Journal of Cellular Biochemistry 81:333–346 (2001)

Characterisation of Nuclear Localisation Signals of the Four Human Core Histones

Matthias Baake, Detlef Doenecke, and Werner Albig*

Institute of Biochemistry, Department of Molecular Biology, University of Goettingen, D-37073 Goettingen, Germany

Abstract The four core histones H2A, H2B, H3 and H4 are transported from the cytoplasm into the nucleus by a receptor-mediated and energy-dependent process. This nuclear transport depends on topogenic signals in the individual histone protein sequences. We have analysed such nuclear localisation signals in the core histones by means of fusion proteins consisting of individual core histones (or fragments thereof) and β-galactosidase as a reporter protein. The results show that each of the four core histones contains several portions that are capable of mediating nuclear transport. One type of topogenic sequences consists of clustered basic amino acids in the amino terminal segments of each of the core histones. The globular portions of the core histones represent a second type of nuclear localisation signals that could only mediate nuclear transport when the whole protein domains were fused to the β-galactosidase reporter. Fragments of the globular domains derived from each of the four core histones could not serve as nuclear localisation signals. We conclude that the nuclear targeting of core histones requires information conferred by the globular domain conformation. J. Cell. Biochem. 81:333–346, 2001. © 2001 Wiley-Liss, Inc.

Key words: core histones; nuclear localisation signal; nuclear transport

The fundamental structural unit of eukaryotic chromatin, the nucleosome, is composed of a nucleosomal core particle consisting of a histone octamer and DNA around its surface, and secondly of linker DNA connecting core particles and H1 histones associated with this linker DNA at the exit–entry site of the nucleosomal DNA [Wolffe, 1998]. The histone octamer forming the protein moiety of the nucleosomal core consists of two sets of the four histone classes H2A, H2B, H3 and H4. The size of core histones varies between 102 (H4) and 135 amino acids (H3). All histones are basic proteins. The linker histone H1 is highly enriched in lysine, H2A and H2B are moderately lysine-rich, and H3 and H4 are rich in arginine.

Histones, as all proteins encoded in nuclear DNA, are synthesised in the cytoplasm and must be transported post-translationally to the nucleoplasm. During the S-phase of the cell cycle, a vast amount of histones is needed for the formation of new nucleosomes on replicated DNA. Thus, histones are among the most abundant substrates for nuclear transport. Due to their small size, histones may be expected to enter the nucleus simply by diffusion, since transport studies with model substrates have shown that molecules smaller than 40 kDa can reach the nucleoplasm by diffusion [Feldherr et al., 1994]. Until now, however, this has not been demonstrated for any nuclear protein. In fact, we and others have shown that H1 histones (23 kDa) are transported to the nucleoplasm by an energy-dependent, receptor-mediated process [Breeuwer and Goldfarb, 1990; Imamoto et al., 1995; Kurz et al., 1997; Schwamborn et al., 1998]. Similarly, Moreland et al. [1987] have identified a nuclear localisation signal (NLS) within yeast H2B that indicates a receptor-mediated process for nuclear import.

The receptor-mediated nuclear transport of proteins depends on topogenic signals in the sequence of karyophilic proteins. The first NLSs that were identified in karyophilic proteins, were that of the large-T-antigen of SV40 virus [Kalderon et al., 1984] and of nucleoplasmin,

Grant sponsor: Deutsche Forschungsgemeinschaft; Grant numbers: SFB 523, Graduiertenkolleg 41.

*Correspondence to: Dr. Werner Albig, Institut für Biochemie, Universität Göttingen, Humboldtallee 23, D-37073 Göttingen, Germany. E-mail: walbig@gwdg.de

Received 23 June 2000; Accepted 24 August 2000

This article published online in Wiley InterScience, January XX, 2001.

an acidic protein involved in chromatin assembly [Robbins et al., 1991]. Both signals consist of clusters of basic amino acids: in the SV40 large-T-antigen, the NLS consists of a heptapeptide containing five basic amino acids, whereas the nucleoplasmin NLS is bipartite and contains two short clusters of basic amino acids separated by a spacer of about 10 amino acids. These two NLS are referred to as classical nuclear localisation signals. In the meantime, other primary structure motifs have been described that facilitate the nuclear transport of proteins, such as the M9 sequence in hnRNP A1 [Pollard et al., 1996] and NLSs of ribosomal proteins, such as L23a [Jäkel and Görlich, 1998] and L5 [Claussen et al., 1999; Rosorius et al., 2000]. In general, these non-classical nuclear targeting signals are longer than the two classical NLS motifs [for review, see Christophe et al., 2000].

We have previously assayed in transfection studies with HeLa cells the nuclear import competence of fusion proteins that were transiently expressed after transfection of plasmids composed of individual portions of the human H1° gene and the bacterial β-galactosidase gene. Using this reporter fusion protein system, we could demonstrate that the H1° protein contains several sequence elements for nuclear targeting [Schwamborn et al., 1998]. In the present study, we have assayed with the in vitro transport assay each of the four human core histones for the characteristics of receptor-mediated and energy-dependent nuclear transport and for sequence motifs involved in nuclear targeting. It is demonstrated that core histones do not enter the nucleus simply by diffusion and that the N-terminal domains of each core histone as well as the remaining, globular portions of the H2A, H2B, H3 and H4 histones can serve as NLSs.

METHODS

Plasmid Constructs

All expression plasmids were constructed by cloning the respective fragments in an expression vector derived from pSVβ (Clontech, Palo Alto, USA). pSVβ is a mammalian expression vector for β-galactosidase with the SV40 early promoter. Since this vector does not contain the appropriate restriction sites needed, we introduced a new multicloning site in front of the β-galactosidase gene as previously described [Schwamborn et al., 1998]. The resulting construct called pKS10 was our original construct into which all different human core histone gene fragments were cloned. The β-galactosidase gene inserted in the plasmid codes for the entire protein sequence except the first eight amino acids and contains the unique cloning sites *Bgl*II and *Nhe*I. The coding sequences for human core histones and fragments thereof that were inserted into pKS10 were obtained by PCR amplification from plasmids [compiled in Albig and Doenecke, 1997] containing the genes H2A/d (Acc. No. Z83739), H2B/d (Acc. No. Z83336), H3/k (Acc. No. Z83735), H4/j (Acc. No. Z80787), except those fragments that resulted in pMB8, pMB9, pMB22 and pMB29. The PCR products were then digested with *Bgl*II and *Nhe*I and cloned into the *Bgl*II/*Nhe*I restriction sites of pKS10. Inserts for pMB8, pMB9, pMB22 and pMB29 were generated by annealing two oligonucleotides containing the coding strand and the complementary sequence of the corresponding core histone fragment with overhanging *Bgl*II and *Nhe*I sites for cloning into the pKS10-vector.

The proper reading frame of each construct was verified by sequencing, and in every construct generation of obvious NLS motifs at the transition site of fused core histone and β-galactosidase gene sequences was carefully avoided.

Mutations

Mutations in pMB22 and pMB29 were generated by using oligonucleotides containing one or two substituted nucleotides, respectively, whereas PCR amplification using a primer with two nucleotide exchanges led to pMB32. Nucleotide exchange in pMB53 and pMB60 was done by site-directed mutagenesis using the method of Ho et al. [1989].

Cell Culture

HeLa cells were obtained from the German Collection of Microorganisms and Cell Cultures (DSM, Braunschweig, Germany). They were grown in Modified Eagle's Medium (MEM; Biochrom, Berlin, Germany) containing 10% (v/v) fetal calf serum (FCS, Biochrom). Cultures were maintained in a humidified incubator with 5% CO_2 atmosphere at 37°C. Cells were removed from plastic dishes by trypsinisation and replated at a density of 5×10^5 cells/ml MEM on glass coverslips 48 h before use in the permeabilisation assay.

Transfection

Transfection experiments were done by electroporation with the Easyject system (Eurogentec, Seraing, Belgium). A total of 8×10^5 cells and 25 μg plasmid DNA were electroporated in electroporation cuvettes in a total volume of approximately 1 ml MEM. Electroporation parameters were: 240 V, 1350 μF, 156 Ω, pulse time 211 ms. After electroporation the cells were seeded in 5 ml MEM in 6-well plates and grown on coverslips. After 24 h the subcellular localisation of the transport substrate was detected by indirect immunofluorescence.

Immunodetection

For detection by indirect immunofluorescence the cells were washed two times in phosphate-buffered saline (PBS, 0.14 M NaCl, 0.25 M KCl, 8 mM Na_2HPO_4, pH 7.2). For fixation the cells were incubated with 3% paraformaldehyde in PBS for 15 min at room temperature, followed by two washes with PBS and a permeabilisation step with 0.5% (v/v) Triton X-100 in PBS for 15 min. After two washes with PBS unspecific binding was blocked with blocking buffer (3% bovine serum albumin in PBS) for 15 min. The cells were then incubated in a 1:1000 dilution of the polyclonal antibody against β-galactosidase (Sigma, St. Louis, MO) in blocking buffer for 1 h at 37°C. For the visualisation of β-galactosidase the cells were washed with PBS and incubated for 1 h at 37°C with a 1:200 dilution in blocking buffer of an anti-mouse IgG antibody conjugated to TRITC, i.e., tetramethylrhodamine isothiocyanate (Dianova, Hamburg, Germany).

DNA was visualised by the DNA-binding substrate diamino-2-phenylindol (DAPI, Sigma) that was part of the embedding medium (90% glycerol, 10% PBS, 1 mg/ml phenylenediamin, 1 μg/ml DAPI). Immunofluorescence was evaluated with a Zeiss microscope (Axioscope 20) using a 100× objective lens (Plan Neofluar 100). Pictures were obtained by digital imaging.

Preparation of Labelled Import Substrates

Core histone H2A was labelled with TRITC (Sigma). Protein concentration was estimated by SDS-polyacrylamide gel electrophoresis. A 0.5 ml aliquot of a 1 mg/ml calf thymus H2A solution (Roche, Mannheim, Germany) in 0.1 M sodium carbonate buffer (pH 9.0) was mixed with 25 μl of a 1 mg/ml TRITC stock solution in dimethyl sulfoxide. After 3 h incubation at room temperature the TRITC-modified H2A was purified by gel filtration on Sephadex G25 (Sephadex™; Pharmacia, Uppsala, Sweden). H2B was labelled with ALEXA™546 (Molecular Probes, Eugene), H3 and H4 were labelled with ALEXA™488 (Molecular Probes). The procedure followed was as described in the manufacturer's manual with the exception that precipitated H3 and H4 conjugates were dissolved in 10 mM HCl to a maximum concentration of 1 mg/ml. Using a control substrate it was ensured that this concentration of hydrochloric acid had neither a detectable effect on nuclear transport nor on the permeabilised cells in the in vitro import assay (not shown).

Permeabilisation Assay

Nuclear import was studied using the permeabilised HeLa cell assay [Adam et al., 1990]. HeLa cells were grown on 10 mm coverslips, washed one time with ice-cold transport buffer (20 mM HEPES, pH 7.3, 110 mM potassium acetate, 2 mM magnesium acetate, 1 mM EGTA, 2 mM DTT, 1 μg/ml each of aprotinin, leupeptin and pepstatin) and permeabilised for 5 min on ice with 40 μg/ml digitonin (Calbiochem, Bad Soden, Germany) in transport buffer. Subsequently, the cells were kept on ice for efficient cytosol depletion and were washed twice with transport buffer every 5 min. The coverslips were incubated with 20 μl of import mixture for 30 min at 37°C. The import mixture contained 1 μl transport substrate at a final concentration of either 500 nM (H2A, H2B) or 750 nM (H3, H4) fluorescence-labelled core histone, 10 μl nuclease-treated reticulocyte lysate (Promega, Madison, WI) as cytosol source, 1 mg/ml BSA, energy regenerating system consisting of 1 mM ATP, 10 mM creatine phosphate, 20 U/ml creatine kinase (final concentrations) and transport buffer in a total volume of 20 μl. Transport reactions were stopped by fixation with 3% paraformaldehyde (w/v) in PBS, washed with PBS and mounted with embedding medium containing DAPI for DNA visualisation. Monitoring of the subcellular localisation of the import substrates was done as described above.

Inhibition experiments were done as indicated either by substitution of cytosol by transport buffer, incubation of the import mixture at 4°C, simultaneous omission of ATP and addition of apyrase (Sigma) at a final concentration of 100 U/ml, pre-incubation for 15 min

at room temperature with 50 μg/ml wheat germ agglutinin (WGA; Calbiochem, a Jolla, CA) or addition of 1 mM GTPγS or GMP-PNP (both from Roche).

RESULTS

Core Histones do Not Enter the Nucleus Simply by Diffusion

Core histones are small proteins with a size of 11 (H4) to 15 kDa (H3) that form the nucleosome octamer of eukaryotic chromatin and thus are localised in the nucleus. As all nuclear proteins that are synthesised in the cytoplasm they must enter the nucleus through the nuclear pore complex. This complex contains an aqueous channel that allows molecules up to 40 kDa in size to diffuse into the nucleus [Paine et al., 1975; Feldherr et al., 1984]. Since core histones are smaller than this exclusion size, they might be expected to be translocated into the nuclear compartment by diffusion. To study the import characteristics of the core histones H2A, H2B, H3 and H4 we used the digitonin permeabilisation assay [Adam et al., 1990] and tested core histone transport under varied conditions (Fig. 1).

Under standard conditions (37°C, addition of cytosol and ATP regenerating system) we observed predominantly nuclear staining indicating nuclear transport of fluorescence-labelled calf thymus core histones H2A, H2B, H3 and H4. Replacement of cytosol by transport buffer led to an exclusion of fluorescence-labelled histones from the nucleus. Thus, the import must depend on cytoplasmic components that are involved in the transport reaction into the nucleus. When the temperature was lowered to 4°C, fluorescence was excluded from the nucleus. This strong temperature dependence shows that the nuclear import of core histones is not based on diffusion through the nuclear pore. In order to analyse the energy dependence of the import, the ATP regenerating system was omitted and apyrase was added, which depletes endogenuous ATP. These conditions inhibited the nuclear import of core histones. The import was also blocked when incubating the permeabilised cells with WGA. This lectin binds to sugar residues of nucleoporins and thus inhibits the translocation of the transport substrates through the nuclear pore complex [Burglin and de Robertis, 1987]. Furthermore, the import of core histones is sensitive to the non-hydrolysable GTP analogues GMP-PNP and GTPγS (the latter not shown), indicating the dependence of the translocation step on GTP hydrolysis. Due to an extensive washing of the digitonin-treated cells after performing the transport reaction, the import substrate was washed out from the cytoplasm in some cases and therefore no cytoplasmic staining was observed.

As shown in the first row of Figure 1, nuclear import of each of the core histones resulted in an accumulation of labelled histones in the nucleoplasm and in nucleoli. Nucleolar staining may be due to the fact that just few unoccupied binding sites for core histones are available in the pre-existing interphase chromatin of the permeabilised cells. There they may bind as basic proteins to the ribosomal RNA accessible in the nucleolus.

In conclusion, the data shown in Figure 1 demonstrate that core histones are transported into the nucleus by an active and receptor-mediated process and do not enter the nucleus simply by diffusion.

Each Core Histone Contains Two Different Types of NLS Structure

After we had shown that all core histones enter the nucleus in a receptor-mediated manner, we searched in each core histone for sequence elements that are functional for nuclear targeting, i.e., the NLSs. Classical NLS motifs have been described as clusters of basic amino acids such as in SV40 large-T-antigen [Kalderon et al., 1984] or nucleoplasmin [Robbins et al., 1991], but there also are non-classical nuclear targeting motifs that contain clusters of basic amino acids such as the NLS of ribosomal proteins [Jäkel and Görlich, 1998; Claussen et al., 1999; Rosorius et al., 2000].

Since the core histones are very basic proteins rich in lysine and arginine several potential NLSs may be identified within the primary structure. To determine which portions of the different core histones can function as an NLS, a series of core histone-β-galactosidase fusion constructs were generated. These were constructed from partial sequences of each core histone gene fused downstream of the SV40 promoter to the 5′ end of the bacterial β-galactosidase gene in a mammalian expression vector. We used the bacterial β-galactosidase as a reporter protein since it is a large protein, devoid of an NLS and is thus excluded

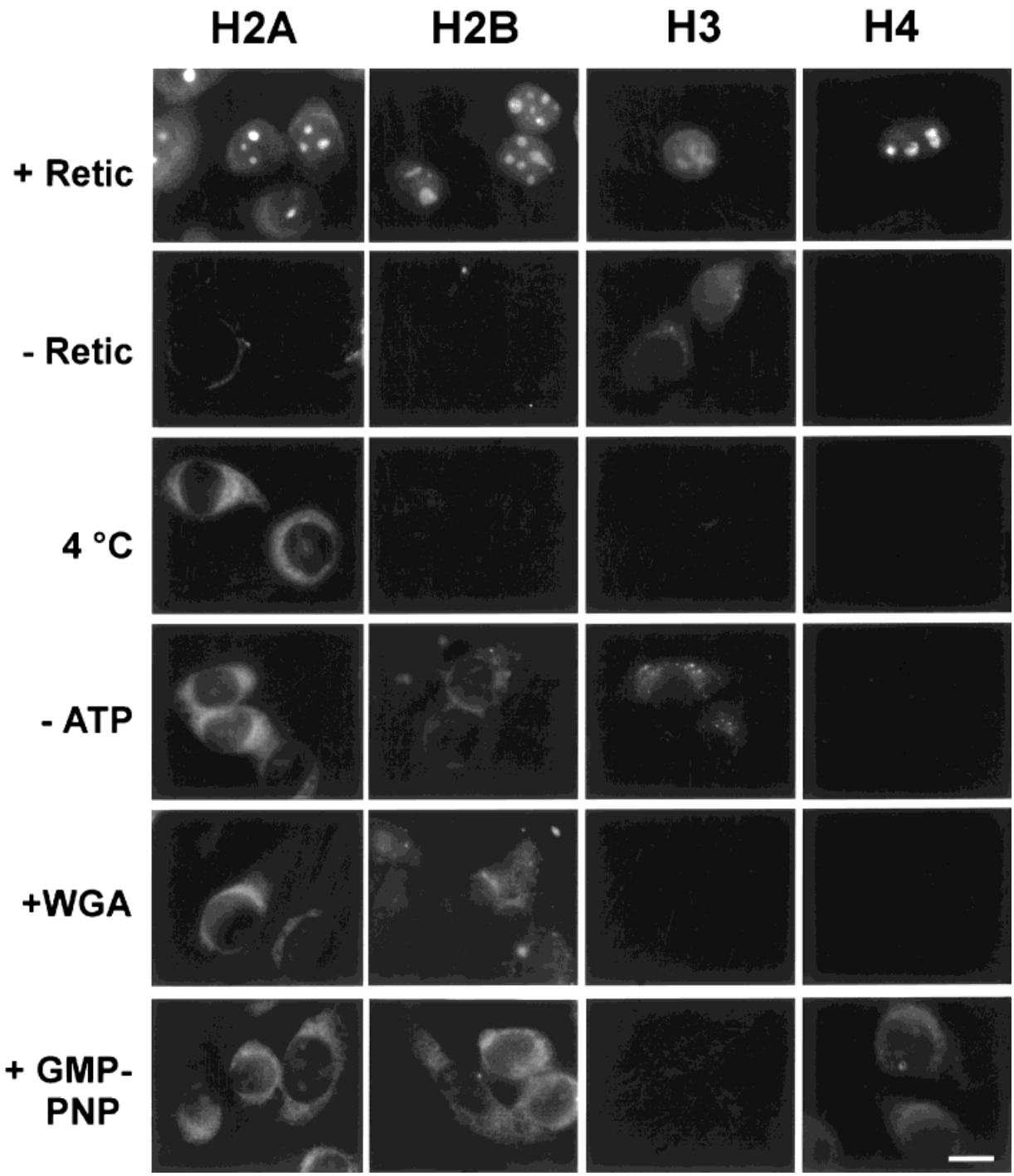

Fig. 1. Transport of core histones into HeLa cell nuclei and analysis of inhibitory conditions. Analysis of fluorescence-labelled calf thymus core histone H2A, H2B, H3 and H4 in digitonin-permeabilised HeLa cells. Permeabilised HeLa cells were incubated with fluorescence-labelled histones, reticulocyte lysate and an energy regenerating system at 37°C for 30 min (control conditions, first row). Second row: as in control, but reticulocyte lysate was replaced by transport buffer. Third row, as in control, but at 4°C. Fourth row, as in control, but ATP-regenerating system was omitted and HeLa cells were pre-incubated with apyrase. Fifth row, as in control, but cells were pre-incubated with WGA. Last row: as in control, but in the presence of GMP-PNP. Note that labelled core histones that are transported into the nucleus preferentially bind to nucleolar material (apparently due to a lack of available binding sites in the pre-existing chromatin). Bar represents 10 μm.

from the nucleus [Newmeyer and Forbes, 1988; Nelson and Silver, 1989; Schmolke et al., 1995]. Each construct was transiently expressed in HeLa cells and the subcellular localisation of the fusion proteins was monitored by indirect immunofluorescence with a polyclonal antibody against β-galactosidase. The results of these experiments are compiled in Figure 2. Examples obtained with individual constructs are shown in Figures 3 and 4.

The fusion proteins generated from full length core histone sequences and β-galactosidase are represented by pWA355 (H2A), pWA352 (H2B), pWA356 (H3) and pWA357

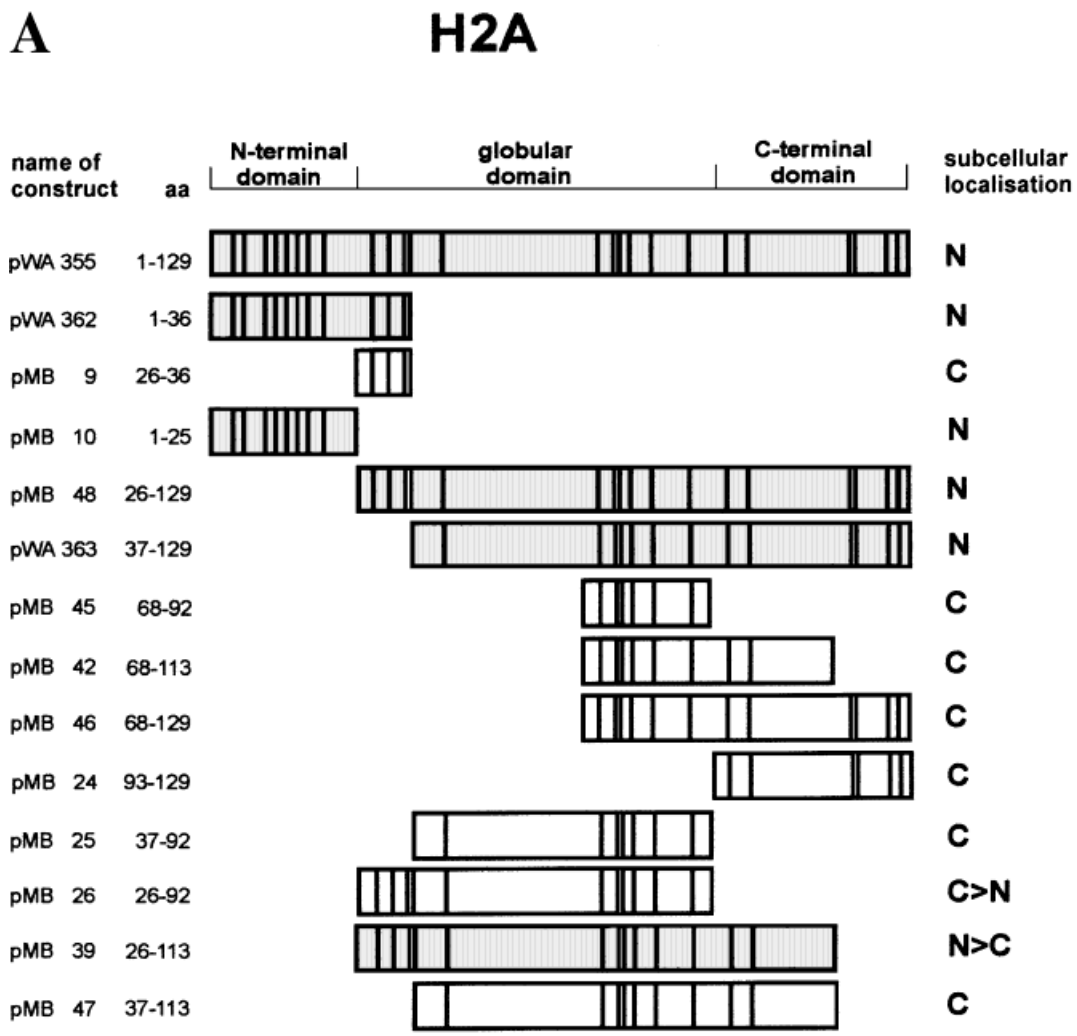

Fig. 2. Nuclear transport of constructs consisting of core histones or core histone fragments fused to β-galactosidase. Summary of the transfection experiments: Expression vectors containing gene fusions of the core histone gene or gene portions with the β-galactosidase gene were transfected into HeLa cells and the subcellular localisation of the expressed fusion proteins was monitored by indirect immunofluorescence. Each histone fragment was cloned upstream of the β-galactosidase gene and downstream of the SV40 early promoter within pKS10 [6]. On the left, the names of the constructs and the amino acids contained in the constructs are listed. The core histone fragments are drawn as boxes in relation to the full length protein as indicated on top. Arginine and lysine residues are indicated as vertical lines within each box. The subcellular localisation of each fusion protein is listed on the right (N, nuclear localisation; C, cytoplasmic localisation; N > C, localisation in both compartments, but predominant staining in the nucleus; C > N, indicates the reverse). Nuclear localisation of constructs is in addition indicated by shading of the respective box. **A**. Nuclear transport of H2A or H2A fragments fused to β-galactosidase. **B**. Nuclear transport of constructs consisting of H2B or fragments of H2B fused to β-galactosidase. In the cases of pMB27 and pMB28 we have observed fluorescence of fusion proteins either in the nucleus alone or in both nucleus and cytoplasm (C; C/N). In the case of pMB22, fluorescence was found in nuclei or cytoplasm or in both compartments (N; C; C/N). **C**. Nuclear transport of constructs consisting of H3 or fragments of H3 fused to β-galactosidase. **D**. Nuclear transport of constructs consisting of H4 or fragments of H4 fused to β-galactosidase.

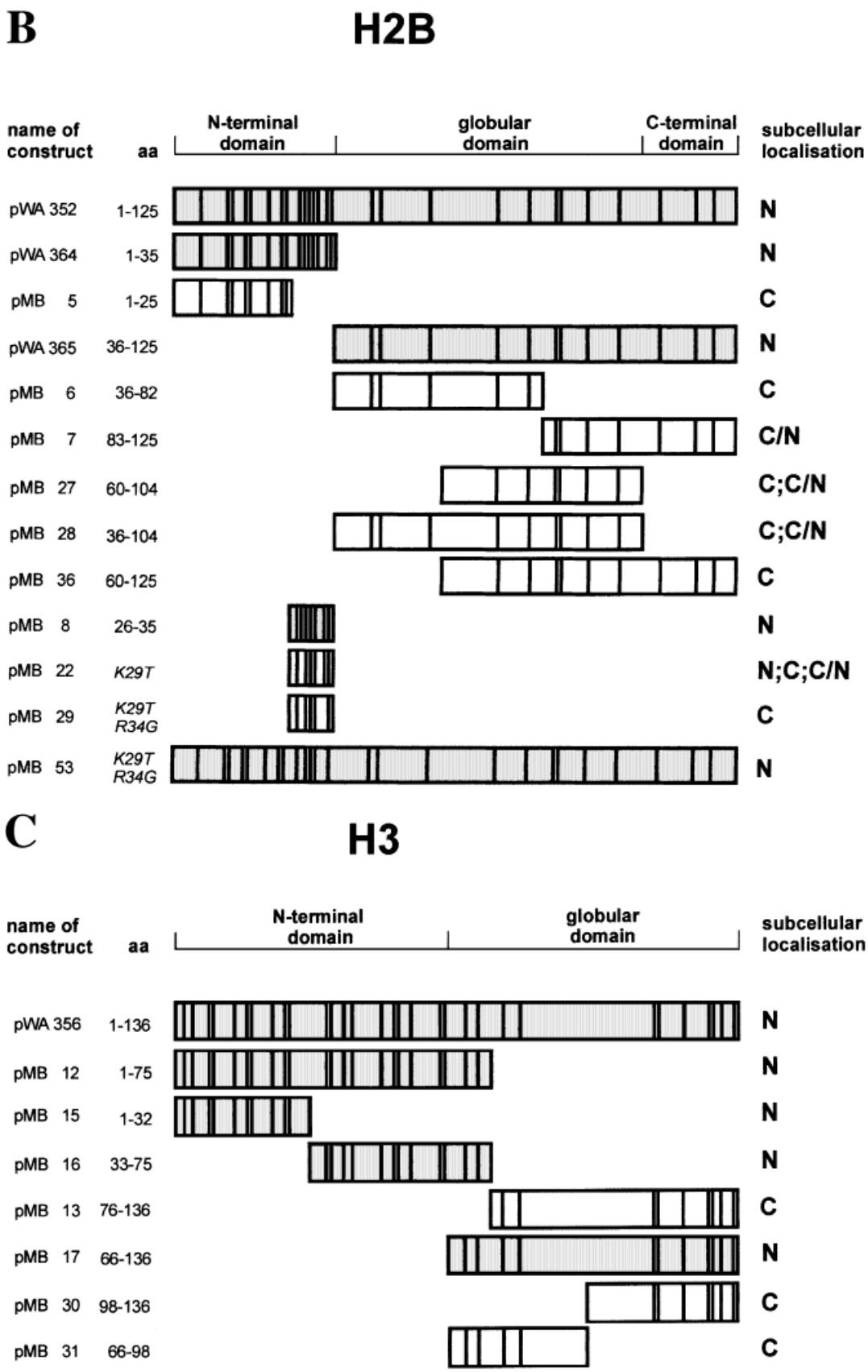

Fig. 2. (*Continued*)

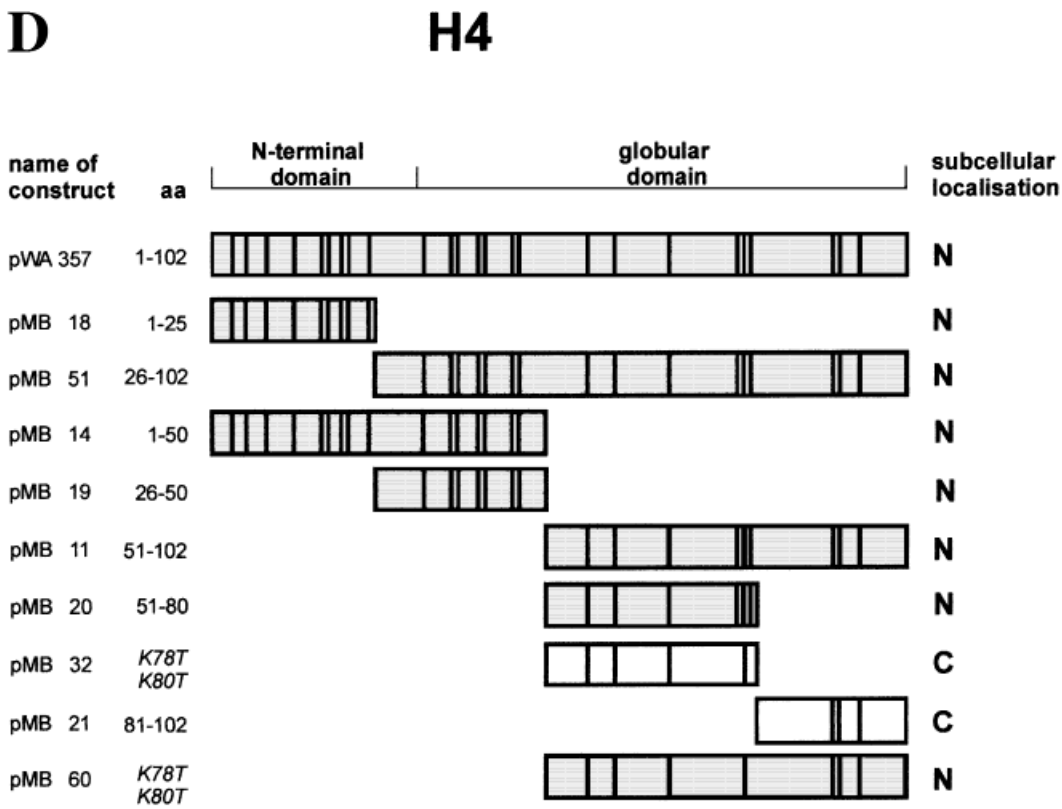

Fig. 2. (*Continued*)

(H4). As shown in the first lines of Figure 2A–D, each core histone fused to β-galactosidase is imported into the nucleus. We then analysed individual segments of each of the core histones fused to the reporter protein. Core histones show a three-domain structure consisting of a highly charged, unstructured amino terminal domain, a short carboxy terminal tail, and a conserved globular domain.

First we examined the NLS function of the individual domains of each core histone (Fig. 2A–D). The core histone H2A contains nearly equal amounts of arginines and lysines that are clustered in the N-terminal segment (residues 1–36) and again at the C-terminus (residues 118–129). The highly conserved core region is comparatively hydrophobic and is involved in the interaction with other histones, whereas the basic protein segments bind to DNA. The construct pMB10 containing the amino terminal domain of H2A in front of β-galactosidase was efficiently imported into the nucleus. Within this H2A fragment a basic motif is recognisable that resembles the classical NLS type. The region of H2A containing the globular and the C-terminal domain fused to β-galactosidase (pMB48) was also detected in the nucleus. Surprisingly, its nuclear import behaviour changed when we divided it into two parts representing the globular (pMB26) and the C-terminal domains (pMB24), respectively: both are predominantly localised in the cytoplasm. To ensure that we did not split an NLS motif we tested several overlapping fusion proteins (pMB42, pMB45, pMB46, pMB47) that should not be interrupted within their potential NLS. The cytoplasmic localisation of the fusion proteins with either pMB47, pMB26, or pMB25 and the increasing import competence of fragments extending towards the amino or C-terminal direction (pMB39, pWA363) demonstrates that the NLS within the segment 26–129 (pMB48) is different from that in the amino terminal domain and suggests that the recognition of this signal may be based on the three-dimensional structure of the globular domain. This appears to form a signal region rather than a primary structure motif.

The N-terminal third of human H2B histones, when compared with H2B from other

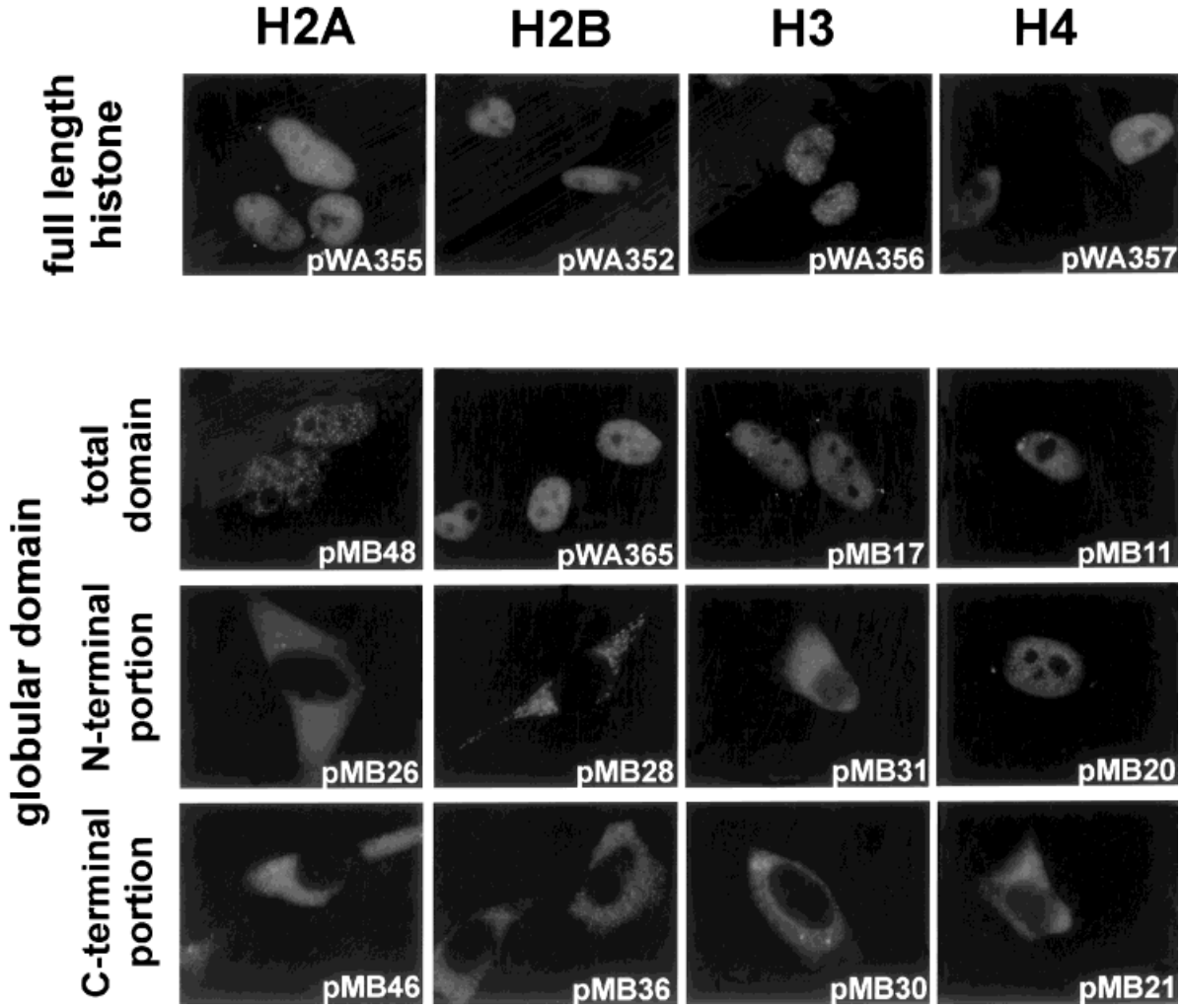

Fig. 3. Transient transfection with core histone β-galactosidase fusion constructs in HeLa cells: examples for subcellular localisation of fusion proteins as detected by immunofluorescence using a monoclonal antibody against β-galactosidase and a fluorescence-labelled second antibody. Data obtained with fusion proteins of total core histones, full length globular domains and portions of the globular domain with β-galactosidase, respectively, are shown. For details of individual constructs, see Figure 2A–D.

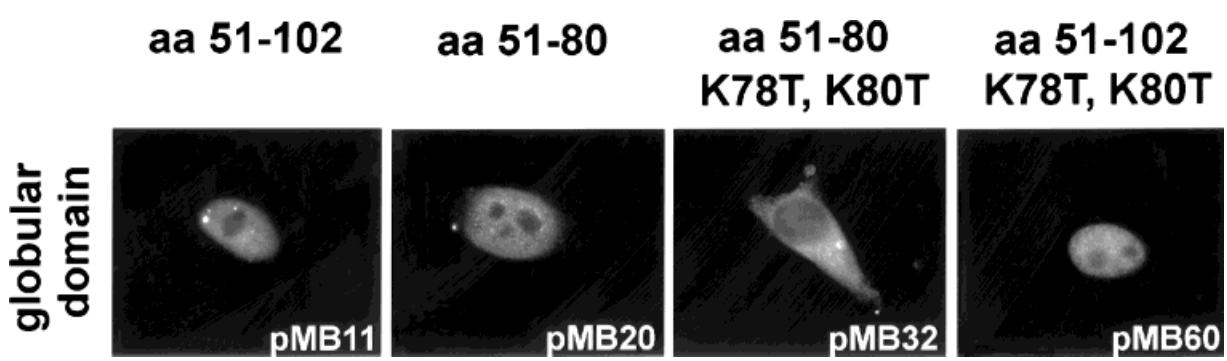

Fig. 4. Subcellular localisation of fusion constructs of wild-type and mutated H4 globular domain histone fragments fused to β-galactosidase. The globular domain (pMB11) or the N-terminal part of the globular domain (pMB20) can guide the fusion protein to the nucleus, but mutation of two basic amino acids in the N-terminal subdomain leads to exclusion from the nucleus (pMB32). Mutation of the same amino acid residues in the entire globular domain, however, does not preclude nuclear targeting of the fusion construct (pMB60).

species, is the least conserved part of the core histones. Within this domain positively charged amino acids are clustered around the postions 27–34, whereas the region 36–125 (pWA365) representing the globular and C-terminal domain is relatively hydrophobic. We could show that the amino terminal domain fused to β-galactosidase is imported into the nucleus (pWA364) and we furthermore identified a basic region of nine amino acids that can function as an NLS (pMB8), whereas the amino terminal region 1–25 (pMB5) did not show any import competence (Fig. 2B). Exchange of one basic amino acid to an uncharged one diminished the import competence of the pMB8 fragment (pMB22) and mutation of a second basic amino acid resulted in a complete cytoplasmic localisation of the β-galactosidase fusion protein (pMB29). This result shows that the nuclear import competence of this basic region (aa 26–35; pMB8) of the histone H2B represents an NLS motif that resembles a classical type such as in SV40 large-T-antigen. Mutation of these two basic aa to uncharged aa within the whole H2B gene does not reduce the import rate of the β-galactosidase fusion protein (pMB53). This observation indicates that the core histone H2B has at least one second NLS motif. The region comprising residues 36–125 (pWA365) containing the globular and C-terminal domains features 12 basic amino acids, and a typical NLS-like motif is not recognisable within this region. Surprisingly, this H2B fragment guides the fusion protein into the nucleus. Subdivision of this region into two parts (pMB6 and pMB7) resulted in a change of the import behaviour in a way that the proximal part remained completely cytoplasmic and the distal part was uniformly localised in the cytoplasm as well as in the nucleus. Fragments that lack the C-terminal domain (pMB27, pMB28) or the aa 36–59 of the globular domain (pMB36) were not efficiently imported into the nucleus when fused to β-galactosidase. This result shows that the NLS is formed by the entire region of aa 36–125 (pWA365). This signal is different from the classical type of NLS and it may be assembled from different parts of the protein by its conformation, but may not depend just on a primary structure element.

Analysis of the core histone H3 (Fig. 2C) revealed two transport-competent sequence elements within the amino terminal domain (pMB15, pMB16), that contain multiple stretches of basic amino acids and thus suggest that these elements form NLSs resembling a classical NLS-type. The C-terminus which is rich in positively charged residues is not imported into the nucleus when fused to β-galactosidase (pMB30). On the other hand the globular domain (pMB17) can function as an NLS, but deletion of the first 10 amino acids that contain two basic amino acids abolishes the import competence of this fragment (pMB13). These 10 amino acid residues are not able to target β-galactosidase into the nucleus (pMB31) showing that they may only form a recognition signal in the context of the whole globular domain.

In case of H4 (Fig. 2D) we found the amino terminal domain to be able to guide β-galactosidase into the nucleus (pMB18). Within the globular domain (pMB51) we could identify two import competent regions (pMB19 and pMB20) and a portion towards the C-terminus (pMB21) that failed to guide the fusion protein into the nucleoplasm. As shown in Figure 2D, the region covered by pMB19 (aa 26–50) contains three basic dipeptides and pMB20 covers a region with three single basic amino acids and a basic tripeptide (aa 78–80). Exchanging the two lysines 78 and 80 in pMB20 to uncharged amino acids (resulting in pMB32) we observed a total loss of the ability of this fagment to target β-galactosidase into the nucleus suggesting that the two basic residues 78 and 80 are involved in the formation of an NLS within this construct (Figs. 2D, 3, 4). In contrast, mutation of these residues within the fusion construct comprising the major part of the globular domain (pMB60) did not exclude the fusion protein from the nucleus. Thus we suggest that the nuclear transport of pMB20 is due to an artificial NLS resulting from fusion of this particular histone H4 fragment with the β-galactosidase.

In conclusion, we could show that each core histone contains within the amino terminal domain a (basic) NLS that resembles the classical NLS-type. For example, in case of H2B we identified a sequence element enriched in basic amino acids that functions as an NLS and shows high agreement with the SV40 large-T-antigen NLS. Furthermore, we demonstrate that each core histone contains another type of NLS that is different from the classical type and resides within the globular domain of the core histones.

DISCUSSION

During the S-phase of the cell cycle, the formation of chromatin on replicating DNA requires a vast amount of core and linker histones. Since histones are synthesised in the cytoplasmic compartment, an efficient nuclear transport system must guarantee an adequate supply of histone proteins during the S-phase. Nuclear transport of histones simply based on diffusion may in principle be possible due to the low molecular weight of these basic proteins. However, diffusion would not suffice to provide, in a regulated way, the S-phase nucleus with the high amount of histones needed during the short period of DNA replication. Hence, it has been shown [Breeuwer and Goldfarb, 1990; Kurz et al., 1997] that the nuclear transport of H1 histones is a receptor-mediated, energy-dependent process.

Nuclear transport depends on topogenic signals in the protein sequence as binding sites for the respective transport factors. Nuclear translocation of proteins with classical type NLS proceeds via importin β and importin α, the latter being the adaptor protein interacting with the classical-type NLS. A type of NLS that differs from the classical type has been described in the hnRNP A1 protein. This sequence termed M9 consists of 38 amino acids with no basic amino acid clusters. It is the binding site of the soluble cytoplasmic receptor protein transportin, which is a member of the importin β family [Pollard et al., 1996].

Histones are nuclear proteins enriched in basic amino acids. Hence, their nuclear transport may be expected to be a classical NLS-mediated process. This is supported by Moreland et al. [1987] who have shown that in yeast H2B a sequence similar but not identical to the SV40 large-T-antigen NLS determines the nuclear localisation of this histone. This was shown with histone–β-galactosidase fusions that were directed towards the nucleus when the whole histone sequence or fragments thereof fused to the β-galactosidase were analysed in transient transfection assays in yeast. Using a similar assay in HeLa cells, we have shown that the H1° histone contains multiple sequence elements for nuclear targeting [Schwamborn et al., 1998]. Fragments enriched in basic amino acids (but not identical to one of the two classical NLS motifs) derived from the C-terminal domain of the protein showed nuclear import competence in the fusion constructs. Remarkably, the central globular domain of the protein could also serve as a nuclear targeting signal, although it contains no basic amino acid clusters. Since subfragments of this globular domain could not serve as NLS, we concluded that the conformation of this part of the protein may contribute to the assembly of a topogenic signal for nuclear targeting. We have recently shown that this H1 transport depends on cooperation of import receptors importin β and importin 7 [Jäkel et al., 1999]. It is, however, not yet known which segments of the H1 histone protein interact with the importin 7 or the importin β protein binding site.

In the present study we have searched for NLSs in the sequences of the four core histones. All four core histones essentially have a very basic N-terminal domain, but just a short or no C-terminal domain and a central globular portion with a well-defined tertiary structure [for review, see Bradbury, 1992]. In our search for nuclear localisation signals in each of the four core histones, we have constructed fusions between the individual histone genes (or fragments thereof) and the β-galactosidase gene.

The N-terminal domains of each of the four core histones could serve as a signal for the nuclear targeting of the fusion protein. In the case of H2B, this was expected since Moreland et al. [1987] had previously shown that a heptapeptide (GKKRSKA) in the yeast H2B N-terminal domain could serve as an NLS, and the respective sequence portion in human H2B (KKRKRSRK) is even more similar to the classical SV40 large T-antigen-type NLS. The N-terminal domains of H2A, H3 and H4, however, could mediate nuclear targeting although the basic amino acid residues enriched in these protein domains are not arranged as in classical-type NLS. Thus, the nuclear transport competence of these portions of the core histones may be mediated by a similar, but not identical transport pathway as for proteins with classical-type NLS.

Analysis of the remaining, globular portions of the core histones indicated that these could serve as nuclear targeting signals when taken in their entirety. Fusion proteins constructed with subfractions missing parts of the globular domains of H2A, H2B or H3 were not transported to the nucleus. Since none of these globular protein domains contains clusters of

basic amino acids, we again must conclude that no classical NLS motif mediates nuclear targeting of this part of the protein. The primary structures of the globular domains of the individual core histone classes show no obvious sequence similarity, but share conformation motifs. The globular domains, which have initially been defined by NMR spectroscopy and controlled proteolysis [Bradbury, 1992], comprise the now-established histone fold motif. This motif is involved in both histone–histone and histone–DNA interactions within the nucleosome [Luger et al., 1997]. Other DNA-binding proteins also contain the histone fold motif [Baxevanis et al., 1995]. These histone fold proteins are distantly related, with a relatively small degree of absolute sequence similarity.

Since the globular protein domains of each core histone form a similar motif, we suggest that the conformation of this part of the protein may be recognised by a common receptor protein mediating nuclear transport. We cannot exclude that the nuclear transport competence of the globular domains includes formation of histone oligomers. Since the histone fold motif is necessary to form H2A–H2B dimers or $(H3–H4)_2$ tetramers [Freeman et al., 1996], it may be that the globular domain of the core histone contained in the fusion construct may interact with the corresponding endogenous newly synthesised partner core histone in the cytoplasm of the transfected cell. In this case the dimeric complex would be imported into the nucleus on the basis of the NLS within the N-terminal domain of the endogenous histone. This possibility is supported by the results of Kolodrubetz et al. [1982] and Schuster et al. [1986]. They have studied in vivo the viability of yeast mutants with deletions in H2A and H2B histones. Deletions of the N-terminal domain of either H2A or H2B did not affect the viability of the yeast cells. In addition, the respective regions were interchangeable between H2A and H2B, but simultaneous deletions of the N-termini of both histones were lethal. Therefore the authors also discuss the possibility that the N-terminal ends may be involved in nuclear targeting of H2A–H2B dimers.

Similarly, our data cannot exclude that interaction of globular domains with a protein that contains a bona fide NLS may mediate the nuclear transport of core histones. Given the necessity to transport a vast amount of histones during the S-phase of the cell cycle, such a histone-binding protein would require a high transport capacity and should be able to efficiently shuttle between nucleus and cytoplasm.

Several core histone-binding proteins have been described; often they are involved in chromatin assembly and form subunits of the chromatin assembly factors (CAF). The nucleosome assembly factors N1/N2 bind preferentially H3 and H4, whereas nucleoplasmin and members of the NAP family bind H2A and H2B, even in the cytoplasm [Ito et al., 1996]. This observation and the fact that NAP-1 contains a classical NLS, led to the proposal that NAP-1 acts as a core histone shuttle that delivers the histones by a piggy-back mechanism from the site of synthesis in the cytoplasm to the chromatin assembly machinery in the nucleus [Ito et al., 1996]. Since until now no experimental data exist for this model in terms of nuclear transport and since we have identified multiple NLS within the core histones, an involvement of such factors in nuclear transport appears rather hypothetical. It remains to be shown whether such oligomeric complexes can be transported towards the nucleoplasm.

The fusion proteins derived from fragments of the globular portion of H4 showed differences compared with other core histone globular domains. The N-terminal domain in its entirety mediated nuclear targeting, but also a fragment overlapping with the N-terminal domain (pMB19) and the fragment pMB20 served as an element for nuclear targeting. In the case of pMB20, a cluster of three basic amino acids appeared to be involved in this nuclear targeting, since mutation of two of these basic amino acids suppressed this NLS function. On the other hand, the same mutations when introduced into the long globular domain construct (pMB11 mutated to pMB60) did not abolish the nuclear targeting competence of this segment of the protein (Fig. 4). These data suggest that trimming of the fragment of aa 51–102 (pMB11) to aa 51–80 (pMB20) may have created a conformational change that assigned transport competence to this fusion protein. This, however, would be an artifact of the fusion construct.

In conclusion, our transport data suggest that nuclear transport of core histones varies from classical NLS-mediated transport. As shown by mutation analysis within the H2B

N-terminal region, the NLS formation of this domain appears to depend on the distribution of basic amino acids within that part of the protein. Mutation of this targeting element between amino acids 26 and 35 (as in fusion construct pMB29), however, had no effect on the transport competence of the entire H2B protein, but just on the ability of the dekapeptide 26–35 to mediate nuclear targeting of the fusion protein. Data obtained with the globular domains of each of the four core histones suggest that nuclear targeting by these regions depends on conformational information. Further studies now must identify the import factors involved in core histone transport. Dissection of the histone proteins and binding studies with recombinant import receptors (members of the importin β family) will reveal differential binding of importin subtypes to individual core histone fragments.

The nuclear transport of histones varies from the transport of all nuclear proteins in several ways. First, the need to transfer a vast amount of proteins is exceptionally high in the case of newly synthesised histones during the S-phase of the cell cycle. Secondly, histones vary from other proteins in their amino acid composition. Their high content of basic amino acids may require a shielding against electrostatic interactions with other proteins before being deposited onto newly replicated DNA. Thirdly, core histones can specifically interact with each other and with other proteins during nucleosome assembly. These requirements may necessitate complex interactions during cytoplasmic assembly, translocation and nuclear disassembly of histone transport complexes.

Both core and H1 histones contain multiple sequence elements for nuclear targeting. Basic amino acid residues in extended domains of the proteins may mediate transport by factors similar (but not identical) to the classical pathway, whereas the globular histone domains appear to require interactions with transport receptors on the basis of conformational rather than short sequence information.

REFERENCES

Adam SA, Sterne-Marr R, Gerace L. 1990. Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. J Cell Biol 111:807–816.

Albig W, Doenecke D. 1997. The human histone gene cluster at the D6S105 locus. Hum Genet 101:284–294.

Baxevanis AD, Arents G, Moudrianakis EN, Landsman D. 1995. A variety of DNA-binding and multimeric proteins contain the histone fold motif. Nucl Acids Res 23:2685–2691.

Breeuwer M, Goldfarb DS. 1990. Facilitated nuclear transport of histone H1 and other small nucleophilic proteins. Cell 60:999–1008.

Bradbury EM. 1992. Reversible histone modifications and the chromosome cell cycle. BioEssays 14:9–16.

Burglin TR, de Robertis EM. 1987. The nuclear migration signal of *Xenopus laevis* nucleoplasmin. EMBO J 6:2617–2625.

Christophe D, Christophe-Hobertus C, Pichon B. 2000. Nucler targeting of proteins: how many different signals? Cellular Signalling 12:337–341.

Claussen M, Rudt F, Pieler T. 1999. Functional modules in ribosomal protein L5 for ribonucleoprotein complex formation and nucleocytoplasmic transport. J Biol Chem 274:33951–33958.

Feldherr CM, Kallenbach E, Schultz N. 1984. Movement of a karyophilic protein through the nuclear pores of oocytes. J Cell Biol 99:2216–2222.

Freeman L, Kurumizaka H, Wolffe AP. 1996. Functional domains for assembly of histones H3 and H4 into the chromatin of *Xenopus* embryos. Proc Natl Acad Sci USA 93:12780–12785.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77:51–59.

Imamoto N, Tachibana T, Matsubae M, Yoneda Y. 1995. A karyophilic protein forms a stable complex with cytoplasmic components prior to nuclear pore binding. J Biol Chem 270:8559–8565.

Ito T, Bulger M, Kobayashi R, Kadonaga JT. 1996. Drosophila NAP-1 is a core histone chaperone that functions in ATP-facilitated assembly of regularly spaced nucleosomal arrays. Mol Cell Biol 16:3112–3124.

Jäkel S, Görlich D. 1998. Importin β, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. EMBO J 17:4491–4502.

Jäkel S, Albig W, Kutay U, Bischoff FR, Schwamborn K, Doenecke D, Görlich D. 1999. The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1. EMBO J 18:2411–2423.

Jackson V. 1990. In vivo studies on the dynamics of histone–DNA interaction, evidence for nucleosome dissolution during replication and transcription and a low level of dissolution independent of both. Biochemistry 29:719–731.

Kalderon B, Roberts BL, Richardson WD, Smith AE. 1984. A short amino acid sequence able to specify nuclear location. Cell 39:499–509.

Kolodrubetz D, Rykowski MC, Grunstein M. 1982. Histone H2A subtypes associate interchangeably in vivo with histone H2B subtypes. Proc Natl Acad Sci USA 79:7814–7818.

Kurz M, Doenecke D, Albig W. 1997. Nuclear transport of H1 histones meets the criteria of a nuclear localisation signal-mediated process. J Cell Biochem 64:573–578.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389:251–260.

Moreland RB, Langevin GL, Singer RH, Garcea RL, Hereford LM. 1987. Amino acid sequences that determine the nuclear localization of yeast histone 2B. Mol Cell Biol 7:4048–4057.

Nelson M, Silver P. 1989. Context affects nuclear protein localization in *Saccharomyces cerevisiae*. Mol Cell Biol 9:384–389.

Newmeyer DD, Forbes DJ. 1988. Nuclear import can be separated into distinct steps in vitro: nuclear pore binding and translocation. Cell 52:641–653.

Paine PL, Moore LC, Horowitz SB. 1975. Nuclear envelope permeability. Nature 254:109–114.

Pollard VW, Michael WM, Nakielny S, Siomi MC, Wang F, Dreyfuss G. 1996. A novel receptor-mediated nuclear protein import pathway. Cell 86:985–994.

Robbins J, Dilworth SM, Laskey RA, Dingwall C. 1991. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell 64:615–623.

Rosorius O, Fries B, Stauber RH, Hirschmann N, Bevec D, Hauber J. 2000. Human ribosomal protein L5 contains defined nuclear localization and export signals. J Biol Chem 275:12061–12068.

Schmolke S, Drescher P, Jahn G, Plachter B. 1995. Nuclear targeting of the tegument protein pp65 UL83 of human cytomegalovirus: an unusual bipartite nuclear localization signal functions with other portions of the protein to mediate its efficient nuclear transport. J Virol 69:1071–1078.

Schuster T, Han M, Grunstein M. 1986. Yeast histone H2A and H2B amino termini have interchangeable functions. Cell 45:445–451.

Schwamborn K, Albig W, Doenecke D. 1998. The histone H1(0) contains multiple sequence elements for nuclear targeting. Exp Cell Res 244:206–217.

Wolffe AP. 1998. Chromatin, structure and function, 3rd ed., London: Academic Press.

Worcel A, Han S, Wong ML. 1978. Assembly of newly replicated chromatin. Cell 15:969–977.