3830–3838 *Nucleic Acids Research, 2000, Vol. 28, No. 19*

Cloning and characterization of the histone-fold proteins YBL1 and YCL1

Fabrizio Bolognese, Carol Imbriano, Giuseppina Caretti and Roberto Mantovani[1,*]

Dipartimento di Genetica e Biologia dei Microrganismi, Università di Milano, Via Celoria 26, 20133 Milano, Italy and [1]Dipartimento di Biologia Animale, Università di Modena e Reggio, Via Campi 213, 41100 Modena, Italy

Received February 4, 2000; Revised May 26, 2000; Accepted July 24, 2000 DDBJ/EMBL/GenBank accession nos AF230805, AF230806

ABSTRACT

Histones are among the most conserved proteins in evolution, sharing a histone fold motif. A number of additional histonic proteins exist and are involved in the process of transcriptional regulation. We describe here the identification, cloning and characterization of two small members of the H2A–H2B sub-family (YBL1 and YCL1) related to the NF-YB and NF-YC subunits of the CCAAT-binding activator NF-Y and to the TATA-binding protein (TBP) binding repressor NC2. Unlike the latters, YBL1 and YCL1 have no intrinsic CCAAT or TATA-binding capacity. In nucleosome reconstitution assays, they can form complexes with histones in solution and on DNA and they are part of relatively large complexes, as determined by glycerol gradient experiments. Our data support the idea that YBL1 and YCL1 are divergent with respect to NF-YB and NF-YC for specific functions, but have coevolved the capacity to interact with nucleosomal structures.

INTRODUCTION

Chromatin structure plays a fundamental role in cellular physiology, enabling DNA to be conveniently packed in the relatively small nuclear space. At the same time, chromatin can be a formidable barrier to the cellular processes that involve DNA, such as gene transcription, repair, replication and recombination; recent advances have clearly established that chromatin plays an active role in regulating these functions (1). The fundamental chromatin unit is the nucleosome, a DNA–protein complex formed by 146 bp of DNA wrapped around core histones H2A, H2B, H3 and H4. Histones, probably the most conserved proteins in evolution, are formed by variable N-terminal tails, subject to several levels of post-translational regulation such as acetylation, phosphorylation, methylation (2) and by the quasi-invariant C-terminal sequences. Analysis of the latter domains revealed a 65 amino acid histone fold motif (HFM) shared by all histone proteins, with low sequence identity (14/18%) and high structural resemblance (3). Crystallographic analysis detailing the nucleosomal structure showed that the histone fold motif is composed of three α-helices separated by short loops/strand regions, enabling histones to heterodimerize in a compact form and bind DNA in a non-sequence-specific way (4). Histone-like proteins containing the HFM are also found in *Archaea* and are thought to homodimerize and complex DNA (5). Recently, many proteins involved in the basic transcriptional mechanisms have been added to the histone-like list (6,7) as follows. (i) Many of the TATA-binding protein (TBP) associated factors ($hTAF_{II}80$, $hTAF_{II}31$, $hTAF_{II}18$, $hTAF_{II}28$, $hTAF_{II}20$ and $hTAF_{II}135$) that are part of the TFIID complex known to mediate activated transcription and constitute a link between gene-specific activators and the general transcription machinery (8–13). Indeed core histones and TAF_{II}s can interact through their HFMs (10). (ii) The two subunits α and β of the TBP-binding general repressor NC2 (also termed Dr1/DRAP1) a global repressor of basal transcription that interferes with TFIIA and TFIIB association (14–17). (iii) SPT3, also a TBP-binding protein involved in activation (11,18). (iv) PAF65α, an H4-like subunit of the P/CAF complex, a multi-subunit protein with histone acetylase activity that comprises other HFM proteins, $hTAF_{II}31$ and $hTAF_{II}20$ (19). (v) The B and C subunits of the sequence-specific CCAAT-binding factor, NF-Y (20,21).

NF-Y (also called CBF and HAP2/3/5 in yeast) is a ubiquitous protein, composed of NF-YB, NF-YC and NF-YA, a third non-HFM subunit absolutely required for CCAAT-binding. Consistent with their HFM nature, NF-YB and NF-YC dimerize tightly, offering a complex surface for NF-YA association, comprising residues in NF-YC α1, NF-YB α2 and a short region at the C-terminal of NF-YC α3 (20). The resulting trimer has an extremely high affinity and sequence-specificity for DNA. NF-YB, NF-YC and NC2α-β share common features: their HFMs are of the H2A–H2B type and have a particularly high resemblance (30/35% identity, 60–65% similarity) clearly higher than other HFM proteins (16,17). Sequence homology is perfectly colinear across species and clearly visible beyond the three helices of the histone motifs within the larger yeast/human conserved domains. We recently showed that the NF-YB–NF-YC dimer can exist in the absence of NF-YA, that NF-YB and NF-YC can interact with TBP and that they are partially associated with TFIID; the TBP-contacting parts of NC2β, NF-YC and NF-YB reside in the same short sequence between α3 and αC (15,22). Unlike NC2, which has

*To whom correspondence should be addressed. Tel: +39 (0)59 2055542; Fax: +39 (0)59 2055548; Email: mantor@mail.unimo.it

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

transcriptional repression potential in an alanine-rich region of NC2β/Dr1, the NF-Y trimer and the HFM dimer has activation capacity in GAL4 assays, thanks to the two Q-rich domains of NF-YA–NF-YC (23,24). In keeping with its HFM nature, NF-Y is peculiarly efficient in targeting the CCAAT box in a nucleosomal context, through specific interactions of NF-YB–NF-YC with histones H3 and H4 (25,26).

We decided to use the powerful means of BLAST algorithms to identify other genes containing potential HFMs similar to NF-YB–NF-YC. Here we describe the cloning and characterization of YBL1 and YCL1, two novel members of the H2A–H2B subfamily.

MATERIALS AND METHODS

Cloning of YBL1 and YCL1

Amplification reactions contained 100 ng of DNA from a mouse M12 (B lymphoma) cDNA library, 1× PCR buffer (1.5 mM $MgCl_2$), 200 µM of each dNTP, 250 ng of each primer and 2.5 U of AmpliTaq (Perkin-Elmer) in a final volume of 50 µl. After the first denaturation step of 2 min 30 s at 94°C, the reaction underwent 35 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 2 min. The following primers specific for the YBL1: 5′-ACGGATCCACTGGCGGAGAGGCCCGAG, 3′-ACAACGTTGTTGTCTATTTCCTCCTCTTCGTTCTGG; and YCL1: 5′-ACGGATCCTGCCGTGGGGAAGGAAAAGTG, 3′-ACAACGTTAGCCAGGGCTTCACCAAGGTC were used. PCR products were cloned into the Pinpoint T vector (Promega) and entirely sequenced on both strands. YCL1 (amino acids 5–145) and full-length YBL1 were cloned in frame into the PET32b vector (Novagen), containing Thioredoxin, ProtS and His-tags, and in PET3b. In PET32–YBL1 and YCL1 vectors Thioredoxin and ProtS tags were eliminated by cutting with *Nde*I, excising the insert and religating.

Northern blot analysis was performed with identical amounts of RNA (10 µg) from different mouse cell lines. Hybridization and washing was done according to standard procedures with YBL1 and YCL1 probes labeled by random priming.

Production of recombinant proteins and antibodies

Fusion proteins containing His-tags were produced in BL21(LysS) bacteria by inducing an exponential growth culture (OD 0.6) with 1 mM IPTG for 3 h. Purifications were performed by incubating crude bacterial extracts in BC300 (100 mM KCl, 20 mM HEPES pH 7.9, 10% glycerol, 5 mM imidazole, 5 mM β-mercaptoethanol) with NTA-agarose; the column was washed with BC300 and BC1000, containing 300 mM and 1 M KCl, respectively. Proteins were eluted in BC100 containing 300 mM imidazole and dialyzed against BC100. YBL1 from inclusion bodies (50–70% pure) was solubilized in 6 M GnCl, dialyzed against PBS and used for immunizing rabbits. Positive sera were purified over a YBL1-Sepharose antigen column. For western blots, 50–100 µg of nuclear extracts were loaded on a 15% SDS gel, blotted on nitrocellulose, blocked with 3% non-fat dry milk for 2–4 h at room temperature. The samples were then incubated with the purified antibody (1:2000 dilution) in TBS (1 mg/ml BSA) for 2 h at room temperature or overnight at 4°C. Filters were washed three times for 10 min in TBS and developed with an anti-rabbit peroxidase-conjugated secondary antibody (Pierce).

Protein–protein interactions

Pure histones and His-tagged YBL1 or YCL1 (10 µg each in 80 µl) were incubated together in BC2000 (2 M KCl, 20 mM Tris–HCl pH 7.5, 1 mM β-ME, 0.05% NP-40, 100 µg/ml BSA, 0.25 mM PMSF) and step diluted with BC100 (same buffer, containing 100 mM KCl) over a period of 2 h, until a salt concentration of 0.35 M KCl was reached; 20 µl of nickel NTA-agarose column was added, rocked for 1 h and washed twice with 1 ml of BC500; all procedures were performed at 4°C. Bound proteins were eluted with 300 mM imidazole in BC100 and analyzed in 17% gels, stained with Coomassie blue.

The denaturation–renaturation experiment described in Figure 3B was performed by incubating 500 µg of CH27 nuclear extracts with 20 µg of His-tagged YCL1 in 500 µl of 6 M GnCl. Proteins were slowly renatured by dialysis and passed on 50 µl of NTA-agarose, which was washed with 1 M KCl and eluted with 0.3 M imidazole.

Immunoprecipitations were performed with NF-YA (500 ng) and an equivalent amount of YBL1–YCL1 or NF-YB–NF-YC; trimers were added to 25 µl of Protein-G–Sepharose to which 7.5 µg of the purified anti-NF-YA Mab7 monoclonal antibody had been previously bound. Incubation was pursued for 2 h on ice, unbound material recovered after centrifugation and the beads washed with NDB100 with the addition of 0.1% NP-40. SDS buffer was added, the samples boiled at 90°C for 5 min and loaded on SDS gels. Western blots were performed according to standard procedures. For multiple interactions, the filter was stripped, blocked with non-fat dry milk and re-hybridized.

EMSAs with TATA and with the 39mer CCAAT oligonucleotides were detailed previously (22).

Nucleosome reconstitutions

Nucleosome reconstitutions and MNase I assays are described by Caretti *et al.* (25). The probe used for nucleosome reconstitution was fragment 2 of PE3 plasmid (–115/+60) containing the MHC class II Eα promoter (25). Briefly, 1 µg of renatured core histones, or 600 ng H3–H4 or H2A–H2B dimers, were incubated with 500 ng competitor DNA and 2 ng of labeled DNA (10^5 c.p.m.) in a final volume of 10 µl in 1.5 M NaCl, 10 mM Tris–HCl pH 7.8, 500 ng/µl BSA and 1 mM β-mercaptoethanol for 30 min at 20°C. The reaction was serially diluted to 0.8, 0.67, 0.57, 0.5 and 0.1 M NaCl, by addition of TE buffer (10 mM Tris–HCl pH 7.6, 1 mM EDTA) every 15 min. Finally 20 µl of the samples were loaded on a polyacrylamide gel. When indicated, equivalent amounts of purified YBL1 and YCL1, devoid of ProtS and Trx tags, were added in the reconstitutions.

Glycerol gradient experiments (15–40%) were performed as described by Bellorini *et al.* (22), except that CH27 nuclear extracts or recombinant YBL1 and YCL1 were used.

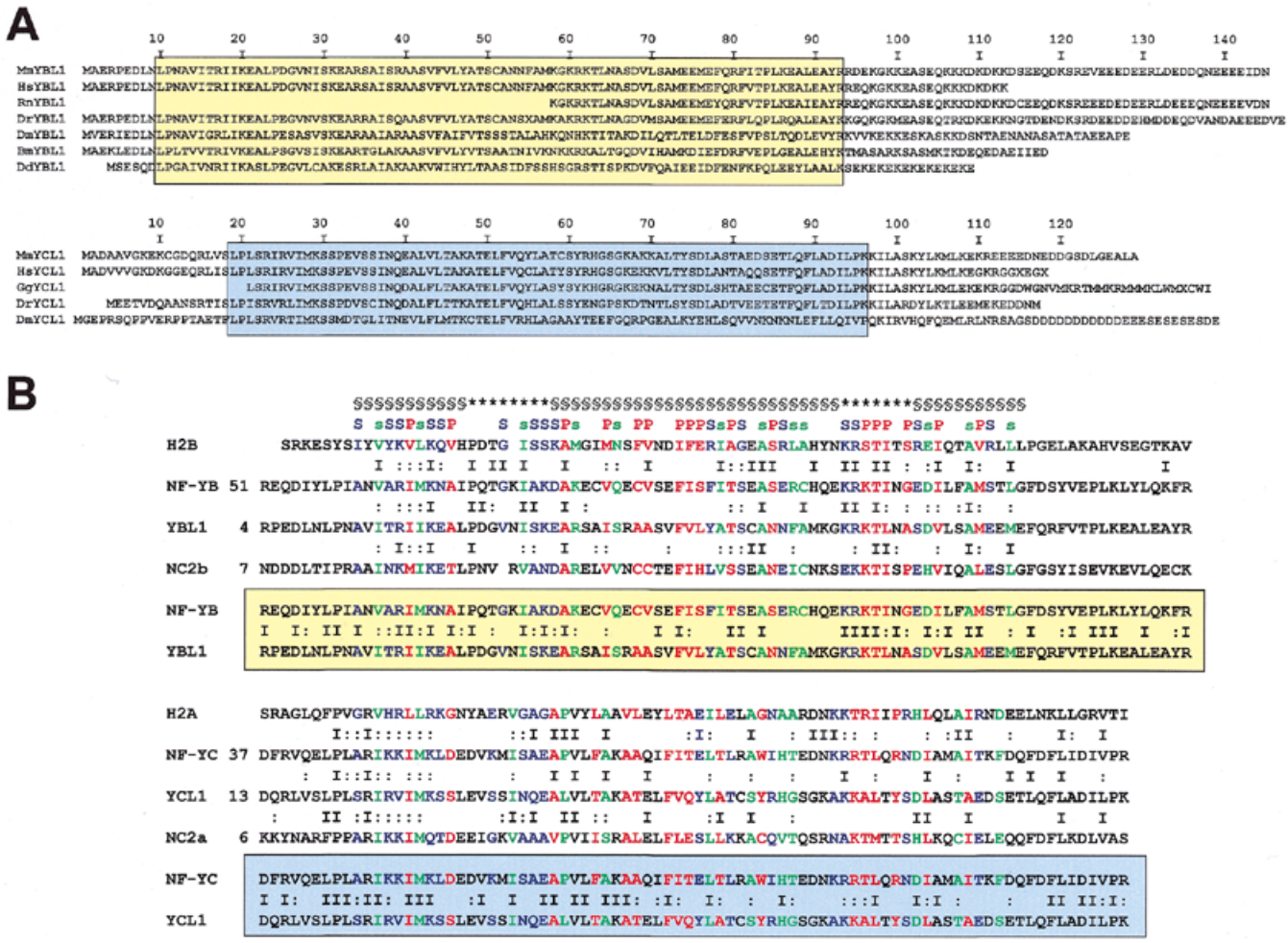

Figure 1. Sequence comparisons of YBL1 and YCL1. (**A**) Alignments of YBL1 and YCL1 sequences from mouse (GenBank accession no. AA839465 for YBL1, AI195704 and AI931369 for YCL1), man (AW246427 for YBL1, R55966 for YCL1), rat (AA818074 for YBL1), zebrafish (AI416300 for YBL1, AI437035 for YCL1), chicken (AI980453 for YCL1), *Drosophila* (AI237921 for YBL1, AI061700 for YCL1), *B.mori* (AU003161 for YBL1) and *D.discoideum* (C83909 for YBL1). (**B**) Alignments of H2B and H2A-like sequences. §, α-Helical regions; *, loops; S (blue), s (green) and P (red), surface, self and Pair residues, respectively, as discussed by Arents and Moudrianakis (3); I, identity; :, similarity.

RESULTS

Identification, cloning and characterization of YBL1 and YCL1

To identify new genes of the HFM family, we searched the EST databases by TBLASTN with two probes corresponding to the 65 amino acid core HFM domains of NF-YB and NF-YC. Several overlapping ESTs were found in human, mouse and other species corresponding to two distinct genes showing high resemblance (e^{-11}/e^{-14}) with NF-YB and NF-YC. We ordered the mouse EST clones, derived oligonucleotides and amplified cDNAs from a mouse M12 B cell library; PCR products of the expected size were obtained, cloned in a T-vector and sequenced, indeed confirming the original EST sequences. Because of the high homology with NF-Y subunits, we named these genes YBL1 and YCL1 (YB-like 1 and YC-like 1, respectively; Fig. 1A). The mYBL1 cDNA contain optimal Kozak sequences around the presumed ATG; multiple mouse YCL1 ESTs align well with humans and one (AI931369) apparently contains the entire ORF, as suggested from the identification of the human genomic sequence, located on chromosome 8q11-q12 (not shown). The cDNA of mYBL1 codes for a highly acidic protein (pI 4.4) of 145 amino acids with an expected molecular weight of 17 kDa; mYCL1 is 129 amino acids (14 kDa) and is also acidic (pI 4.8). Analysis of YBL1 from human, mouse, *Drosophila*, *Bombyx mori*, *Dictyostelium discoideum* and zebrafish, and of YCL1 from human, mouse, chicken, *Drosophila* and zebrafish evidences a remarkable degree of phylogenetic conservation, extending beyond the 65 amino acid HFMs and colinear among different species. Outside the homology domains, both proteins have essentially short N-terminals and a C-terminus that is rich in lysine and acidic residues (Fig. 1A).

Sequence comparisons with other histone-like proteins clearly indicate that these are two new members of the H2A–H2B subgroup (Fig. 1B). Several features of bona fide HFM proteins are indeed present: basic residues in α1, L1, L2, as well as two important alanines in α2 and α3, alternating charged hydrophobics in α2 and the capping α1 proline (3; see Discussion). Homology with NF-YB–NF-YC is strongest and extends to a

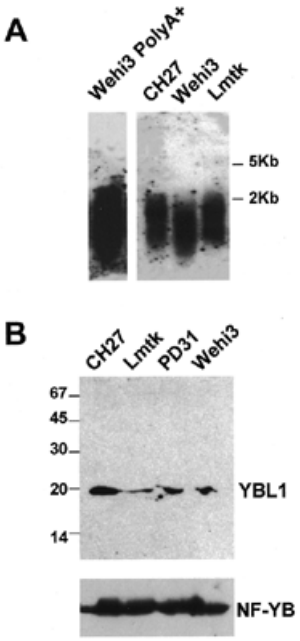

Figure 2. Expression of YBL1. (**A**) Northern blot analysis using 10 µg of total RNA or 3 µg of polyA+ RNA with the indicated cell lines. (**B**) Western blot analysis using 10 µg of nuclear extracts from the indicated cell line. An anti-NF-YB antibody was used to normalize protein concentrations.

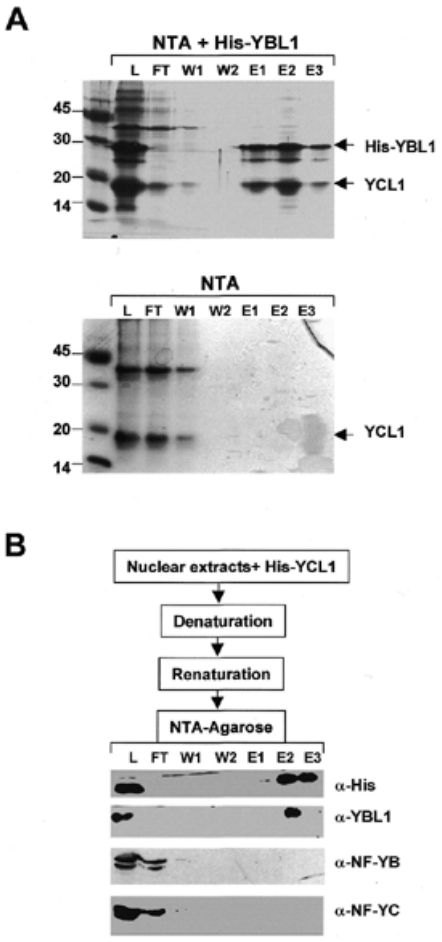

Figure 3. Heterodimerization of YBL1–YCL1. (**A**) Recombinant His-tagged YBL1 was renatured with YCL1 and purified over nickel NTA-agarose. L, load; FT, flow through; W1, 0.3 M KCl wash; W2, 1 M KCl wash; E, eluates with 0.3 M imidazole. Lower panel, YCL1 was processed without His–YBL1. (**B**) Same as (A), except that His–YCL1 was used with nuclear extracts from CH27 as a source of YBL1, NF-YB and NF-YC. The proteins were revealed by western blots with the indicated antibodies.

region at the C-terminus of α3 and into the putative αC, shown to be important for TBP-interactions in NF-YB and NC2β (15,22) and for TBP and NF-YA-binding in NF-YC (22,27).

We checked the mRNA levels of the two genes in different mouse cell lines by northern blot analysis: expression is ubiquitous, as shown in Figure 2A from the presence of transcripts of 1.2 kb for YBL1 in different cell lines and of equivalent size for YCL1 (not shown). Moreover, ESTs from a wide range of mouse and human tissues and cell lines were detected for both genes: testis, mammary gland, diaphragm, blastocyst, fetal heart in mouse, neuronal NT2, melanocyte, colon carcinoma, Jurkat T cells and 8-week embryo in humans. Western blot analysis of nuclear extracts of CH27 (B lymphocytes), Lmtk (fibroblasts), PD31 (Pre-B lymphocytes) and Wehi-3 (macrophages) with anti-YBL1 purified rabbit polyclonal antibodies confirmed that the mRNA is translated at equivalent levels in different cell lines (Fig. 2B). The band of 20 kDa is higher than the expected molecular weight: this might be due to the large amounts of E and D, that repels SDS, and K, that neutralises the SDS charge; alternatively, YBL1 might be post-translationally modified *in vivo*, by acetylation, for example, as is the case for NF-YB (28).

Production and interactions of YBL1 and YCL1

The identification of two new HFM proteins prompted us to verify whether they represent interacting partners. To this aim, we exploited the protein–protein interaction assays on NTA-agarose developed in our lab (29): we cloned the full-length YBL1 and a YCL1 mutant containing the HFM (amino acids 5–129) into the PET3b and PET32b bacterial expression vectors, that produce proteins lacking and containing His-tags, respectively. The His–YBL1 was first purified on NTA columns and incubated with His-less YCL1 from inclusion bodies in GnCl and slowly renatured by removal of the denaturing agent, a procedure shown to be highly efficient in reconstitution of functional HFM proteins (8,29); the proteins were then loaded on NTA-agarose, washed with high salts and eluted with imidazole. Figure 3A shows that YCL1 indeed co-elutes with the His–YBL1, whereas on a control column, in which His–YBL1 was absent, YCL1 is found in the unbound material (Fig. 3A, lower panel). The same type of result was obtained when the reverse experiment (His–YCL1 and YBL1) was tried (not shown, see below). To confirm this and investigate whether NF-YB–NF-YC can cross-heterodimerize and/or bind as dimers to YBL1–YCL1, we mixed an excess of His–YCL1 from inclusion bodies with proteins from CH27 nuclear extracts, denatured the mix in GnCl, slowly renatured by dialysis and ran an NTA column: Figure 3B indicates that the endogenous YBL1, but not NF-YB nor NF-YC, is retained. We conclude that YBL1–YCL1 proteins bind to each other, most

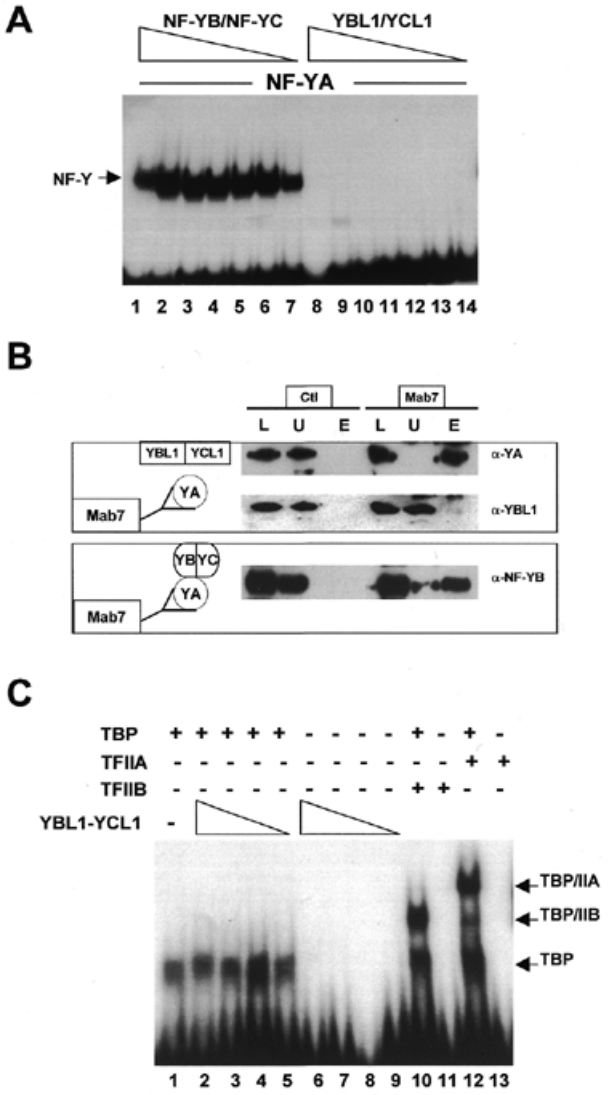

Figure 4. YBL1–YCL1 is not a CCAAT or TATA interacting dimer. (**A**) EMSA analysis of decreasing concentrations (300, 100, 30, 10, 3, 1 and 0.3 ng) of NF-YB–NF-YC (lanes 1–7) or YBL1–YCL1 (lanes 8–14), in the presence of 100 ng of NF-YA and a labeled CCAAT-box oligonucleotide (MHC class II Ea Y box). (**B**) Immunoprecipitations with an anti-NF-YA monoclonal-Mab7 or a control (anti-GATA) antibody after incubation of NF-YA with YBL1–YCL1 (upper panel) or NF-YB–NF-YC (lower panel). L, load; U, unbound; E, bound. (**C**) Same as (A), except that the oligonucleotide used was AdML TATA box with decreasing concentrations of YBL1–YCL1 (100, 30, 10 and 3 ng) and TBP (100 ng) in lanes 2–5, or without TBP (lanes 6–9). Control incubations with purified endogenous TFIIA (lanes 10 and 11) or recombinant TFIIB (lanes 12 and 13) were run with (lanes 10 and 12) or without TBP (lanes 11 and 13).

likely by heterodimerizing, and are incapable of heterodimerizing with NF-YB–NF-YC.

Because of the relationship with NF-YB–NF-YC, we investigated the possibility that the YBL1–YCL1 dimer could bind to NF-YA and that the resulting trimer might be a CCAAT-binding complex. We performed EMSA with a labeled CCAAT-box oligonucleotide in the presence of NF-YA and of increasing doses of YBL1–YCL1 or NF-YB–NF-YC. Figure 4A shows that, unlike NF-YB–NF-YC, YBL1–YCL1 is unable to generate a DNA-bound complex, even at very high concentrations. This could be due either to a lack of NF-YA association, or to the NF-YA–YBL1–YCL1 trimer being unable to interact with DNA. To discriminate between these possibilities, we performed immunoprecipitations of a mix containing the latter proteins with an anti-NF-YA monoclonal antibody. Figure 4B shows that YBL1 and YCL1 were not associated to NF-YA, unlike the NF-YB–NF-YC control that was found in the bound fraction. We therefore conclude that YBL1 and YCL1 are incapable of interacting with NF-YA and this explains the lack of CCAAT binding. An alternative possibility is that they might be more related to NC2 and hence TBP/TATA-binding proteins. To investigate this possibility, we used YBL1–YCL1 with recombinant TBP on a labeled TATA-box oligonucleotide. Figure 4C shows that unlike TFIIA and TFIIB (or NC2, not shown), even high doses (300 ng) of recombinant YBL1–YCL1 are incapable of generating a further retarded complex in this assay. We conclude that in terms of DNA-binding, YBL1 and YCL1 are clearly divergent from NF-YB and NF-YC or NC2α-β.

Interactions of YBL1–YCL1 with histones and nucleosomes

We have recently shown that NF-YB–NF-YC binds to H3–H4 and forms complexes with core histones during nucleosome formation (25); we decided to verify whether this would be true for YBL1–YCL1. First, we ascertained whether YBL1 or YCL1 could interact with core histones. H2A, H2B, H3 and H4 were produced in bacteria as inclusion bodies (Fig. 5A) and singularly incubated them with equivalent amounts of recombinant His–YBL1 or His–YCL1 in the NTA protein–protein interaction assays described before (Fig. 5B and C). YBL1 interacts with H2B and H3, but not with H2A and H4, as shown by the fact that they co-elute on the columns, whereas YCL1 does not interact with any of the core histones in this assay, which are all found in the unbound fractions. Note that most of the H3 reproducibly precipitated during this procedure when incubated with YCL1, but not YBL1.

Next, we assayed the new H2A–H2B-like proteins in nucleosome reconstitutions. A labeled MHC class II Ea fragment was incubated in the presence of cold competitor DNA and 1.5 M NaCl with different combinations of histonic proteins. Salt concentration was then progressively lowered over 15 min periods and aliquots of the samples were finally loaded on a running polyacrylamide gel. The resulting patterns give indications of the binding of the different complexes to DNA and are shown in Figure 6A for H2A–H2B (lane 1), H3–H4 (lane 5), H2A–H2B–H3–H4 (lane 9) or with increasing concentrations of YBL1–YCL1, either alone (lanes 10–12), with H3–H4 (lanes 2–4) or with H2A–H2B and H3–H4 (lanes 6–8). Two faint bands are observed with H3–H4: they have been observed in other studies and most likely represent the H3–H4 tetramer and di-tetramers species detailed in glycerol gradient experiments (30). Upon reconstitution with YBL1–YCL1 and H3–H4 (lanes 2–4), two distinct bands are generated, showing a dose-dependent increase in intensity and with electrophoretic mobilities different from that of the nucleosome or of the H3–H4 tetramers (Fig. 6A, compare lane 4 with lanes 5 and 8). Interestingly, these bands are observed at sub-optimal concentrations of H3–H4 and of YBL1–YCL1, the latter being unable to form complexes of similar mobility on their own. At high

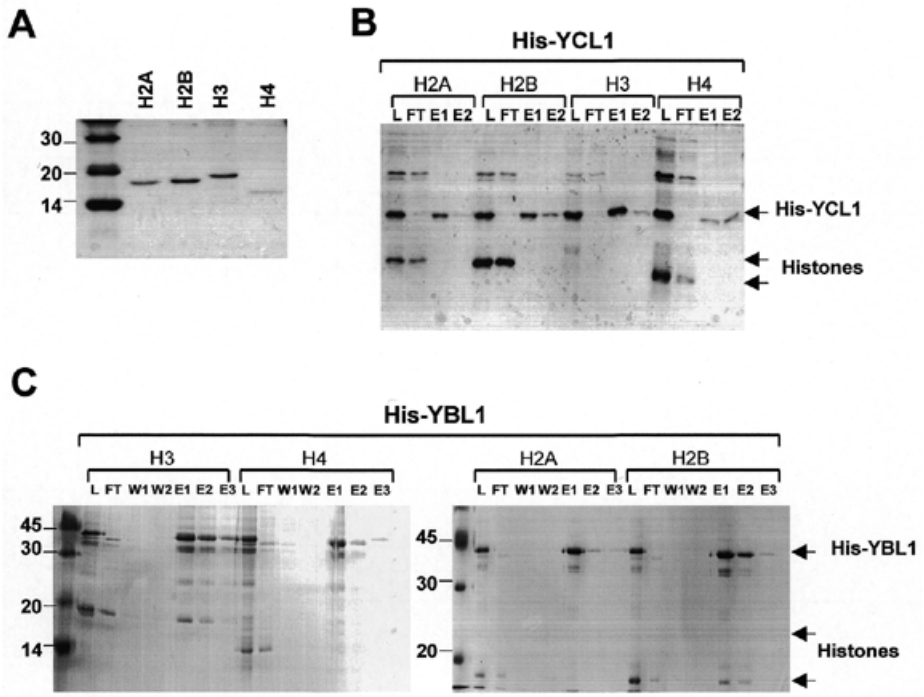

Figure 5. Interactions of YBL1–YCL1 with histones. (**A**) Coomassie blue stained SDS gel of recombinant histones purified from inclusion bodies. (**B**) His–YCL1 was incubated with the indicated histones, renatured and passed over a nickel NTA-agarose column. Gels were stained with Coomassie blue. (**C**) Same as (B), except that His–YBL1 was used.

concentrations faint bands of slower mobility are observed (marked by asterisks). We then verified whether such complexes contain YBL1 by performing supershift experiments with the anti-YBL1 antibody: the results shown in Figure 6B show that addition of a specific antibody decreases and upshifts YBL1–YCL1-containing complexes with H3–H4 or nucleosomes (Fig. 6B, lanes 8 and 14), whereas it has no effect on H3–H4 tetramers and nucleosomes (Fig. 6B, lanes 5 and 11). A control anti-GATA1 antibody has no effect on any complex; when assayed on the faint YBL1–YCL1 complexes generated in the absence of histones, the anti-YBL1 antibody confirms that they do contain the HFM protein (Fig. 6B, lanes 1–3). Similar to what we have recently performed with NF-YB–NF-YC histone complexes (25), we decided to investigate whether the YBL1–YCL1-containing complexes are similar to octameric structures; we therefore purified them on glycerol gradients and assayed them with micrococcal nuclease: compared to the H3–H4 tetramers or nucleosomes, H3–H4–YBL1–YCL1 shows a more prominent hypersensitive site (Fig. 6C, compare lanes 2 and 3); YBL1–YCL1-nucleosome also increases a site and shows protection (Fig. 6C, lanes 4 and 5); however, the nucleosome pattern is clearly distinct from that obtained when H3–H4–YBL1–YCL1 was used (Fig. 6C, compare lanes 3 and 4). We conclude that YBL1–YCL1 can associate with histones H3–H4 without conferring to the hybrid complex a bona fide nucleosomal status. In this respect, they are similar to NF-YB–NF-YC (25).

YBL1 is complexed with other proteins

We have previously shown that NF-YB and NF-YC, unlike NF-YA, exist in high molecular weight complexes (22). We decided to verify whether this was also the case for YBL1–YCL1 by running crude CH27 extracts, extracted from nuclei with 1 M KCl, on glycerol gradients. The resulting fractions were checked in western blots with the anti-YBL1 antibody (Fig. 7). The protein is distributed in a range between <100 and 400 kDa (upper panel, fractions 13–21). On the other hand, a glycerol gradient containing only the recombinant YBL1–YCL1 dimer indicates that the latter is found in fractions of low molecular weight, with one peak in fraction 11 (30 kDa), consistent with the molecular weight of a dimer (30/40 kDa). Thus, the fact that the YBL1–YCL1 dimer is found in fractions of higher weight in the extracts suggests that these HFM proteins are present in complexes with other polypeptides.

DISCUSSION

We report the cloning and characterization of two novel histone-like proteins of the H2A–H2B sub-type, showing conspicuous homology to the HFM subunits of the CCAAT-binding factor NF-Y, a protein that activates 25% of mammalian promoters. YBL1–YCL1 are ubiquitously expressed, complexed with other polypeptides and show typical features of HFM proteins, including the ability to bind core histones in solution and on the DNA.

Structure/function considerations

Comparison of YBL1 with NF-YB shows a higher degree of homology in α1, L1 and α3 (Fig. 1B). The long α2, which contains most of the information for heterodimerization (31), is relatively divergent: only seven of 28 residues are identical. In particular, it is noteworthy that of the two NF-YB residues

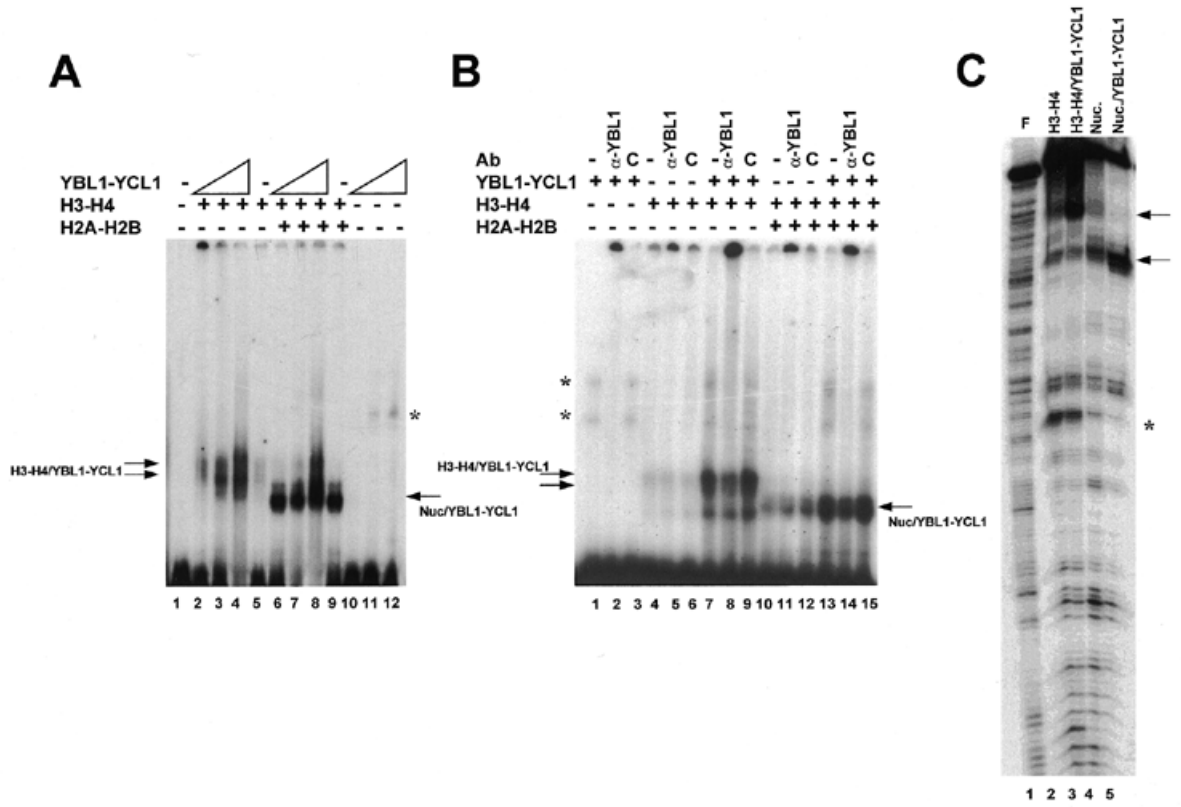

Figure 6. Nucleosome reconstitutions with YBL1–YCL1. (**A**) Increasing concentrations of YBL1–YCL1 dimer (0.3, 1 and 3 µg) were used with 1 µg of the indicated H2A–H2B, H3–H4 combinations. Arrows, complexes formed with histones; *, faint complexes generated by YBL1–YCL1. (**B**) Antibody challenge of EMSA complexes containing YBL1–YCL1. H3–H4 tetramers (lanes 1–6) and nucleosomes (lanes 7–12) were reconstituted without YBL1–YCL1 (lanes 1–3 and 7–9) or with (lanes 4–6 and 10–12) in the absence of antibodies (lanes 1, 4, 7 and 10) or with anti-YBL1 (lanes 2, 5, 8 and 11) or anti-GATA1 control antibodies (lanes 3, 6, 9 and 12). In lanes 13–15, YBL1–YCL1 only was used in reconstitutions and similarly challenged with antibodies. (**C**) MNase I digestions of H3–H4 tetramers (lane 2), H3–H4–YBL1–YCL1 (lane 3), nucleosome (lane 4) and nucleosome–YBL1–YCL1 (lane 5). F, free DNA digested with Mnase I; arrows, increased sensitivity sites generated by the presence of YBL1-YCL1; *, protections.

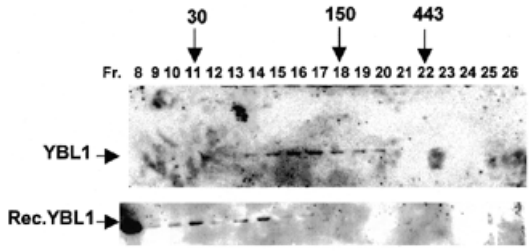

Figure 7. YBL1–YCL1 is found in relatively large complexes. Fractions from glycerol gradients with CH27 nuclear extracts were checked in a western blot with the anti-YBL1 antibody (upper panel). The molecular weight markers, run on parallel gradients, are indicated. Similarly, fractions generated from a gradient containing only the recombinant YBL1–YCL1 dimer were checked in westerns and shown in the lower panel. In the first lane 10 ng of recombinant YBL1 was loaded.

required for NF-YA binding (E90 and S97) only the serine (S50 in YBL1) is conserved, unlike E90, that is strictly conserved in all 26 NF-YBs identified so far (G.Gusmaroli, C.Tonelli and R.Mantovani, manuscript submitted), S97 varies among species. It is possible that the E to V difference alone might explain why the YBL1–YCL1 heterodimer is incapable of binding to NF-YA, either in solution or on DNA. It is known, in fact, that several positively charged residues of NF-YA are required for binding to the HFM dimers, at least one of these presumably contacting this particular NF-YB glutamic acid (32,33). The residues shown to be involved in NF-YC dimerization are mostly conserved or have conservative substitutions, with the exception of F118 in α3, which is a serine in YBL1. It is difficult to say whether this difference is sufficient to explain the lack of cross-heterodimerization between NF-YB–NF-YC and YBL1–YCL1, as F118R was only assayed in conjunction with Y128G (31). Among residues that are involved in CCAAT-binding, many (R64, A77, K78) are conserved in YBL1, but not the crucial N61, which is present in all NF-YBs and was shown, by swapping analysis, to be important for CCAAT specificity (29). This residue corresponds to a non-conserved valine in YBL1. Finally, we remark that several residues in the TBP-binding subregion between α3 and αC are identical.

On the NF-YC–YCL1 side, the residues shown to be important for DNA-binding, residing mainly in α1, are conserved. Interestingly, the NF-YC residues shown to be involved in NF-YB dimerization are either conserved (L68, A70, L80) or, as in the case of I77, I101, A104, V58, L82 and T85, in YCL1, conservatively changed. However, other

residues in the NF-YC region 71–80 required for this function are rather different in YCL1 and thus could explain the dimerization specificity. As for NF-YA association, mapped in two distant regions of NF-YC α1 and at the C-terminal of α3 we remark that they are quite conserved, particularly the 110–120 region. Altogether, the lack of heterodimerization between NF-YB–YCL1 and NF-YC–YBL1 and the incapacity of YBL1–YCL1 heterodimer to bind NF-YA lend experimental support to these structural considerations and indicate that in functional terms, NF-YB–NF-YC and YBL1–YCL1 are clearly divergent. Analogous considerations can be put forward about the relationship between YBL1–YCL1 and NC2α-β, although less is known about single residues involved in dimerization and DNA-binding of the latter. However, the lack of TBP–TATA interactions by YBL1–YCL1 in EMSA is not in favor of a role of the two proteins in the basic TATA-centered mechanisms of transcription. In addition, transfections of GAL4–YBL1 fusions yielded none of the repressor activities described in GAL4 experiments with NC2β/Dr1 (F.Bolognese and R.Mantovani, unpublished data).

Phylogenetic considerations

The abundance of NF-YB and NF-YC genes in plants does not reflect an equivalent abundance of YBL1–YCL1. In fact, we were unable to retrieve bona fide YBL1 or YCL1 genes from *Arabidopsis thaliana* ESTs or genomic sequences, or from any other plant species, despite the advanced stages of the sequencing projects. Note that 23 NF-Y and several NC2 genes were identified in *Arabidopsis* alone. Despite the emergence of possible candidate genes showing some similarity, we could not find *Saccharomyces cerevisiae* genes that are obvious counterparts of YBL1–YCL1 in this organism (note that yeast HAP2/3/5 and NF-Y proteins are 75% conserved). On the other hand, YBL1–YCL1 homologs are found in *Drosophila*, the only organism in which NF-Y genes are apparently absent, a fact that matches quite well the notable lack of CCAAT boxes as important *cis*-acting elements in *Drosophila* promoters. Thus, YBL1–YCL1 and NF-YB–NF-YC (and NC2α-β) genes obviously coevolved, but apparently took partially separate turns in evolution, independently falling into oblivion in some species, but not vertebrates.

YBL1–YCL1 function

The data on protein–protein interactions with core histones and nucleosome reconstitutions suggest that common features between YBL1–YCL1 and NF-YB–NF-YC do exist and are probably relevant to their functions. We have recently shown, in fact, that the HFM subunits of NF-Y are capable of binding DNA during nucleosome formation and in the absence of NF-YA and form complexes containing H3–H4 with distinct electrophoretic behavior. However, these are not octameric structures. We indeed suggested that this might be a common theme for other histone fold proteins, particularly of the H2A–H2B subtype. When tested in these assays, YBL1–YCL1 show a highly similar behavior, reinforcing the notion that, while other aspects of the HFM diverged, maintainance of core histones relationships coevolved. Moreover, the H2B-like proteins are apparently responsible for this function, as we previously showed for NF-YB, and as is the case for the H2B-like $hTAF_{II}20$, both interacting with H4 and with the H4-like $hTAF_{II}80$ (10). A difference worth noting is that YBL1 interacts with H3, rather than H4. On the other hand, the H2A-like proteins NF-YC and YCL1 are apparently incapable of interacting directly with histones and other HFMs (10,25).

The glycerol gradient experiments strongly suggest that YBL1 and YCL1 are complexed with other polypeptides *in vivo*. This is hardly a novelty for HFM proteins. (i) HFM TAF_{II}s, in addition to being part of TFIID, are also present in other multi-protein complexes such as TFTC and the P/CAF complex (9,19,34). (ii) The acetylase STAGA complex contains HFM $hTAF_{II}31$ and hSPT3 (35). (iii) PAF65α is thought to be the $hTAF_{II}31$-interacting partner in the P/CAF complex (19). (iv) NF-YB–NF-YC and NC2 are found associated with TFIID (14,22). These complexes are involved in either direct transcriptional activation or modifications of chromatin structure by histone acetylation. One possibility is therefore that by locally perturbing histone H2A–H2B deposition and through formation of hybrid sub-complexes less stable and/or tightly bound to DNA than octamers, YBL1 and YCL1 facilitate the access of other activities influencing transcription (positively or negatively). In fact, as these small polypeptides are apparently devoid of activating or repressing functions on their own, they might have other crucial roles. Indeed, the recent biochemical characterization and the cloning of the genes of the two small subunits of the chromatin remodeling complex CHRAC identified the *Drosophila* and human equivalent of YBL1–YCL1 (36,37). This raises the exciting possibility that these two HFM proteins are an important interface between core histones and energy-utilizing machines that alter chromatin structure. Establishing the exact nature of YBL1–YCL1 interacting partners will clearly be required to start understanding their role in cellular physiology.

NOTE ADDED IN PROOF

While this manuscript was under revision, Corona *et al.* (36) and Poot *et al.* (37) reported the cloning of YBL1 and YCL1 genes from human and *Drosophila* as part of the CHRAC chromatin remodeling complex.

ACKNOWLEDGEMENTS

We thank P. Becker (University of Munich D, Germany) and P. Varga-Weisz (Marie Curie Research Institute, Oxford, UK) for communicating results prior to publication. C.I. and G.C. were recipients of a Telethon fellowship. This work was supported by grants from MURST (PRIN-'Nucleic acid–proteins interactions') and CNR to R.M.

REFERENCES

1. Workman,J.L. and Kingston,R.E. (1998) *Annu. Rev. Biochem.*, **67**, 745–779.
2. Strahl,B.D. and Allis,C.D. (2000) *Nature*, **403**, 41–45.
3. Arents,G. and Moudrianakis,E.N. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 11170–11174.
4. Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) *Nature*, **389**, 251–260.
5. Starich,M.R., Sandman,K., Reeve,J.N. and Summers,M.F. (1996) *J. Mol. Biol.*, **255**, 187–203.
6. Baxevanis,A.D., Arents,G., Moudrianakis,E.N. and Landsman,D. (1995) *Nucleic Acids Res.*, **23**, 2685–2691.
7. Baxevanis,A.D. and Landsman,D. (1998) *Nucleic Acids Res.*, **26**, 372–375.
8. Xie,X., Kokubo,T., Cohen,S.L., Mirza,U.A., Hoffmann,A., Chait,B.T., Roeder,R.G., Nakatani,Y. and Burley,S.K. (1996) *Nature*, **380**, 316–322.

9. Wieczorek,E., Brand,M., Jacq,X. and Tora,L. (1998) *Nature*, **393**, 187–191.
10. Hoffman,A., Chiang,C.-M., Oelgeschlager,T., Xie,X., Burley,S.K., Nakatani,Y. and Roeder,R.G. (1996) *Nature*, **380**, 356–359.
11. Birck,C., Poch,O., Romier,C., Ruff,M., Mengus,G., Lavigne,A.-C., Davidson,I. and Moras,D. (1998) *Cell*, **94**, 239–249.
12. Lavigne,A.C., Gangloff,Y.-G., Carr,L., Mengus,G., Birck,C., Poch,O., Romier,C., Moras,D. and Davidson,I. (1999) *Mol. Cell. Biol.*, **19**, 5050–5060.
13. Gangloff,Y.-G., Werten,S., Romier,C., Carre,L., Poch,O., Moras,D. and Davidson,I. (2000) *Mol. Cell. Biol.*, **20**, 340–351.
14. Inostroza,J.A., Mermelstein,F.H., Ha,I., Lane,W.S. and Reinberg,D. (1992) *Cell*, **70**, 477–489.
15. Yeung,K.C., Inostroza,J.A., Mermelstein,F.H., Kannabiran,C. and Reinberg,D. (1994) *Genes Dev.*, **8**, 2097–2109.
16. Mermelstein,F., Yeung,K., Cao,J., Inostroza,J.A., Erdjument-Bromage,H., Eagelson,K., Landsman,D., Tempst,P. and Reinberg,D. (1996) *Genes Dev.*, **10**, 1033–1048.
17. Goppelt,A., Steltzer,G., Lottspeich,F. and Meisterernst,M. (1996) *EMBO J.*, **15**, 3105–3116.
18. Dudley,A.M., Rougeulle,C. and Winston,F. (1999) *Genes Dev.*, **13**, 2940–2945.
19. Ogryzko,V.V., Kotani,T., Zhang,X., Schiltz,R.L., Howard,T., Yang,X.-J., Howard,B.H., Qin,J. and Nakatani,Y. (1998) *Cell*, **94**, 35–44.
20. Maity,S. and de Crombrugghe,B. (1998) *Trends Biochem. Sci.*, **23**, 174–178.
21. Mantovani,R. (1999) *Gene*, **239**, 15–27.
22. Bellorini,M., Lee,D.-K., Dantonel,J.-K., Zemzoumi,K., Roeder,R.G., Tora,L. and Mantovani,R. (1997) *Nucleic Acids Res.*, **25**, 2174–2181.
23. diSilvio,A., Imbriano,C. and Mantovani,R. (1999) *Nucleic Acids Res.*, **27**, 2578–2584.
24. Yeung,K., Kim,S. and Reinberg,D. (1997) *Mol. Cell. Biol.*, **17**, 36–45.
25. Caretti,G., Motta,M.C. and Mantovani,R. (1999) *Mol. Cell. Biol.*, **19**, 8591–8603.
26. Motta,M.C., Caretti,G., Badaracco,F. and Mantovani,R. (1999) *J. Biol. Chem.*, **274**, 1326–1333.
27. Kim,I.-S., Sinha,S., de Crombrugghe,B. and Maity,S.N. (1996) *Mol. Cell. Biol.*, **16**, 4003–4013.
28. Li,Q., Herrler,M., Landsberger,N., Kaludov,N., Ogryzko,V.V., Nakatani,Y. and Wollfe,A.P. (1998) *EMBO J.*, **17**, 6300–6315.
29. Zemzoumi,K., Frontini,M., Bellorini,M. and Mantovani,R. (1998) *J. Mol. Biol.*, **286**, 327–337.
30. Spangenberg,C., Eisfeld,K., Stuckel,W., Luger,K., Flaus,A., Richmond,T.J., Truss,M. and Beato,M. (1998) *J. Mol. Biol.*, **278**, 725–739.
31. Sinha,S., Kim,I.-S., Sohn,K.Y., de Crombrugghe,B. and Maity,S.N. (1996) *Mol. Cell. Biol.*, **16**, 328–337.
32. Xing,Y., Fikes,J.D. and Guarente,L. (1993) *EMBO J.*, **12**, 4647–4655.
33. Xing,Y., Zhang,S., Olesen,J.T., Rich,A. and Guarente,L. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 3009–3013.
34. Brand,M., Yamamoto,K., Staub,A. and Tora,L. (1999) *J. Biol. Chem.*, **274**, 18285–18289.
35. Martinez,E., Kundu,T.K., Fu,J. and Roeder,R.G. (1998) *J. Biol. Chem.*, **273**, 23781–23785.
36. Corona,F.V., Eberharter,A., Budde,A., Deuring,R., Ferrari,S., Varga-Weisz,P., Wilm,M., Tamkun,J. and Becker,P. (2000) *EMBO J.*, **19**, 3049–3059.
37. Poot,R.A., Dellaire,G., Hulsmann,B., Grimaldi,M.A., Corona,D.F.V., Becker,P.B., Bickmore,W.A. and Varga-Weisz,D. (2000) *EMBO J.*, **19**, 3377–3387.