain comprising 11 zinc-fingers and a C-terminal domain ranging from aa 578 to 728 [3]. As fusions of either domain to Gal4 led to repression of reporter gene activity in various cell lines [3,4], autonomous repression domains seem to be distributed throughout the protein. Co-precipitation experiments identified SIN3A [4], YB1 [5], nucleophosmin and CTCF itself [6] as interaction partners. The presence of SIN3A, which is part of a large co-repressor complex containing HDAC1 and HDAC2 [7], suggests that transcriptional repression by CTCF is mediated by histone deacetylation. Others, however, were unable to confirm an association of CTCF with HDAC1 and HDAC2 in human MCF-7 breast cancer cells [8]. So far, no further interaction partners of CTCF have been published nor have repression motifs within CTCF been mapped and characterized.

We have analyzed subdomains of CTCF for repression activity in a double transient, doxycycline (dox) dependent enhancer silencing assay. We mapped repression activity to an element spanning residues CTCF(150–170) in HeLa cells. This motif displays no obvious sequence similarity to any other regulatory protein. Furthermore, it is both necessary and sufficient for CTCF mediated repression of enhancer activity in various cell lines and also represses the activity of a core promoter.

2. Materials and methods

2.1. Construction of reporter plasmids

All reporter plasmids are based on the pGL3 vector series (PROMEGA). The plasmid pWHE206 encodes a luciferase gene under control of an SV40 promoter/enhancer with $(tetO)_7$ elements flanking the enhancer (Ameres et al., submitted). Introduction of DNA spacers into pWHE206 was carried out as follows: A unique *Eco*RV restriction site was introduced 3′ to the *Pst*I site with the primer "EcoRV-SV40enh" (GTGGTAAACTGCAGAAGGATATCTGAACGATGGAGCGG). 500 bp spacer fragments were created by digesting pUHD16-1 [9] with either *Bcl*I/*Eco*47III or *Bcl*I/*Cla*I. Blunt ends were generated using T4-DNA-Polymerase. These spacer fragments were integrated into the blunt ended restriction site *Bam*HI or *Eco*RV, respectively, resulting in pWHE295. A wild-type $(tetO)_7$ box was amplified from pUHC13-3 [9] and inserted into the *Bgl*II site of pGL3-Promoter resulting in pWHE248. pWHE279 was constructed as follows: The hCMV core promoter (+75/−53) [9] was amplified from pWHE120(sB) [10] and introduced into pGL3-Promoter using *Bgl*II and *Nco*I restriction sites, thereby replacing the SV40 core promoter. The $(tetO)_7$ element was isolated from pWHE201 and introduced into the *Bgl*II site. The pCMV enhancer (−54/−675) [9] element was amplified from pWHE120(sB) using the primers "CMV-SalI(-54)" and "CMV-PstI(full)" and was introduced into the *Xho*I site resulting in pWHE279.

2.2. Construction of Tet-transregulators

The coding sequence for full length CTCF(2–728) was amplified from pBSKII-CTCF-FL(2–728) and inserted into pWHE120(sB) [10] using restriction sites for *Ngo*MIV and *Bgl*II, resulting in pWHE210. The *sc tetR* allele was excised from pWHE120(sB + B) [10] and introduced into pWHE210 using the *Spe*I and *Ngo*MIV sites, thereby replacing *tetR* resulting in pWHE219. The coding sequences of all CTCF subdomains were amplified by PCR or overlap extension PCR from pBSKII-CTCF-FL(2–728) and inserted into pWHE219 using

* Corresponding author. Fax: +49-9131-8528082.
E-mail address: cberens@biologie.uni-erlangen.de (C. Berens).

Abbreviations: CTCF, CCCTC binding factor; dox, doxycycline; sc TetR, single chain Tet repressor; SV, simian virus; CMV, cytomegalovirus

0014-5793/$22.00
doi:10.1016/j.febslet.2004.07.027

restriction sites for *Ngo*MIV and *Age*I, thereby replacing the *CTCF(2–728)* allele.

The sequences of the oligonucleotides used and the plasmids constructed in this study are available upon request.

2.3. Transient transfections

Transfection of HeLa, HEK293 or COS-7 cells were performed with PerFectin™ according to the instructions of the producer (PEQLAB). The DNA mix used to examine the regulatory properties of different regulator/reporter combinations contained 400 ng of pUHD16-1 [9] constitutively expressing β-galactosidase to monitor transfection efficiency, 100 ng pGL3-Promoter or equimolar amounts of other reporter plasmids, 25 ng of the regulator plasmid control pWHE355, coding for sc TetR* lacking a regulatory domain (Ameres et al., submitted), or equimolar amounts of other regulator expressing plasmids and pWHE121 [10] as non-specific DNA to a total amount of 1 μg per well. After 4 h incubation, transfection reactions were stopped by adding 1 ml of D-MEM medium supplemented with 20% FBS. For induction, dox (Sigma) was added to a final concentration of 1 or 5 μg/ml.

2.4. Luciferase assay

Cells were harvested after 24 h incubation and lysed with 100 μl of 25 mM Tris-phosphate, 2 mM EDTA (pH 8.0), 5% glycerol, 1% Triton X-100 and 20 mM DTT. A 30 μl aliquot was used to determine luciferase activity in 100 mM potassium-phosphate-buffer (pH 7.8) containing 15 mM magnesium-sulfate, 5 mM ATP and 0.18 mM D-Luciferin (Boehringer). Luciferase activity was normalized for protein content (Bio-Rad) and β-galactosidase activity [11] and represented by the means of triplicate samples with standard deviation given in corrected light units (corr. RLU) per μg of total cell protein.

2.5. Western blot analysis of transregulators in HeLa cells

20 μg total protein from cell extracts prepared for the luciferase assays was separated on a Novex 4–12% precast Tris-Glycine polyacrylamide (PAA) gradient gel (Invitrogen) and transferred to a polyvinyldifluoride membrane (Amersham). The membrane was blocked with I-Block (Tropix). A polyclonal TetR serum from rabbit (laboratory stock) was used to detect Tet-transregulators. Signal intensities were normalized to the β-actin signals (Sigma). Bound antibodies were visualized using the ECL Plus kit (Amersham).

3. Results

3.1. The N-terminal domain of CTCF harbors repression activity

A transient test assay based on components of the Tet system (reviewed in [12]) was constructed to determine if dox-dependent CTCF binding is capable of silencing an SV40-enhancer. The assay consists of two plasmids, one containing a dox-dependent reporter gene and the other constitutively expressing a TetR-based transregulator.

The reporter plasmid pWHE295 encodes a luciferase reporter gene under control of an SV40-promoter which is activated by an SV40-enhancer located 2 kb downstream from it. Since we use covalently closed plasmids in our assay, the enhancer is also located 2.8 kb upstream of the promoter. In order to silence this enhancer in a dox-dependent manner, it was consequently flanked with heptamerized *tet* operators [9]. To rule out enhancer-silencing by "quenching" [13], 500 bp spacer elements were inserted between the SV40-enhancer and the $(tetO)_7$ elements. The regulator plasmids encode dox-controlled transregulators, which consist of full length (728 aa) or truncated versions of CTCF [4,14] (Fig. 1) fused to the C-terminal end of sc TetR (442 aa) [10]. This setup fuses a single regulatory domain to a monomerized TetR, the DNA-binding and effector-responding unit. This enables us to evaluate the activity of an isolated domain without potential interference from the presence of a nearby second identical

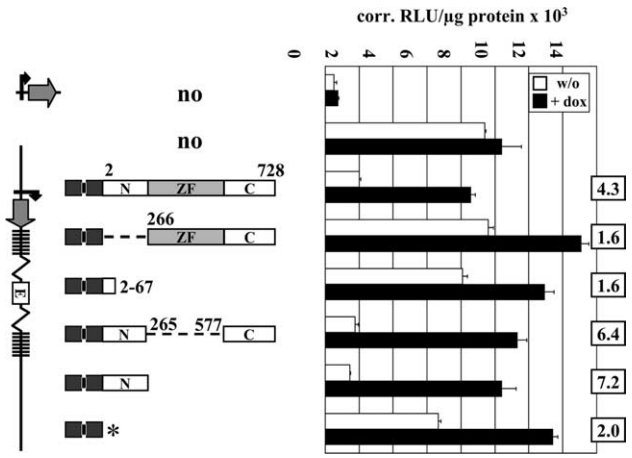

Fig. 1. Characterization of CTCF-domains in an SV40-based enhancer-silencing assay. HeLa cells were transiently transfected with pGL3-Promoter, lacking the SV40-enhancer (E) and *tet* operators to determine unenhanced luciferase activity, or pWHE295 carrying a luciferase gene (gray arrow) under control of an SV40 core promoter (small black arrow) and an SV40-enhancer flanked with $(tetO)_7$ elements (array of seven black boxes). pWHE295 was cotransfected with plasmids coding for sc TetR-CTCF, -CTCF(267–728), -CTCF(2–67), -CTCF(2–265/577–728), -CTCF(2–265) or sc TetR*. A schematic view of the transfected transregulators is depicted. Repression factors are indicated next to the bars.

domain, as is the case in a setup using the Gal4 DNA binding domain. This is noteworthy since CTCF interacts with itself [6,15]. We additionally deleted the zinc-finger domain as the region of interaction between CTCF and the co-repressor SIN3A [4]. Also, an N-terminal aa motif CTCF(2–67) that harbors a barrier element active in *Saccharomyces cerevisiae* was analyzed [14]. These TetR-based transregulators bind with high specificity and affinity to the *tetO* elements of the reporter plasmid and, if they contain an active regulatory domain, will silence enhancer activity. Addition of dox leads to a conformational change in the TetR moiety resulting in the loss of *tetO* binding and subsequent restoration of enhancer-mediated reporter gene expression. Transregulator expressing plasmids were transiently cotransfected with pWHE295 in HeLa cells (Fig. 1). In this experiment, the three transregulators carrying an entire N-terminal domain of CTCF silence an SV40-enhancer in a dox-dependent manner. Since Western blot analysis shows that all CTCF-based transregulators are present (data not shown), we attribute repression activity solely to the N-terminal domain, excluding residues 2–67.

3.2. Mapping a repression motif within the N-terminal domain of CTCF

As the first 67 aa of the N-terminal domain do not repress luciferase activity, the N-terminal domain was further characterized in HeLa cells by fusing the sections CTCF(2–190), CTCF(130–265) and CTCF(130–190) to sc TetR. All three subdomains repress luciferase activity in a dox-dependent manner similar to full length CTCF or the isolated N-terminal domain. The repression motif was thus narrowed down to aa 130–190 of CTCF. Next, CTCF(130–170), CTCF(150–190), CTCF(130–150), CTCF(150–170) and CTCF(170–190) were fused to sc TetR and analyzed. CTCF(130–170), CTCF(150–190) and CTCF(150–170) repress luciferase

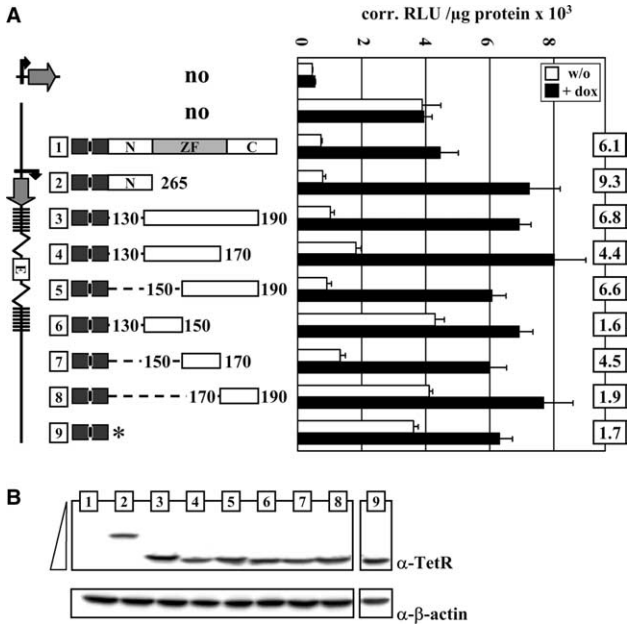

Fig. 2. Identification of a 21 aa repression motif within the N-terminal domain of CTCF. (A) HeLa cells were transiently transfected with pGL3-Promoter, lacking the SV40-enhancer (E) and *tet* operators to determine unenhanced luciferase activity, or pWHE295 carrying a luciferase gene (gray arrow) under control of an SV40 core promoter (small black arrow) and an SV40-enhancer flanked with $(tetO)_7$ elements (array of seven black boxes). pWHE295 was cotransfected with plasmids coding for sc TetR-CTCF, -CTCF(2–265), -CTCF(130–190), -CTCF(130–170), -CTCF(150–190), -CTCF(130–150), -CTCF(150–170), -CTCF(170–190) or sc TetR*. A schematic view of the transfected transregulators is depicted. Repression factors are indicated next to the bars. (B) Expression of CTCF-based transregulators in transiently transfected HeLa cells. Crude cell lysates from transiently transfected HeLa cells were separated on a 4–12% gradient PAA gel and subjected to Western blot analysis. The blot was probed with anti-TetR or anti-β-actin antibodies.

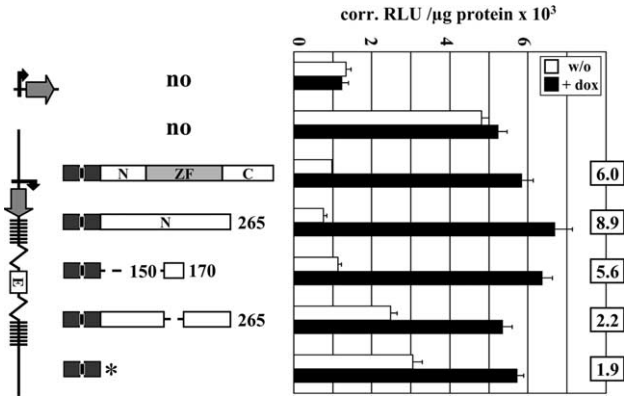

Fig. 3. Deletion of CTCF(150–170) out of the N-terminal domain of CTCF results in an inactive transregulator. HeLa cells were transiently transfected with pGL3-Promoter, lacking the SV40-enhancer (E) and *tet* operators to determine unenhanced luciferase activity, or pWHE295 carrying a luciferase gene (gray arrow) under control of an SV40 core promoter (small black arrow) and an SV40-enhancer flanked with $(tetO)_7$ elements (array of seven black boxes). pWHE295 was cotransfected with plasmids coding for sc TetR-CTCF, -CTCF(2–265), -CTCF(150–190), -CTCF(2–149/171–265) or sc TetR*. A schematic view of the transfected transregulators is depicted. Repression factors are indicated next to the bars.

activity similar to full length CTCF (Fig. 2A). Less than 2-fold repression of luciferase activity was measured for CTCF(130–150), CTCF(170–190) and the sc TetR* control. Western blot analysis of crude cell lysates from the transient transfection (Fig. 2B) shows that all CTCF-based transregulators except sc TetR-CTCF are present at similar levels. We were able to detect sc TetR-CTCF in Western blots by loading twice the amount of total protein in combination with long exposure times (data not shown). Lack of expression is thus not responsible for the failure of the respective transregulators to repress luciferase expression in the transient transfection experiment. We then deleted the aa 150–170 segment from the N-terminal domain of CTCF. The resulting construct was inactive, like the sc TetR* control (Fig. 3). Taken together, the 21 aa motif CTCF(150–170) is both sufficient and necessary for silencing activated transcription from an SV40-enhancer in HeLa cells.

3.3. CTCF(150–170) represses the activity of an SV40 core promoter

We then targeted CTCF based transregulators to a promoter proximal $(tetO)_7$ element in order to test their ability to repress the basal activity of an SV40 core promoter (Fig. 4A). In the absence of dox, the isolated N-terminal domain and CTCF(150–170) repress luciferase activity to levels below the background represented by pWHE248 alone. Less than 2-fold repression of luciferase activity was determined for CTCF(170–190) and the sc TetR* control. Upon administration of dox, repression levels are relieved to the level represented by pWHE248. Therefore, the 21 aa long, N-terminal repression motif CTCF(150–170) is not only limited to silencing an SV40-enhancer, but also represses the activity of an SV40 core promoter.

3.4. CTCF(150–170) represses the activity of a CMV-enhancer in various cell lines

Since the N-terminal domain of CTCF was attributed to harbor cell-type specific repression activity [4], we further tested CTCF based transregulators in transiently transfected HEK293 and COS-7 cells. For this, we used a CMV-based reporter plasmid to determine the ability of CTCF subdomains to repress a stronger enhancer (Fig. 4B). In this experiment full length CTCF, the isolated N-terminal domain, CTCF(130–190) and CTCF(150–170) repress the proximal CMV-enhancer in both cell lines in a dox-dependent manner. Taken together, the N-terminal repression motif CTCF(150–170) silences activated transcription from a distal SV40-, or a proximal CMV-enhancer, represses basal transcription of an SV40 core promoter and is active in HeLa, HEK293 and COS-7 cell lines.

4. Discussion

The multivalent transcription factor CTCF not only participates in transcriptional activation and enhancer blocking, but also in transcriptional repression. Gal4-based CTCF fusions revealed several autonomous regulatory domains, some of which act in a cell-type specific manner (for reviews see [1,16]). CTCF binds its cognate target sites, associated with diverse CTCF activities, by combinatorial use of the eleven zinc fingers possibly resulting in alternative protein confor-

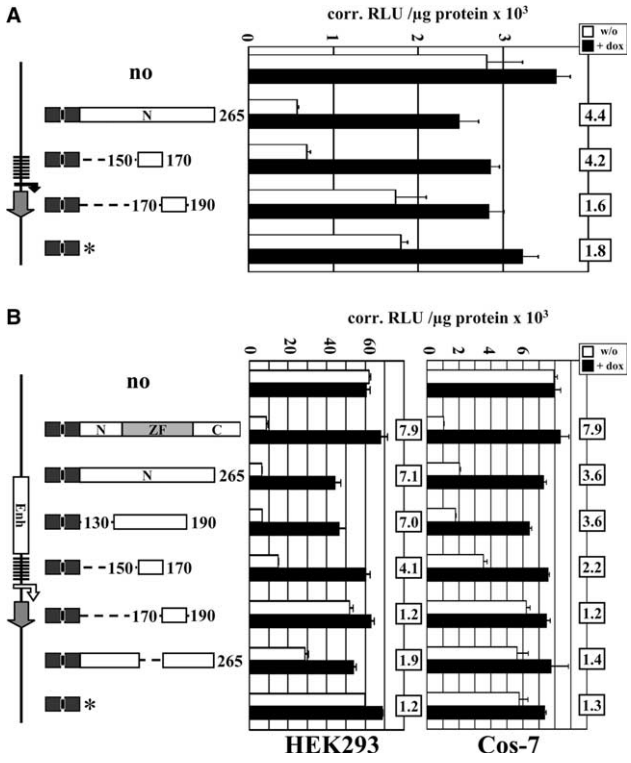

Fig. 4. (A) CTCF(150–170) represses the activity of an SV40 core promoter. HeLa cells were transiently cotransfected with pWHE248, carrying a luciferase gene (gray arrow) under control of an SV40 core promoter (small black arrow) and a promoter proximal $(tetO)_7$ element (array of seven black boxes), and plasmids coding for sc TetR-CTCF(2–265), -CTCF(150–170), -CTCF(170–190) or sc TetR*. A schematic view of the transfected transregulators is depicted. Repression factors are indicated next to the bars. (B) The identified repression motifs are active in HEK293 and COS-7 cells. HEK293 and COS-7 cells were transiently cotransfected with pWHE279, carrying a luciferase gene (gray arrow) under control of a CMV core promoter (small white arrow)/enhancer (Enh) interspersed with a $(tetO)_7$ element (array of seven black boxes), and plasmids coding for sc TetR-CTCF, sc TetR-CTCF(2–265), sc TetR-CTCF(130–190), sc TetR-CTCF(150–170), sc TetR-CTCF(170–190), sc TetR-CTCF(2–149/171–265) or sc TetR*. A schematic view of the transfected transregulators is depicted. Repression factors are indicated next to the bars.

mations that present or mask different regulatory motifs [1,14]. By fusing subdomains of CTCF to sc TetR, we were able to characterize the activity of single, defined CTCF subdomains without interference from a second identical domain nearby, as might be the case in a setup using dimeric Gal4 DNA binding domains. This is particularly noteworthy, since CTCF can interact with itself [6,15]. Using our transient SV40-based enhancer silencing assay, we mapped repression activity to the N-terminal domain of CTCF (Fig. 1) [3,4]. In contrast, we did not observe repression activity in the zinc finger and C-terminal domains [3,4] in HeLa cells.

The first 65 aa of CTCF display barrier activity in yeast [14], but failed to repress luciferase activity in our assay (Fig. 1). However, recent studies have demonstrated that many proteins, including the isolated DNA-binding domain of the prokaryotic repressor LexA, exhibit barrier function in yeast [17,18]. Moreover, CTCF does not display barrier activity in vertebrates [19,20]. Utilizing our enhancer silencing assay in HeLa cells, we instead mapped repression activity to the 21 aa motif CTCF(150–170) (Fig. 2). This motif is sufficient and necessary for silencing a distal SV40- (Fig. 3), as well as a proximal CMV-enhancer (Fig. 4B). Moreover, it represses the basal activity of an SV40 core promoter (Fig. 4A) and is active in HeLa, HEK293 and COS-7 cell lines (Fig 4B), demonstrating that this 21 aa element represents a bona fide autonomous repression motif. Except for the first residue, this element is identical in the CTCF proteins from *Xenopus*, mouse, chicken and man. In eukaryotes, transcription factors frequently mediate repression by recruiting co-repressors via short amino acid motifs [21–23], like "Pro–Xaa–Asp–Leu–Ser", a motif recognized by the co-repressor family of "C-terminal binding proteins" [24] or the four amino acid motif "Trp–Arg–Pro–Trp" that recruits "Transducin-like Enhancers of Split" as co-repressors [25]. Since BLAST or FASTA analysis of CTCF(150–170) revealed no identical clusters to those motifs or to any other regulatory proteins (data not shown), we suggest that we have identified a yet unknown co-repressor recruiting motif.

Acknowledgements: We thank Drs. Renkawitz and Lutz, University of Giessen, for kindly providing pBSKII-CTCF-FL(2–728), J. Guthmann, C. Krueger and J. Daam for the construction of CTCF based transregulators. This work was supported by the Deutsche Forschungsgemeinschaft through SFB 473, the Graduiertenkolleg 805 and the Fonds der Chemischen Industrie Deutschlands.

References

[1] Ohlsson, R., Renkawitz, R. and Lobanenkov, V. (2001) Trends Genet. 17, 520–527.
[2] Lutz, M., Burke, L.J., LeFevre, P., Myers, F.A., Thorne, A.W., Crane-Robinson, C., Bonifer, C., Filippova, G.N., Lobanenkov, V. and Renkawitz, R. (2003) EMBO J. 22, 1579–1587.
[3] Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J. and Lobanenkov, V.V. (1996) Mol. Cell. Biol. 16, 2802–2813.
[4] Lutz, M., Burke, L.J., Barreto, G., Goeman, F., Greb, H., Arnold, R., Schultheiß, H., Brehm, A., Kouzarides, T., Lobanenkov, V. and Renkawitz, R. (2000) Nucleic Acids Res. 28, 1707–1713.
[5] Chernukhin, I.V., Shamsuddin, S., Robinson, A.F., Carne, A.F., Paul, A., El-Kady, A.I., Lobanenkov, V.V. and Klenova, E.M. (2000) J. Biol. Chem. 275, 29915–29921.
[6] Yusufzai, T.M., Tagami, H., Nakatani, Y. and Felsenfeld, G. (2004) Mol. Cell 13, 291–298.
[7] Fleischer, T.C., Yun, U.J. and Ayer, D.E. (2003) Mol. Cell. Biol. 23, 3456–3467.
[8] Dunn, K.L., Zhao, H. and Davie, J.R. (2003) Exp. Cell Res. 288, 218–223.
[9] Gossen, M. and Bujard, H. (1992) Proc. Natl. Acad. Sci. USA 89, 5547–5551.
[10] Krueger, C., Berens, C., Schmidt, A., Schnappinger, D. and Hillen, W. (2003) Nucleic Acids Res. 31, 3050–3056.
[11] Miller, J.H. (1972) Experiments in Molecular Genetics, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.
[12] Berens, C. and Hillen, W. (2003) Eur. J. Biochem. 270, 3109–3121.
[13] Nibu, Y., Senger, K. and Levine, M. (2003) Mol. Cell. Biol. 23, 3990–3999.
[14] Defossez, P.A. and Gilson, E. (2002) Nucleic Acids Res. 30, 5136–5141.
[15] Pant, V., Kurukuti, S., Pugacheva, E., Shamsuddin, S., Mariano, P., Renkawitz, R., Klenova, E., Lobanenkov, V. and Ohlsson, R. (2004) Mol. Cell. Biol. 24, 3497–3504.
[16] Dunn, K.L. and Davie, J.R. (2003) Biochem. Cell. Biol. 81, 161–167.
[17] Bi, X., Yu, Q., Sandmeier, J.J. and Zou, Y. (2004) Mol. Cell. Biol. 24, 2118–2131.
[18] Yu, Q., Qiu, R., Foland, T.B., Griesen, D., Galloway, C.S., Chiu, Y.H., Sandmeier, J., Broach, J.R. and Bi, X. (2003) Nucleic Acids Res. 31, 1224–1233.

[19] Recillas-Targa, F., Pikaart, M.J., Burgess-Beusse, B., Bell, A.C., Litt, M.D., West, A.G., Gaszner, M. and Felsenfeld, G. (2002) Proc. Natl. Acad. Sci. USA 99, 6883–6888.
[20] Yao, S., Osborne, C.S., Bharadwaj, R.R., Pasceri, P., Sukonnik, T., Pannell, D., Recillas-Targa, F., West, A.G. and Ellis, J. (2003) Nucleic Acids Res. 31, 5317–5323.
[21] Lai, E.C. (2002) EMBO Rep. 3, 840–845.
[22] Christian, M., Tullet, J.M. and Parker, M.G. (2004) J. Biol. Chem. 279, 15645–15651.
[23] Eilers, A.L., Billin, A.N., Liu, J. and Ayer, D.E. (1999) J. Biol. Chem. 274, 32750–32756.
[24] Turner, J. and Crossley, M. (2001) Bioessays 23, 683–690.
[25] Grbavec, D. and Stifani, S. (1996) Biochem. Biophys. Res. Commun. 223, 701–705.