SUPPLEMENTAL MATERIAL

Supplemental Text

Accumulation of Oregon Green/pCMVβ-DTS into nuclei of digitonin-permeabilised HeLa cells. Before using the digitonin-permeabilised cell technique to assay pDNA nuclear import we verified the ability of permeabilised HeLa cells to accumulate fluorescent-labelled pDNA into their nucleus. As expected, both fluorescein-labelled human serum albumin (NLS-HSA-FITC) and the Oregon Green /pCMVβ-DTS accumulate in the nuclei of permeabilised HeLa cells in the presence of exogenously added cytoplasmic extract in a WGA-inhibitable manner (Supplemental Fig. S1*A* and S1*B*, respectively).

Analysis of plasmid DNA nuclear import active fraction by two-dimensional gel electrophoresis and identification by sequence analysis using mass spectrometry. Proteins eluted from a pDNA-affinity column were resolved using 2-D SDS-PAGE. Pooled concentrates of elution fractions were initially separated using a pH 3-10 broad range IPG in the IEF first dimension, coupled to 12% SDS-PAGE in the second dimension (hereafter referred to as pH 3-10/2-D SDS-PAGE) (Supplemental Fig. S2). For comparison, native HeLa cytoplasmic extract, control PNA-Sepharose column flow-through concentrates and pCMVβ-DTS/PNA-Sepharose column flow-through (100 μg proteins each) were resolved in the same way. The proteins were visualized from each 2-D SDS-PAGE gels after silver staining and the gels then digitised by laser densitometry and analysed using dedicated informatics software. The pH 3-10 broad range gels detected 590 ± 102, 426 ± 41, 333 ± 30, and 40 ± 22 (mean ± SEM from N=3 gels from 3 separate experiments) protein spots in the native extract, flow-through concentrates from both the pCMVβ-DTS/PNA-Sepharose and control PNA-Sepharose columns, and eluate fractions, respectively (Supplemental Fig. S2*E*). Proteins with alkaline isoelectric points (p*I*), which include most DNA-binding proteins, are poorly resolved using 2-D SDS-PAGE. To increase the likelihood of identifying additional DNA-binding proteins we, therefore, further separated the native extract, column flow-throughs and eluate concentrate using a pH 6-9 narrow basic IPG in the IEF first dimension and a 12% SDS-PAGE second dimension (denoted pH 6-9/2-D SDS-PAGE) (Supplemental Fig. S3). Supplemental Figure S3*E* shows that pH 6-9/2-D SDS-PAGE gels detected 342 ± 10, 262 ± 44, 226 ± 57, and 76 ± 6 (mean ± SEM) protein spots in the native extract, flow-through concentrates from both the pCMVβ-DTS/PNA-Sepharose and control PNA-Sepharose columns, and eluate fractions, respectively. Putative DNA-binding shuttle proteins were identified by sequence analysis using mass spectrometry. Using electrophoretic coordinates based on the silver-stained eluate fraction reference pH 3-10 (Supplemental Fig. S2*D*) and pH 6-9/2-D SDS-PAGE gels (Supplemental Fig. S3*D*), we manually removed 22 and 40 protein spots of different intensities from their respective preparative native HeLa cell cytoplasmic extract 2-D SDS PAGE gels, and subjected them to mass spectrometric analysis.

Functional classification and subcellular localization of proteins identified in the plasmid DNA nuclear import active fraction. The identified proteins were categorized by their biological function using the literature indicated in Supplemental Table S1. Where annotations were not clear, a Basic Local Alignment search (BLAST) query against the SWISSPROT database was undertaken for functional information.

HeLa cell G_1/S-block and cell cycle flow cytometric analysis. Cell cycle status was analysed by FACS analysis. Supplemental Figure S5 shows histograms of relative fluorescence representing the amount of genome DNA, versus the number of propidium iodide-stained HeLa cells. Serum-deprivation and aphidicolin-treatment of HeLa cells did not increase the number of cells in G_1 DNA content. However, the second peak indicating 4*n* DNA content (G_2/M) disappeared. In dividing cells, the nuclear envelope is disassembled at the M phase of division, and has been shown not to be a barrier to DNA nuclear entry (1). Consistent with this notion, Supplemental Figure S5*C* shows that arresting HeLa cells at the G_1/S border does indeed limit ($P<0.001$) the transfer of β-galactosidase reporter gene compared to that in normal dividing cells.

Supplemental Materials and Methods

Nuclear import and quantitation- Nuclear import into the digitonin-permeabilised HeLa cell assay using NLS-HSA-FITC (25 μg/ml) or fluorescent DNA (10 μg/ml of Oregon Green 488-/pCMVβ-DTS as import subtrate was performed as described previously (2, 3). HeLa cells were grown in DMEM media (supplemented with 10% FCS, 1 mM glucose, 5 mM L-glutamine, 100 IU/ml penicillin and 100 μg/ml streptomycin) to 60% confluence on 10-mm glass coverslips for at least 18 h before the assay. Each coverslip was placed briefly into a 35-mm dish containing 1 ml of wash buffer (consisting of 20 mM HEPES (pH 7.3), 100 mM potassium acetate, 2 mM magnesium acetate, 1 mM EGTA, 0.25 M sucrose) before permeabilisation. Each coverslip was then placed in 1 ml of permeabilisation buffer (consisting of wash buffer supplemented with 40 μg/ml digitonin, 2 mM DTT, and 1 μg/ml aprotinin, leupeptin and pepstatin) on ice. After 5 min, each coverslip was washed, rinsed for 5-10 min with 1 x 1 ml of ice-cold import buffer (consisting of wash buffer supplemented with 2 mM DTT). After draining the excess import buffer onto a paper towel, the assay was initiated by inverting washed coverslips onto 50 μl of import reaction mixture (consisting of import buffer supplemented with 1 μg/ml protease inhibitors, defatted bovine serum albumin (1 mg/ml), an ATP-regenerating system (2 mM ATP, 1 mM GTP, 10 mM creatine phosphate, 20 units/ml creatine phosphokinase, native purified exogenous HeLa cytoplasmic extract (5 mg/ml), and fluorescent import substrate) placed on a Parafilm-lined petri dish. Where indicated, 100 μg/ml wheat germ agglutinin (WGA) was added for 20 min at room temperature prior to initiation of the import assay. The cells were incubated in a humidified chamber at 37^{o}C for 6 h, after which reactions were terminated by washing the cells in wash buffer, fixed in 4% paraformaldehyde and mounted onto

microscope slides. Eight optical sections through the z-axis of the nuclei (~0.5-μm step size) on each field (randomly chosen from each of the 6 coverslips per treatment) were captured within a linear range of fluorescence intensity using a Leica TCS CLSM, equipped with an X40 objective (plan apo; NA = 1.25). Identical exposure times and laser output levels were used for all samples within an experiment. For presentation of representative images, medial sections were mounted by using Adobe Photoshop software. Nuclear import of fluorescent labelled pDNA was quantitated as follows: nuclei were carefully traced on the mid-section phase contrast images using MetaMorph imaging software (Universal Imaging, Molecular Devices, PA). These regions of interest (ROI) were then superimposed onto the corresponding mid-section fluorescent image and background subtracted to obtain nuclear fluorescent intensity. Viability of the nuclei after digitonin-permeabilised HeLa cell nuclei was routinely verified by monitoring exclusion of 70-kDa tetramethyl rhodamine dextran. Fluorescein conjugated human serum albumin (HSA-FITC; Sigma-Aldrich) was conjugated to NLS peptides as described previously (2). Both NLS-HSA-FITC and the Oregon Green/pCMVβ-DTS accumulate in the nuclei of permeabilised HeLa cells in the presence of exogenously added cytoplasmic extract in a WGA-inhibitable manner (Supplemental Fig. S1*A*,*B*).

PNA coupling to cyanogen bromide-activated Sepharose and column chromatography- Cyanogen bromide (CNBr)-activated Sepharose-4B (Amersham Pharmacia Biotech, UK) (1 g dry weight) was swollen in 1 mM HCl for 15 min then washed under vacuum on a sintered glass filter. PNA (1 nmol) was coupled to CNBr-activated Sepharose 4B in 1 ml coupling buffer (consisting of 0.1 M $NaCO_3$, pH 8.3 and 0.5 M NaCl) by overnight end-over-end mixing at 4^{o}C. Residual unreacted groups were blocked by incubation for 2 h with buffer containing 1 M Tris, pH 8.0, 0.5 M NaCl at room temperature. Non-covalently bound PNA was removed by washing alternatively with 0.1 M acetate, pH 4.0, 0.5 M NaCl, followed by coupling buffer (4 x 5 ml of each). The resultant PNA-Sepharose was washed and stored in hybridisation buffer at 4^{o}C until use. The wash solutions were pooled, concentrated using Centriprep YM-3 microconcentration devices (Millipore, UK), and the binding capacity of the CNBr-Sepharose for PNA estimated spectrophotometrically at 260 nm by quantification of potential unreacted PNA. No unbound PNA was detected. Plasmid DNA (200 μg) was hybridised to the Sepharose 4B conjugated PNA by incubation in TE (pH 6.5) buffer at 37^{o}C for 4 h followed by gentle end-over-end mixing at 4^{o}C overnight. The washed pCMVβ-DTS /PNA-Sepharose 4B media was then packed into a C10/10 column (Amersham Pharmacia Biotech, UK) and washed with 20 column volumes of buffer consisting of 20 mM HEPES (pH 7.3), 110 mM potassium acetate, 2 mM magnesium acetate, 1 mM EGTA, 2 mM dithiothreitol (DTT) and 1 mM zinc sulphate (column buffer). Affinity capture of pDNA-binding proteins from purified HeLa subcellular fractions were carried out at 4^{o}C under non-denaturing conditions in column buffer. DNA content in each of the elution fractions was measured by UV absorption at 260 nm and the protein concentration estimated by the method of Bradford using bovine serum albumin as the standard (4). Salt gradients were used

for elution and the concentration of salt in the collected fractions determined using a Corning PS 17 conductivity meter (Corning, USA). Conductivity of the diluted column fractions was corrected to NaCl concentration (M) by comparison to standard curve generated by diluting solutions of known NaCl concentrations made in column buffer with deionised water. The PNA-Sepharose was regenerated by washing with 50 ml TE buffer, pH 8 (containing 1 M NaCl). After equilibration with hybridisation buffer, the PNA-Sepharose media was re-charged with pCMVβ-DTS at 37^{o}C for 4 h followed by gentle end-over-end mixing at 4^{o}C overnight.

Two-dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis- 2-D SDS-PAGE was performed as described previously (5). Briefly, protein samples were solubilised in a buffer (9.5 M urea, 2% CHAPS, 1% DTT (Sigma, Poole, UK) 0.8% pharmalyte) containing Mini Complete Protease Inhibitor Cocktail (Roche Molecular Biomolecules, Lewes, UK; 1 tablet per 10 ml buffer). For analytical gels 100 μg of total protein, made up to 450 μl with rehydration solution (8 M urea, 0.5% CHAPS, 0.2% DTT and 0.2% Pharmalyte was applied to nonlinear broad- and narrow-range immobilized pH gradients (pH 3-10 and pH 6-9, 18 cm long, respectively) using an in-gel rehydration method. Proteins were focused using a Multiphor-II electrophoresis unit at 0.05-mA/IPG strip for 60 kVh at 20^{o}C. Prior to the second dimension separation, the IPG strips were equilibrated for 15 min in equilibration buffer (6 M urea, 30% glycerol (VWR, Poole, UK)) 2% SDS, 50 mM Tris-HCl (pH 8.8) containing 1% DTT). Strips were then re-equilibrated for 15 min in the same buffer containing 4.8% iodoacetamide (IAA) (Sigma, Poole, UK). For the second dimension, 12%T SDS polyacrylamide gels were used. After separation, protein detection was carried out by silver staining using the "Owl" silver staining kit (Insight Biotechnology, Wembley, UK). The gels were then scanned using a laser densitometer and the digital images were qualitatively analysed using ImageMaster 2D Elite (V3.1) software according to the following protocol: spot detection parameters were determined and applied to all the images; background subtraction (lowest value pixel on spot boundary) and vertical streaks were removed from each gel image and spots digitised to give a synthetic image; a reference gel image was then nominated and 'test' gel images matched to the reference gel. In order for ImageMaster to carry out automatic determination of M_r values, it was first necessary to calibrate the digitised images with a co-electrophoresed molecular weight marker ladder. The proteins of interest were located on preparatively loaded gels (400 μg; visualized by MS- compatible silver stain, "Plus-One Silver Stain Kit"), and excised for identification by mass spectrometry. All chemicals and equipment supplied by GE Healthcare, Amersham, unless stated otherwise.

Protein identification by mass spectrometry- In-gel reduction, alkylation and digestion with trypsin were performed prior to subsequent analysis by mass spectrometry. Cysteine residues were reduced with DTT and derivatised by treatment with iodoacetamide to form stable carbamidomethyl derivatives. Trypsin digestion was carried out overnight at room temperature after initial 2-h incubation at 37^{o}C. Initially the digested protein samples (4 μl) were desalted using ZipTip C18

microtips (Millipore). Peptides were eluted in 4 μl 50% acetonitrile / 0.1% trifluoroacetic acid. 0.5 μl was then loaded onto a target plate with 0.5 μl matrix (α-Cyano-4-hydroxy-cinnamic acid). Peptide mass fingerprints were acquired using a Voyager DE-PRO MALDI-TOF mass spectrometer (Applied Biosystems). The mass spectra were acquired in reflectron mode with delayed extraction. An autolytic tryptic peptide giving a molecular ion of m/z 2163.0569 was then used as an internal lock mass calibrant to achieve a mass accuracy of better than 50ppm. The identity of the protein(s) present in each sample was established by searching the Peptide Mass Fingerprinting dataset against the Swiss Prot and NCBI non-redundant databases using Mascot software (Matrix Science, UK). The data was searched using specific amino acid modification parameters, i.e. variable cysteine carbamidomethylation (resulting from reduction and alkylation reaction) and variable methionine oxidation.

In cases where it was not possible to identify the protein using this approach further analysis involving LC/MS/MS was undertaken. Here peptides were extracted from the gel pieces by a series of acetonitrile and aqueous washes. The extract was pooled with the initial supernatant and lyophilised. Each sample was then resuspended in 23 μl of 50 mM ammonium bicarbonate and analysed by LC/MS/MS. Chromatographic separations were performed using an Ultimate LC system (Dionex, UK). Peptides were resolved by reversed phase chromatography on a 75 μm C18 PepMap column. A gradient of acetonitrile in 0.05% formic acid was delivered to elute the peptides at a flow rate of 200 nl/min. Peptides were ionised by electrospray ionisation using a Z-spray source fitted to a QTof-micro (Waters Corp.). The instrument was set to run in automated switching mode, selecting precursor ions based on their intensity, for sequencing by collision-induced fragmentation. The MS/MS analyses were conducted using collision energy profiles that were chosen based on the m/z and the charge state of the peptide. Resulting data was pre-processed to create packing list (.pkl) files and these were searched against the Swiss Prot and NCBI non-redundant databases using Mascot software (Matrix Science, UK). The data was searched using the same modification parameters indicated previously.Transfection of growth-arrested HeLa cells with plasmid and quantitative analysis of β-galactosidase expression.

Following 24 hours serum starvation and a further 24 hours incubation in aphidicolin to block cells at the G_1/S border, HeLa cells were transiently transfected with DNA (1 μg) Lipofectamine 2000 (2.5 μl) complexes according to the manufacturer's instructions. Briefly, the blocking medium was removed and the transfection complex added. After 6 h, the culture medium was removed before the addition of Optimem media (supplemented with 10% FCS) and incubated overnight. The culture media was replaced with DMEM media (supplemented with 10% FCS, 1 mM glucose, 5 mM L-glutamine, 100 IU.ml penicillin and 100 μg/ml streptomycin) prior to β-galactosidase expression analysis. For the quantitative analysis of β-galactosidase expression, 0.2 ml of lysis buffer (100 mM Potassium phosphate, pH 7.8, and 0.2% Triton-X 100) was added to each well. The cell suspensions

were transferred to micro centrifuge tubes, rapidly freeze thawed three times and centrifuged. The lysates were transferred to micro centrifuge tubes and stored at −80 °C for use in subsequent assays. β-galactosidase activity was determined using a chemiluminescent β-galactosidase reporter system kit (Clontech, UK) using an Appliskan luminometer (Thermo Electron Corporation, MA, USA). Protein concentration was determined by a modified Folin–Lowry assay (4).

HeLa cell synchronisation, G_1/S block and cell cycle flow cytometric analysis- HeLa cells (1 x 10^5) were seeded in each of six replicate wells and cultured normally for 24 h in DMEM media (supplemented with 10% FCS, 1 mM glucose, 5 mM L-glutamine, 100 IU/ml penicillin and 100 μg/ml streptomycin). The cells were synchronised by serum-deprivation in serum-free DMEM media (Sigma, UK) for 24 h. To block cell growth at the G_1/S border, cells were then treated with 5 μg/ml aphidicolin in fresh DMEM medium and incubated for another 24 h. Flow cytometric analysis of cell cycle progression was done essentially as described by Al-Mohanna *et al*., 2001 (6). Data was acquired using Cell Quest Pro software (version 5.2.1; Becton Dickson, San Jose, CA) on a Becton Dickinson FACSCalibur flowcytometer. Analysis was performed offline using FlowJo software (Version 4.6.1; Treestar.com, Ashland, Oregon). Recombinant proteins. Recombinant NM23-H1 and NM23-H2 were expressed under lacZ induction in the Escherichia coli strain BL21 (DE3) using the vector strain pET3C (Novogen, USA) and purified using sequential ammonium sulphate precipitation, DEAE-Sephacel and hydroxylapatite chromatographic steps as described previously (7). To verify their viability, NM23-H1 and -H2 nucleoside diphophokinase (NDPK) enzyme activity was assayed spectrophotometrically as described previously (8) by following the formation of ADP from ATP in a coupled pyruvate-lactate dehydrogenase system containing dTDP, NADH, phosphoenol pyruvate, lactate dehydrogenase, and pyruvate kinase. The rate of NADH oxidation was followed by measuring the decrease in absorbance at 340 nm. One unit of NDPK activity was defined as the amount of activity required to oxidize 1 μmole of NADH/min/ml. NM23-H2 and NM23-H1 displayed activities of 712.5 and 604.0 units/mg protein and compare well with literature values for human erythrocyte NM23-H2 and NM23-H1 proteins (9). Assays were performed in quadruplicate. Escherichia coli expressed recombinant human proteins; his-tagged-importin-α, -β, Ran (Calbiochem, USA), histone H2B (Upstate Cell Signalling, USA) and Hsc70 (Stressgen, Canada) were obtained from commercial sources. Recombinant NM23-H1 and -H2 were labelled with Alexa Fluor 555 (Molecular Probes). Hsc70 conjugated with Cy5 using the Cy5 mono-Reactive Dye (GE Healthcare), in all cases following the manufacturers instructions.

DNA-binding assays- DNA binding to recombinant NM23-H2 and H1 protein was assessed essentially as detailed previously (7). Briefly, double-stranded oligonucleotide 5'-CTCCCCACCTTCCCCACCCTCCCCACCCTCCCCA-3' (10 nM) comprising the c-*myc* NHE -150 to -115 relative to the P1 transcription initiation site, with or without labelled biotin, was mixed with recombinant NM23 protein (16 nM) binding buffer (50 mM Tris-HCl, pH 7.9, 0.5 mM DTT, 2 mM

$MgCl_2$, 100 mM KCl and 50 μg/ml BSA. The entire reaction was loaded on a pre-run 5% native polyacrylamide gel in TBE (45 mM Tris-borate (pH 8.3), 1.25 mM EDTA). Gels were electrophoresed at 100 V for 3 h at 4°C in 0.5x TBE. After the electrophoresis, the gels were blotted on PVDF membranes and the bands were visualized using enhanced chemiluminescence (ECL) according to the manufacturer's instructions with the Pierce Lightshift chemiluminescence EMSA kit. Recombinant histone H2B protein DNA-binding was assessed as described previously (10). Briefly, 1 ng of biotin-labelled double-stranded 5'-GAT-CCG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ATC-3' oligonucleotide probe (11) was incubated with recombinant histone H2B protein (1-3 μg of total protein) in binding buffer (20 mM HEPES, pH 7.9, 150 mM KCl, 1 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 0.1 mg/ml bovine serum albumin, 5% glycerol, 100-200 ng of denatured *Hae*III digest of *E. coli* DNA). After incubation for 30 min at room temperature, the probes were loaded on a 6% polyacrylamide gel prepared in 0.5x TGE buffer (50 mM Tris base, 50 mM glycine, 2 mM EDTA). After electrophoresis, the gels were blotted on PVDF membranes and the bands visualized using ECL as above.

Supplemental References

1. Tseng, W.C., Haselton, F.R., and Giorgio, T.D. (1999) *Biochim Biophys Acta* **1445**(1), 53-64.
2. Wilson, G. L., Dean, B. S., Wang, G., and Dean, D. A. (1999) *J Biol Chem* **274**(31), 22025-22032.
3. Hillery, E., Munkonge, F. M., Xenariou, S., Dean, D. A., and Alton, E. W. F. W. (2006) *Anal Biochem* **352**(2), 169-175.
4. Peterson, G.L. (1977) *Anal Biochem* **83**(2), 346-56.
5. Weekes., J., Wheeler, C.H., Yan, J.X., Weil, J., Eschenhagen, T., Scholtysik, G., and Dunn MJ. (1999) Electrophoresis **20**(4-5), 898-906.
6. Al-Mohanna, M. A., Al-Khodairy, F. M., Krezolek, Z., Bertilsson, P. A., Al-Houssein, K. A., and Aboussekhra, A. (2001) *Carcinogenesis* **22**(4), 573-578.
7. Postel, E. (1999) *J Biol Chem* **274**(32), 22821-22829.
8. Postel, E.H., and Ferrone, C.A. (1994) *J Biol Chem* **269**(12) 8627-8630.
9. Gilles, A.M., Presecan, E., Vonica, A., and Lascu, I. (1991). *J Biol Chem* **266**(14), 8784-8789.
10. Zalensky, A.O., Siino, J.S., Gineitis, A.A., Zalenskaya, I.A., Tomilin, N.V., Yau, P., Bradbury, E.M. (2002) *J Biol Chem* **277**(45), 43474-43480.
11. Fraga, M.F., Ballestar, E., Esteller, M. (2003) *J Chromatogr B Analy. Technol Biomed Life Sci* **789**(2), 431-435.
12. Krummrei, U., Bang, R., Schmidtchen, R., Brune, K., and Bang, H. (1995) *FEBS Lett* **371**(1), 47-51.
13. Ryffel, B., Woerly, G., Greiner, B., Haendler, B., Mihatsch, M.J., and Foxwell, B.M. (1991) *Immunology* **72**(3), 399-404.
14. Wang, B.B., Hayenga, K.J., Payan, D.G., and Fisher, F.M. (1996) *Biochem J* 314(1), 313-319.
15. Vishwanatha JK, Jindal HK, and Davis RG. (1992) *J Cell Sci* **101**(1), 25-34.
16. Wang, W., Wang, L., Endoh, A., Hummelke, G., Hawks, C.L., and Hornsby, P.J. (2005) *J Endocrinol* **184**(1), 85-94.
17. Okuno, Y., Imamoto, N., and Yoneda, Y. (1993) *Exp Cell Res* **206**(1), 134-142.
18. Kim, Y.S., Ko, J., Kim, I.S., Jang, S.W., Sung, H.J., Lee, H.J., Lee, S.Y., Kim, Y., and Na, D.S. (2003) *Eur J Biochem* **270**(2), 4089-4094.
19. Schwamborn, K., Albig, W., and Doenecke, D. (1998) *Exp Cell Res* **244**(1), 206-217.
20. Bauer, A., Huber, O., and Kemler, R. (1998) *Proc Natl Acad Sci USA* **95**(25), 14787-14792.
21. Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007) *Nucleic Acids Res* 35(Web Server Issue) W585-W587.
22. Nair, R., and Rost, B. (2004) *Nucleic Acids Res* **32**(Web Server Issue), W517-W521.

23. Postel, E.H., Berberich, S.J., Rooney, R.W., and Kaetzel, D.M. (2000) *J Bioenerg Biomembr* **32**(3), 277-284.
24. Kraeft, S.K., Traincart, F., Mesnildrey, S., Bourdais, J., Véron, M., and Chen, L.B. (1996) *Exp Cell Res* **227**(1), 63-69.
25. Pinon, V.P., Millot, G., Munier, A., Vassy, J., Linares-Cruz, G., Capeau, J., Calvo, F., and Lacombe, M.L. (1992) *Exp Cell Res* **246**(2) 355-367.
26. Bryant, J.A., Brice, D.C., Fitchett, P.N, and Anderson, L.E. (2000) *J Exp Bot* **51**(352), 1945-1947.
27. Ghosh, A.K., Steele, R., and Ray, R.B. (1999) *Mol Cell Biol* **19**(4), 2880-2886.
28. Feo, S., Arcuri, D., Piddini, E., Passantino, R., and Giallongo, A. (2000) *FEBS Lett* **473**(1), 47-52.
29. Hirata, A., and Hirata, F. (1999) *Biochem Biophys Res Commun* **265**(1), 200-204.
30. Svendsen, P.C., and McGhee, J.D. (1995) *Development* **121**(5), 1253–1262.
31. Liu., I.S., Chen, J., Ploder, L., Vidgen, D., van der Kooy, D., Kalnins, V.I., and McInnes, R.R. (1994) *Neuron* **13**(2), 377–393.
32. Levine, E.M., Hitchcock, P.F., Glasgow, E., and Schechter, N. (1994) *J Comp Neurol* **348**(4), 596–606.
33. Ferda Percin, E., Ploder, L.A., Yu, J.J., Arici, K., Horsford, D.J., Rutherford, A., Bapat, B., Cox, D.W., Duncan, A.M., Kalnins, V.I., Kocak-Altintas, A., Sowden, J.C., Traboulsi, E., Sarfarazi, M., and McInnes, R.R. (2000) *Nat Genet* **25**(4), 397-401.
34. Levine EM, Hitchcock PF, Glasgow E, and Schechter N. (1994) *J. Comp. Neurol. J Comp Neurol* **348**(4), 596–606.
35. Nacheva, G.A., Guschin, D.Y., Preobrazhenskaya, O.V., Karpov, V.L., Ebralidse, K.K., and Mirzabekov, A.D. (1989) *Cell* **58**(1), 27-36.
36. Moreland, R.B., Langevin, G.L., Singer, R.H., Garcea, R.L., and Hereford, L.M. (1987) *Mol Cell Biol* **7**(11), 4048-4057.
37. Baake, M., Doenecke, D., and Albig, W. (2001) *J Cell Biochem* **81**(2), 333–346.
38. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997) *Nature* **389**(6648), 251-260.
39. Chan CK, Hübner S, Hu W, and Jans DA. (1998) *Gene Therapy* **5**(9), 1204-1212.
40. Pemberton, L.F., Rosenblum, J.S., and Blobel, G. (1999) *J Cell Biol* **145**:(7), 1407-1417.
41. Forwood, J.K., Harley, V., and Jans, D.A. (2001) *J Biol Chem* **276**(49), 46575-46582.
42. Lacasse, E.C., and Lefebvre, Y.A. (1995) *Nucleic Acids Res* **23**(10), 1647-1656.
43. Kanemaki, M., Makino, Y., Yoshida, T., Kishimoto, T., Koga, A., Yamamoto, K., Yamamoto, M., Moncollin, V., Egly, J.M., Muramatsu, M.,Tamura, T., and Kanemaki, M. (1997) *Mol Biophys Res Com* **235**(1), 64–68.

Supplemental Figure Legends

Supplemental Fig. S1. Fluorescein-labelled NLS-conjugated human serum albumin (NLS-HSA-FITC) and Oregon Green/pCMVβ-DTS nuclear import into digitonin-permeabilised HeLa cell assay. Digitonin-permeabilised HeLa cells were incubated at 37 °C with NLS-HSA-FITC or Oregon Green/pCMVβ-DTS for 30 min and 6 h, respectively. The reaction was initiated by addition of cytoplasmic extract in the absence (*top panels*) or presence (*bottom panels*) of WGA and an ATP-regenerating system. After incubation, cells were processed and visualised using a Leica TCS CLSM. Representative central section fluorescent images matched and superimposed with their respective phase contrast images are shown. Scale bar: 10 μm. Representative of 6 similar experiments.

Supplemental Fig. S2. Silver-stained pH 3-10/2-D SDS-PAGE gel images of native HeLa cytoplasmic extract (*A*), control PNA-Sepharose column (*B*) and pCMVβ-DTS/PNA-Sepharose column flow-through (100 μg protein each) (*C*) compared with pCMVβ-DTS/PNA-Sepharose eluate (20 μg proteins) (*D*). Proteins were separated on non-linear narrow basic pH 3-10 range IPG strips in the first dimension followed by second dimension resolution on 12% SDS-PAGE. Positions of molecular size markers (kilo Daltons) are shown on the left. Data are representative images of 3 gels run with samples from 3 independent experiments. The identities assigned to the protein spots indicated in the pH 3-10/SDS-PAGE elution gel are listed in Supplemental Table S1, entries numbers 19-49. (*E*) Quantification of protein spots on matched pH 3-10/2-D SDS-PAGE gel separation of native HeLa cytoplasmic extract, PNA-Sepharose control column, pCMVβ-DTS/PNA-Sepharose column flow-throughs and eluate concentrates.

Supplemental Fig. S3. Silver-stained pH 6-19/2-D SDS-PAGE gel images of native HeLa cytoplasmic extract (*A*), control PNA-Sepharose column (*B*) and (pCMVβ-DTS/PNA-Sepharose column flow-through (100 μg protein each) (*C*) compared with pCMVβ-DTS/PNA-Sepharose eluate (20 μg proteins) (*D*). Proteins were separated on non-linear narrow basic pH 6-9 range IPG strips in the first dimension followed by second dimension resolution on 12% SDS-PAGE. Positions of molecular size markers (kilo Daltons) are shown on the left. Data are representative images of 3 gels run with samples from 3 independent experiments. The identities assigned to the protein spots indicated in the pH 6-9/SDS-PAGE elution gels are listed in Table S1 (*E*) Quantification of protein spots on pH 6-19/2-D SDS-PAGE gel separation of native HeLa cytoplasmic extract, PNA-Sepharose control column, pCMVβ-DTS/PNA-Sepharose column flow-throughs and eluate concentrates.

Supplemental Fig. S4. Classification of identified proteins (*A*)**.** Distribution of identified proteins among different protein functional classes. Shown are percentages of pCMVβ-DTS-affinity purified proteins from various functional classes identified by means of mass spectrometry. Example of identification of human histone H2B using tandem mass spectrometry (MS/MS) (*B*)**.** The MS/MS fragmentation spectrum for Leu- Leu- Leu-Pro-Gly-Glu- Leu-Ala-Lys of M_r 952.6 Da derived from tryptic fragments of spot number 38 (Supplemental Figure S3*D*) with labelling of the major

fragmentation sites is shown. Another tryptic fragment was also identified (data not shown), and both are indicated. (*C*) Together both mapped peptides covered 19% of the polypeptide.

Supplemental Fig. S5. Flow cytometric analysis of HeLa cells blocked at the G_1/S border. Propidium iodide histograms indicating the fluorescence intensity of HeLa cell cultures grown in serum-free medium for 24 h, followed by further 24 h incubation in the absence (*A*) or presence (*B*) of 5 μg/ml aphidicolin. Results are representative of 3 independent experiments. (*C*) The effect of cell cycle phase on transfer of β-galactosidase gene in HeLa cells in culture. β -galactosidase expression (RLU/mg protein) of cells blocked at the G_1/S border (*dark bars*) was compared with control free-cycling cells (*grey bars*) forty eight hours after transfection with pCMVβ and pCMVβ-DTS. UT indicates β-galactosidase expression in negative control untransfected cells. Results are presented as means ± SEM (error bars) (N=6 independent experiments).

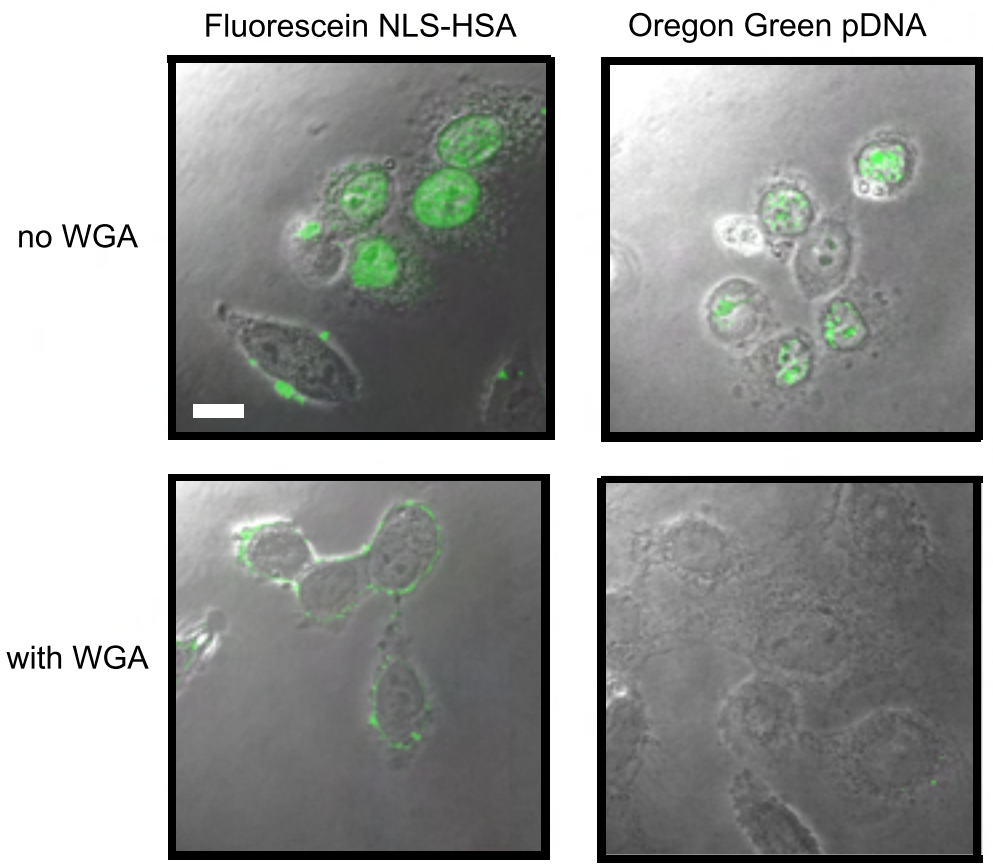
Supplemental Fig. S6. Coomassie blue stained SDS-PAGE of purified *E. coli*-expressed wild-type (WT) recombinant human NM23-H2, Alexa Fluor 555-labeled NM23-H2 and NM23-H1 from the final hydroxylapatite chromatographic purification step (*A*); wild-type and Alexa Fluor 555-labelled recombinant human histone H2B (*B*); wild-type (WT) and Cy5-labelled recombinant human Hsc70 (*C*); His-tagged Ran, importin-β and -α (*D*). Samples (5-10 μg protein each) were boiled in SDS-PAGE loading buffer and electrophoresed in 15% (*A, B, D*) or 10% (C) SDS-PAGE gels. Mobility of intact NM23-H2, NM23-H1, histone H2B, Hsc70, Ran, importin-α and -β proteins indicated by arrowheads correspond to molecular masses of ~17, ~18, ~70, ~15, ~70, ~30, ~97, and ~60 kDa, respectively. Positions of molecular size markers (kilo Daltons) are shown on the right. (*E*) EMSA showing the relative abilities of wild-type (WT) and fluorescent-labelled recombinant NM23-H2 proteins to bind to a double-stranded 34-bp biotin-labelled oligonucleotide comprising the *c-myc* NHE fragment. Alexa Fluor 555-labeled NM23-H2 or NM23-H1 proteins (16 nM) were incubated with the biotin-labelled 34-bp *c-myc* NHE oligonucleotide (10 nM), followed by electrophoresis on a 5% polyacrylamide gel. Arrows on the right side of the panel indicate complexed, free and cleaved c-myc NHE oligonucleotides. Lane 1 contains no recombinant protein. Lanes 2 and 4 contain the NM23-H2 and Alexa Fluor 555-labelled NM23-H2, respectively. In lane 3, recombinant NM23-H2 protein was incubated in the presence of a 100-fold molar excess of an unlabelled 34-bp *c-myc* NHE oligonucleotide. (*F*) EMSA showing histone H2B binding to double-stranded biotin-labelled 5'-GAT-CCG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ACG-ATC-3' oligonucleotide probe. All reactions contained biotin-labelled oligonucleotide probe. The reaction in lane 1 contains no recombinant histone H2B. Reaction shown in lane 3 was incubated in the presence of a 100-fold molar excess of an unlabeled oligonucleotide probe. Arrows on the right side of the panel indicate positions in lane 2 of the free DNA probe and the histone H2B protein/DNA complexes.

Supplemental Table Legends

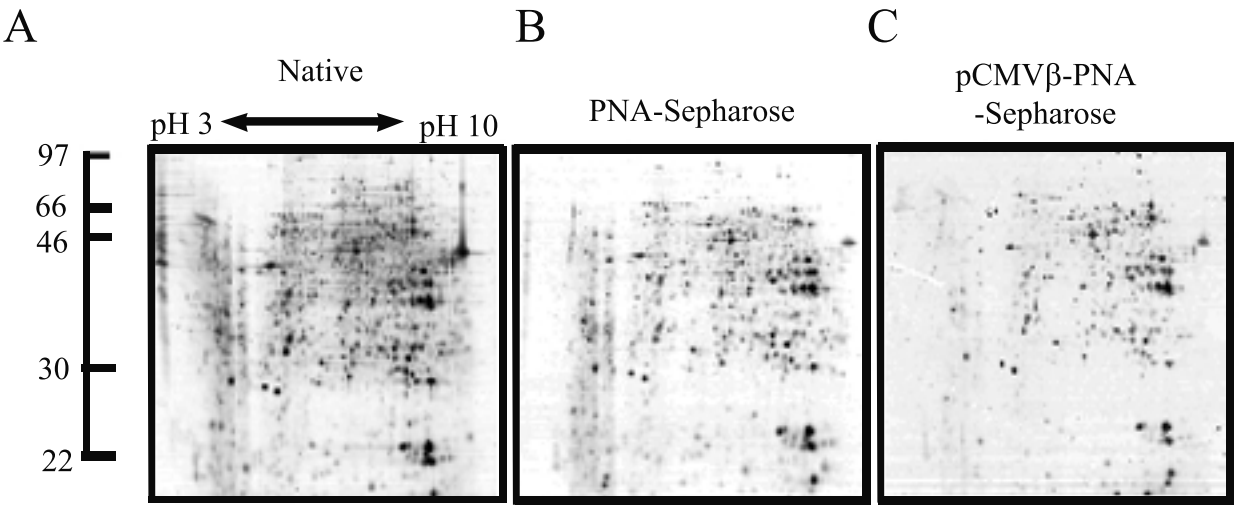
Supplemental Table S1. DNA-binding proteins with nuclear targeting capabilities identified in HeLa cytoplasmic extracts after affinity-purification on a pCMVβ-DTS/PNA-Sepharose column, and separation by two-dimensional gel electrophoresis, and mass spectrometry.

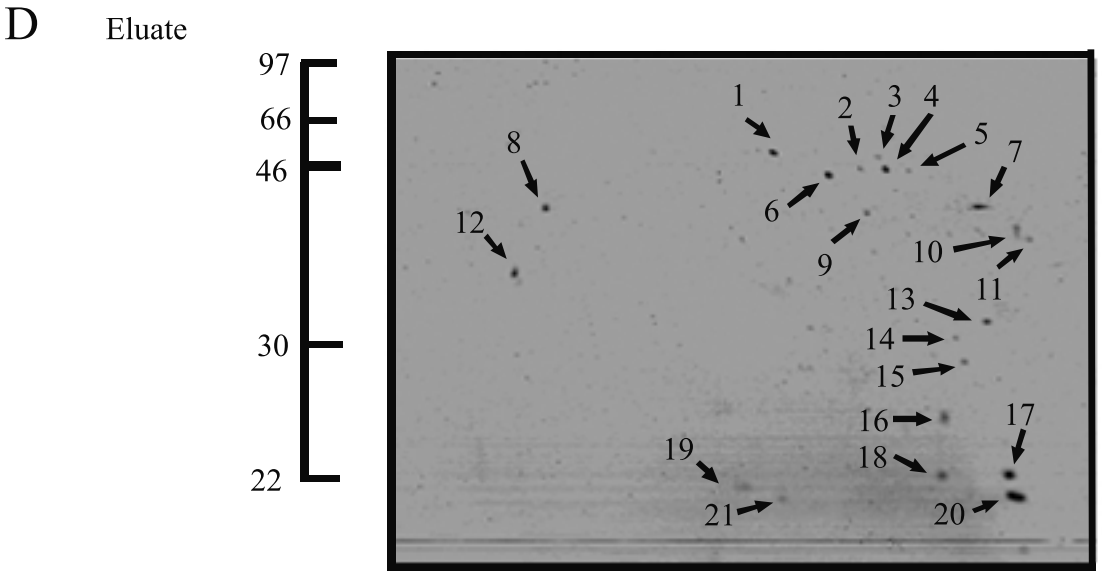
Supplemental Table S2. Combined results from the identification by MALDI-TOF and mass fingerprinting and LC/MS/MS of HeLa cytoplasmic extract after affinity-purification on a pCMVβ-DTS/PNA-Sepharose column separated by pH 3-10 (*A*) and pH 6-9 (*B*) two-dimensional gel electrophoresis.

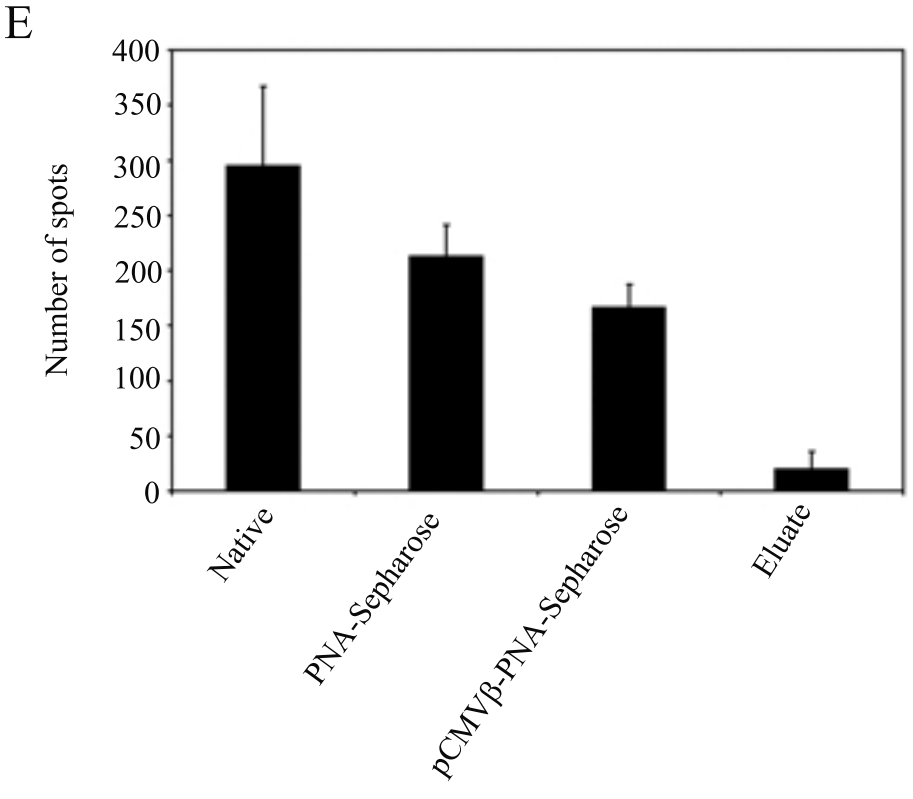
Supplemental Figure S1



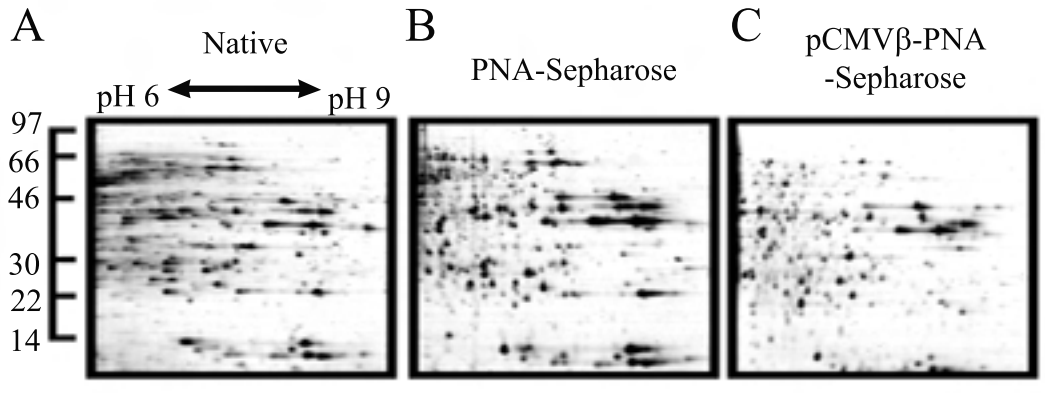
Supplemental Figure S2

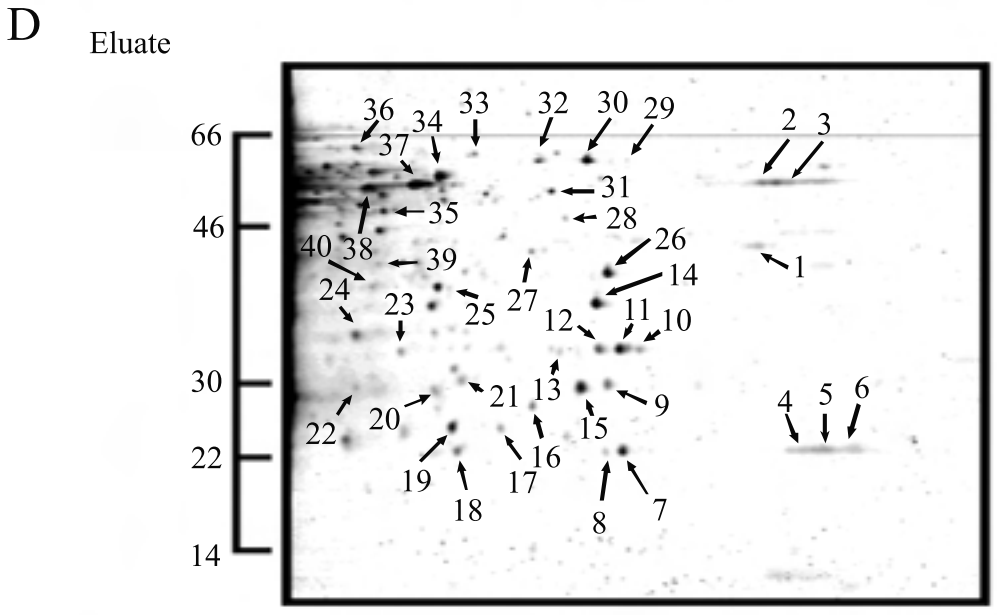


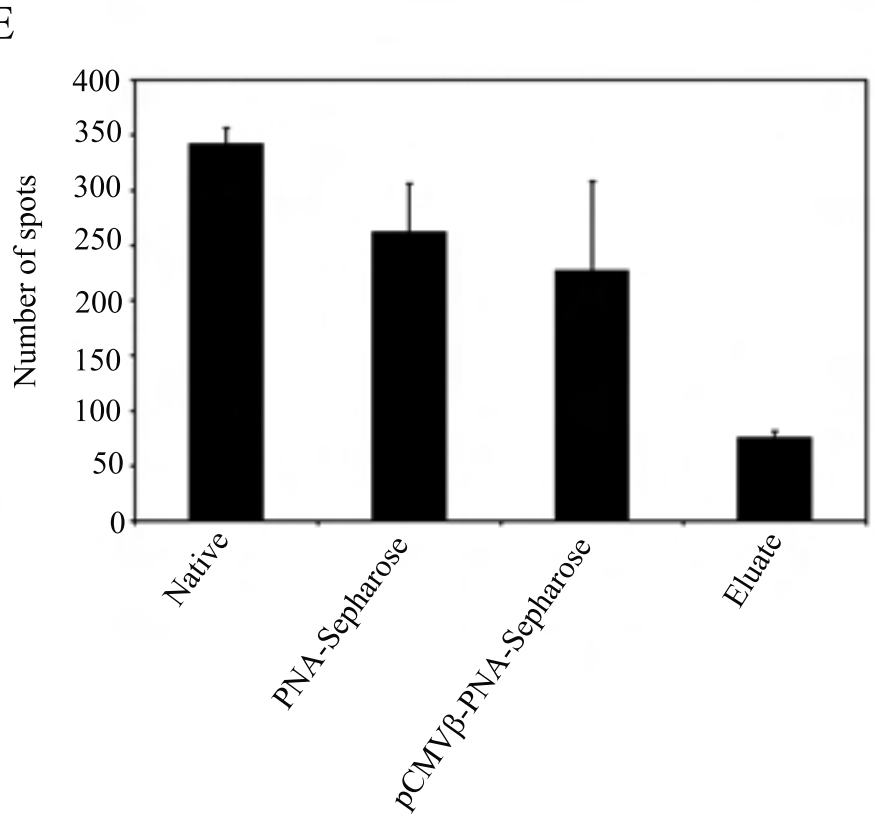




Supplemental Figure S3

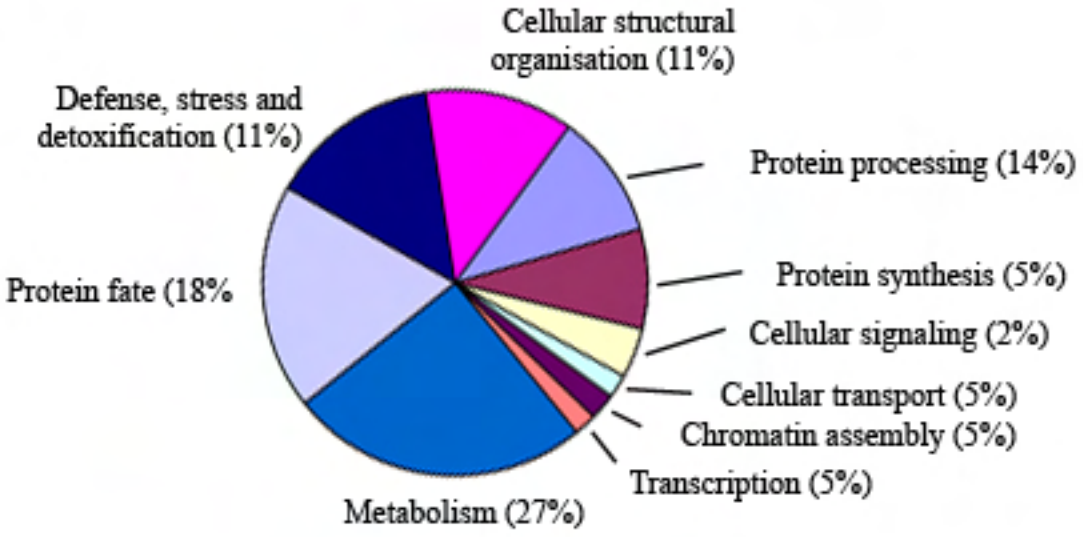




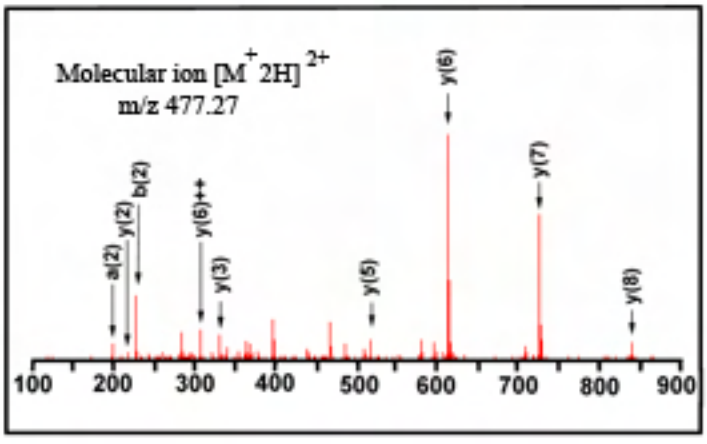


Supplemental Figure S4

A



B

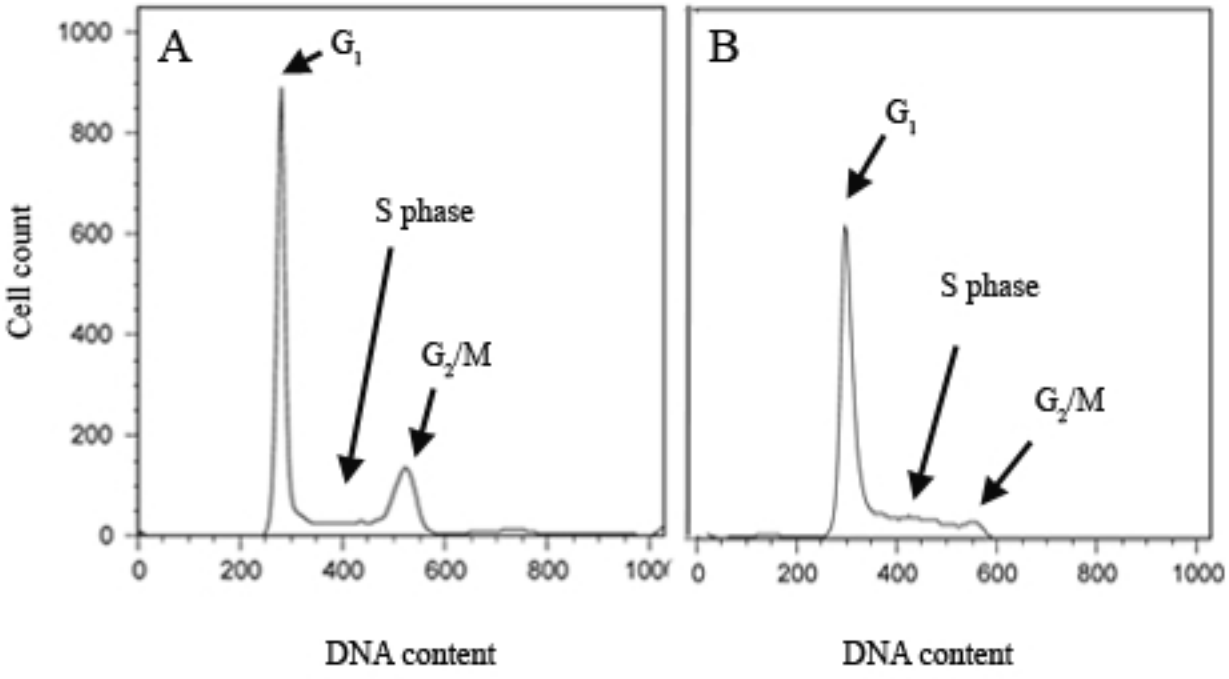


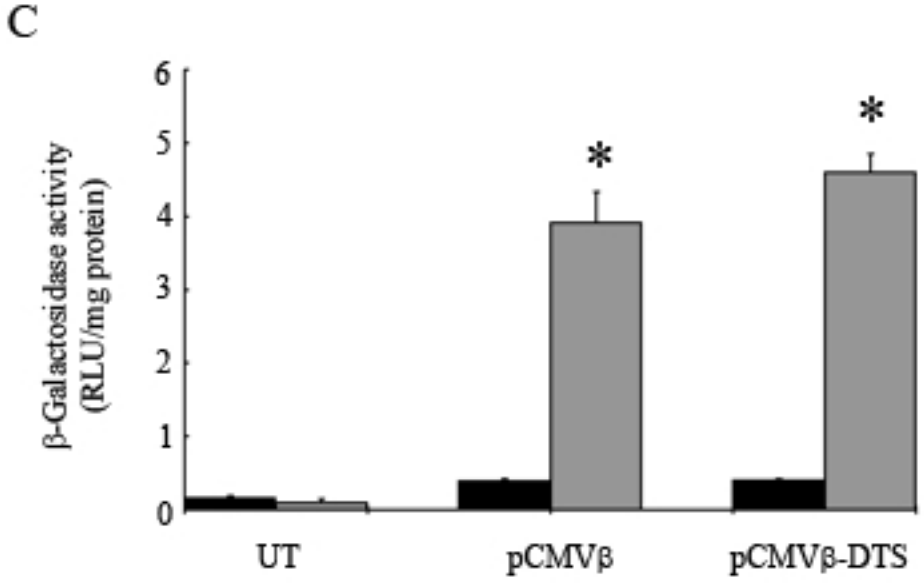
#	a	a++	b	b++	Seq.	y	y++	y*	y*++	#
1	86.10	43.55	114.09	57.55	L					9
2	199.18	100.09	227.18	114.09	L	840.52	420.76	823.49	412.25	8
3	312.27	156.64	340.26	170.63	L	727.44	364.22	710.41	355.71	7
4	409.32	205.16	437.31	219.16	P	614.35	307.68	597.32	299.17	6
5	466.34	233.67	494.33	247.67	G	517.30	259.15	500.27	250.64	5
6	595.38	298.19	623.38	312.19	E	460.28	230.64	443.25	222.13	4
7	708.47	354.74	736.46	368.73	L	331.23	166.12	314.21	157.61	3
8	779.50	390.26	807.50	404.25	A	218.15	109.58	201.12	101.07	2
9					K	147.11	74.06	130.09	65.55	1

C

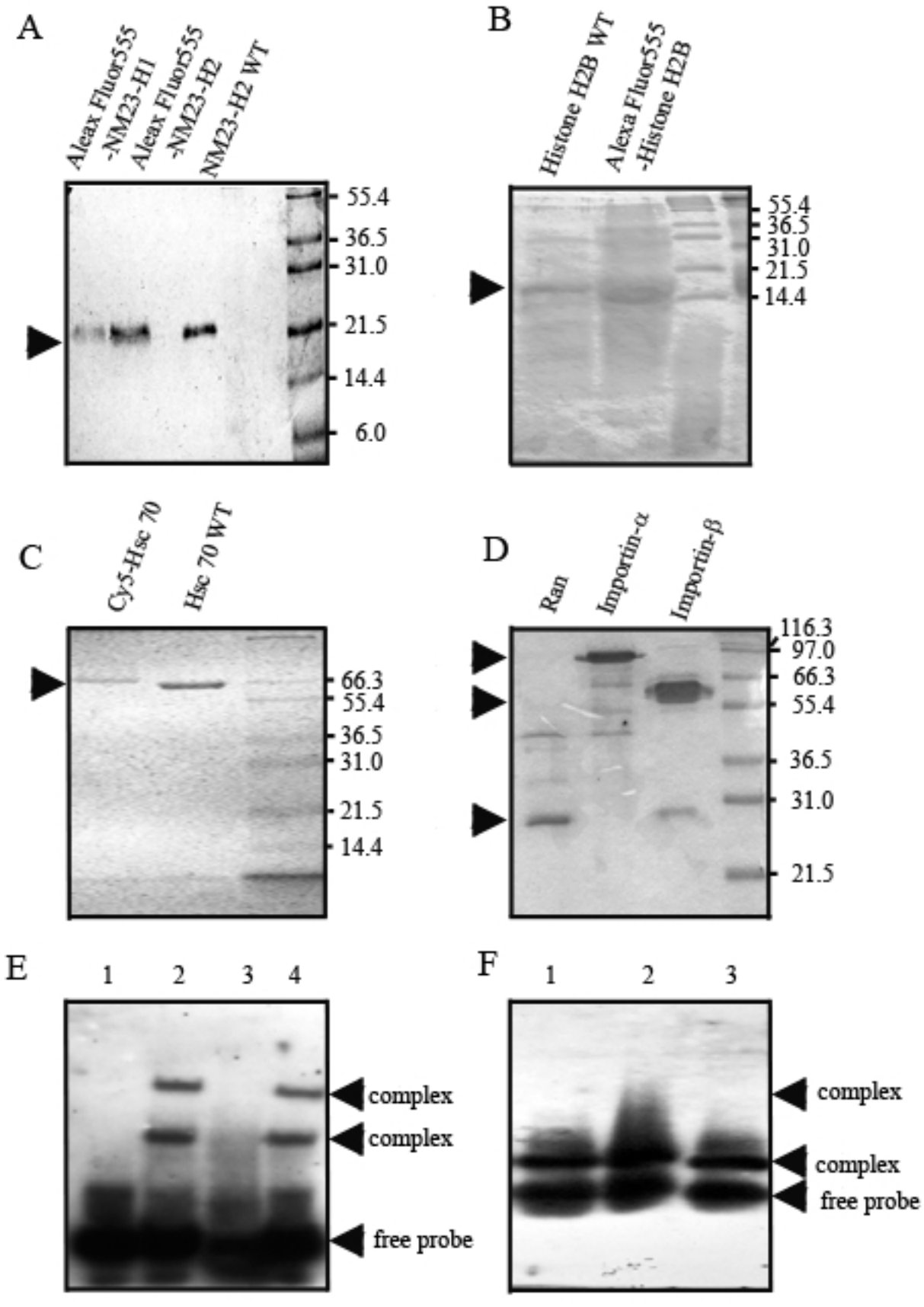
```
  1   PEPAKSAPAP KKGSKKAVTK AQKKDGRKRK RSRKESYSVY VYKVLKQVHP
 51   DTGISSKAMG IMNSFVNDIF ERIAGEASRL PHYNKRSTIT SREIQTAVRL
101   LLPGELAKHA VSEGTKAVTK YTSAK
```

Supplemental Figure S5





Supplemental Figure S6



Supplemental Table S1.
Combined results from the identification by MALDI-TOF and LC/MS/MS of HeLa cytoplasmic extract after affinity-purification on a pCMVβ-DTS/PNA-Sepharose column separated by pH 3-10 (A) and pH 6-9 (B) two-dimensional gel electrophoresis.

(A) pH 3-10 two-dimensional gel electrophoresis; MALDI-TOF

[a]Spot no.	[b]Protein identification	[b]Accession number	[c]Functional classification	[d]Matching peptides/ions	[d]Coverage (%)	[d]MB_{rB} (Da)	[d]p*I* (pH)	[c]Subcellular location
1	Cyclophilin A	P62937	Defense, stress and detoxification	9	74	17732	8.37	nuc/cyt (12,13,14)
2	Cofilin	P23528	Cellular structural organisation	9	65	18521	8.22	nuc/cyt
3	NM23-H2	P22392	Transcription	10	75	17850	6.92	nuc/cyt
4	Ubiquitin-conjugating enzyme E2	P68036	Protein fate	9	77	17850	8.68	nuc/cyt
5,7	Peroxiredoxin 1	Q06830	Defense, stress and detoxification	16,14	71,61	22095	8.27	cyt
6	Glutathione S-transferase P	P09211	Defense, stress and detoxification	9	52	23293	7.30	cyt
8	GTP-binding nuclear protein Ran	P62826	Cellular transport and transport mechanisms	16	50	24408	7.01	nuc/cyt
9,10	Triosephosphate isomerase	P50174	Metabolism	20,14	77,54	26773	6.51	cyt
11	Fructose-bisphosphate aldolase	P04075	Metabolism	18	67	39196	8.39	cyt
12,13	Phosphoglycerate kinase	P00558	DNA synthesis and processing/metabolism	12,15	50,55	44395	7.52	nuc/cyt (15)
14	Hypothetical protein FLJ10849	Q9NVA2	DNA synthesis and processing	19	47	49367	6.36	cyt
15,17	Fascin	Q61558	Cellular structural organisation	22,19	48,50	54398	6.81	cyt
16	Phosphoglycerate	O04375	Metabolism	21	48	56457	6.28	cyt

	dehydrogenase							
18	α-enolase	P06733	DNA synthesis and processing/metabolism	18	48	46955	6.16	nuc/cyt (16)
19	Heat shock cognate (Hsc) 70-kDa protein	P11142	Protein fate	21	36	70827	5.37	nuc/cyt (17)
20	Glutathione synthetase	P48637	Defense, stress and detoxification	24	55	52312	5.48	cyt
21	Fertility protein SP22 (DJ-1 protein)	PO0558	DNA synthesis and processing/metabolism	13	70	19961	6.32	nuc/cyt
22	Protein disulfide isomerase A6 precursor	Q15084	Defense, stress and detoxification	16	45	47191	4.95	cyt

(B) pH 6-9 two-dimensional gel electrophoresis; MALDI-TOF including identification by LC/MS/MS*

[a]Spot no.	[b]Protein identification	[b]Accession number	[c]Functional classification	[d]Matching peptides/ions	[d]Coverage (%)	[d]MB_{rB} (Da)	[d]p*I* (pH)	[c]Subcellular location
1	Phosphoglycerate kinase	P00558	DNA synthesis and processing/metabolism	18	59	44568	8.30	nuc/cyt (15)
2	*T-complex protein-1, eta-subunit (HIV-1 Nef interacting protein)	P87153	Protein fate	12	29	59842	7.55	nuc/cyt
3	T-complex protein-1, eta-subunit	Q99832	Protein fate	10	25	59842	7.55	nuc/cyt
4,5,6,7,8	Peroxiredoxin 1	Q06830	Defense, stress and detoxification	13,16,9,9,16	58,60,55,43,59	22096	8.27	cyt
9	Proteasome subunit alpha-type 4	P25789	Protein fate	14	55	29465	7.57	nuc/cyt
10,11	Annexin I	P04083	Cellular structural organisation	14,20	46,60	38674	6.64	nuc/cyt (18)

12	Guanine nucleotide-binding protein-beta subunit	P63244	Cellular signalling	16	58	35055	7.60	nuc/cyt
13,23	Elongation factor 2	P13639	Protein synthesis	10	15	95146	6.42	cyt
14	Fructose-bisphosphate aldolase A	P04075	Metabolism	10	28	39196	8.39	cyt
15	Homeobox protein Chx10	P58304	Transcription	21	21	39386	7.11	nuc/cyt
16,19	Triosephosphate isomerase	P50174	Metabolism	20,6	72,28	26522	6.51	cyt
17	GTP-binding nuclear protein Ran	P62826	Cellular transport and transport mechanisms	9	34	24408	7.01	nuc/cyt
18	*Proteasome subunit, alpha-type	P25787	Protein fate	8	39	25910	6.92	nuc/cyt(19)
20,22	Phosphoglycerate mutase	P18669	Metabolism	9,11	50,50	28655	6.75	cyt
21	*Tyrosyl-tRNA synthetase	P54577	Protein synthesis	8	14	59106	6.61	cyt
24	Aldose reductase	P15121	Metabolism	11	37	35643	6.28	cyt
25	Aldose reductase-related protein 2	O60218	Metabolism	8	30	35967	5.98	cyt
26	Phosphoserine aminotransferase	Q9Y617	Metabolism	17	42	40447	8.15	cyt
27	Aspartate aminotransferase	P17174	Metabolism	11	25	46168	6.30	cyt
28	Lissencephaly-1 (LIS-1)	P43034	Cellular structural organisation	13	38	46509	7.03	cyt
29,30, 32	Pyruvate kinase, M1 isozyme)	P14618	Metabolism	9,30	19,56	57806	7.40	cyt
31	Adenylyl cyclase-associated protein-1	Q01518	Cellular structural organisation	14	32	51542	7.16	cyt
33	Bifunctional purine biosynthesis protein	P31939	Metabolism	15	29	64117	6.30	cyt

34	*Heat shock protein 90-beta	P08238	Protein fate	9	17	72828	5.07	cyt
35	Fascin	Q16658	Cellular structural organisation	18	41	54348	6.81	cyt
36	Hsc70/Hsp90-organising protein	P31948	Cellular transport and transport mechanisms	19	28	62599	6.40	nuc/cyt
37	[*]Fascin	Q16658	Cellular structural organisation	27	52	54348	6.81	cyt
37	*TATA box-binding protein-interacting protein (TIP49)	Q9Y265	DNA synthesis and processing	17	49	50196	6.02	nuc/cyt(20)
37	*T-complex protein-1, beta-subunit	P78371	Protein fate	8	24	57452	6.01	nuc/cyt
37	*alpha-enolase	P06733	Transcription	4	15	46955	6.16	nuc/cyt
37	Phosphoglycerate dehydrogenase	O43175	Metabolism	15	31	56457	6.28	cyt
38	*Methylthioadenosine phosphorylase	L40432	Metabolism	6	28	31202	6.75	cyt
38	*Histone H2B	O60814	Chromatin assembly and DNA binding	2	19	13898	10.3	nuc/cyt(19)
39	*Mitotic checkpoint BUB3	O43684	DNA synthesis and processing	12	50	36931	6.36	cyt
39	*Cyclophilin D	Q08752	Protein fate	15	35	40738	6.77	cyt
40	*Similar to acetyl-Coenzyme A acetyltransferase 2	Q9BWD1	Metabolism	7	14	41352	6.47	cyt

A. Entries represent spots from a native HeLa cytoplasmic extract gel using coordinates from the corresponding pCMVβ-DTS/PNA-Sepharose column eluate silver stained-gel separated on an immobilised pH 3-10 non-linear gradient strip followed by separation on a 12% SDS-PAGE (supplemental Fig. S3*D*). Proteins were identified using MALDI-TOF mass fingerprinting.

B. Entries represent spots from a native HeLa cytoplasmic extract gel using coordinates from the corresponding pCMVβ-DTS/PNA-Sepharose column eluate silver stained-gel separated on an immobilised pH 6-9 non-linear gradient strip followed by separation on a 12% SDS-PAGE (supplemental Fig. S4*D*). Proteins were identified using MALDI-TOF mass fingerprinting.

[a]indicates spots from the pH 6-9/2-D SDS-PAGE that were refractory to identification by MALDI-TOF and therefore re-analysed by *de novo* peptide sequencing using LC/MS/MS.
[b]Spot no. corresponds to the position marked on the pH 3-10 (supplemental Fig. S3*D*) and pH6-9 2-D SDS-PAGE pCMVβ-DTS/PNA-Sepharose column eluate gels shown in supplemental Fig. S4*D*.
[c]Protein name and accession no. derived from the Swiss Prot or NCBI nonredundant database (limited to *Homo sapiens*).
[c]Information on function and subcellular location was obtained from indicated references and SwissProt or NCBI, WoLF PSORT (21), LOCtarget (22) prediction programs.
[d]Number of peptides indicates the number of analysed peptides that matched the identified protein by sequence and/or mass; their coverage of the entire amino acid sequence is listed under sequence coverage (%). Molecular mass/isoelectric point (Mr/p*I*) values were calculated from the amino acid sequence.

Supplementary Table S2. DNA-binding proteins with nuclear targeting capabilities identified in HeLa cytoplasmic extract after affinity-purification on a pCMVβ-DTS/PNA-Sepharose column, separation by two-dimensional gel electrophoresis, and protein identification by mass spectrometry

Protein names	DNA- binding capability[a]	Nuclear targeting capability[a]
Cyclophilin A	(12)	(13,14)
NM23-H2	(23)	(24,25)
Phosphoglycerate kinase	(26)	(15,26)
α-enolase	(16,27,28)	(16)
Annexin I	(18,29)	(15,18)
Homeobox protein Chx10	(30,31,32,33)	(31,32,33)
Histone H2B	(34)	(19,35,36)
49-kDa TBP-interacting protein	(37,38,39,40)	(20,41,42,43)

[a]Information on DNA-binding and nuclear targeting was obtained from indicated references, DNA Prot (21) and LOCtarget (22) prediction programs in this study.